Year : 2022

Structure-function analysis of disaggregating chaperones and their evolution across the Tree of Life - Public health, treatment controversy and meta-analyses on COVID-19

Rebeaud Mathieu Marius Jean Edouard

Rebeaud Mathieu Marius Jean Edouard, 2022, Structure-function analysis of disaggregating chaperones and their evolution across the Tree of Life - Public health, treatment controversy and meta-analyses on COVID-19

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive http://serval.unil.ch
Document URN : urn:nbn:ch:serval-BIB_A3314E4135DB6

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Département de Biologie Moléculaire Végétale (DBMV)

Structure-function analysis of disaggregating chaperones and their evolution across the Tree of Life - Public health, treatment controversy and meta-analyses on COVID-19

Thèse de doctorat ès sciences de la vie (PhD)

Présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

Par

Mathieu Marius Jean Edouard REBEAUD

Ingénieur en Biotechnologies – Master HES-SO en Biosciences Appliquées

Jury

Prof. Christophe Dessimoz, Président
Prof. Pierre Goloubinoff, Directeur de thèse
Prof. Caroline Samer, Experte
Prof. Anat Ben-Zvi, Experte
Prof. Abdussalam Azem, Expert

Lausanne
(2022)

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Département de Biologie Moléculaire Végétale (DBMV)

Structure-function analysis of disaggregating chaperones and their evolution across the Tree of Life - Public health, treatment controversy and meta-analyses on COVID-19

Thèse de doctorat ès sciences de la vie (PhD)

Présentée à la

Faculté de biologie et de médecine
de l'Université de Lausanne

Par

Mathieu Marius Jean Edouard REBEAUD

Ingénieur en Biotechnologies – Master HES-SO en Biosciences Appliquées

Jury

Prof. Christophe Dessimoz, Président
Prof. Pierre Goloubinoff, Directeur de thèse
Prof. Caroline Samer, Experte
Prof. Anat Ben-Zvi, Experte
Prof. Abdussalam Azem, Expert

Lausanne
(2022)

UNIL | Université de Lausanne
Faculté de biologie et de médecine

Ecole Doctorale
Doctorat ès sciences de la vie

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président·e	Monsieur	Prof.	Philippe	**Reymond**
Directeur·trice de thèse	Monsieur	Prof.	Pierre	**Goloubinoff**
Expert·e·s	Madame	Prof.	Caroline	**Samer**
	Monsieur	Prof.	Anat	**Ben-Zvi**
	Monsieur	Prof.	Abdussalam	**Azem**

le Conseil de Faculté autorise l'impression de la thèse de

Mathieu Rebeaud

Master - Msc en biosciences appliquées, HES-SO Valais, Suisse

intitulée

Structure-function analysis of disaggregating chaperones and their evolution across the Tree of Life - Public health, treatment controversy and meta-analyses on COVID-19

Lausanne, le 29 juillet 2022

pour le Doyen
de la Faculté de biologie et de médecine

Prof. Philippe Reymond

From so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved.
On the Origin of Species (1859)
Charles Darwin

What can change the nature of a man?
Planescape Torment

What a brilliant idea you had! Idiot, why did you not think of it earlier?
Pierre Goloubinoff

A Smile Better Suits a Hero…
Haurchefant Greystone

Stay awhile and listen!
Deckard Cain

The cake is a lie
Portal

To my parents and family.

À mes parents et ma famille.

Acknowledgments

I would like to thank Prof. Pierre Goloubinoff for having taken the risk to choose me to carry out a Ph.D. in his laboratory and for believing in me despite my atypical background of a “certified colocynth -planter”. I am fortunate to have been hosted by Pierre in his laboratory from August 2017 to July 2022, until his compulsory retirement in Switzerland.

I especially thank my parents and close siblings, who endured, tolerated, and assisted me during all these years, in particular until the age of 20, when I was confused about the direction my life should take. I know I could always count on you, whether it was to encourage me or to kick me in the behind when it was time to do so. Thank you, I love you.

I would like to warmly thank all my current and former colleagues of the DBMV, for the good atmosphere, the parties, the interesting meetings and discussions, and the good mood despite the pandemic in the last years. I would like to thank my colleagues of the last years in Pierre's laboratory, namely Anthony Guihur, John Perrin, Bruno Fauvet, Satyam Tiwari, Tatiana Fomekong Tsoutezo Mbefo, and Baptiste Bourgine. I enjoyed great scientific discussions, and received encouragement when needed, but also some good cold beers, invigorating wines, and delicious aperitifs. Apparently, we also published some pretty nice papers together.

I would like to thank my sweet caffeine, my painkillers, and other medications that helped me.

I acknowledge mistermv (Xavier Dang) who, through his funny and explosive personality, has managed to make me laugh and keep me going during the difficult times for more than 10 years with his funny streams and sometimes touching moments. Thank you “vieux monsieur”.

What would life (and the internet) be without cats? And what would I be without Gilgamesh and Mailikki, the two little Night Furies who come and sit on my keyboard and make mischief while I write this manuscript? We need cats to be good human beings.

Finally, I would like to thank, in no particular order, several people who have been with me for more or less a long time and who are either family, old friends, or new acquaintances: Kusa, Kit, Fabien, Trahreg, Dansi, Noémie, Marion, Ana, Eric, Florian, David, Christian, Laure, Michaël, Vincent, Pierre, several others and finally the juiced-up youngsters: Nathan, Thibault, Yahya, Matthieu, and Anthony.

Table of Contents

Abstract

For space constraints, this summary focuses only on the six core chapters addressing the main research theme of my thesis: The Structure-function analysis of disaggregating chaperones and their evolution across the Tree of Life. Protein homeostasis is regulated by molecular chaperones and proteases. The apparition of the main chaperone families can be traced at least 3 billion years ago. Previous publications showed evidence that HSP20 and HSP60 already existed in the last common universal ancestor. **Chapter 1** was written as an introduction to address the thermodynamic dilemma of cells, between chaperone-mediated protein repair and protease-mediated protein degradation. We found that the genomes of the simplest free-living bacteria and archaea, likely representing a snapshot of the very distant past, contained a full set of complex proteases, but only two out of the five conserved chaperone families (HSP20 and Hsp60). This indicates that at the beginning of life protein degradation likely preceded protein-repair mechanisms.

In **Chapter 2**, I sought to further address the evolutionary history of protein homeostasis, by analyzing proteomes and their chaperomes in the genomes of ~200 distant present-day living organisms, representing the major clades on the tree of life. We found that before eukaryotisation, the 5 main chaperone families (HSP20, 60, 70, 90, and 100) were already present in terrestrial bacteria and complex archaea, at least 2.5 billion years ago. To maintain effective protein homeostasis in the much more complex eukaryotic proteomes, co-chaperones needed to diversify, in particular the JDP family (HSP40), to recruit increasingly complex native and aggregated proteins to be processed by the HSP70 unfoldases. We found that HSP70 acts as a central hub, coordinating all other chaperone families in the network.

In **Chapter 3** we performed an *in vitro* study of the eukaryotic nucleotide exchange factor, SSE1 in yeast, and HSP110 in mammals. We created various mutants and observed that interdomain communication is required for its co-disaggregase action with SSA1 (HSP70). Interdomain communication is also required to suppress an otherwise intrinsically unleashed elevated ATPase activity of the HSP110 molecule.

Chapter 4 is a preprint. I created Swap mutants of the J-domain between the main yeast class A DnaJ (Ydj1) and class B DnaJ (Sis1). The aim was to investigate *in vitro* the ability of WT and the SWAP JDPs to prevent protein aggregation and to target and induce HSP70 (Ssa1) to unfold stably misfolded substrates, into natively refolded products. We found that the SWAP mutants were more efficient at preventing aggregation and refolding of misfolded substrates, albeit at a higher ATP cost, suggesting the presence of a stop-and-start mechanism in eukaryotic JDPs, that reduces futile ATP hydrolysis by Hsp70 in the absence of protein substrates.

Chapter 5: With global warming, we aimed here to review here the molecular mechanisms by which plants which are sessile organisms, can adapt and survive to heat stress, relying on specific heat-responsive ion channels as sensors, and the accumulation of thermoprotective metabolic compounds and molecular chaperones.
Chapter 6: is a forum article that draws a parallel between animal TRPVs and plants CNGCs. The two heat sensors CNGC2 for plants and TRPV1 for animals similarly respond to heat stress and have a very similar structure.

In addition, in parts two and three of this thesis, I included several published, peer-reviewed research and review articles, of which I am a co-author, on various pharmaco-medical, ethical and societal aspects of the COVID-19 pandemic.

Résumé

Pour des raisons d'espace, ce résumé se concentre uniquement sur les six chapitres principaux qui traitent du thème de recherche principal de ma thèse : l'analyse de la structure et fonction des chaperonnes désagrégatrices et leur évolution à travers l'Arbre de la Vie. L'homéostasie des protéines est régulée par les chaperonnes moléculaires et les protéases. L'apparition des principales familles de chaperonnes peut être retracée il y a au moins 3 milliards d'années. Des publications précédentes ont montré que HSP20 et HSP60 existaient déjà chez le dernier ancêtre universel commun. Le **Chapitre 1** a été rédigé comme une introduction pour aborder le dilemme thermodynamique des cellules, entre la réparation des protéines médiée par les chaperonnes et la dégradation des protéines médiée par les protéases. Nous avons découvert que les génomes des bactéries et des archées autonomes les plus simples, représentant probablement un instantané d'un passé très lointain, contenaient un ensemble complet de protéases complexes, mais seulement deux des cinq familles de chaperonnes conservées (HSP20 et Hsp60). Cela indique qu'au début de la vie, la dégradation des protéines a probablement précédé les mécanismes de réparation des protéines.

Dans le **Chapitre 2**, j'ai cherché à approfondir l'histoire évolutive de l'homéostasie des protéines, en analysant les protéomes et leurs chaperonnes chez près de 200 organismes vivants actuels éloignés, représentant les principaux clades de l'arbre de la vie. Nous avons découvert qu'avant l'eucaryotisation, il y a environ 2 milliards d'années, les 5 principales familles de chaperonnes (HSP20, 60, 70, 90 et 100) étaient déjà présentes chez les bactéries terrestres et les archées complexes, il y a au moins 2,5 milliards d'années. Afin de maintenir une homéostasie protéique efficace dans les protéomes eucaryotes beaucoup plus complexes, les co-chaperonnes ont dû se diversifier, en particulier la famille des JDP (HSP40), pour recruter des protéines natives et agrégées de plus en plus complexes à traiter par les dépliants HSP70. Nous avons découvert que la HSP70 agit comme le hub central des autres familles de chaperonnes.

Dans le **Chapitre 3**, nous avons réalisé une étude *in vitro* du facteur d'échange de nucléotides eucaryotes, SSE1 chez la levure, HSP110 chez les mammifères, et observé qu'une communication inter domaine est nécessaire pour son action de co-disaggregase avec SSA1 (HSP70). Nous avons créé divers mutants et remarqué que cette communication est nécessaire pour l'activité co-disaggregase de HSP70, et aussi pour supprimer l'activité ATPase intrinsèquement élevée exercée par la molécule HSP110.

Le **Chapitre 4** est un preprint. J'ai créé des mutants en échangeant (SWAP) le J-domain entre la principale DnaJ de classe A (Ydj1) et la principale DnaJ de classe B (Sis1). L'objectif était d'étudier *in vitro* la capacité des JDPs WT et SWAP à prévenir l'agrégation des protéines et à cibler et induire HSP70 (Ssa1) à déplier un substrat mal plié en produits nativement repliés. Nous avons constaté que

les mutants SWAP étaient plus efficaces pour prévenir l'agrégation et le repliement des substrats mal repliés, bien qu'à un coût plus élevé en ATP, ce qui suggère la présence d'un mécanisme stop-and-start dans les JDPs eucaryotes, qui réduit l'hydrolyse futile de l'ATP par Hsp70 en l'absence de substrats protéiques.

Chapitre 5 : Avec le réchauffement climatique, nous avons voulu ici passer en revue les mécanismes moléculaires par lesquels les plantes, qui sont des organismes sessiles, peuvent s'adapter et survivre au stress thermique, en s'appuyant en particulier sur les canaux ioniques thermosensibles comme capteurs, et sur l'accumulation de composés métaboliques thermoprotecteurs et de chaperonnes moléculaires.

Le **Chapitre 6** est un article de forum qui établit un parallèle entre les TRPV des animaux et les CNGC des plantes. Les deux capteurs de chaleur CNGC2 pour les plantes et TRPV1 pour les animaux répondent au stress thermique de manière similaire et ont une structure très proche.

En outre, en tant que deuxième et troisième partie de cette thèse, j'ai inclus un certain nombre d'articles de recherche et de revue publiés, évalués par les pairs et dont je suis co-auteur, sur divers aspects pharmaco-médicaux, éthiques et sociétaux à propos de la pandémie de COVID-19.

Introduction

Foreword

It is not simple to write an introduction to a 5-year research project of which two were under restrictive pandemic conditions. Consequently, the thesis has 3 distinct parts. Part 1, which is the original core of the project, addresses the biochemistry of chaperone mechanisms. Part 2, which is related to pharmaco-medical aspects of the Covid-19 pandemic, and Part 3 which addresses the problem of predatory journals giving an increased chance for wrong and fake science to unjustly earn credibility and gain a veneer of unmerited authority in the media and the lay audience.

In the following introduction, I summarize the different interconnected chapters and parts composing this amalgamated but coherent Ph.D. thesis. Most of this is now published in international peer-reviewed journals. Each chapter comprises a cover page describing my specific contribution, followed by the relevant published article.

Additional research projects have produced significant results that are not yet mature enough to be included in a thesis composed exclusively of published articles. One of the more mature unpublished research projects will be shortly described in the general discussion and perspectives.

Part 1 of the thesis on the Structure-function analysis of disaggregating chaperones and their evolution across the Tree of Life is the core of my work in Prof. Goloubinoff's laboratory.
Anfinsen et al., (*1*) demonstrated that under optimal *in vitro* conditions, such as low temperature and low protein concentrations, an unfolded polypeptide chain of ribonuclease can spontaneously reach its native three-dimensional structure without assistance from other macromolecules. In cells of aging mammals, however, various toxic protein aggregates can form and accumulate and become the primary cause for cell death, leading to tissue degeneration, as in the case of Huntington's and Parkinson's diseases (*2-4*). To maintain proper protein homeostasis in cells, all organisms house a network of molecular chaperones that can efficiently maintain, at least in youth, low levels of inactive and potentially harmful misfolded and aggregated proteins below their toxicity threshold. Some chaperones, such as the small HSPs, may only passively bind misfolding intermediates and thereby prevent protein aggregations. The small HSPs may also present the bound misfolded polypeptides to other classes of ATP-consuming chaperones for subsequent active unfolding and refolding (5). Thus, chaperones such as HSP70, HSP100, and HSP60 can use the energy of ATP hydrolysis to forcefully unfold toxic misfolded species and convert them into functional, harmless proteins or into protease targets for degradation (*6*). HSP90 is also suspected to act as an ATP-fueled unfolding nanomachine. Out of about 30'000 expressed ORFs in human or plant cells, 2'500 are the most abundant proteins that contribute 98% of the total protein mass of cells. Emphasizing the key role of the chaperone-

based protein homeostasis network, less than 150 ORFs encoding for conserved molecular chaperones and cochaperones in human cells may contribute between 5% and 10% of the total cellular protein mass (*7-9*). Yet, whereas in young mammals the unfolding/disaggregating chaperones are highly effective, for unclear reasons, beyond reproduction age, the protein quality control machinery gradually fails (*10, 11*), unleashing toxic protein aggregations, cell death, and tissue degeneration (*10, 11*) (Fig1).

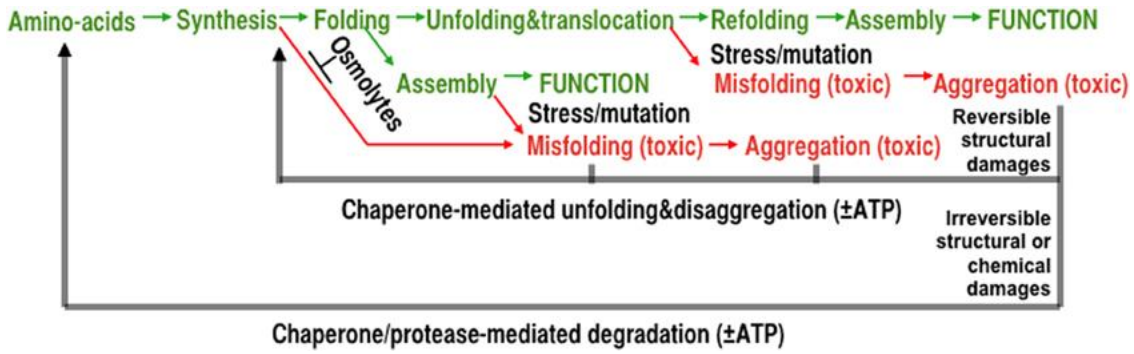


Figure 1 The various mechanisms of protein homeostasis. The maturation process of a polypeptide, from synthesis to the assembly of a native functional oligomer (green) is composed of delicate successive processes that can be affected by mutations, chemical modifications, and stresses such as heat shock. Under stressful conditions, polypeptides may transiently unfold and readily convert into misfolded and increasingly insoluble aggregates lacking specific biological activities. Misfolded conformers exposing highly interactive hydrophobic surfaces can also induce misfolding of other metastable polypeptides. Moreover, they can damage membranes, and in animal cells, cause apoptosis and tissue loss, aging, and degenerative diseases. Whereas molecular chaperones can prevent and actively repair structural damages in reversibly damaged aggregates, ATP-fueled proteases can unfold, degrade, and recycle irreversibly damaged protein conformers. Taken from (*12*).

The central aim of this part of the thesis is to use methods of bioinformatics, molecular biology, biochemistry and biophysics to gain detailed knowledge of the molecular mechanisms of two distinct disaggregating chaperone machineries (*13*) that co-exist in the cytosol of yeast and plants, but not of animals. Both systems are composed of an HSP70 core, acting as an ATP-fueled polypeptide-unfolding nanomachine. In one, HSP70 (SSA1, yeast chaperones in parentheses), together with HSP40 co-chaperones (SIS1, YDJ1), activates a ring-shaped complex belonging to the AAA^+ family (HSP104) that forcefully threads entangled polypeptide loops protruding from aggregates within its central cavity. It is generally believed, albeit without experimental evidences, that owing to the presence of HSP104, plants and yeast don't suffer from degenerative protein misfolding diseases like aging metazoans that lack HSP104. In the other disaggregating system, the HSP70 core, together with HSP40 co-chaperones, closely collaborates with an HSP70-cognate called HSP110 (SSE1) at solubilizing aggregates. Despite strong sequence and structure homologies between HSP70 and HSP110, indicating that the HSP110s stemmed out from an ancestor that was acting as today's HSP70s, the HSP110s are generally thought to serve in the cytosol and the ER of eukaryotes as a mere nucleotide exchange factor (NEF) of the

HSP70s (SSA1), i.e. that they transiently bind HSP70s and by doing so, accelerate ADP- and polypeptide substrate release from HSP70s. HSP110s (SSE1) are thought to have lost during evolution their original ability to act on misfolded polypeptides as a *bona fide* ATP-fueled HSP70-like unfoldases chaperones. Aside from SSE1, there are in yeast several additional candidate cochaperones called Snl1 and Fes1 (as well as the M-domain of Hsp104) that reportedly act as nucleotide exchange factores (NEFs) onto the cytosolic HSP70s (*14, 15*). Given the general principle in evolution "use it or lose it", the question arises: What is so particular about the NEF activity of SSE1 that it strictly maintain during over two billion years of eukaryote evolution, a seemingly unnecessary ability to hydrolyze ATP, whereas Snl1 or Fes1 can also drive nucleotide exchange in SSA1 upon binding to the same region in the nucleotide-binding domain (NBD) of SSA1, without having to bind and hydrolyze ATP by themselves? Even more intriguingly, what is so particular about the NEF activity of SSE1 that has so strictly maintained the seemingly unnecessary ability to bind protein substrates, whereas Snl1 or Fes1 can drive nucleotide exchange in SSA1 without having to bind aggregated protein substrates by themselves? We hypothesized that in addition to its HSP70-specific NEF activity, SSE1 may have kept some sort of ATP-dependent ability to process (unfold) misfolded polypeptides by itself remnant from its HSP70 ancestor.

Chapter 1 serves as a general introduction to the thesis. It is a review article on the thermodynamic dilemma of cells, between the need to repair or to degrade structurally compromised proteins. Cells have evolved molecular chaperones, such as HSP60, HSP70, and HSP100, to actively counteract the misfolding of proteins under out-of-equilibrium stressful conditions. Chaperone-resistant aggregates and misfolded proteins can then become forcefully unfolded and degraded by endo-cellular chaperones and proteases, and by chaperone-mediated autophagy in eukaryotes. Evolution has solved the conundrum between "protein repair" and "protein degradation" by creating a multitude of chaperones and chaperone-gated proteases and by regulating their quantities and rates of action.

This chapter was published in Front. Mol. Biosci. doi: 10.3389/fmolb.2021.768888

Chapter 2 presents a joint computational biology analysis between proteome expansion and the evolution of chaperones and cochaperones across the tree of life. In this paper, Dr. Saurav Malik from the laboratory of late prof Dan S. Tawfik at the Weizmann Institute, and I analyzed nearly 200 genomes of bacteria, archaea, and eukaryotes distantly distributed across all phyla of life, to determine when various chaperone families and their co-chaperones emerged and determine whether there is a likely correlation between genome size, multidomain protein complexity, and the presence of the said chaperones. Our systematic analysis gave an order of magnitude to several parameters: the number of

proteins present in the genome of the simplest archaea to the most complex mammals expanded 200-fold. Multidomain proteins increased 50-fold. During evolution, the number of protein folds (protein folds are islands of discrete structural similarity in which structures share a certain level of sequence similarity) has expanded 5-fold and the combination of these folds is 20-fold. Repeat proteins and beta-rich proteins that are particularly prone to misfold, and aggregation proliferated by a factor of ~600. Proteins predicted to be prone to aggregation became 6-fold more frequent in mammalian proteomes compared to bacterial proteomes. We also found that the core chaperones (HSP20, 60, 70, 90, and 100) were already present in simple bacteria and archaea, a billion years before eukaryotes emerged from the endosymbiosis of archaea with a *Pseudomonadota*. To reconcile the expanding proteomes and proteins, when increasing their cellular abundance did not suffice, core chaperones combined with increasingly diversified cochaperones to form a highly cooperative network in which folding, misfolding, and aggregating intermediates optimally bounce from one chaperone to another, until becoming functional native proteins or being degraded by chaperone-gated proteases.

This chapter was published in PNAS doi.org/10.1073/pnas.2020885118

Chapter 3 is an experimental work focused on the yeast (eukaryotic) nucleotide exchange factor HSP110 (Sse1) and HSP70 (Ssa1) and addresses the role of interdomain communication in HSP110 in particular. Sse1 maintains an open-lid substrate-binding domain (SBD) in close contact with its nucleotide-binding domain (NBD), independent of its ATP hydrolysis status. To better appreciate this comprehensive ATP hydrolysis-independent interaction between two domains of Hsp110s, NBD-SBD chimeras were constructed between Hsp110 (Sse1) and Hsp70 (Ssa1). In the Sse1/Ssa1 chimeras, we observed the uncoupling of two domains leading to a complete loss of the NEF activity of Sse1. Interestingly, the chimeric proteins showed a higher ATPase level of Sse1-NBD, compared to the wild-type protein, implying that in WT SSE1, the intrinsic ATPase activity of the protein is being maintained and repressed, purposefully. In addition to the repression of the strong ATPase activity of its NBD, the interactions between the two domains confer thermal stability to Sse1 and play a critical role in the chaperone function of Sse1 in Ssa1-mediated disaggregation activity. Overall, Sse1 exhibits a unique interdomain interaction, which is essential for its NEF activity, suppression of high intrinsic ATPase activity, co-chaperoning activity in the disaggregation mechanism, and protein stability.

This chapter was published in FEBS doi:10.1111/febs.15045

Chapter 4 is an experimental paper, in which we measured the ATP hydrolysis, protein-binding, protein disaggregation, and refolding activities of the main yeast cytosolic HSP70, SSA1, in the

presence of its two main JDPs, SIS1 (SS), YDJ1 (YY) and two J-domain swap mutants (referred to as SY and YS). Enzymatic activities, polypeptide binding capacity, and limited proteolysis, coupled with mass spectrometry revealed that SY and YS mutants maximally triggered ATP hydrolysis of HSP70, both without and in the presence of a misfolded substrate. In contrast, SS and YY minimally triggered ATP hydrolysis in the absence of a misfolded substrate, suggesting that wild-type JDPs can achieve a self-repressive conformation (a stop-and-start system) in which the J-domain reversibly interacts with the unoccupied substrate-binding region on their own CTD, rather than with HSP70 ATP.

This chapter was included in the thesis as a preprint.

Chapter 5 is a review where we focus on how plants feel the heat and adapt using, among others, heat-accumulated molecular chaperones to survive. Currently, with the growing human population and global warming becoming unavoidable, solutions to preserve plants and means of agriculture need to be found. Plants are sessile organisms that cannot seek the cover of shade on hot summer days. In this paper, we are addressing the mechanisms by which plant cells can sense rising temperatures and timely establish effective molecular defenses, such as molecular chaperones and thermoprotective metabolites, to survive noxious diurnal variations in temperatures and seasonal heat waves. Achieving agricultural productivity under climate change calls for a close examination of the molecular mechanisms of heat-stress resistance in model and crop plants.

This chapter was published in Trends in Biochemical Sciences doi.org/10.1016/j.tibs.2022.05.004

Chapter 6 is a forum article that draws the parallels and differences between heat-sensors: animals TRPVs and plants CNGCs. Land plant CNGC2 and vertebrate TRPV1 similarly respond to heat stress and have a similar structure. The 2021 Nobel Prize in Medicine or Physiology was given to David Julius for his discovery of the temperature-sensitive TRPV1 ion channel in vertebrates. Like CNGC2/4 in plants, TRPV1 is a membrane-embedded Calcium channel that plays a key role in the perception and regulation of temperature modulations. The awarding of a Nobel Prize demonstrates the importance given today to a field of research that has become essential in the context of the global warming crisis. Knowledge beyond metazoans is necessary to understand how crop plants, perceive and adapt to sudden changes in temperature.

This chapter was published in Trends in Plant Science doi.org/10.1016/j.tplants.2022.03.006

Part 2 of the thesis resulted and gained momentum, in terms of published articles, from the COVID-19 pandemic (*16*). In February 2020, anticipating a lockdown, Pierre Goloubinoff advised the members of his laboratory to redirect part of their efforts to address questions of public health related to the pandemic. As of March 2020, and with the lockdown in Switzerland, I took Pierre at his word by taking an interest in COVID-19, in addition to my Protein Biochemistry work. Chapters 7-12 are the product of Pierre's recommendation and are related to the Covid-19 pandemic. Chapter 7 is however closely related to the research theme of the Goloubinoff laboratory.

Chapter 7 is intertwined with the first part, concerning fever-induced proteins, HSP70 in particular. In this exploratory review, a potential mechanism to protect the respiratory system by allowing several moderate fever cycles is presented. Mortality in patients with COVID-19 occurs predominantly after ARDS (acute respiratory distress syndrome) during which programmed cell death of alveolar cells in the lungs is involved. Previous studies of the Goloubinoff laboratory with Prof Yoram Weiss from the Hadassah medical school showed that mortality in a sepsis-induced ARDS rat model is reduced by adenovirus over-expression of the HSP70 chaperone. A natural rise in body temperature during mild fever can naturally accumulate high cellular levels of HSP70 that can arrest apoptosis and protect alveolar lung cells from inflammatory damage. However, beyond 1–2 h of fever, no HSP70 is being further produced and a decrease in body temperature is required to the restore cell's ability to produce more HSP70 in a subsequent fever cycle. In this study, we suggest that antipyretics may be beneficial in COVID-19 patients after several hours of mild advantageous fever, allowing lung cells to accumulate protective HSP70 against damages from an inflammatory response. With age, the ability to develop fever and accumulate HSP70 decreases. This could be ameliorated, when advised to do so, by thermotherapies and/or physical training.

This chapter was published in Front. Med. Doi: 10.3389/fmed.2020.564170

Chapter 8 is an article that was written at the beginning of the pandemic. It addressed the issue of anti-hypertension treatment in the case of COVID-19. In this opinion piece, we hypothesized that the reductions in Angiotensin-Converting Enzyme 2 (ACE-2) observed in cases of hypertension and obesity may explain many of the abnormalities observed in SARS-CoV-2 and question the role of treatments interfering with ACE2. ACE inhibitors (ACEi) and AT1R blockers (ARB) are two classes of drugs that are widely used in medicine to treat hypertension or heart failure. ACEi and ARB upregulate ACE2 expression on the cell surface, and ACE2 activity is not prevented by ACEi. Observing that the downregulation of ACE2 induced by viral binding, leads to increased stimulation of the AT1R, we

concluded that this may be an important element in explaining severe COVID-19. Overall, the ACE2 increase mediated by ACE/ARBs is not deleterious and may even be protective. We also recall the need to conduct robust clinical trials to provide valid answers to this question. At the time of the article's release early in the pandemic, and based on the bibliographic research we made, stopping anti-hypertension treatment because of presumed considerations alone did not appear to be a recommended option, later confirmed by other studies such as (*17*).

This chapter was published in Front. Cardiovasc. Med. doi: 10.3389/fcvm.2020.00071

Chapter 9 is an article about SARS-CoV-2 and the use of (Hydroxy)Chloroquine as an antiviral treatment. Here, we researched the published literature cases of viruses for which chloroquine or hydroxychloroquine have been considered possible antivirals. The medical history is one of the unfulfilled hopes for potentially beneficial drugs having shown, at best, no efficacy or associated with increased adverse effects. The failure to translate in organisms, and the prior in vitro success of CQ on the Chikungunya virus, is another reminder that careful clinical evaluation is needed. At the time of the writing of this article, there was no published data to support the use of (H)CQ in COVID19. As in Chapter 8, we recalled the need to conduct well-designed large-scale clinical trials, unbiased, randomized, and well-controlled, with the most valuable and less subjective Endpoints to establish the safety and efficacy of quinine derivatives like (Hydroxy)Chloroquine as antiviral treatments.

This chapter was published in Front. Med. doi: 10.3389/fmed.2020.00184

Chapter 10 is a systematic review and a meta-analysis including 186 studies carefully chosen based on their proper controls that included a total of 210'447 deaths among 1'304'587 patients with COVID-19. In this study, we calculated the absolute risk of COVID-19 death, and it was found to be increased by 14%, 11%, 12%, and 7% for diabetes, hypertension, obesity, and smoking, respectively. We performed several subgroup analyses and a non-linear dose-response relation between body mass index and COVID-19 mortality. The proportion of deaths accountable to diabetes, hypertension, obesity, and smoking was 8%, 7%, 11%, and 2%, respectively. Our findings suggest that diabetes, hypertension, obesity, and smoking were associated with statistically significantly higher COVID-19 mortality, contributing to nearly 30% of COVID-19 deaths.

This chapter was published in BMJ Open doi:10.1136/bmjopen-2021-052777

Chapter 11 is a systematic review and meta-analysis on the effect of hydroxychloroquine with or without azithromycin on the mortality of COVID-19 where we included 11'932 participants for the hydroxychloroquine group, 8'081 for the hydroxychloroquine with azithromycin group and 12'930 for the control group. We found that hydroxychloroquine was not significantly associated with mortality in COVID-19: pooled relative risk RR=0.83 (95% CI: 0.65-1.06, n=17 studies) for all studies and RR=1.09 (95% CI: 0.97-1.24, n=3 studies) for RCT. Hydroxychloroquine with azithromycin was found to be significantly associated with increased mortality: RR=1.27 (95% CI: 1.04-1.54, n=7 studies). An alternative Bayesian meta-analysis approach produced similar results. This was one of the first meta-analyses clearly showing that HCQ was not a useful treatment in the case of COVID-19.

This chapter was published in Clinical Microbiology and Infection https://doi.org/10.1016/j.cmi.2020.08.022

Chapter 12 is a letter that relates directly to Chapter 11 in which we summarized the campaign of harassment that followed the publication of our peer-reviewed meta-analysis on hydroxychloroquine. We aimed at warning the scientific community that a campaign of harassment in the media has one goal: to scare and silence honest researchers and medical doctors and deter them from challenging the validity of a scientifically unverified general belief held by a majority of scientifically untrained individuals. Yet, silence is the worse response to aggression. It renders societies vulnerable to populism and obscurantism. Citizens deserve transparent and honest medical information. I consider it is central to challenge and improve the quality of the medical communication aimed at a lay audience, to restore its confidence in science-based medicine and well-carried out experimentally established science, as opposed to scientifically unverified so-called alternative medicines based on personal beliefs and convictions.

This chapter was published in Lancet Infectious Diseases doi.org/10.1016/S1473-3099(20)30866-5

Part 3 of the thesis was written as a natural consequence of the Covid-19 crisis: the need to denounce predatory e-journals. The pandemic rapidly highlighted the need for the scientific community to better control the quality of published scientific papers by rigorous peer-reviewing, not only before, but also after publication. There is an increasing need to encourage reviewers to become identified, to deter approval of wrong science by non-experts.

The COVID-19 pandemic produced an overload of publications that included low-quality preprints. It became usual to read in the media promises of care or COVID treatment, based on plainly flawed methodologies, or statistically insignificant conclusions, based on scarce data (*18*).

Although most preprint servers are quick to indicate that the hosted papers are not yet peer-reviewed, journalists, politicians and the public generally don't make the difference with papers approved by bona fide recognized experts.

In itself, an online publication of preprint on servers is a great advance of open science (*19*). This normally allows a publication to be discussed before and during peer-review by other members of the scientific community and sometimes this improves the quality of the paper, validates the research, and promotes the scientific efforts of the authors. Yet, during the pandemic, merely indicating that the hosted papers are not yet peer-reviewed did not deter a minority of boisterous iniquitous scientists to claim plain recognition and fame and didn't dissuade the media from promoting fake news and from raising hopes of treatments in the lay audience and among populist politicians. Moreover, as a consequence, serious scientists who in response needed to criticize the methods and attempted to tame the public hopes for miraculous treatments often became victims of threats by conspiracy adepts.

During the pandemic, a fair number of preprints, which were put forward by journalists, possibly acting in good faith but often lacking scientific background, often promoted unverified methods of care or COVID-19 treatment. Some promoted cases that were outright disinformation, without counterbalancing or putting into perspective the state of the art at the time of the publication.

When some publications that turned out to be of poor quality were eventually retracted, this lead to a decreased trust in public science (*20*) and gave birth to "scientific populism" (*21, 22*).

To restore the public's trust in the scientific process, it is the responsibility of trained scientists to ensure that science communication and dissemination are carried out openly and rigorously (*23*).

During the pandemic, many scientists that were critical of bad science became the target of hate messages, life threats, and harassment online, by email, or via mail to their hierarchy and the directions of their institutions (*24, 25*). Harassment affects everyday life. In extreme persistent cases, this can lead to burnout and the total withdrawal of a scientist from public appearances and avoiding taking an open stand on particularly sensitive issues (*26*). During this thesis, I participated in many online exchanges on the social network Twitter. Whereas this rendered me knowledgeable of the work of colleagues and of different research laboratories, my open vocal pro-science stand on sensitive issues which often questioned the values of particular clinical treatments for Covid-19 exposed me to life threats smear campaigns on Twitter, YouTube, and other media (*21, 27*). The chapters presented in the third part of this thesis (Chapters 13-15) are related to the theme of predatory journals (*28*). I describe a method I used to entrap a predatory journal and demonstrate its nonsense (*29, 30*). The abuses

suffered by some honest scientific curators, that like me encountered severe online harassment following the publication of our meta-analysis on HCQ (Chapter 11), were reported in chapter 12.

Chapter 13 was an invited letter to the editor, explaining a year later, the purpose of a hoax paper (Chapter 14), that we wrote aiming to demonstrate the lack of scientific probity in predatory journals in general. We were asked by the editor to explain why we wrote this hoax article, why the Asian Journal of Medicine and Health was chosen, and what was the consequent press coverage of this exercise. Here we focused on the French-speaking press. It appears that the hoax helped to raise public awareness of 'fake' or plain predatory scientific journals. We suggest that scientists should not regularly engage in such hoaxes, as repetitions would lose their "shock value" and undermine public confidence in scientific publications.

This chapter was published in Therapies doi.org/10.1016/j.therap.2021.10.009

Chapter 14 is the hoax paper. It was generated to raise the public's awareness of wrong science being published in predatory journals. During the pandemic many fanciful or irrelevant statements on COVID-19 cures have been made by a minority of injudicious, careless scientists and by policymakers. Among these, were claims that COVID-19 would kill fewer people than push-scooter accidents and that the idea of a second wave would be unwise, and so on (see https://www.youtube.com/watch?v=00_vy-f22nE). Facing the massive publications of very low-quality studies, with flawed methodological and lacking statistical valid values, I decided with friends to demonstrate the problem of the predatory journals and chose to challenge the quality of peer reviewing by the Asian Journal of Medicine and Health a journal, because it has been chosen by a group of French doctors and politicians, to promote COVID-19 treatment with hydroxychloroquine (*31*). As a consequence also of promotions by the presidents of the USA and Brazil, treatment of COVID-19 symptoms with hydroxychloroquine became prevalent in France (see https://www.youtube.com/watch?v=Pc04cjGDFRw) and even in Switzerland, as described chronologically in the editorial (*21*), we decided to recuperate several quotes that were said and published to make the most idiotic, false, statistically invalid article that intentionally lacked serious references, to demonstrate that the pseudo-scientific veneer offered by this type of journal doesn't fool anyone if one is minimally attentive. We aimed to demonstrate the inconsistency of the peer-review process that is carried out by this kind of predatory journal and be critical of the risky promotion of treatments that have not been rigorously validated and be wary of the scientific veneer and arguments of authority that some want to boast about, to put forward their beliefs. Because of the high risk to my security, I chose to use a preposterous pen name as the first author of this politically loaded

publication. The story was also picked up by Retraction Watch who interviewed me about our intentions regarding this hoax paper. (https://retractionwatch.com/2020/08/16/hydroxychloroquine-push-scooters-and-covid-19-a-journal-gets-stung-and-swiftly-retracts).

This chapter was published (and retracted) as an intentional hoax in AJMAH doi: 10.9734/AJMAH/2020/v18i930232

Chapter 15 is an open letter asking to protect the academy's whistleblowers and advocate for post-publication peer review. This letter, which was written to take a stand in support of Dr. Elisabeth Bik in the face of threats of legal action by the IHU Marseille, was picked up in Nature news to highlight the chilling effect on post-publication peer-review that such legal actions could have (*32*). In this letter, we wrote that as scientists, we believe it is important to be free to address scientific questions in an unbiased manner, free of threats of legal actions from opponents. This article serves as a conclusion to this doctoral thesis. Although it is less on the experimental side, it is still centrally important for science in general. Indeed, I think that every scientist must stay in close touch with the lay audience and ensure that it receives, unavoidably simplified, yet still unbiased information. Scientists should seek to share knowledge and advances outside their field of research and abandon the idea of remaining inaccessible in their comfortable ivory towers. Scientists should never be allowed by their peers to present partial lacunary scientific data and use predatory or low-quality Journals, on TV sets or diffuse their lacunary science on YouTube, without being thoroughly challenged by expert peers. Finally, it is worth noting the importance of people like Dr. Bik, who performs post-publication peer-review, to ensure that already published articles do not contain clumsy errors or outright fraud, thereby restoring the confidence of the lay audience in science.

This chapter was deposited as a preprint on OSF. doi: 10.31219/osf.io/2awsv

References

1. C. B. Anfinsen, Principles that govern the folding of protein chains. *Science* **181**, 223-230 (1973).
2. M. P. Hinault, A. Ben-Zvi, P. Goloubinoff, Chaperones and proteases: cellular fold-controlling factors of proteins in neurodegenerative diseases and aging. *J Mol Neurosci* **30**, 249-265 (2006).
3. M. P. Hinault, A. Farina-Henriquez-Cuendet, P. Goloubinoff, Molecular chaperones and associated cellular clearance mechanisms against toxic protein conformers in Parkinson's disease. *Neurodegener Dis* **8**, 397-412 (2011).
4. C. Soto, S. Pritzkow, Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases. *Nature Neuroscience* **21**, 1332-1340 (2018).
5. L. Veinger, S. Diamant, J. Buchner, P. Goloubinoff, The small heat-shock protein IbpB from Escherichia coli stabilizes stress-denatured proteins for subsequent refolding by a multichaperone network. *J Biol Chem* **273**, 11032-11037 (1998).

6. P. Goloubinoff, A. S. Sassi, B. Fauvet, A. Barducci, P. De Los Rios, Chaperones convert the energy from ATP into the nonequilibrium stabilization of native proteins. *Nat Chem Biol* **14**, 388-395 (2018).
7. A. Finka, P. Goloubinoff, Proteomic data from human cell cultures refine mechanisms of chaperone-mediated protein homeostasis. *Cell Stress Chaperones* **18**, 591-605 (2013).
8. N. Shemesh *et al.*, The landscape of molecular chaperones across human tissues reveals a layered architecture of core and variable chaperones. *Nature Communications* **12**, 2180 (2021).
9. M. E. Rebeaud, S. Mallik, P. Goloubinoff, D. S. Tawfik, On the evolution of chaperones and cochaperones and the expansion of proteomes across the Tree of Life. *Proceedings of the National Academy of Sciences* **118**, (2021).
10. N. Shai, N. Shemesh, A. Ben-Zvi, Remodeling of Proteostasis Upon Transition to Adulthood is Linked to Reproduction Onset. *Curr Genomics* **15**, 122-129 (2014).
11. M. Santra, K. A. Dill, A. M. R. d. Graff, Proteostasis collapse is a driver of cell aging and death. *Proceedings of the National Academy of Sciences* **116**, 22173-22178 (2019).
12. P. Goloubinoff, Recent and future grand challenges in protein folding, misfolding, and degradation. *Front Mol Biosci* **1**, 1 (2014).
13. H. Saibil, Chaperone machines for protein folding, unfolding and disaggregation. *Nat Rev Mol Cell Biol* **14**, 630-642 (2013).
14. N. K. Gowda, G. Kandasamy, M. S. Froehlich, R. J. Dohmen, C. Andreasson, Hsp70 nucleotide exchange factor Fes1 is essential for ubiquitin-dependent degradation of misfolded cytosolic proteins. *Proc Natl Acad Sci U S A* **110**, 5975-5980 (2013).
15. J. L. Abrams, J. Verghese, P. A. Gibney, K. A. Morano, Hierarchical functional specificity of cytosolic heat shock protein 70 (Hsp70) nucleotide exchange factors in yeast. *J Biol Chem* **289**, 13155-13167 (2014).
16. C. Huang *et al.*, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497-506 (2020).
17. J. B. Cohen *et al.*, Continuation versus discontinuation of renin-angiotensin system inhibitors in patients admitted to hospital with COVID-19: a prospective, randomised, open-label trial. *The Lancet Respiratory Medicine* **9**, 275-284 (2021).
18. C. Vlasschaert, J. M. Topf, S. Hiremath, Proliferation of Papers and Preprints During the Coronavirus Disease 2019 Pandemic: Progress or Problems With Peer Review? *Adv Chronic Kidney Dis* **27**, 418-426 (2020).
19. L. Besancon *et al.*, Open science saves lives: lessons from the COVID-19 pandemic. *BMC Med Res Methodol* **21**, 117 (2021).
20. L. Besancon, E. Bik, J. Heathers, G. Meyerowitz-Katz, Correction of scientific literature: Too little, too late! *PLoS Biol* **20**, e3001572 (2022).
21. S. Gould, S. L. Norris, Contested effects and chaotic policies: the 2020 story of (hydroxy) chloroquine for treating COVID-19. *Cochrane Database Syst Rev* **3**, ED000151 (2021).
22. G. Lasco, V. G. Yu, Pharmaceutical messianism and the COVID-19 pandemic. *Soc Sci Med* **292**, 114567 (2022).
23. K. Kupferschmidt, On the trail of bullshit. *Science* **375**, 1334-1337 (2022).
24. C. O'Grady, In the line of fire. *Science* **375**, 1338-1343 (2022).
25. C. Samer, K. Lacombe, A. Calmy, Cyber harassment of female scientists will not be the new norm. *Lancet Infect Dis* **21**, 457-458 (2021).
26. E. Ektorp, Death threats after a trial on chloroquine for COVID-19. *Lancet Infect Dis* **20**, 661 (2020).
27. P. Kory, G. U. Meduri, J. Iglesias, J. Varon, P. E. Marik, RETRACTED: Clinical and Scientific Rationale for the "MATH+" Hospital Treatment Protocol for COVID-19. *Journal of Intensive Care Medicine* **36**, 135-156 (2021).
28. A. Grudniewicz *et al.*, Predatory journals: no definition, no defence. *Nature* **576**, 210-212 (2019).

29. A. de La Blanchardiere, F. Barde, N. Peiffer-Smadja, H. Maisonneuve, [Predatory journals: A real threat for medical research. 1. Identify these journals and understand how they work]. *Rev Med Interne* **42**, 421-426 (2021).
30. A. de La Blanchardiere, F. Barde, N. Peiffer-Smadja, H. Maisonneuve, [Predatory journals: A real threat for medical research. 2 Assess their consequences and initiate a response]. *Rev Med Interne* **42**, 427-433 (2021).
31. V. Guérin *et al.*, Azithromycin and Hydroxychloroquine Accelerate Recovery of Outpatients with Mild/Moderate COVID-19 *Asian Journal of Medicine and Health*, (2020).
32. H. Else, Scientific image sleuth faces legal action for criticizing research papers. *Nature* **594**, 17-18 (2021).

PART1

Protein Biochemistry

Structure-function analysis of disaggregating chaperones and their evolution across the Tree of Life

CHAPTER 1

An introduction to the thesis

Repair or Degrade: the Thermodynamic Dilemma of Cellular Protein Quality-Control

Repair or Degrade: the Thermodynamic Dilemma of Cellular Protein Quality-Control

Bruno Fauvet†, Mathieu E. Rebeaud†, Satyam Tiwari, Paolo De Los Rios and Pierre Goloubinoff

†*Authors contributed equally to this work*

I contributed to this work conceptually and in the manuscript writing.

AUTHOR CONTRIBUTIONS
BF, MR, ST, PD, and PG wrote the article; BF, MR, PD, and PG prepared the figures.

frontiers in Molecular Biosciences

REVIEW
published: 27 October 2021
doi: 10.3389/fmolb.2021.768888

Repair or Degrade: the Thermodynamic Dilemma of Cellular Protein Quality-Control

Bruno Fauvet[1†], *Mathieu E. Rebeaud*[2†], *Satyam Tiwari*[2], *Paolo De Los Rios*[1,3] *and Pierre Goloubinoff*[2*]

[1]Institute of Physics, School of Basic Sciences, École Polytechnique Fédérale de Lausanne—EPFL, Lausanne, Switzerland, [2]Department of Plant Molecular Biology, Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland, [3]Institute of Bioengineering, School of Life Sciences, École Polytechnique Fédérale de Lausanne—EPFL, Lausanne, Switzerland

Life is a non-equilibrium phenomenon. Owing to their high free energy content, the macromolecules of life tend to spontaneously react with ambient oxygen and water and turn into more stable inorganic molecules. A similar thermodynamic picture applies to the complex shapes of proteins: While a polypeptide is emerging unfolded from the ribosome, it may spontaneously acquire secondary structures and collapse into its functional native conformation. The spontaneity of this process is evidence that the free energy of the unstructured state is higher than that of the structured native state. Yet, under stress or because of mutations, complex polypeptides may fail to reach their native conformation and form instead thermodynamically stable aggregates devoid of biological activity. Cells have evolved molecular chaperones to actively counteract the misfolding of stress-labile proteins dictated by equilibrium thermodynamics. HSP60, HSP70 and HSP100 can inject energy from ATP hydrolysis into the forceful unfolding of stable misfolded structures in proteins and convert them into unstable intermediates that can collapse into the native state, even under conditions inauspicious for that state. Aggregates and misfolded proteins may also be forcefully unfolded and degraded by chaperone-gated endo-cellular proteases, and in eukaryotes also by chaperone-mediated autophagy, paving the way for their replacement by new, unaltered functional proteins. The greater energy cost of degrading and replacing a polypeptide, with respect to the cost of its chaperone-mediated repair represents a thermodynamic dilemma: some easily repairable proteins are better to be processed by chaperones, while it can be wasteful to uselessly try recover overly compromised molecules, which should instead be degraded and replaced. Evolution has solved this conundrum by creating a host of unfolding chaperones and degradation machines and by tuning their cellular amounts and activity rates.

Keywords: proteostasis, thermodynamics, protein repair, chaperones, protein degradation

OPEN ACCESS

Edited by:
Eileen M. Lafer, The University of Texas Health Science Center at San Antonio, United States

Reviewed by:
Giorgio Colombo, University of Pavia, Italy
Peter Chien, University of Massachusetts Amherst, United States

***Correspondence:**
Pierre Goloubinoff
Pierre.Goloubinoff@unil.ch

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Protein Folding, Misfolding and Degradation, a section of the journal Frontiers in Molecular Biosciences

Received: *01 September 2021*
Accepted: *13 October 2021*
Published: *27 October 2021*

Citation:
Fauvet B, Rebeaud ME, Tiwari S, De Los Rios P and Goloubinoff P (2021) Repair or Degrade: the Thermodynamic Dilemma of Cellular Protein Quality-Control. Front. Mol. Biosci. 8:768888. doi: 10.3389/fmolb.2021.768888

INTRODUCTION

Protein Folding, Misfolding and Aggregation

On a free energy landscape, an ensemble of amino acids can be found in different energy states. As testified by the spontaneous hydrolysis of polypeptides in the presence of trypsin, polymerized amino acids are higher in free energy than when unpolymerized. Moreover, a single polypeptide chain can be found in different structural states: unfolded, which is in most cases inactive, natively-folded,

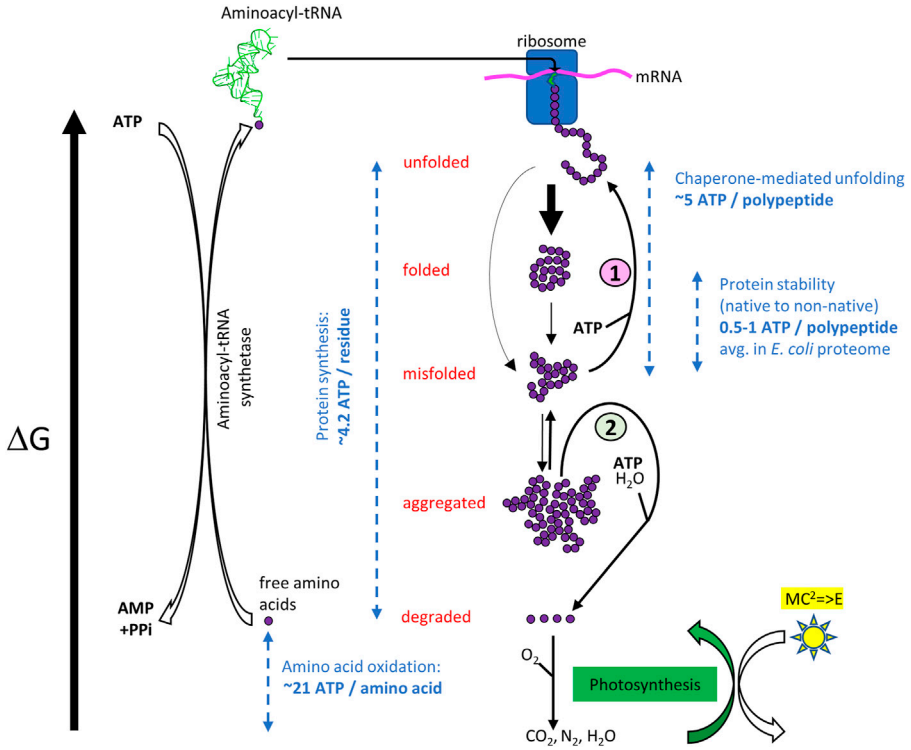


FIGURE 1 | ATP-cost of the open, non-equilibrium free energy cycle of cellular protein homeostasis. Top. When forming in the ribosome, a nascent polypeptide is unfolded, and it has the highest free energy. It may then spontaneously collapse into the native state (folded) or reach metastable misfolded states (misfolded). Heat stress can reduce the kinetic barriers between states. Even without heat stress, the native state can partially unfold and convert into a misfolded state, which in turn converts into stable non-native oligomers (aggregated). (1) The misfolded and aggregated species can be structurally "repaired" into native species again by ATP-fueled unfolding chaperones, or (2) be degraded by proteases gated by ATP-fueled unfolding chaperones (Rothman and Kornberg, 1986). Free amino acids may further degrade into more stable inorganic molecules from which new amino acids can be synthetized, using energy initially originating from the sun, in the form of ATP generated by photosynthesis. ATP-consuming aminoacyl-tRNA synthetases may then load free amino acids with the highest free energy state, which in the ribosomes can spontaneously form less energetic peptide bonds into an unfolded polymer with a lower free energy. Dashed blue arrows indicate typical energy differences between various states, expressed as ATP equivalents (Jungas et al., 1992; Ghosh and Dill, 2010; Sharma et al., 2010; Kaleta et al., 2013).

which is generally biologically active, misfolded, which is inactive and often toxic. Misfolded species may oligomerize into amorphous aggregates and further evolve into increasingly stable and more compact fibrils (**Figure 1**). The unfolded state, such as when a nascent polypeptide exits from the ribosome, has the highest free energy and, as initially shown by Anfinsen et al., (Anfinsen, 1973), can spontaneously collapse into a specific native state without requiring assistance from other molecules. Noticeably, this is not necessarily the most stable state and when kinetic barriers are reduced by mutations or higher temperatures, it may spontaneously undergo transient unfolding and readily collapse into various misfolded states devoid of specific dedicated biological functions, which can be more stable. Misfolded species expose more hydrophobic surfaces to water (Natalello et al., 2013; Sharma et al., 2011) and at high concentrations, tend to further assemble into larger and more compact aggregates that may be more stable than the native state (Hartl and Hayer-Hartl, 2009) (**Figure 1**). By virtue of their exposed hydrophobic surfaces, misfolded polypeptides and aggregates may seek to interact with lipids and affect membrane permeability and activity (Lashuel and Lansbury, 2006; Mahul-Mellier et al., 2015). In metazoans, various aggregates thus cause inflammation and cell death, leading to degenerative diseases and aging (Goloubinoff, 2016). Whereas small aggregates and misfolded polypeptides may be soluble (Diamant et al., 2000), aggregates with many polypeptide chains are often more compact and less soluble. Sustained stress, heat shock in particular, may favor the formation of increasingly larger and more stable aggregates resisting artificial solubilization by urea, or by ATP-fueled disaggregating chaperones, as normally occurring in the cell (**Figure 1**). Because *de novo* synthetized polypeptides exit from the ribosome mostly unfolded, sequentially from the N- to the C-terminal end, they are given an optimal chance to orderly fold, first the N-terminal domains, then the C-terminal domains, leading to proper folding to the native state. In contrast, when a native protein is under heat-stress, misfolding N- and C-terminal domains may concomitantly occur, and improper distal interactions may take place to form stable aggregates.

In the wild, protein misfolding and aggregation may not necessarily be a severe problem for cells living mostly in a quiescent state, as in the case of terminally differentiated adult

TABLE 1 | Absence (NO) or presence (YES) of individual core-chaperones and AAA+ proteases in the genomes of present-day free-living simple Archaea (TACK), simple Bacteria (Aquificae), more complex Archaea (Asgard) and more complex bacteria (Proteobacteria).

Domain	Archaea		Bacteria	
Clade	**TACK**	**ASGARD**	**Aquificae**	**Proteobacteria**
Organism	**Thermogladius Calderae DSM 22663**	**Heimdallarchaeota archaeon (strain LC_2)**	**Desulfurobacterium thermolithotrophum DSM 11699**	***Escherichia coli* K12**
References	Mardanov et al. (2012)	Zaremba-Niedzwiedzka et al. (2017); Rebeaud et al. (2021)	Göker et al. (2011); Rebeaud et al. (2021)	Blattner et al. (1997); Rebeaud et al. (2021)
Chaperones				
HSP90	NO	YES	NO	YES
HSP100/ ClpB	NO	YES	NO	YES
Hsp70	NO	YES	NO	YES
HSP20	NO	YES	YES	YES
HSP60	YES	YES	YES	YES
Proteases				
HslUV	NO	NO	YES	YES
FtsH	NO	NO	YES	YES
ClpAP	NO	NO	YES	YES
ClpXP	NO	NO	YES	YES
Lon	NO	YES	YES	YES
PAN-20S	YES	YES	NO	NO

mammalian cells or of bacteria and archaea living on extremely limited sources of nutrients and energy, for example in the ocean abyss or deep in rocks (Hoshino and Inagaki, 2019; Merino et al., 2019). Under such extremely limiting conditions for growth, most ribosomes are expected to be unoccupied, the few polypeptides to be synthetized being only in replacement of those naturally damaged and degraded. In contrast, when exponentially growing microorganisms are artificially studied in rich media and saturating oxygen, as typically in a laboratory, their protein synthesis and quality control machineries are being maximally challenged and more errors are expected to occur on their folding pathway. Moreover, stresses, such as heat shock, may cause misfolding and aggregation of particularly labile proteins (Mogk et al., 1999). It is thus not fortuitous that point mutations in the bacterial chaperone GroEL, GroES, DnaK, DnaJ and GrpE had been initially identified in a screen of *E. coli* mutants that failed to massively synthetize and properly assemble T4 phage proteins at 37°C, also turned out to be affected in their growth above physiological temperature (Liberek et al., 1988; Wild et al., 1992; Georgopoulos and Welch, 1993).

Evolution of Protein Repair and Degradation Machineries

It has been inferred that more than 3.5 billion years ago, the common ancestor to bacteria and archaea (LUCA) possessed a simple genome that encoded for relatively uncomplicated short proteins, generally uninclined to misfold and aggregate. LUCA likely used only two chaperones, HSP20 that can prevent protein aggregation, and Hsp60 that can use ATP hydrolysis to unfold and repair structurally-damaged small proteins. LUCA likely possessed a single chaperone-gated protease that can use ATP hydrolysis to unfold and degrade proteins that became irreversibly damaged (Rothman and Kornberg, 1986; Rebeaud et al., 2021). The evolutionary history of the sequential buildup of various families of chaperones and proteases can be tentatively evaluated by addressing which ones are currently encoded in the genomes of the simplest and the most complex free-living archaea and bacteria (Rebeaud et al., 2021) (**Table 1**). The genome of the very simple free-living TACK archaeon *Thermogladius calderae* (Mardanov et al., 2012), with only 1,414 genes, encodes for a single Hsp60 chaperone and a single PAN-20S protease (Horwitz et al., 2007). The genome of the very simple free-living Aquificae bacterium *Desulfurobacterium thermolithotrophum* (Göker et al., 2011) with only 1,496 genes, encodes for two chaperones, Hsp20 and Hsp60 and five proteases: ClpAP, ClpXP (Gottesman et al., 1998; Glynn et al., 2009; Zeiler et al., 2013), Lon (Thomas-Wohlever and Lee, 2002), FtsH (Bieniossek et al., 2006) and HslUV (Yoo et al., 1996). By contrast, the genome of the complex ASGARD *Heimdallarchaeota* archaeon (strain LC_2) (Zaremba-Niedzwiedzka et al., 2017), with 4,485 genes and of the complex Gammaproteobacteria *Escherichia coli* (Blattner et al., 1997) with 4,391 genes, both encode for five conserved chaperone families: Hsp20, Hsp60, Hsp70, Hsp90, and Hsp100. In contrast, the picture of the endo-cellular proteases did not evolve much further: a single PAN-20S remained in the most complex archaeon (with Lon in some ASGARDs), and the same complex network of five endo-cellular protease families (ClpAP, ClpXP, Lon, FtsH, HslUV) that were already present in the simplest bacteria, remained present in the most complex ones. This is suggesting that in the evolution of the first

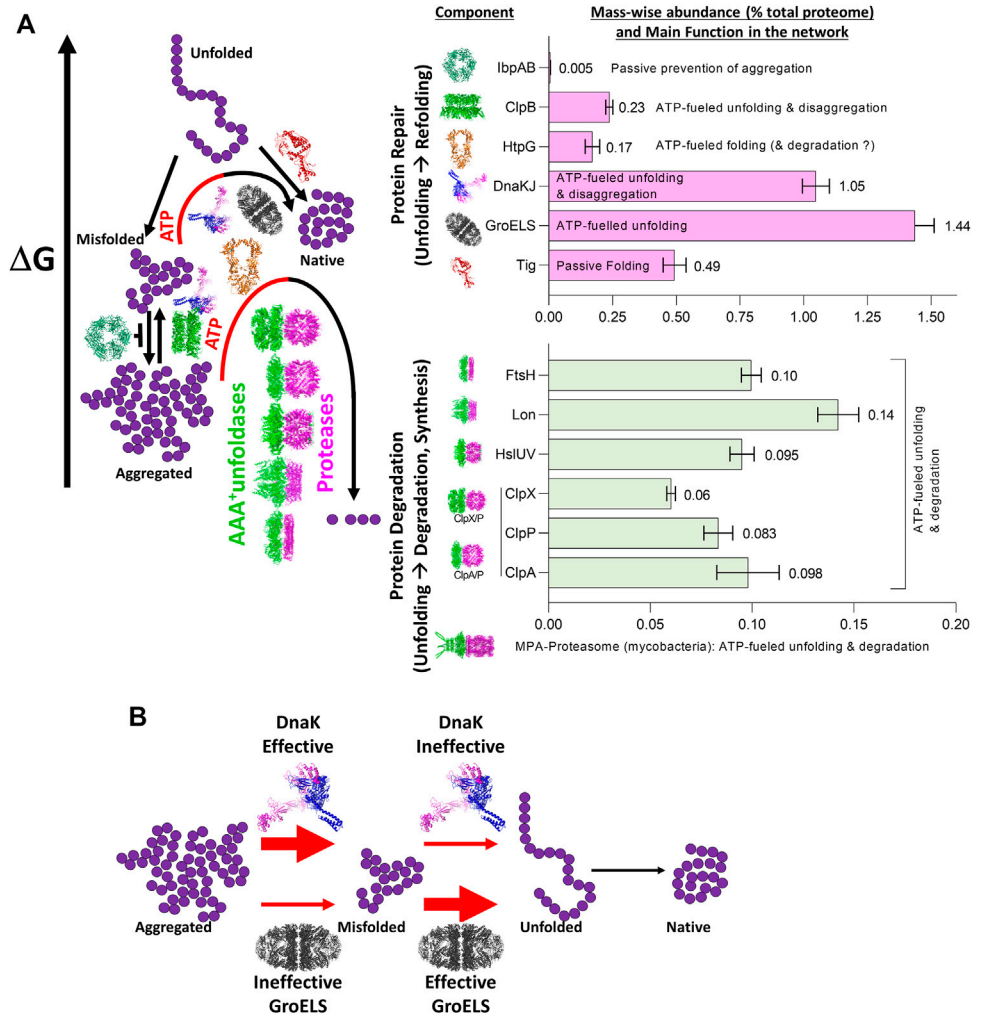


FIGURE 2 | The role of the main bacterial chaperones and proteases. **(A)** Left panel: The main folding states of a polypeptide arranged by free energy, as shown in **Figure 1**. Whereas trigger factor (Tig) passively drives the native folding of unfolded polypeptides exiting the ribosome and HSP20 passively prevents misfolding and aggregation, GroEL, DnaK and ClpB can inject energy from ATP hydrolysis into the expansion of collapsed misfolded and aggregates species. HtpG may use ATP hydrolysis to complete the native folding of slow-folding intermediates from DnaK-DnaJ and failing that, may promote their degradation by proteases (Honoré et al., 2017; Fauvet et al., 2021). The bacterial proteases are gated by N-terminal chaperone-like AAA+ rings (green) that can unfold stable aggregates and thread them into a protease central cylinder (magenta), to be specifically degraded. Right panel: mass-wise abundances of chaperones and proteases from the protein homeostasis network; data from (Fauvet et al., 2021) **(B)** Example of synergism between bacterial HSP70 (DnaK) being effective (red thick arrow) at dismantling aggregates into individual misfolded polypeptides, but ineffective (red thin arrow) at unfolding individual misfolded polypeptides into natively refoldable species, and the bacterial Hsp60 system (GroELS) being ineffective at dismantling aggregates into individual misfolded polypeptides but being highly effective at unfolding individual misfolded polypeptides into natively refoldable species (thick black arrow).

prokaryotes, the "protein degradation toolbox" of microorganisms became fully deployed earlier than a more complex and more versatile "protein repair toolbox," which would have evolved later, hand-in-hand with the proteome complexification. Interestingly, *D. thermolithotrophum* harbors a full set of five different endo-cellular proteases, while lacking HSP70, the central hub of the chaperone network, as well as Hsp90 and ClpB. Therefore, in contrast to ClpB (Hsp100), whose disaggregase activity strictly depends on Hsp70, the endo-cellular proteases are unlikely to depend on HSP70 for their activity. This is particularly the case of ClpA (which is closely related to ClpB), but also the more distantly related AAA+ proteases.

Principal Chaperone Components of the Proteostasis Network

The ATP-consuming chaperones and endo-cellular proteases involved in protein quality control share a common ability to gate their catalytic activities to substrates possessing the required physical characteristics: misfolded proteins in the case of chaperones, and either misfolded or tagged proteins in the case of proteases. Moreover, both need energy from ATP hydrolysis for their function: chaperones to actively unfold and thereby spontaneously *repair* the structure of misfolded substrates into native structure, and proteases to actively

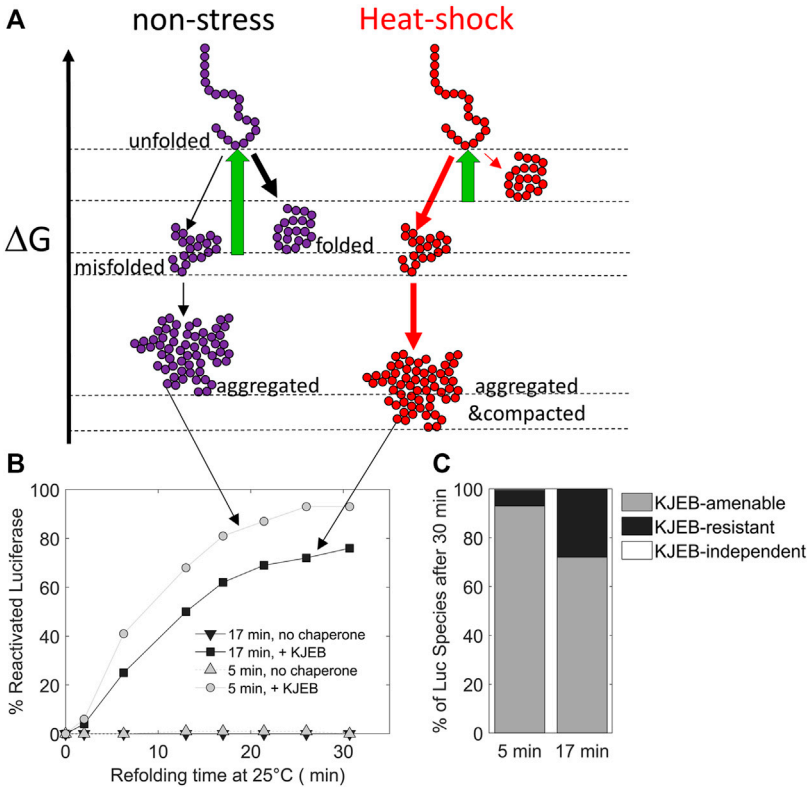


FIGURE 3 | The duration of the heat-stress may change the position of protein conformers on a free energy scale. **(A)** free-energy diagram showing the various folding states of a polypeptide under non-stress (left) and heat-shock conditions (right, red). **(B)** time-dependent refolding of heat pre-denatured firefly luciferase (gray curves: 5 min denaturation at 45°C, black curves: 17 min at 45°C) in the presence or absence of DnaK, DnaJ, GrpE and ClpB (KJEB) at 25°C and ATP. **(C)** Deduced fraction of chaperone-amenable, chaperone-resistant and chaperone-independent luciferase species, determined after 30 min of refolding with KJEB at 25°C and ATP. Data from Goloubinoff (unpublished), based on similar experiments from (Diamant et al., 2000) and (Sharma et al., 2011).

unfold and thereby spontaneously degrade them into peptides. Peptidases will in turn spontaneously degrade peptides into free amino acids, to be used to replace the damaged proteins (*degrade-to-replace*) (**Figure 1**). As mentioned, in *E. coli*, the protein repair part of the proteostasis network is mostly composed of the conserved chaperone families Hsp70s (DnaK, HscA), Hsp60s (GroEL), Hsp90s (HtpG), Hsp100 (ClpB) and Hsp20s (IbpA/B) (bacterial names in brackets), with their main co-chaperones (DnaJ, CbpA, DjlA, HscB, GrpE, GroES) (Finka and Goloubinoff, 2013). Indicating the importance of this chaperone network, its members contribute ~3.3% of the total proteome mass of unstressed *E. coli* cells, compared to ~1% for the endo-cellular proteases, with ~0.6% of the total proteome belonging to quality-control proteases (**Figure 2A**), (Fauvet et al., 2021).

The three distinct chaperone families Hsp60, Hsp70 and Hsp100, which are composed of completely different constituting protomers also greatly differ structurally: Hsp60s are tetra- and hexa-decamers, the Hsp100s are hexamers with a central cavity in which misfolded polypeptides are being unfolded; and the Hsp70s are active as monomers and dimers without a central cavity. Yet, they all use the energy of ATP hydrolysis to extend and partially unfold misfolded structures into transiently bound protein aggregates, thereby providing unfolded segments a renewed chance to spontaneously refold into their native conformation, even under non-equilibrium conditions that are inauspicious for the native state (Sharma et al., 2011; Goloubinoff et al., 2018) (**Figure 2**). But why has evolution favored the appearance of several different families of ATP-fueled protein unfolding chaperones, which are structurally and mechanistically so different?

Misfolded Proteins Are Structurally and Functionally Diverse

Due to the structural diversity and complexity of the proteome, misfolded proteins can adopt a wide spectrum of non-native conformations, and depending on the nature and duration of stresses, a single native polypeptide may adopt a large array of misfolded and aggregated species with different properties (**Figure 3**).

Therefore, a single chaperone might not be able to process with the same efficiency all the possible different substrates. Instead, individual chaperone systems may preferentially bind to, and apply forceful unfolding onto, different kinds of misfolded intermediates, with various efficacies (Mapa et al., 2012; Tiwari et al., 2013). Thus, different chaperone systems expressed in the

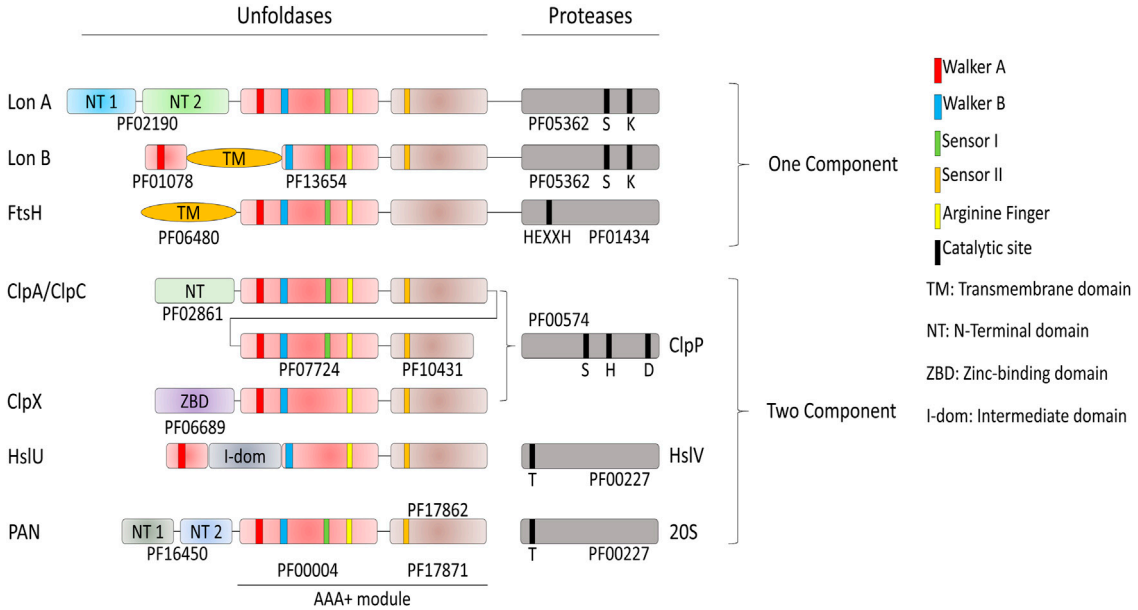


FIGURE 4 | Domain organization of the AAA+ proteases of Bacteria and Archaea. Proteases have different catalytic sites; HslV and the archaeal 20S peptidase use an N-terminal threonine as the active-site nucleophile. Lon proteases have a lysine-serine dyad, ClpP uses a serine-histidine-arginine catalytic triad. FtsH is a Zn^{2+}-dependent peptidase, with the specific HEXXH motif acting as zinc ligand. The AAA+ module is composed of the conserved PFAM domains PF00004 and PF17871 for every protease if not specified by another PFAM. ClpA and ClpC are composed of two AAA+ modules.

same cellular compartment can function synergistically, sequentially processing misfolded substrates, each by their corresponding most efficient chaperone. This is exemplified by the Hsp60 and Hsp70 systems from bacteria: while Hsp70 (DnaK) can efficiently disentangle highly compact preheated MDH aggregates into small, partially misfolded intermediates, but remains rather ineffective at converting them into native proteins, Hsp60 is totally unable to disaggregate stable MDH aggregates, but is highly effective at converting monomeric misfolded MDH intermediates into native proteins. Therefore, when both Hsp60 and Hsp70 are together, the conversion of MDH aggregates into native is most effective and rapid (Veinger et al., 1998) (**Figure 2B**). Similarly, whereas Hsp70, but not Hsp90, can disaggregate and unfold misfolded species, Hsp90 but not Hsp70 can accelerate the maturation of stalled Hsp70-intermediates, into native proteins (Genest et al., 2011; Morán Luengo et al., 2018; Morán Luengo et al., 2019).

Principal Components of the Proteolytic Network

The *degrade-to-replace* section of the *E. coli* protein homeostasis network is composed of AAA+ proteases, namely FtsH, Lon, HslUV (also referred as ClpYQ), ClpXP and ClpAP, with the addition, in actinobacteria and Archaea of the eukaryotic-like MPA-20S proteasome (Neuwald et al., 1999; Iyer et al., 2004; Olivares et al., 2016) (**Figure 4**).

In contrast to the conserved chaperone families that vary extensively in structure and molecular mechanism, the conserved families of quality-control proteases share a common ancestor with the same structural features and a similar molecular mechanism to specifically degrade tagged or misfolded proteins: All are hollowed cylinders, with on their N-terminal side a hexameric ring of AAA+ domains, as in the case of Lon and FtsH, or a hexameric ring of independent AAA+ subunits gating a heptameric protease chamber, as in the case of ClpX, ClpA and HslU and PAN-20S (**Figure 2**). Proteases can specifically recognize and bind few specific alternatively folded, and in general misfolded protein substrates to be degraded, some in collaboration with N-terminal adaptors, such as ClpS in the case of the ClpAP protease (Alhuwaider and Dougan, 2017). Because of the high similarity of sequence and of oligomeric organization, the proteases likely use aa very similar mechanism of action: the energy of ATP hydrolysis is first harnessed to the forceful pulling, and thereby local unfolding, of stably misfolded or alternatively folded loops that are protruding from oligomers or aggregates. This is followed by the threading of the loops within the central cavity of the hexameric AAA+ cylinder, and their forceful extension (Deville et al., 2019), thereby activating and feeding an otherwise inaccessible catalytic proteolytic chamber, which is either found at the C-terminus (LonA and FtsH), or as a separate polypeptide (ClpAP and ClpXP systems, **Figure 2**). Thus, the participation of ATP-fueled unfoldases is a common feature to both the protein-repair and protein degrade-to-replace sides of the protein homeostasis network, in an attempt to unfold and spontaneously recover aggregated proteins into functional native proteins, or alternatively to unfold, thread and degrade them into peptides and free amino acids products of the peptidases, to be reused for the synthesis of new polypeptides.

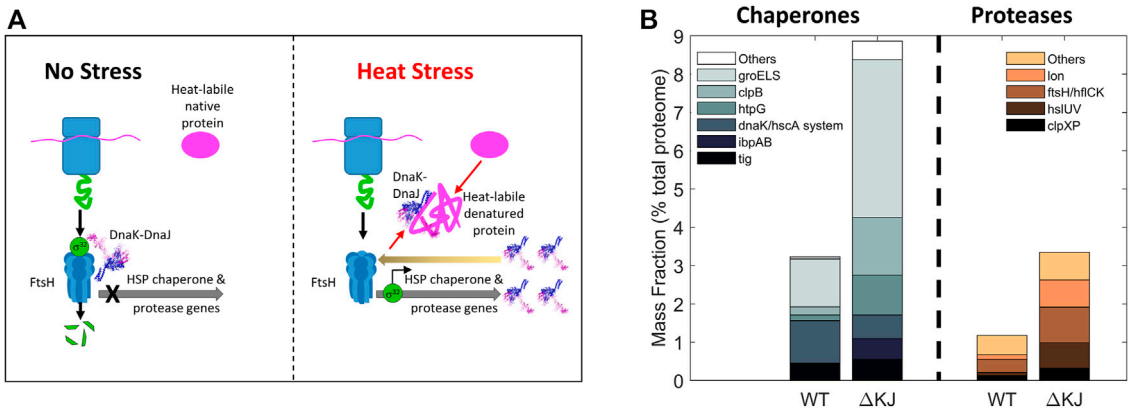


FIGURE 5 | Collaboration between Hsp70 and a protease. **(A)** Bacterial mechanism for the heat-induced expression HSP-chaperones and proteases. **Left**: In the absence of stress, the transcription factor σ^{32} is bound by DnaK/DnaJ as is presented to the protease FtsH to be rapidly degraded, thereby preventing σ^{32}-mediated transcription of heat-shock genes. **Right**: under heat stress, misfolded heat-labile proteins recruit DnaK/DnaJ (thick black arrows) and thus displace them from binding to σ^{32}, allowing it to accumulate and resulting in the up-regulation of heat-shock proteins. The resulting excess of DnaK may then cause the degradation of σ^{32} and arrest excess HSP synthesis (back arrow). **(B)** Deletion of DnaK-DnaJ upregulates the steady state cellular levels of the chaperone-base repair and protease-based degradation machineries. Mass fraction (% of total proteome) of the chaperone and co-chaperones (left) and of proteases (right) in wild type and a delta DnaK-DnaJ *E. coli* strain. Protein quantities are shown as mass fractions (in % of the total proteome) of individual bacterial molecular chaperones and proteases (Data from (Fauvet et al., 2021)).

"Alter-Native States": Chaperones and Proteases Recognize More Than Misfolded Proteins

Diversification of biological processes occurs in different cellular environments, such as the cytosol where both chaperones and proteases are present and the endoplasmic reticulum where chaperones control the quality of the proteins to be secreted. It is thus not surprising that some physiological cellular pathways unrelated to protein misfolding have become "chaperone addicted" in the course of evolution and behave as chaperone-dependent substrates under particular conditions. In eukaryotes, this is the case of clathrin-cages that, following endocytosis, become specifically dismantled into their triskelia constituents by the cytosolic Hsc70 chaperones (Jiang et al., 2003; Sousa et al., 2016). Similarly, the protein IscU of the iron-sulfur cluster assembly pathway can assume two conformations whose populations are regulated by members of the Hsp70 family (Cupp-Vickery et al., 2004; Hoff et al., 2000). Thus, in the course of evolution, some proteins have evolved to use chaperones to switch between different "alter-native" and native conformations, none of which being per se, either structurally or functionally compromised. The bacterial heat shock transcription factor σ^{32} from *E. coli* is another example of such an "alternatively"-folded protein that is both a chaperone- and a protease-substrate. At non-heat shock temperatures, cellular σ^{32} levels are very low because it is constantly produced but it is also constantly unfolded by DnaK/DnaJ and delivered to the FtsH protease for degradation (Herman et al., 1995), resulting in a relatively low basal cellular expression level of the HSP genes (Taylor et al., 1984; Straus et al., 1987; Grossman et al., 1987). During heat shock, DnaJ and DnaK become recruited by labile heat-denaturing proteins, thereby reducing the amount of DnaK that can bind σ^{32} and present it to the protease FtsH for degradation. Consequently, σ^{32}, which is less degraded by FtsH (Herman et al., 1995), binds more HSP promoters and activates the transcription of HSP genes (**Figure 5**). The consequent cellular accumulation of DnaK, DnaJ and FtsH in turn restores the degradation of excess σ^{32} and reduces the expression of HSPs in the long term (Straus et al., 1989). Here, the combined action of an unfolding chaperone and a protease is used to control their own expression. In eukaryotic cells, the major conserved orchestrator of the heat shock response, HSF1, functions in a similar manner to σ^{32}. In unstressed cells, its "alternatively-folded" state is monomeric, maintained in this form in the cytosol by binding to multiple chaperones, notably HSP60s (Neef et al., 2014), HSP70s (Shi et al., 1998) and HSP90s (Ali et al., 1998). Upon stress, they release HSF1 upon to perform their stress-remediating activities. The released HSF1 is then able to reach its native trimeric form, that translocates to the nucleus to activate genomic heat shock elements (Trinklein et al., 2004).

In eukaryotic cells, a substantial fraction of the substrates that are being hydrolyzed by the ATP-dependent Ubiquitin-Proteasome system (UPS) (Glickman and Ciechanover, 2002) consists of natively folded functional proteins, which do not present large hydrophobic patches on their surfaces through which they could interact with chaperone and protease unfoldases. These short and medium half-life proteins, like transcription factors, cyclins or second intracellular messengers, have long been considered to be the main substrates of the proteasomal degradation system, rather than long half-life and damaged proteins (Lecker et al., 2006). To cope with that, a system had to evolve to recognize with extremely high efficiency and selectivity, the few proteins that must be rapidly degraded (Prakash et al., 2004; Yu et al., 2016). Therefore, this

system labels them with a degradation tag (polyubiquitin) that can associate with specific receptors on the 19S cap of the regulatory particles gating the access to the 20S core particle in the 26S Proteasome (Marshall and Vierstra, 2019; Martinez-Fonts et al., 2020). Specifically, three proteasome subunits, Rpn1, Rpn10, and Rpn13 recognize the polyubiquitin chains. In a recent work, Cresti and colleagues demonstrated that the conformational changes of the 26S Proteasome control the unfolding capacity (Cresti et al., 2021). This selectivity has a considerable energy cost (synthesis of ubiquitin molecules, and then of the isopeptide linkage (Ciechanover et al., 1980; Wilkinson et al., 1980; Hershko et al., 1983)) and the degradation of these folded and functioning proteins has a very high energy cost, both for unfolding and for ubiquitin binding. Yet this is necessary as prolonged accumulation of these proteins would cause perturbation of cellular homeostasis, rendering it incompatible with life.

What are the thermodynamic implications of the protein repair and degrade-to-replace networks? With an amino acid pool (free or inserted into polypeptides), which is roughly constant (Hans et al., 2003; Roe et al., 1998)), the degrade-to-replace machinery is necessary to adapt the concentrations of different proteins to internal and external cues. Yet, whereas under stationary conditions errors of folding may be rare, under exponential growth, more spontaneous protein misfolding and aggregation may occur, and higher concentrations of chaperones and proteases may be needed to cope with this rare condition in nature. Purely relying on proteases to disassemble wrongly folded polypeptides would be both risky and energetically costly. On the one hand, some misfolded conformations, and even more so aggregates, might resist the unfolding necessary for threading through the AAA+ gate to be able to reach the proteolytic chamber, leading to progressive accumulation of waste in the cell that can be cytotoxic. For this case, eukaryotes have developed chaperone-mediated autophagy and micro-autophagy (Dong et al., 2020; Schuck, 2020). On the other hand, synthesizing, degrading and re-synthesizing proteins that have misfolded, so to maintain their native concentration at the correct cellular level, requires energy from ATP and GTP hydrolysis by several components of the degrade-to-replace pathway: the AAA+ that gate the proteases hydrolyze ATP to unfold their substrates; the aminoacyl-tRNA transferases that hydrolyze ATP to load amino acids onto their specific tRNAs and form high energy containing aminoacyl-tRNAs that bring the amino acids one-by-one to the ribosome for spontaneous oligomerization; the translation initiation factors that consume GTP molecules. Taking furthermore into account the ATP molecules that are necessary for the ubiquitination of proteasome substrates, and for pupylation in some bacteria (Pearce et al., 2008), degrading and re-synthesizing a new protein can cost a great number of nucleotide molecules (considering the energy from GTP hydrolysis similar to the one from ATP) that is at least equal to, but likely significantly greater than the free energy contained in the peptide bonds. As such, it is estimated that *de* novo protein synthesis costs about 4.2 ATP equivalents per residue (including the energy costs of mRNA synthesis) (Kaleta et al., 2013) (**Figure 1**). In contrast the hydrolysis of as little as five ATPs has been shown to suffice for one bacterial Hsp70 chaperone to convert one stably misfolded luciferase polypeptide containing 564 peptide bonds, into a stable, natively refolded enzyme (Sharma et al., 2010). Thus, the targeted degradation and replacement of damaged proteins necessitates at least two orders of magnitude more ATP than repairing damaged protein conformations by unfolding chaperones. We can thus expect that protein degradation should occur at a rate slow enough to allow a few tens of ATP-fueled chaperone protein-repair cycles per misfolded protein, but not significantly longer, to avoid wasting energy by trying to repair unrecoverable polypeptides.

Following proteolysis, exercising or starving animals may further degrade free amino acids into water, CO_2, urea and/or ammonia, rather than recycling them into new proteins, thereby utilizing them as energy sources via oxidation (**Figure 1**) (Jungas et al., 1992). Extending the proposed view, these simpler compounds represent states of even lower free-energy for the atoms that were part of the amino acids. Therefore, an external energy source is needed to bring them back together into high-energy organic molecules, from which a series of mostly spontaneous metabolic reactions can replenish cells with the needed amino acids: photosynthesis was thus placed on the general free energy landscape of proteins (**Figure 1**). Photosynthetic organisms, such as cyanobacteria, can directly use the energy from the Sun and produce reduced amino acids from water, CO_2 and N_2, to be used by them and other organisms in the food chain, for the synthesis of their own polypeptides. Other non-photosynthetic organisms must rely on high-energy molecules, which they can take from the environment as a source of energy, to drive their metabolic processes that lead to the synthesis of amino-acids and nucleotides.

Protein conformational homeostasis in the cell is thus a series of nested cycles, comprising both spontaneous exergonic and energy-consuming endergonic processes (**Figure 1**). The degrade-to-replace cycle requires amino acids to be "pumped" up into higher-energy aminoacyl-tRNAs, and then polymerized into unfolded polypeptides in the ribosomes. From there, the inevitable tendency of all systems to progress toward their free-energy minimum leads to protein folding, but also, on a different path, to protein misfolding and aggregation, and to degradation of wrongly folded proteins, thus back to free amino acids, whose pool can be further reduced by amino acid degradation. In this case, re-synthesis of amino acids would be part of the degrade-to-replace cycle. This cycle is extremely energy-consuming, with a number of hydrolyzed nucleotides growing at least linearly with the length of the proteins to be degraded and replaced. The protein-repair cycle is nested within the degrade-to-replace one, and its role is, from an energetic perspective, to reduce the necessity of obligatory degradation, by rescuing misfolded and aggregated proteins and giving them an opportunity to refold properly. The unfolding action of chaperones is likely related to the intrinsic stability of individual misfolded domains, and it is thus able to produce an unfolded polypeptide at a lesser energy cost than its *de novo* synthesis. The repair machinery thus represents a more parsimonious approach for protein conformational homeostasis. Nonetheless, *degrade-to-replace*

may become unavoidable and ultimately more energetically advantageous when chemical modifications of proteins, such as glycation, oxidation, and unwanted partial proteolysis could prevent efficient chaperone-driven structural repair, leading to many wasteful ATPase cycles (Chondrogianni et al., 2014). In Cuba under embargo, 60 years of increasingly expensive iterative cycles of repair of old American cars have likely cost their owners much more than if they would have been given the possibility to replace them, even by pricy new cars (which in their case were unavailable).

CONCLUSION

The thermodynamic dilemma of the proteostasis machinery of all organisms is whether it is more energetically convenient to repair or to replace stress-damaged proteins. The protein homeostasis network is composed of abundant "holding" and ATP-fueled unfolding chaperones that can respectively prevent aggregation and actively repair misfolded proteins into native ones, at a relatively low ATP cost. Yet, when proteins are irreversibly damaged, either chemically or structurally, ineffective attempts of repair would waste ATP, which is to be avoided. ATP-fueled unfolding proteases, although unlikely to be activated by chaperones, may yet specifically recognize the chaperone-stalled, resistant misfolded species, bind them, unfold and funnel them into the proteolysis chambers of the proteases for degradation, while maintaining untouched a large excess of surrounding functional proteins which are native. Although degradation is to be followed by costly ATP and GTP consuming re-synthesis, proteins are like used cars: There always comes a point where the cost of cumulative repairs exceeds that of buying a new car.

AUTHOR CONTRIBUTIONS

BF, MR, ST, PD, and PG wrote the article; BF, MR, PD, and PG prepared the figures.

FUNDING

This project was financed in part by the University of Lausanne and by the Swiss National Science Foundation Grants 140512/1, 31003A_156948; by Grant C15.0042 from the Swiss State Secretariat for Education, Research and Innovation; and by the Swiss Federal Institute of Technology.

REFERENCES

Alhuwaider, A. A. H., and Dougan, D. A. (2017). AAA+ Machines of Protein Destruction in Mycobacteria. *Front. Mol. Biosci.* 4, 49. doi:10.3389/fmolb.2017.00049

Ali, A., Bharadwaj, S., O'Carroll, R., and Ovsenek, N. (1998). HSP90 Interacts with and Regulates the Activity of Heat Shock Factor 1 in Xenopus Oocytes. *Mol. Cel Biol* 18, 4949–4960. doi:10.1128/mcb.18.9.4949

Anfinsen, C. B. (1973). Principles that Govern the Folding of Protein Chains. *Science* 181, 223–230. doi:10.1126/science.181.4096.223

Bieniossek, C., Schalch, T., Bumann, M., Meister, M., Meier, R., and Baumann, U. (2006). The Molecular Architecture of the Metalloprotease FtsH. *Proc. Natl. Acad. Sci.* 103, 3066–3071. doi:10.1073/pnas.0600031103

Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., et al. (1997). The Complete Genome Sequence of *Escherichia C* K-12. *Science* 277, 1453–1462. doi:10.1126/science.277.5331.1453

Chondrogianni, N., Petropoulos, I., Grimm, S., Georgila, K., Catalgol, B., Friguet, B., et al. (2014). Protein Damage, Repair and Proteolysis. *Mol. Aspects Med.* 35, 1–71. doi:10.1016/j.mam.2012.09.001

Ciechanover, A., Heller, H., Elias, S., Haas, A. L., and Hershko, A. (1980). ATP-Dependent Conjugation of Reticulocyte Proteins with the Polypeptide Required for Protein Degradation. *Proc. Natl. Acad. Sci.* 77, 1365–1368. doi:10.1073/pnas.77.3.1365

Cresti, J. R., Manfredonia, A. J., Bragança, C. E., Boscia, J. A., Hurley, C. M., Cundiff, M. D., et al. (2021). Proteasomal Conformation Controls Unfolding Ability. *Proc. Natl. Acad. Sci. USA* 118, e2101004118. doi:10.1073/pnas.2101004118

Cupp-Vickery, J. R., Peterson, J. C., Ta, D. T., and Vickery, L. E. (2004). Crystal Structure of the Molecular Chaperone HscA Substrate Binding Domain Complexed with the IscU Recognition Peptide ELPPVKIHC. *J. Mol. Biol.* 342, 1265–1278. doi:10.1016/j.jmb.2004.07.025

Deville, C., Franke, K., Mogk, A., Bukau, B., and Saibil, H. R. (2019). Two-Step Activation Mechanism of the ClpB Disaggregase for Sequential Substrate Threading by the Main ATPase Motor. *Cel Rep.* 27, 3433–3446. doi:10.1016/j.celrep.2019.05.075

Diamant, S., Ben-Zvi, A. P., Bukau, B., and Goloubinoff, P. (2000). Size-Dependent Disaggregation of Stable Protein Aggregates by the DnaK Chaperone Machinery. *J. Biol. Chem.* 275, 21107–21113. doi:10.1074/jbc.m001293200

Dong, S., Aguirre-Hernandez, C., Scrivo, A., Eliscovich, C., Arias, E., Bravo-Cordero, J. J., et al. (2020). Monitoring Spatiotemporal Changes in Chaperone-Mediated Autophagy *In Vivo*. *Nat. Commun.* 11, 645. doi:10.1038/s41467-019-14164-4

Fauvet, B., Finka, A., Castanié-Cornet, M.-P., Cirinesi, A.-M., Genevaux, P., Quadroni, M., et al. (2021). Bacterial Hsp90 Facilitates the Degradation of Aggregation-Prone Hsp70-Hsp40 Substrates. *Front. Mol. Biosci.* 8, 653073. doi:10.3389/fmolb.2021.653073

Finka, A., and Goloubinoff, P. (2013). Proteomic Data from Human Cell Cultures Refine Mechanisms of Chaperone-Mediated Protein Homeostasis. *Cell Stress and Chaperones* 18, 591–605. doi:10.1007/s12192-013-0413-3

Genest, O., Hoskins, J. R., Camberg, J. L., Doyle, S. M., and Wickner, S. (2011). Heat Shock Protein 90 from *Escherichia C* Collaborates with the DnaK Chaperone System in Client Protein Remodeling. *Proc. Natl. Acad. Sci.* 108, 8206–8211. doi:10.1073/pnas.1104703108

Georgopoulos, C., and Welch, W. J. (1993). Role of the Major Heat Shock Proteins as Molecular Chaperones. *Annu. Rev. Cel. Biol.* 9, 601–634. doi:10.1146/annurev.cb.09.110193.003125

Ghosh, K., and Dill, K. (2010). Cellular Proteomes Have Broad Distributions of Protein Stability. *Biophysical J.* 99, 3996–4002. doi:10.1016/j.bpj.2010.10.036

Glickman, M. H., and Ciechanover, A. (2002). The Ubiquitin-Proteasome Proteolytic Pathway: Destruction for the Sake of Construction. *Physiol. Rev.* 82, 373–428. doi:10.1152/physrev.00027.2001

Glynn, S. E., Martin, A., Nager, A. R., Baker, T. A., and Sauer, R. T. (2009). Structures of Asymmetric ClpX Hexamers Reveal Nucleotide-Dependent Motions in a AAA+ Protein-Unfolding Machine. *Cell* 139, 744–756. doi:10.1016/j.cell.2009.09.034

Göker, M., Daligault, H., Mwirichia, R., Lapidus, A., Lucas, S., Deshpande, S., et al. (2011). Complete Genome Sequence of the Thermophilic Sulfur-Reducer Desulfurobacterium Thermolithotrophum Type Strain (BSAT) from a Deep-Sea Hydrothermal Vent. *Stand. Genomic Sci.* 5, 407–415. doi:10.4056/sigs.2465574

Goloubinoff, P. (2016). Mechanisms of Protein Homeostasis in Health, Aging and Disease. *Swiss Med. Wkly* 146, w14306. doi:10.4414/smw.2016.14306

Goloubinoff, P., Sassi, A. S., Fauvet, B., Barducci, A., and De Los Rios, P. (2018). Chaperones Convert the Energy from ATP into the Nonequilibrium Stabilization of Native Proteins. *Nat. Chem. Biol.* 14, 388–395. doi:10.1038/s41589-018-0013-8

Gottesman, S., Roche, E., Zhou, Y., and Sauer, R. T. (1998). The ClpXP and ClpAP Proteases Degrade Proteins with Carboxy-Terminal Peptide Tails Added by the SsrA-Tagging System. *Genes Develop.* 12, 1338–1347. doi:10.1101/gad.12.9.1338

Grossman, A. D., Straus, D. B., Walter, W. A., and Gross, C. A. (1987). Sigma 32 Synthesis Can Regulate the Synthesis of Heat Shock Proteins in *Escherichia C. Genes Develop.* 1, 179–184. doi:10.1101/gad.1.2.179

Hans, M. A., Heinzle, E., and Wittmann, C. (2003). Free Intracellular Amino Acid Pools during Autonomous Oscillations inSaccharomyces Cerevisiae. *Biotechnol. Bioeng.* 82, 143–151. doi:10.1002/bit.10553

Hartl, F. U., and Hayer-Hartl, M. (2009). Converging Concepts of Protein Folding *In Vitro* and *In Vivo*. *Nat. Struct. Mol. Biol.* 16, 574–581. doi:10.1038/nsmb.1591

Herman, C., Thevenet, D., D'Ari, R., and Bouloc, P. (1995). Degradation of Sigma 32, the Heat Shock Regulator in *Escherichia C*, Is Governed by HflB. *Proc. Natl. Acad. Sci.* 92, 3516–3520. doi:10.1073/pnas.92.8.3516

Hershko, A., Heller, H., Elias, S., and Ciechanover, A. (1983). Components of Ubiquitin-Protein Ligase System. Resolution, Affinity Purification, and Role in Protein Breakdown. *J. Biol. Chem.* 258, 8206–8214. doi:10.1016/s0021-9258(20)82050-x

Hoff, K. G., Silberg, J. J., and Vickery, L. E. (2000). Interaction of the Iron-Sulfur Cluster Assembly Protein IscU with the Hsc66/Hsc20 Molecular Chaperone System of Escherichiacoli. *Proc. Natl. Acad. Sci.* 97, 7790–7795. doi:10.1073/pnas.130201997

Honoré, F. A., Méjean, V., and Genest, O. (2017). Hsp90 Is Essential Under Heat Stress in the Bacterium Shewanella Oneidensis. *Cel Rep.* 19, 680–687. doi:10.1016/j.celrep.2017.03.082

Horwitz, A. A., Navon, A., Groll, M., Smith, D. M., Reis, C., and Goldberg, A. L. (2007). ATP-Induced Structural Transitions in PAN, the Proteasome-Regulatory ATPase Complex in Archaea. *J. Biol. Chem.* 282, 22921–22929. doi:10.1074/jbc.m702846200

Hoshino, T., and Inagaki, F. (2019). Abundance and Distribution of Archaea in the Subseafloor Sedimentary Biosphere. *Isme J.* 13, 227–231. doi:10.1038/s41396-018-0253-3

Iyer, L. M., Leipe, D. D., Koonin, E. V., and Aravind, L. (2004). Evolutionary History and Higher Order Classification of AAA+ ATPases. *J. Struct. Biol.* 146, 11–31. doi:10.1016/j.jsb.2003.10.010

Jiang, J., Taylor, A. B., Prasad, K., Ishikawa-Brush, Y., Hart, P. J., Lafer, E. M., et al. (2003). Structure–Function Analysis of the Auxilin J-Domain Reveals an Extended Hsc70 Interaction Interface,. *Biochemistry* 42, 5748–5753. doi:10.1021/bi034270g

Jungas, R. L., Halperin, M. L., and Brosnan, J. T. (1992). Quantitative Analysis of Amino Acid Oxidation and Related Gluconeogenesis in Humans. *Physiol. Rev.* 72, 419–448. doi:10.1152/physrev.1992.72.2.419

Kaleta, C., Schäuble, S., Rinas, U., and Schuster, S. (2013). Metabolic Costs of Amino Acid and Protein Production inEscherichia Coli. *Biotechnol. J.* 8, 1105–1114. doi:10.1002/biot.201200267

Lashuel, H. A., and Lansbury, P. T., Jr. (2006). Are Amyloid Diseases Caused by Protein Aggregates that Mimic Bacterial Pore-Forming Toxins? *Quart. Rev. Biophys.* 39, 167–201. doi:10.1017/s0033583506004422

Lecker, S. H., Goldberg, A. L., and Mitch, W. E. (2006). Protein Degradation by the Ubiquitin-Proteasome Pathway in Normal and Disease States. *J. Am. Soc. Nephrol.* 17, 1807–1819. doi:10.1681/asn.2006010083

Liberek, K., Georgopoulos, C., and Zylicz, M. (1988). Role of the *Escherichia C* DnaK and DnaJ Heat Shock Proteins in the Initiation of Bacteriophage Lambda DNA Replication. *Proc. Natl. Acad. Sci.* 85, 6632–6636. doi:10.1073/pnas.85.18.6632

Mahul-Mellier, A.-L., Vercruysse, F., Maco, B., Ait-Bouziad, N., De Roo, M., Muller, D., et al. (2015). Fibril Growth and Seeding Capacity Play Key Roles in α-Synuclein-Mediated Apoptotic Cell Death. *Cell Death Differ* 22, 2107–2122. doi:10.1038/cdd.2015.79

Mapa, K., Tiwari, S., Kumar, V., Jayaraj, G. G., and Maiti, S. (2012). Information Encoded in Non-Native States Drives Substrate-Chaperone Pairing. *Structure* 20, 1562–1573. doi:10.1016/j.str.2012.06.014

Mardanov, A. V., Kochetkova, T. V., Beletsky, A. V., Bonch-Osmolovskaya, E. A., Ravin, N. V., and Skryabin, K. G. (2012). Complete Genome Sequence of the Hyperthermophilic Cellulolytic Crenarchaeon "Thermogladius Cellulolyticus" 1633. *J. Bacteriol.* 194, 4446–4447. doi:10.1128/jb.00894-12

Marshall, R. S., and Vierstra, R. D. (2019). Dynamic Regulation of the 26S Proteasome: From Synthesis to Degradation. *Front. Mol. Biosci.* 6, 40. doi:10.3389/fmolb.2019.00040

Martinez-Fonts, K., Davis, C., Tomita, T., Elsasser, S., Nager, A. R., Shi, Y., et al. (2020). The Proteasome 19S Cap and its Ubiquitin Receptors Provide a Versatile Recognition Platform for Substrates. *Nat. Commun.* 11, 477. doi:10.1038/s41467-019-13906-8

Merino, N., Aronson, H. S., Bojanova, D. P., Feyhl-Buska, J., Wong, M. L., Zhang, S., et al. (2019). Living at the Extremes: Extremophiles and the Limits of Life in a Planetary Context. *Front. Microbiol.* 10, 780. doi:10.3389/fmicb.2019.00780

Mogk, A., Tomoyasu, T., Goloubinoff, P., Rudiger, S., Roder, D., Langen, H., et al. (1999). Identification of Thermolabile *Escherichia C* Proteins: Prevention and Reversion of Aggregation by DnaK and ClpB. *EMBO J.* 18, 6934–6949. doi:10.1093/emboj/18.24.6934

Morán Luengo, T., Kityk, R., Mayer, M. P., and Rüdiger, S. G. D. (2018). Hsp90 Breaks the Deadlock of the Hsp70 Chaperone System. *Mol. Cel* 70, 545–552. doi:10.1016/j.molcel.2018.03.028

Morán Luengo, T., Mayer, M. P., and Rüdiger, S. G. D. (2019). The Hsp70-Hsp90 Chaperone Cascade in Protein Folding. *Trends Cel Biol.* 29, 164–177. doi:10.1016/j.tcb.2018.10.004

Natalello, A., Mattoo, R. U. H., Priya, S., Sharma, S. K., Goloubinoff, P., and Doglia, S. M. (2013). Biophysical Characterization of Two Different Stable Misfolded Monomeric Polypeptides that Are Chaperone-Amenable Substrates. *J. Mol. Biol.* 425, 1158–1171. doi:10.1016/j.jmb.2012.12.025

Neef, D. W., Jaeger, A. M., Gomez-Pastor, R., Willmund, F., Frydman, J., and Thiele, D. J. (2014). A Direct Regulatory Interaction between Chaperonin TRiC and Stress-Responsive Transcription Factor HSF1. *Cel Rep.* 9, 955–966. doi:10.1016/j.celrep.2014.09.056

Neuwald, A. F., Aravind, L., Spouge, J. L., and Koonin, E. V. (1999). AAA+: A Class of Chaperone-Like ATPases Associated with the Assembly, Operation, and Disassembly of Protein Complexes. *Genome Res.* 9, 27–43.

Olivares, A. O., Baker, T. A., and Sauer, R. T. (2016). Mechanistic Insights into Bacterial AAA+ Proteases and Protein-Remodelling Machines. *Nat. Rev. Microbiol.* 14, 33–44. doi:10.1038/nrmicro.2015.4

Pearce, M. J., Mintseris, J., Ferreyra, J., Gygi, S. P., and Darwin, K. H. (2008). Ubiquitin-Like Protein Involved in the Proteasome Pathway of *Mycobacterium T. Science* 322, 1104–1107. doi:10.1126/science.1163885

Prakash, S., Tian, L., Ratliff, K. S., Lehotzky, R. E., and Matouschek, A. (2004). An Unstructured Initiation Site Is Required for Efficient Proteasome-Mediated Degradation. *Nat. Struct. Mol. Biol.* 11, 830–837. doi:10.1038/nsmb814

Rebeaud, M. E., Mallik, S., Goloubinoff, P., and Tawfik, D. S. (2021). On the Evolution of Chaperones and Cochaperones and the Expansion of Proteomes across the Tree of Life. *Proc. Natl. Acad. Sci. U S A.* 118, e2020885118. doi:10.1073/pnas.2020885118

Roe, A. J., McLaggan, D., Davidson, I., O'Byrne, C., and Booth, I. R. (1998). Perturbation of Anion Balance during Inhibition of Growth of *Escherichia C* by Weak Acids. *J. Bacteriol.* 180, 767–772. doi:10.1128/jb.180.4.767-772.1998

Rothman, J. E., and Kornberg, R. D. (1986). Cell Biology: An Unfolding Story of Protein Translocation. *Nature* 322, 209–210. doi:10.1038/322209a0

Schuck, S. (2020). Microautophagy - Distinct Molecular Mechanisms Handle Cargoes of many Sizes. *J. Cel Sci* 133, jcs246322. doi:10.1242/jcs.246322

Sharma, S. K., De los Rios, P., Christen, P., Lustig, A., and Goloubinoff, P. (2010). The Kinetic Parameters and Energy Cost of the Hsp70 Chaperone as a Polypeptide Unfoldase. *Nat. Chem. Biol.* 6, 914–920. doi:10.1038/nchembio.455

Sharma, S. K., De Los Rios, P., and Goloubinoff, P. (2011). Probing the Different Chaperone Activities of the Bacterial HSP70-HSP40 System Using a Thermolabile Luciferase Substrate. *Proteins* 79, 1991–1998. doi:10.1002/prot.23024

Shi, Y., Mosser, D. D., and Morimoto, R. I. (1998). Molecularchaperones as HSF1-Specific Transcriptional Repressors. *Genes Develop.* 12, 654–666. doi:10.1101/gad.12.5.654

Sousa, R., Liao, H.-S., Cuéllar, J., Jin, S., Valpuesta, J. M., Jin, A. J., et al. (2016). Clathrin-Coat Disassembly Illuminates the Mechanisms of Hsp70 Force Generation. *Nat. Struct. Mol. Biol.* 23, 821–829. doi:10.1038/nsmb.3272

Straus, D. B., Walter, W. A., and Gross, C. A. (1989). The Activity of Sigma 32 Is Reduced Under Conditions of Excess Heat Shock Protein Production in *Escherichia C. Genes Develop.* 3, 2003–2010. doi:10.1101/gad.3.12a.2003

Straus, D. B., Walter, W. A., and Gross, C. A. (1987). The Heat Shock Response of *E. C* Is Regulated by Changes in the Concentration of σ32. *Nature* 329, 348–351. doi:10.1038/329348a0

Taylor, W. E., Straus, D. B., Grossman, A. D., Burton, Z. F., Gross, C. A., and Burgess, R. R. (1984). Transcription from a Heat-Inducible Promoter Causes Heat Shock Regulation of the Sigma Subunit of *E. C* RNA Polymerase. *Cell* 38, 371–381. doi:10.1016/0092-8674(84)90492-6

Thomas-Wohlever, J., and Lee, I. (2002). Kinetic Characterization of the Peptidase Activity of *Escherichia C* Lon Reveals the Mechanistic Similarities in ATP-Dependent Hydrolysis of Peptide and Protein Substrates. *Biochemistry* 41, 9418–9425. doi:10.1021/bi0255470

Tiwari, S., Kumar, V., Jayaraj, G. G., Maiti, S., and Mapa, K. (2013). Unique Structural Modulation of a Non-Native Substrate by Cochaperone DnaJ. *Biochemistry* 52, 1011–1018. doi:10.1021/bi301543g

Trinklein, N. D., Murray, J. I., Hartman, S. J., Botstein, D., and Myers, R. M. (2004). The Role of Heat Shock Transcription Factor 1 in the Genome-Wide Regulation of the Mammalian Heat Shock Response. *Mol. Biol. Cel* 15, 1254–1261. doi:10.1091/mbc.e03-10-0738

Veinger, L., Diamant, S., Buchner, J., and Goloubinoff, P. (1998). The Small Heat-Shock Protein IbpB from *Escherichia C* Stabilizes Stress-Denatured Proteins for Subsequent Refolding by a Multichaperone Network. *J. Biol. Chem.* 273, 11032–11037. doi:10.1074/jbc.273.18.11032

Wild, J., Altman, E., Yura, T., and Gross, C. A. (1992). DnaK and DnaJ Heat Shock Proteins Participate in Protein export in *Escherichia C. Genes Develop.* 6, 1165–1172. doi:10.1101/gad.6.7.1165

Wilkinson, K. D., Urban, M. K., and Haas, A. L. (1980). Ubiquitin Is the ATP-Dependent Proteolysis Factor I of Rabbit Reticulocytes. *J. Biol. Chem.* 255, 7529–7532. doi:10.1016/s0021-9258(19)43857-x

Yoo, S. J., Seol, J. H., Shin, D. H., Rohrwild, M., Kang, M.-S., Tanaka, K., et al. (1996). Purification and Characterization of the Heat Shock Proteins HslV and HslU that Form a New ATP-Dependent Protease in *Escherichia C. J. Biol. Chem.* 271, 14035–14040. doi:10.1074/jbc.271.24.14035

Yu, H., Kago, G., Yellman, C. M., and Matouschek, A. (2016). Ubiquitin-Like Domains Can Target to the Proteasome but Proteolysis Requires a Disordered Region. *Embo J.* 35, 1522–1536. doi:10.15252/embj.201593147

Zaremba-Niedzwiedzka, K., Caceres, E. F., Saw, J. H., Bäckström, D., Juzokaite, L., Vancaester, E., et al. (2017). Asgard Archaea Illuminate the Origin of Eukaryotic Cellular Complexity. *Nature* 541, 353–358. doi:10.1038/nature21031

Zeiler, E., List, A., Alte, F., Gersch, M., Wachtel, R., Poreba, M., et al. (2013). Structural and Functional Insights into Caseinolytic Proteases Reveal an Unprecedented Regulation Principle of Their Catalytic Triad. *Proc. Natl. Acad. Sci.* 110, 11302–11307. doi:10.1073/pnas.1219125110

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

CHAPTER 2

On the evolution of chaperones and cochaperones and the expansion of proteomes across the Tree of Life

On the evolution of chaperones and cochaperones and the expansion of proteomes across the Tree of Life

Mathieu E. Rebeaud†, Saurav Mallik†, Pierre Goloubinoff and Dan S. Tawfik
†*Authors contributed equally to this work*

I contributed to this work conceptually and in the manuscript writing.

Author contributions: M.E.R., S.M., P.G., and D.S.T. conceived and conceptualized the research; M.E.R. performed the phylogenetic analysis shown in Fig. 3A; M.E.R. and S.M. performed the chaperone copy number analysis shown in Figs. 3 B and E and 4B; S.M. performed the Tree of Life and proteome expansion analysis shown in Figs. 1, 2, and 4A, the chaperone expression, and abundance analysis shown in Figs. 3 C and D and 4B; S.M. also prepared all the figures and wrote the paper; M.E.R., S.M., P.G., and D.S.T. edited versions of the paper; and D.S.T. and P.G. acquired funding and supervised the study.

On the evolution of chaperones and cochaperones and the expansion of proteomes across the Tree of Life

Mathieu E. Rebeaud[a,1], Saurav Mallik[b,1], Pierre Goloubinoff[a,2], and Dan S. Tawfik[b,3]

[a]Department of Plant Molecular Biology, Faculty of Biology and Medicine, University of Lausanne, CH-1015 Lausanne, Switzerland; and [b]Department of Biomolecular Sciences, The Weizmann Institute of Science, 7610001 Rehovot, Israel

Edited by Lila M. Gierasch, University of Massachusetts, Amherst, MA, and approved April 2, 2021 (received for review October 6, 2020)

Across the Tree of Life (ToL), the complexity of proteomes varies widely. Our systematic analysis depicts that from the simplest archaea to mammals, the total number of proteins per proteome expanded ∼200-fold. Individual proteins also became larger, and multidomain proteins expanded ∼50-fold. Apart from duplication and divergence of existing proteins, completely new proteins were born. Along the ToL, the number of different folds expanded ∼5-fold and fold combinations ∼20-fold. Proteins prone to misfolding and aggregation, such as repeat and beta-rich proteins, proliferated ∼600-fold and, accordingly, proteins predicted as aggregation-prone became 6-fold more frequent in mammalian compared with bacterial proteomes. To control the quality of these expanding proteomes, core chaperones, ranging from heat shock proteins 20 (HSP20s) that prevent aggregation to HSP60, HSP70, HSP90, and HSP100 acting as adenosine triphosphate (ATP)-fueled unfolding and refolding machines, also evolved. However, these core chaperones were already available in prokaryotes, and they comprise ∼0.3% of all genes from archaea to mammals. This challenge—roughly the same number of core chaperones supporting a massive expansion of proteomes—was met by 1) elevation of messenger RNA (mRNA) and protein abundances of the ancient generalist core chaperones in the cell, and 2) continuous emergence of new substrate-binding and nucleotide-exchange factor cochaperones that function cooperatively with core chaperones as a network.

Tree of Life | expansion of proteomes | core chaperones | cochaperones | chaperone network

All cellular life is thought to have stemmed from the last universal common ancestor (LUCA) (1, 2), that emerged more than 3.6 billion y ago. Two major kingdoms of life diverged from LUCA: bacteria and archaea, which about 2 billion y later merged into the eukaryotes (3). Since the beginning of biological evolution, life's volume has increased on a grand scale: The average size of individual cells has increased ∼100-fold from prokaryotes to eukaryotes (4), the number of cell types has increased ∼200-fold from unicellular eukaryotes to humans (5), and average body size has increased ∼5,000-fold from the simplest sponges to blue whales (6).

This expansion in organismal complexity and variability was accompanied by an expansion in life's molecular workforce, proteomes in particular, which in turn presented a challenge of reaching and maintaining properly folded and functional proteomes. Most proteins must fold to their native structure in order to function, and their folding is largely imprinted in their primary amino acid sequence (7–9). However, many proteins, and especially large multidomain polypeptides, or certain protein types such as all-beta or repeat proteins, tend to misfold and aggregate into inactive species that may also be toxic (10). Life met this challenge by evolving molecular chaperones that can minimize protein misfolding and aggregation, even under stressful out-of-equilibrium conditions favoring aggregation (11, 12). Chaperones can be broadly divided into core and cochaperones. Core chaperones can function on their own, and include ATPases heat shock protein 60 (HSP60), HSP70, HSP100, and HSP90 and the adenosine triphosphate (ATP)-independent HSP20. The basal protein holding, unfolding, and refolding activities of the core chaperones are facilitated and modulated by a range of cochaperones such as J-domain proteins (13–15).

Starting from LUCA, as proteomes expanded, so did the core chaperones and their respective cochaperones. Indeed, chaperones have been shown to facilitate the acquisition of destabilizing mutations and thereby accelerate protein evolution (16–18). However, the coexpansion of proteomes and of chaperones, underscoring a critical balance between evolutionary innovation and foldability, remains largely unexplored. We thus embarked on a systematic bioinformatics analysis that explores the evolution of both proteomes and chaperones, and of both core and their auxiliary cochaperones, along the Tree of Life.

Results

A Tree of Life Analysis of the Expansion of Proteomes and Chaperones. We aimed to explore, systematically, across the Tree of Life (ToL) the expansion of proteomes and compare it with the chaperone composition and level. To this end, we collected proteome sequences from representative organisms belonging to all the major bacterial, archaeal, and eukaryotic clades and constructed a ToL

Significance

Across the Tree of Life, life's phenotypic diversity has been accompanied by a massive expansion of the protein universe. Compared with simple prokaryotes that harbor thousands of proteins, plants and animals harbor hundreds of thousands of proteins that are also longer, multidomain, and comprise a variety of folds and fold combinations, repeated segments, and beta-rich architectures that make them prone to misfolding and aggregation. Surprisingly, the relative representation of core chaperones, those dedicated to maintaining the folding quality of these increasingly complex proteomes, did not change from prokaryotic to mammalian genomes. To reconcile the expanding proteomes, core chaperones have rather increased in cellular abundance and evolved to function cooperatively as a network, combined with their supporting workforce, the cochaperones.

Author contributions: M.E.R., S.M., P.G., and D.S.T. conceived and conceptualized the research; M.E.R. performed the phylogenetic analysis shown in Fig. 3*A*; M.E.R. and S.M. performed the chaperone copy number analysis shown in Figs. 3 *B* and *E* and 4*B*; S.M. performed the Tree of Life and proteome expansion analysis shown in Figs. 1, 2, and 4*A*, the chaperone expression, and abundance analysis shown in Figs. 3 *C* and *D* and 4*B*; S.M. also prepared all the figures and wrote the paper; M.E.R., S.M., P.G., and D.S.T. edited versions of the paper; and D.S.T. and P.G. acquired funding and supervised the study.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]M.E.R. and S.M. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: pierre.goloubinoff@unil.ch.

[3]Deceased May 4, 2021.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2020885118/-/DCSupplemental.

Published May 17, 2021.

(Dataset S1). The overall topology of our tree was borrowed from TimeTree (19) and we also adhered to their order of divergence, which is based on molecular dating and geological records. TimeTree also provides putative dates of emergence and these are provided as branch lengths, yet because our analysis is primarily comparative, branch lengths were only used here as a graphical aid (Fig. 1*A*).

The Tree of Life begins with LUCA at the root (Fig. 1*A*). The edges of the ToL represent the extant three kingdoms—archaea, plotted throughout in black; bacteria, plotted in blue; and eukaryotes. The latter emerged by endosymbiosis of an *Alphaproteobacterium* and an Asgard-like archaeon (20, 21). The emergence of green algae and subsequently of plants occurred with a secondary endosymbiosis of a *Cyanobacterium* into a nonphotosynthetic eukaryote (22). The major eukaryotic clades therefore comprised unicellular, early-diverging eukaryotes (in orange), fungi (gray), plants (green), Metazoa (invertebrate animals; red), and Chordata (vertebrate animals; wine). Overall, our analysis was based on comparing the proteomes of 188 representative organisms, covering 56 major clades of bacteria, archaea, and eukaryote (Dataset S1). The various proteome parameters analyzed below were initially derived for each representative organism in the core tree. The representative organisms of each clade were then pulled together to calculate the clade average and the SD for this average. The clade average values were subsequently plotted using the order of divergence for the *x* axis. Accordingly, these plots also broadly divide into prokaryotes (the left part) and eukaryotes (the right part), and the latter's right edge comprises Chordata including Mammalia (Fig. 1*B* and the following figures).

The Expansion of Proteome Size. Initially, we scrutinized the expansion of proteome size by examining 1) the total number of proteins per proteome in a given clade; 2) the median protein length; and 3) the number of multidomain proteins in the proteome. The clade average values of these three parameters are plotted in Fig. 1 *B–D*, with colors of points matching the branch colors in Fig. 1*A*.

The total number of proteins per proteome expanded ~200-fold. Proteomes that comprise a larger number of proteins unavoidably present a greater challenge for their protein quality control chaperone machinery. To examine the expansion in the number of proteins per proteome, proteome sequences of the 188 representative organisms were obtained. Across the ToL, the number of proteins per proteome expanded roughly 200-fold (Fig. 1*B*) from ~700 proteins in the simplest free-living DPANN (23) archaea to

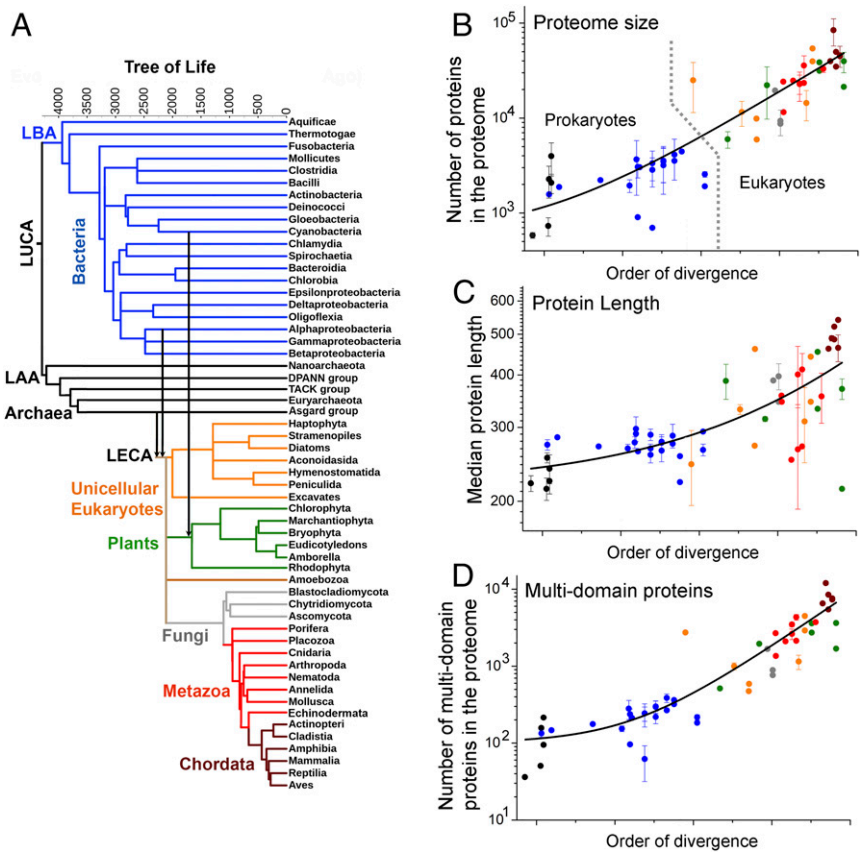


Fig. 1. Expansion of proteome size across the Tree of Life. (*A*) The ToL used in this study. Leaves represent extant phylogenetic clades, while internal nodes represent their presumed ancestors. Branch lengths are in million years, as available from TimeTree (19), and they refer to the relative order of divergence of the corresponding clades rather than the absolute dates of their emergence. The major phylogenetic groups in the tree (Bacteria, Archaea, unicellular eukaryotes, plants, Fungi, Metazoa, and Chordata) are highlighted in different colors. Vertical arrows highlight the two major endosymbiosis events: the Alphaproteobacterial origin of mitochondria and the cyanobacterial origin of plastids. LAA, last archaeal ancestor; LBA, last bacterial ancestor. (*B–D*) The average per-clade values for various proteome size parameters (*y* axis, in log scale) are plotted against their order of divergence (along the *x* axis in linear scale). These parameters include proteome size (*B*), median protein length (*C*), and multidomain proteins in the proteome (*D*). In these scatterplots, the colors of data points represent their major phylogenetic group in the tree (*A*). Error bars represent the clade SD (no error bars relate to clades comprising only one representative organism). The lines were derived by a fit to an exponential equation, and are provided merely as visual guides. Prokaryotic and eukaryotic organisms are separated by a dashed line in *B*.

~120,000 proteins in humans (Dataset S2). Proteome size is similar across prokaryotes, on the order of 3,000 proteins, although the smallest proteomes belong to the earliest-diverging free-living DPANN archaea and *Aquificae* bacteria (23, 24) (not counting parasites and symbionts). Eukaryote proteomes are substantially larger and, considering free-living organisms only, the smallest eukaryotic proteomes harboring ~10,000 proteins belong to Amoebozoa, one of the earliest-diverging eukaryotes (25, 26). Land plants and metazoans comprise hundreds of thousands of proteins per proteome. However, as described later, this dramatic increase in the number of proteins in eukaryotes occurred not only by duplication of preexisting proteins but also by the emergence of completely new domains and folds.

The median protein length increased ~2-fold. The longer the polypeptides are the more prone they are to misfold and aggregate instead of readily reaching their native functional state (27, 28). Analyzing the lengths of all proteins in each representative proteome (*SI Appendix*), we found that compared with ~250 residues across prokaryotes, median protein length increased about 2-fold in multicellular eukaryotes (Fig. 1*C*), with ~400 residues in plants and ~500 residues in Chordata (Dataset S2). Longer proteins were found primarily in multicellular eukaryotes (average lengths of the top 10% largest proteins were roughly 1,300 residues in plants, ~1,500 residues in metazoans, and ~2,150 residues in mammals). The longest polypeptides in mammals are predominantly muscle proteins, including different variants of titin (>34,000 residues) or adhesins (>5,000 residues).

There are different ways by which proteins can increase in size. First, the domains themselves can grow larger by decorating an ancestral core domain with additional segments. Second, the fusion of multiple domains can result in a larger multidomain protein. Third, domain-flanking regions (C- and N-terminal segments, and interdomain linkers), that are typically disordered, can expand. A systematic analysis of 38 distinct folds that are conserved across the ToL (including parasites and symbionts) showed that lengths of individual domains increased mildly, nearly 1.5-fold, across the ToL (*SI Appendix*, Fig. S1*A*). The expansion of domain-flanking segments was also modest, nearly 3-fold, from prokaryotes to multicellular eukaryotes (*SI Appendix*, Fig. S1*B*). Indeed, as elaborated below, length expansion primarily stemmed from the increase in the fraction of multidomain proteins.

Multidomain proteins expanded ~50-fold. Multidomain proteins are inherently more prone to misfolding and aggregation than single-domain proteins, and may therefore demand more chaperone holding–unfolding–refolding action (29–31). To examine their expansion, domain annotations of all proteins in the 188 representative organisms were obtained from Pfam (32). Across the ToL, multidomain proteins comprising ≥3 Pfam-annotated domains have expanded ~50-fold (Fig. 1*D*), from ~100 proteins per proteome in prokaryotes to ~5,000 in plants and animals (Dataset S2). Further, multidomain proteins have expanded beyond the expansion of proteome size, to become nearly 3-fold more frequent in eukaryotic proteomes compared with prokaryotes (*SI Appendix*, Fig. S1*C*) with the expected corresponding shrinkage of proteins comprising one or two domains (*SI Appendix*, Fig. S1*D*). As described later, this expansion occurred not only by duplication of preexisting multidomain proteins but foremost by the emergence of new domain combinations.

Proteome Expansion by Innovation. Most proteins emerge by duplication and divergence of a preexisting protein. The outcome is paralogous proteins with the same overall fold and domain arrangement (for multidomain proteins). Thus, duplication and "local" divergence (point mutations and short insertions or deletions) certainly increase proteome size (the total number of proteins) but do not dramatically change proteome composition or complexity. The latter relates primarily to the birth of completely new proteins possessing new folds, and to the emergence of multidomain proteins with new fold combinations. Additions of new folds and fold combinations likely impose an additional burden on the chaperone machinery. We thus analyzed additional proteome parameters that represent expansion by innovation, rather than by mere duplication and divergence, as detailed below.

Fold types expanded ~5-fold. To assess the emergence of new folds, we used evolutionary classification of domains (ECOD)—a hierarchical classification of protein folds that uses both sequence and structural similarities and clusters all domains with known structures into independently evolved lineages, termed "X-groups" (33). For each representative organism, the Pfam-annotated domains were mapped to ECOD X-groups (*SI Appendix*). We found that from prokaryotes to eukaryotes, the number of unique folds (i.e., unique ECOD X-groups) per proteome expanded about 5-fold (Fig. 2*A*), from ~150 in prokaryotes to ~450 in metazoans (Dataset S3).

New fold combinations expanded ~20-fold. As shown above, multidomain proteins expanded nearly 3-fold (their fraction out to the total number of proteins) alongside a parallel shrinkage of proteins comprising one or two domains (*SI Appendix*, Fig. S1 *C* and *D*). This expansion of multidomain proteins occurred not only by duplication, namely by amplifying preexisting multidomain proteins, but also via the emergence of new combinations. To assess the latter, we examined the number of unique combinations of domains per proteome, with domains being assigned by ECOD X-groups. It appeared that new domain combinations arose throughout evolution and, from prokaryotes to eukaryotes, the number of unique combinations per proteome increased ~20-fold (Fig. 2*B*), from ~100 combinations in bacteria to ~2,000 combinations in Chordata (Datasets S4 and S5).

All-beta and beta-rich folds expanded up to ~600-fold. Proteins that are beta-rich are known to be prone to misfolding and aggregation (34). In the simplest free-living bacteria and archaea, proteins comprising the ancient all-alpha and alpha-beta architectures are the most frequent. Remarkably, upon the emergence of eukaryotes, and in metazoans especially, all-beta or beta-rich architectures (beta superfold) expanded massively, nearly 600-fold (Fig. 2*C*). Beta-rich proteins, in proportion to the total number of proteins, became nearly 6-fold more frequent in mammalian proteomes, as compared with bacteria and archaea (*SI Appendix*, Fig. S2 and Dataset S6). The immunoglobulin fold had a major contribution to this expansion, owing to its diverse roles in immunity, multicellularity, and signaling (35).

Repeat sequences expanded ~700-fold. Proteins comprising tandem repeats of nearly identical sequences emerge readily yet are prone to misfolding and aggregation. We identified proteins with repeated sequences of the size of a single "foldon" unit, ~20 amino acids (aa) (9), with ≥90% sequence similarity (Dataset S7). Most of the early-diverging archaea and bacteria do not possess repeat proteins. Indeed, repeat proteins appear in more recently diverged prokaryotes and foremost in eukaryotes (a similar trend was described in ref. 36). Indeed, metazoan proteomes contain large proteins with long repeated segments, for example, *Drosophila* Ank2p (21 Ankyrin repeats, in total 836 residues) or human Dmbt1p (11 cysteine-rich repeats of a total 1,419-residue length). The cumulative length of repeat sequences in metazoan proteomes can be up to 100,000 residues. Overall, from prokaryotes to eukaryotes, a 700-fold expansion (Fig. 2*D*) of repeated sequences was observed along the ToL (*SI Appendix*, Fig. S3). Repeated sequences also expanded beyond the expansion of proteome size—the percentage of total proteome length that comprises repeats increased nearly 7-fold from prokaryotes to metazoans (Fig. 2*D*).

Proteins predicted as aggregation-prone became ~6-fold more frequent in the proteome. To further examine the expansion of aggregation-prone proteins, for each representative organism, we identified how many proteins in the proteome are predicted to have an

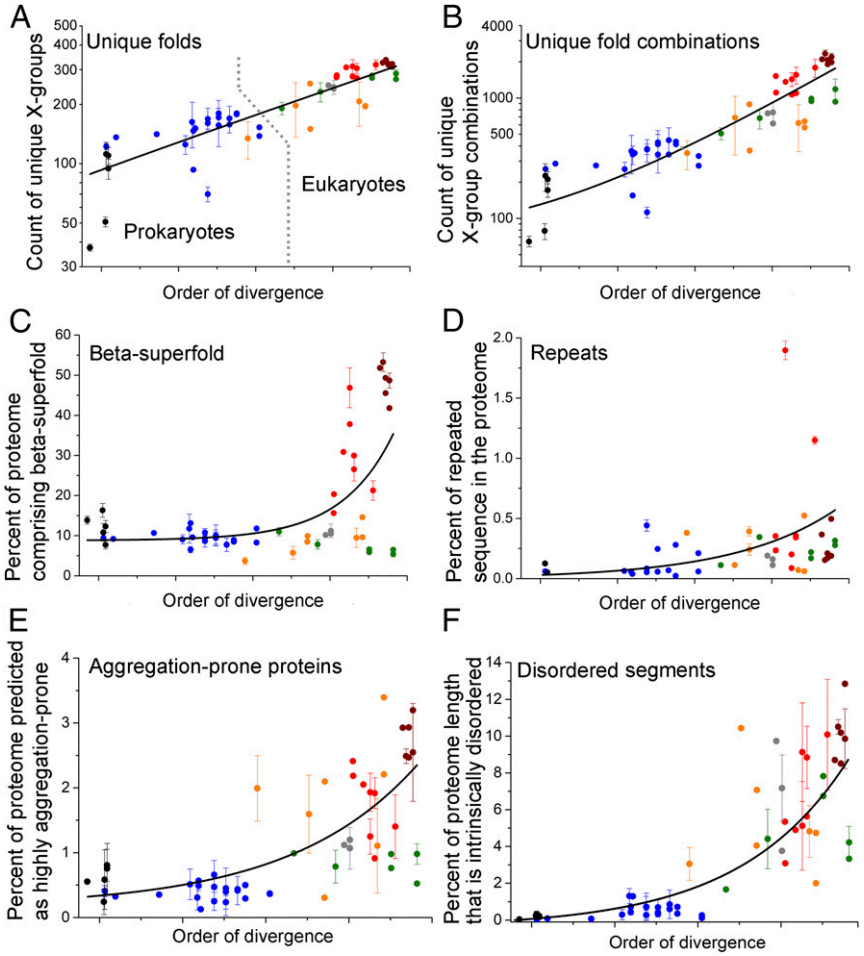


Fig. 2. Expansion of proteomes by innovations. Figure features follow those of Fig. 1. (*A*) Shown on the *y* axis (log scale) is the average number of unique folds [ECOD X-groups (33)] in each phylogenetic clade. The lines were derived by a fit to an exponential equation, and are provided merely as visual guides. Prokaryotic (black and blue dots) and eukaryotic organisms (orange, gray, green, red, and wine dots) are largely separated by a dashed line in *A*. (*B*) Same as *A*, for the count of unique fold combinations. (*C*) Same as *A*, for the percentage of proteins in the proteome comprising at least one beta-superfold domain (note the linear scale). (*D*) Same as *A*, for the percentage of total proteome length that is repeated (also on a linear scale). (*E*) Same as *A*, for the percentage of proteins in the proteome predicted to have ≥20 aggregation hotspots per proteome (also on a linear scale; see also Dataset S10). (*F*) Same as *A*, for the percentage of proteome length that is intrinsically disordered (also on a linear scale).

unusually high number of "aggregation hotspots" (defined as a poorly soluble protein segment of ≥5 aa in length, with solubility predicted from the sequence). The threshold for comparison was set at ≥20 hotspots per protein (at this threshold, ≤3% of proteins are aggregation-prone; the same trend was observed with lower thresholds; Dataset S8). We then calculated what percentage of the entire proteome these aggregation-prone proteins represent. This prediction is restricted by the fact that some of the predicted segments actually reside in the hydrophobic cores of stably folded proteins and hence do not comprise aggregation hotspots (37). However, such segments are likely to be as frequent in prokaryote and eukaryote proteomes (or even less frequent in the latter, where disordered proteins are abundant). Overall, our results indicate that compared with prokaryotes, aggregation-prone proteins have become nearly 6-fold more frequent in eukaryotes (Fig. 2*E*), with the highest frequency seen in Chordata proteomes (Fig. 2*E*).

Intrinsically disordered regions became ~20-fold more frequent in the proteome. From prokaryotes to eukaryotes, did all the changes in proteome composition demand more chaperone action in the latter? The expansion of intrinsically disordered regions, that in principle would not demand increased chaperone action, could be an exception. Though explored before (38, 39), to have this expansion on the same scale and set of organisms used for analyzing all other proteome factors, each protein of the representative proteomes was scanned to infer disordered segments ≥100 aa long (Dataset S9). As plotted in Fig. 2*F*, from prokaryotes to eukaryotes, the percentage of proteome length that is disordered has expanded nearly 20-fold.

Overall, it appears that although gene and whole-genome duplications dominate, and in particular along with the evolution of eukaryotes (40), dramatic changes in proteome composition have occurred owing to bona fide innovations. Specifically, concerning the burden on the chaperone machinery, proteome compositions have changed massively with respect to new folds, beta superfolds, repeat proteins, fold combinations, and aggregation propensity.

The Evolutionary History of Chaperones. In parallel to estimating the expansion of proteome size and composition, we investigated the evolutionary history of chaperones, aiming to date their emergence and their expansion along the ToL. To that end, absence or presence, and copy numbers, of the core chaperones (HSP20, HSP60, HSP70, HSP100, and HSP90) was determined for the representative proteomes. Subsequently, protein trees were generated and compared with the ToL, to account for gene loss and horizontal transfer events. Protein sequences of all core-chaperone

families were extracted from the representative proteomes (Dataset S10). These sequences were aligned and used to generate maximum-likelihood, midpoint-rooted protein trees which were then compared with the ToL (*SI Appendix*, Table S1).

The core chaperones emerged in early-diverging prokaryotes. Our analysis traced the origin of all five core-chaperone families in early-diverging prokaryotes. The phylogenetic tree of a single protein, typically of a few hundred amino acids in length, often lacks the resolution required to reliably date the emergence, especially when horizontal transfer events are frequent. Dating emergence to LUCA is particularly challenging. We followed the recommendations of Berkemer and McGlynn (41) and demanded that for a chaperone family to be assigned to LUCA, there must be a single split between bacterial and archaeal sequences at the root of the protein tree, with strong bootstrap support for this split, and that interkingdom branches would be longer than the intrakingdom branches. These criteria assigned the emergence of only one core chaperone, HSP60—a cage-like ATP-fueled unfoldase (11, 42), to LUCA (Fig. 3*A*). The protein tree of HSP60 further indicated an ancient horizontal gene transfer (HGT) from archaea to Firmicutes, as previously noted (43).

The protein tree of HSP20—an antiaggregation "holding" chaperone—depicted a clear single split of bacterial and archaeal domains at the root, albeit with weak bootstrap support, and interkingdom branch lengths were shorter than intrakingdom branch lengths (*SI Appendix*, Table S1). Similar uncertainties were noted for the majority of protein families formerly assigned to LUCA (41). Previous studies assigned HSP20 to LUCA (44, 45), which, despite the above uncertainties, we concur with (Fig. 3*A*), although later emergence and HGT is an alternative. Indeed, in accordance with a previous study (46), the HSP20 protein tree suggests multiple HGT events, though given the weak bootstrap support it was difficult to distinguish between phylogenetic uncertainty and actual HGT events, let alone to assign donor and acceptor clades.

The remaining core-chaperone families, HSP70, HSP90, and HSP100, appear to have emerged in bacteria, though phylogenetic uncertainties, and probably extensive horizontal transfer between different bacterial clades and between bacteria and archaea, prevent the reliable assignment of their points of origin (Fig. 3*A*). The ATP-dependent core chaperone HSP70—that controls protein unfolding, disaggregation, and degradation (47, 48)—was detected in the earliest-diverging bacterial clades *Aquificae* and *Thermotogae*.

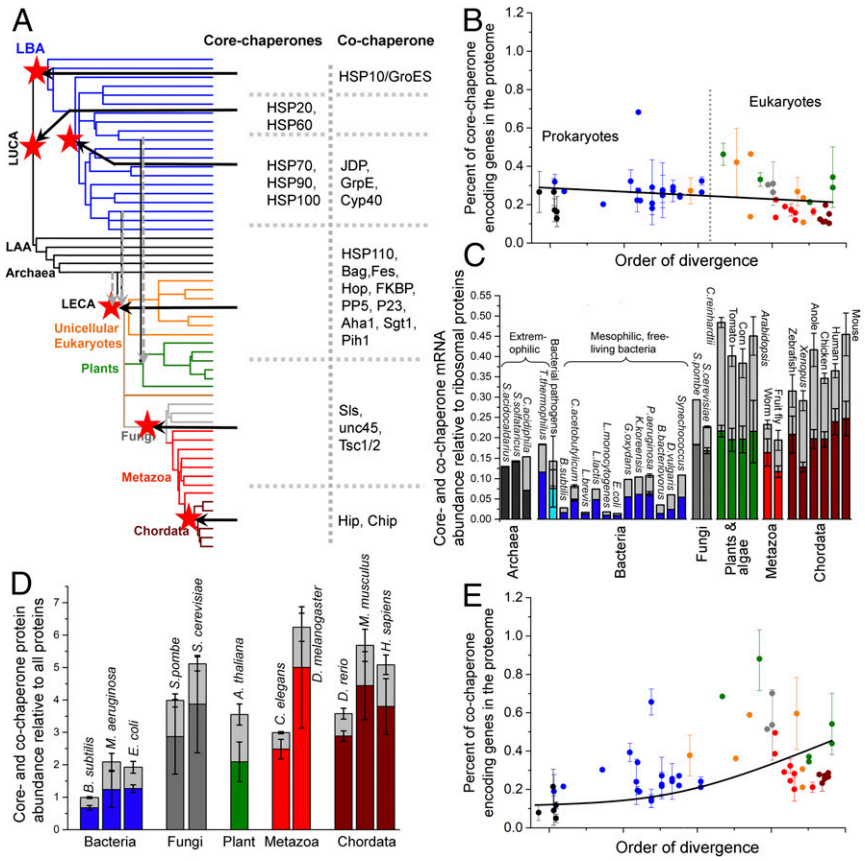


Fig. 3. Evolutionary history of core and cochaperones. (*A*) The de novo emergence of core- and cochaperone families is summarized on the ToL. The ToL is the same as in Fig. 1*A*; clade names are omitted for clarity. Ancestral nodes in which a chaperone family emerged are marked with red stars and the core- and cochaperone families emerged in that node are listed. The dashed gray arrows reflect the endosymbiotic integration of archaeal and bacterial chaperone systems in LECA and eukaryotic and cyanobacterial chaperone systems in photosynthetic algae. (*B*) The percentage of core-chaperone genes per proteome (shown is the average percentage for each phylogenetic clade). Figure features follow Fig. 1. The line was derived by a fit to a linear equation and is provided merely as a visual guide. Prokaryotic (black and blue dots) and eukaryotic organisms (orange, gray, green, red, and wine dots) are largely separated by a dashed line. (*C*) Core- and cochaperone gene expression, relative to ribosomal proteins, in cells. Plotted are model organisms for which sufficient, processed, and reliable expression data were available. The columns include core chaperones (colors represent the phylogenetic clades in Fig. 1*A*) and cochaperones (light gray color). The error bars represent the SD among different nonredundant abundance datasets. (*D*) Same as *C*, for the relative basal abundance of core and cochaperones in cells. (*E*) Same as *B*, for cochaperone genes per proteome. The line was derived by a fit to an exponential equation and is provided merely as a visual guide.

However, the protein tree clustered their HSP70 sequences with those from the later-diverging bacterial lineages *Deltaproteobacteria*, *Clostridia*, and *Bacilli*, with bootstrap values being too low to distinguish between phylogenetic uncertainty and true HGT events (*SI Appendix*, Table S1). Thus, HSP70 appears to have a bacterial origin, around or after the emergence of terrestrial bacteria, that is, around the divergence of *Fusobacteria* (49) (Fig. 3*A*). Following its emergence, HSP70 was likely horizontally transferred to archaea, but the current analysis could not reliably assign donor and acceptor clades.

The protein trees of both HSP90 and HSP100 depict a similar scenario. Both seem to have emerged in bacteria around or after terrestrial bacteria emerged (Fig. 3*A*). Although a reliable point of origin could not be assigned for either of these two chaperones, biochemical assays show that whereas the activity of HSP90 or HSP100 strictly depends on the presence of HSP70, HSP70 itself can act independently. Further, across the ToL, every organism that harbors genes for HSP90 and/or HSP100 also harbors genes for HSP70, but not vice versa. Thus, it is likely that HSP90 and HSP100 have both emerged after HSP70. Similar to HSP70, HSP90 and HSP100 were likely horizontally transferred to archaea. While our protein trees do indicate such trends, the bootstrap values are low.

The archaeal and bacterial core chaperones were integrated into the last eukaryotic common ancestor (LECA), and no new core chaperones emerged with the birth of eukaryotes (Fig. 3*A*). Chaperones of archaeal origin mostly continued to function in their original compartment, the cytosol. Although most Alphaproteobacterial endosymbiont genes were transferred to the nucleus, most of the chaperones of bacterial origin evolved to translocate back to the compartment from which they originated, namely to the mitochondria (50–53). Chaperone evolution in eukaryotes involved gene loss as well; for example, cytosolic and mitochondrial HSP100s have been lost in metazoans (53).

The expansion of core chaperones. Whereas no new core-chaperone family emerged in eukaryotes, gene copy numbers of the existing families did increase via gene duplication, to support expanding proteomes, for condition-specific expression, and also to cater for the emergence of multiple subcellular compartments. Bacteria and archaea typically harbor the same five core-chaperone families as eukaryotes. In any bacterial or archaeal genome, gene copy numbers of individual chaperone families range between 1 and 4, summing up to an average of 8 core-chaperone genes per proteome (Dataset S10). In comparison, the number of core-chaperone genes in higher plants, which are among the most complex eukaryotes, increased ∼30-fold for HSP20, ∼50-fold for HSP60, ∼40-fold for HSP70, ∼20-fold for HSP90, and ∼10-fold for HSP100 (*SI Appendix*, Fig. S4). Parasitic microbes, such as *Mycoplasma pneumoniae*, *Plasmodium falciparum*, and *Entamoeba histolytica*, and photosynthetic bacteria, algae, and plants often harbor unusually high chaperone gene copy numbers, likely to counter the immune response of the host (54, 55) and the oxidative stress (56, 57). However, when proteome size is accounted for, it is evident that the expansion of core chaperones largely coincides with the overall expansion of proteome size. In fact, core chaperones comprise ∼0.3%, that is, 3 out of 1,000 proteins, in all proteomes, from the simplest free-living prokaryotes to mammals (Fig. 3*B*). Further, the expansion of core chaperones occurred by gene duplication only, with no bona fide innovation, as all five core-chaperone families seem to have preexisted in prokaryotes.

The cellular abundance of core chaperones increased ∼6-fold. As described above, the relative representation of core-chaperone genes is roughly the same in the genomes of prokaryotes and eukaryotes. However, gene expression levels could vary, and higher cellular levels of chaperones could support the increasingly complex eukaryotic proteomes. To assess expression levels, the messenger RNA (mRNA) and protein abundance of core or cochaperones was compared in prokaryotic and eukaryotic cells/tissues not subjected to stress or genetic modifications. Curated expression data could be obtained for 30 free-living model organisms spanning 14 major clades along the ToL (Dataset S11). The mRNA abundance of core-chaperone genes, relative to that of ribosomal proteins (as a proxy for the overall level of protein synthesis), has elevated ∼6-fold in chordates compared with mesophilic bacteria (Fig. 3*C*). Protein abundance data available for 11 model organisms along the ToL (Dataset S12) indicate the very same trend (Fig. 3*D*). Further, chaperone levels in eukaryotes are >3-fold compared with extremophilic prokaryotes and pathogenic bacteria that in turn show >2-fold higher levels than those of mesophilic nonpathogenic bacteria.

Cochaperones expanded ∼9-fold. In eukaryotes, as the gene copy numbers of core chaperones expanded, and their protein abundance also increased, what happened to their auxiliary workforce, the cochaperones? The number of unique cochaperone families per proteome expanded from ∼3 in prokaryotes to ∼20 in humans. Most cochaperones are eukaryote-specific (Dataset S10), and therefore have likely emerged relatively recently, and only a few cochaperones are found in all three domains of life. To date their emergence, protein trees were generated. These suggest that, as expected, these cochaperones emerged after the core chaperone they work with. HSP60 is assigned to LUCA, while its bacterial cochaperone HSP10/GroES appears to have emerged later along with the emergence of bacteria (Fig. 3*A*). HSP70's cochaperones, the J-domain proteins (JDPs) and GrpE, and HSP90's cochaperone, Cyp40, appear to have emerged at the same node as their respective core chaperones. Given the phylogenetic uncertainties, possible extensive horizontal transfers, and the resolution that protein trees allow, a more precise dating could not be performed. It appears, however, that JDPs, GrpE, and Cyp40 have all emerged after the emergence of terrestrial bacteria (Fig. 3*A*), and therefore likely emerged after their respective core chaperones. Several cochaperone families emerged in eukaryotes (Fig. 3*A*), including HSP110 that diverged by duplication of HSP70 (58), and Pih1, Aha1, and Chip that harbor eukaryote-specific folds and hence likely emerged de novo. Overall, it is evident that cochaperones tail core chaperones, and not vice versa. With the birth of several cochaperones in eukaryotes, the percentage of genes encoding for cochaperones in the proteome expanded ∼5-fold from prokaryotes to eukaryotes (Fig. 3*E*). Notably, the JDPs, cochaperones of HSP70, are the major contributor to this copy-number expansion (17). Further, alongside the increase in copy numbers and the emergence of new families, the protein expression levels of cochaperones also increased across the ToL (cochaperone protein abundance is ∼2-fold higher in chordates compared with mesophilic bacteria and mRNA abundance is ∼4-fold higher; Fig. 3 *C* and *D*).

Our analysis, therefore, suggests that evolutionary innovation occurred primarily at the level of cochaperones that facilitated the basal core-chaperone activity, thus expanding the chaperone network to meet the challenges of newly emerging protein folds and increasingly complex proteomes.

Discussion

How did the proteomes expand across the ToL, and how did chaperones evolve to support this expansion? To address this question, we compiled data from multiple sources and analyzed them under one roof, thus allowing a systematic, quantitative comparison, as summarized in Fig. 4. From the simplest free-living prokaryotes to plants and animals, proteomes have continuously expanded by both duplications and innovations. It is primarily due to the latter that proteome "complexity" has continuously increased in various ways that demand increased chaperone action. Across the ToL and especially when comparing prokaryotes with eukaryotes, we see a larger number of proteins per proteome (Fig. 4*A*) as well as larger proteins. The latter relates to multidomain proteins being increasingly represented. The

number of different folds increases as well as of unique combinations of folds in multidomain proteins. Proteomes also contain a larger fraction of protein types that are prone to misfolding, such as repeat proteins and proteins comprising beta sheets, and those that are predicted to be highly aggregation-prone. The birth of a new fold, or of a new domain combination, likely results in poor foldability. With time, mutation and selection would improve foldability (59) and could ultimately render a newly born chaperone-independent protein. Nonetheless, the cumulative impact of newly evolved proteins, and of certain protein types (repeat or beta-rich proteins), likely demands increased chaperone capacity.

This dramatic increase in proteome complexity, and hence the demand for chaperone action, has not been met by the emergence of new core chaperones. Eukaryotes possess the same five core-chaperone families as prokaryotes, and metazoans and chordates have in fact lost HSP100 (Fig. 4*B*). Further, the relative representation of core-chaperone genes does not vary between prokaryotes and eukaryotes. Rather, the need for increased chaperone action was met in two ways: first, by an increased cellular abundance of core chaperones, and second, though, by the emergence of new cochaperones—while in bacteria ∼4 cochaperone families are found, in eukaryotes their number increased to 15 or even 20 in Mammalia.

Most of the expanding proteome features are likely the outcome of adaptive evolution (e.g., the emergence of new folds and domain combinations). However, expansion may also occur by drift, that is, by fixation of genetic changes by chance, due to population bottlenecks. Indeed, the effective population size (N_e) has dropped from prokaryotes (typically $>10^8$) to unicellular

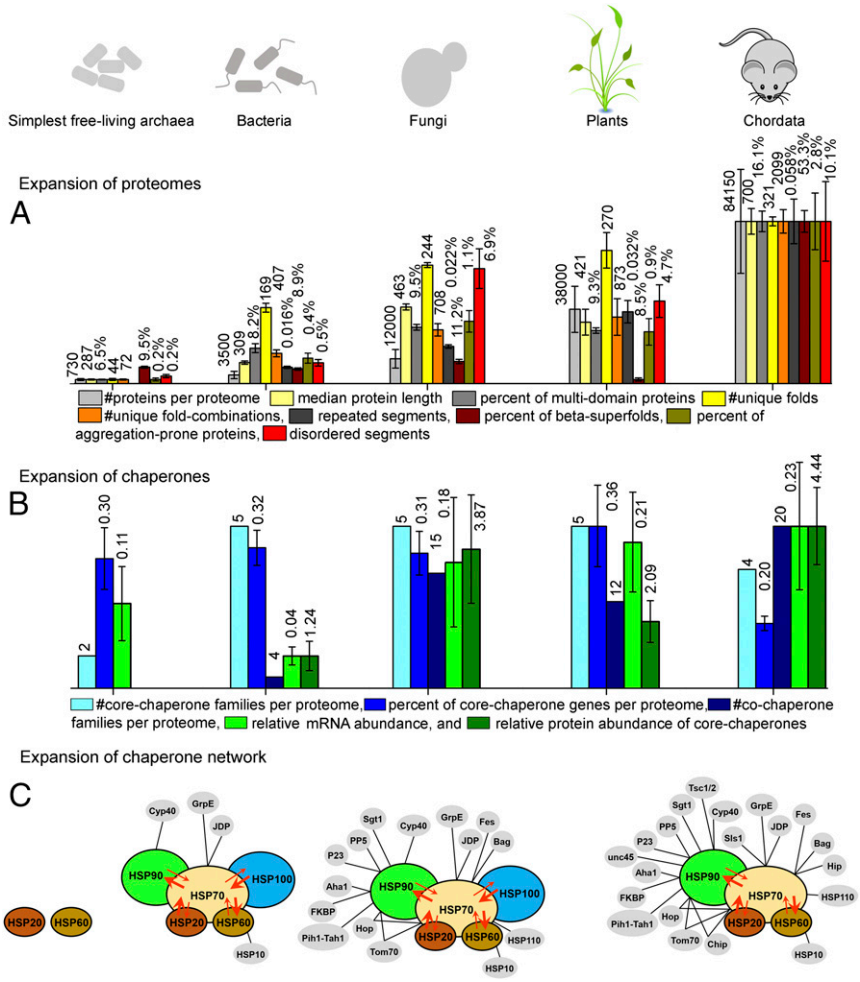


Fig. 4. Summary figure describing the parallel expansion of proteomes and chaperones. Bar heights (*y* axis) were scaled such that the highest value per parameter assumed the same height (the absolute values are listed above the bars). (*A*) Bar graphs describing the expansion of proteomes in a nutshell. For the simplest free-living archaea, bacteria, fungi, plants, and chordates, plotted are the number of proteins per proteome (light gray), median protein length (light yellow), number of unique folds (gray), number of unique fold combinations per proteome (yellow), percentage of multidomain proteins (out of all proteins in the proteome; orange), percentage of proteome length that corresponds to repeat proteins (calculated by residue length; dark gray), percentage of proteins that have the beta-superfold architecture (wine), percentage of proteins predicted as highly aggregation-prone (dark yellow), and percentage of proteome length predicted as intrinsically disordered (red). (*B*) Same as *A*, for the expansion of chaperones. Plotted are the number of core- (cyan) and cochaperone families per proteome (navy), percentage of core-chaperone genes in the proteome (blue), relative mRNA abundance of core chaperones compared with ribosomal proteins (green), and relative protein abundance of core chaperones compared with all other proteins (dark green). (*C*) A schematic description of the expansion of the integrated chaperone network. Core chaperones are shown in various colors and with black outlines, while cochaperones are in gray with no outline. Cochaperones of HSP60, HSP70, and HSP90 are connected to their respective core chaperone by black lines. Cooperativity between core chaperones is represented by overlaps between circles, and substrate sharing between different core chaperones is shown by red arrows. Arrow direction and width represent the direction and magnitude of substrate sharing. Note that the network is shown for the simplest free-living archaea, bacteria, fungi, and chordates.

eukaryotes (~10^7), invertebrates and land plants (~10^6), and chordates (~10^5) (60). Consequently, neutral, or even mildly deleterious mutations that would be purged in prokaryotes, might readily fix in multicellular eukaryotes. For example, drift may have driven the accumulation of hydrophobic residues on protein surfaces regardless of their protein–protein interaction potential, thus leading to lower protein stability and oligomerization but also increased aggregation propensity in eukaryotic proteins (61). Similarly, insertions fixed by drift could elongate disordered segments (62) and repeat proteins. The higher chaperone levels in eukaryotes (Fig. 3 *C* and *D*) may relate to the mitigation of the deleterious effects of such accumulating mutations (63). Pathogenic bacteria often experience severe population bottlenecks, and their chaperone expression levels are comparable to those of extremophiles (Fig. 3*C* and Dataset S11). Overall, the impact of drift on proteome and chaperone evolution merits further investigation.

The above trends highlight two features that comprise hallmarks of the chaperone machinery: the generalist nature of core chaperones, and their ability to act in a cooperative mode alongside cochaperones as an integrated network. HSP60, HSP70, HSP90, and HSP100 are core chaperones acting as generalist unfolding–refolding machineries that work on a broad range of differently misfolded and aggregated protein substrates, largely regardless of size [except HSP60 (64)], structure, and function (11, 31). While core chaperones can exert high-affinity binding to few specific substrates at their native folded state, they generally tend to bind misfolded and aggregated polypeptides that abnormally expose hydrophobic surfaces (31, 65, 66). The main driving force for duplication and specialization is a functional tradeoff—optimization of one function comes at the expense of other functions (67). However, given a "generalist" mode of function, the quality control of increasingly large and complex proteomes could be achieved by an elevated abundance of existing core chaperones, rather than by the emergence of new core-chaperone families. Indeed, although gene copy numbers of core chaperones have indeed increased by gene duplication, their relative representation compared with proteome size remained constant (Fig. 3*B*) and the resulting paralogous copies have mostly relocalized to different subcellular compartments or are expressed under different stress conditions (68). In parasitic microbes and photosynthetic organisms, duplicates of HSP70 and HSP90 have subspecialized to resist host immune responses and oxidative stress (54–57). However, consistent with their generalist nature, the challenge of maintaining large, complex proteomes (Fig. 4*A*) has primarily been met by increased abundances of preexisting core chaperones rather than by the de novo emergence of new ones.

In healthy cells, an integrated chaperone network, comprising both core and cochaperones, controls protein quality (69–71). In this network (Fig. 4*C*), the highly abundant core chaperones operate cooperatively, namely they not only share, and exchange incompletely processed misfolded or unfolded protein substrates, but also trigger the activities of one another. HSP70 plays a critical role in this network by mediating cooperative communications between the other core chaperones. For example, HSP70 triggers the disaggregase activity of HSP100, and jointly they disaggregate aggregated proteins and promote their subsequent refolding (72–74). In another example, HSP20 can transfer misfolded substrates to HSP70 for ATP-driven unfolding, from which they can be further transferred to HSP60 for final refolding to the native state (75). Likewise, HSP90 can promote the maturation of incompletely processed HSP70 substrates (76, 77). Cooperativity and substrate sharing between core chaperones are schematically represented in Fig. 4*C*. Together, these generalist, cooperative core chaperones constitute the core of an integrated chaperone network that has emerged from a simple two-component system in LUCA (Fig. 4*C*).

Alongside the expansion of proteome complexity, the chaperone network has also expanded—primarily by the emergence of cochaperones (Fig. 4*C*). This expanding array of cochaperones augmented the ability of core chaperones to efficiently share substrates and to function cooperatively. In contrast to the generalist core chaperones, cochaperones are more diverse and accordingly seem to subspecialize in specific roles, including cochaperones that handle specific proteins. Examples include UNC45, a cochaperone that emerged in Fungi, and facilitates HSP90-mediated maintenance of myosin in metazoan skeletal and cardiac muscles (78). Another Fungi-born cochaperone, the Tsc1/2 heteromer, specializes in recruiting kinase and some nonkinase substrates to HSP90 (79). Other cochaperones mediate protein transport; examples include Tom70 and P23 that facilitate protein trafficking through Golgi and mitochondrial membranes (80–82). The specialist mode of function of cochaperones coincides with how they expanded, namely by duplication and divergence of ancient prokaryote-born cochaperones but also via bona fide innovations, namely by the emergence of completely new specialized cochaperones in eukaryotes. As shown here, the emergence of new cochaperones coincides with the emergence of new proteins (i.e., by de novo emergence rather than by duplication of preexisting proteins). However, co-occurrence does not mean coevolution—indeed, we know very little about the latter. Did certain cochaperones emerge to support the de novo emergence of a specific protein or protein class? If so, does chaperone dependency persist, hence making codependency a "selfish" irreversible trait? Alternatively, as some newly emerged proteins evolved further, their foldability improved, allowing them to become chaperone-independent.

Thus, across the Tree of Life, proteomes have massively expanded, not just by duplication of preexisting proteins but also by the emergence of completely new ones. Eukaryotic proteomes became particularly large and specifically richer in repeat, beta-rich, and aggregation-prone proteins whose folding is inherently challenging. These changes in proteome size and composition intensified the demand for chaperone action. Curiously, however, no new core chaperones emerged in response to this increased demand. Instead, they increased in abundance relative to all other proteins in the cell. Foremost, an entire network of cochaperones had evolved that facilitate the basal core-chaperone activity.

Materials and Methods

For details, see *SI Appendix, Methods*.

Proteome Size and Median Protein Length. A nonredundant set of 188 prokaryotic and eukaryotic organisms was collected from the TimeTree database (19) (listed in Dataset S1) and their proteome sequences (sequences of all proteins including splice variants, if relevant) were obtained from the National Center for Biotechnology Information genome database (83). For each organism, the total numbers of proteins in the respective proteome, and their lengths, were computed.

Multidomain Proteins. For each protein, domain annotations were collected from Pfam (32). The number of proteins comprising <3 and those comprising ≥3 domains were then counted per each proteome.

Number of Unique Folds and Fold Combinations. Proteins were clustered into independently evolved lineages using the ECOD database (which considers both sequence and structural similarities) where independently evolved lineages are termed X-groups (33). For each representative organism, the Pfam-assigned domains were mapped to their corresponding ECOD X-groups. The numbers of unique X-groups identified in a given proteome were considered as a measure of the number of unique folds. Similarly, we counted the X-group combinations, considering also their order along the polypeptide chain (AB ≠ BA) present in each protein. The total number of X-group combinations identified in a given proteome was considered as a measure of the total number of fold combinations.

Beta Superfolds in the Proteome. We counted how many proteins in the proteome contain at least one domain annotated as all-beta fold (annotated in ECOD's top hierarchy groups: beta barrel, beta meander, beta sandwich, beta duplicates or obligate multimers, and beta-complex topology). By normalizing this number by the total number of proteins in the proteome, we derived the fraction of beta-superfold proteins per proteome.

Repeated Sequences. Each protein in the representative proteomes was scanned by the T-REKS repeat-identifier program (84) to detect repeats that are ≥20 aa long and exhibit ≥90% sequence similarity. The percentage of proteome length that is repeated was subsequently derived (the sum of the number of residues of all repeated segments multiplied by 100, divided by the total length of all proteins).

Aggregation-Prone Proteins. An aggregation hotspot was defined as a "poorly soluble" protein segment of ≥5 aa in length, with solubility predicted from the protein's sequence using CamSol v2.1 (85). For each representative organism, we computed the percentage of proteins in the proteome that contain ≥20 aggregation hotspots.

Intrinsically Disordered Regions. Intrinsically disordered segments were identified by scanning all proteins in the representative proteomes by IuPred2A (86). Disordered segments ≥100 residues were considered. The percentage of proteome length that is disordered was subsequently derived (the total number of residues assigned to disordered segments multiplied by 100, divided by the sum of the length of all proteins).

Evolutionary History of Chaperones. To determine the evolutionary appearance and expansion of the core-chaperone and cochaperone families, we identified their occurrences in the 188 representative organisms, using two complementary methods. The first method involved manual curation of annotated chaperones in model organisms that were subsequently used as queries to find orthologous sequences in the other organisms by protein–protein BLAST (87). The second method involved identifying the Pfam-assigned domain combinations of the various known chaperones in model organisms. Subsequently, any protein in the representative proteomes comprising these domain combinations was assigned as a member of the corresponding chaperone family. The orthologous and paralogous sequences for each chaperone family were aligned using MUSCLE v3.8.31 (88). Maximum-likelihood phylogenetic trees were generated by MEGA X (89). To date the emergence of individual chaperone families, the protein trees were manually compared with the ToL to assign the node of emergence and possible HGT events.

Chaperone Abundance Analysis. To quantify the variation in chaperone mRNA abundance, RNA sequencing (RNA-seq)-based expression data for various model organisms were collected from available resources. Only processed RNA-seq data were considered where the normalized abundances of mRNA transcripts are provided as transcripts per million (TPMs). For each experiment, the sum of the TPM values of the core and the cochaperones was divided by the sum of TPMs of all ribosomal proteins. The average and SD over all experiments per given organism were computed.

To quantify protein abundance, mass spectrometry-based protein abundance data were collected for various model organisms from PaxDb (90). For each dataset, the sum of abundance values of all core chaperones and cochaperones was normalized by the sum of abundance values of all other proteins. The average and SD over all experiments per given organism were computed.

Data Availability. All study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. We thank Jagoda Jablonska for helping to establish the representative ToL, and Ita Gruic-Sovulj for proposing the chaperone abundance analysis. We thank Rina Rosenzweig, Harm H. Kampinga, Matthias P. Mayer, Sudip Kundu, Marc Robinson-Rechavi, and Agnieszka Klosowska for valuable suggestions regarding the manuscript. This work was supported by a Minerva Foundation grant (to D.S.T.) and Grant 31003A_175453 from the Swiss National Fund (to P.G. and M.E.R.). S.M. was supported by a postdoctoral fellowship, provided by Israel's Council for Higher Education Planning and Budgeting Committee. P.G. is an associate professor in the Department of Plant Molecular Biology at the University of Lausanne. D.S.T. is the Nella and Leon Benoziyo Professor of Biochemistry at The Weizmann Institute of Science.

1. C. Woese, The universal ancestor. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6854–6859 (1998).
2. W. F. Doolittle, Phylogenetic classification and the universal tree. *Science* **284**, 2124–2129 (1999).
3. L. Eme, A. Spang, J. Lombard, C. W. Stairs, T. J. G. Ettema, Archaea and the origin of eukaryotes. *Nat. Rev. Microbiol.* **15**, 711–723 (2017).
4. G. M. Cooper, *The Cell: A Molecular Approach* (Sinauer Associates, 2000).
5. D. Arendt, The evolution of cell types in animals: Emerging principles from molecular studies. *Nat. Rev. Genet.* **9**, 868–882 (2008).
6. N. A. Heim *et al*., Hierarchical complexity and the size limits of life. *Proc. Biol. Sci.* **284**, 20171039 (2017).
7. C. B. Anfinsen, Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).
8. K. A. Dill, H. S. Chan, From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **4**, 10–19 (1997).
9. S. W. Englander, L. Mayne, The nature of protein folding pathways. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15873–15880 (2014).
10. J. H. Han, S. Batey, A. A. Nickson, S. A. Teichmann, J. Clarke, The folding and evolution of multidomain proteins. *Nat. Rev. Mol. Cell Biol.* **8**, 319–330 (2007).
11. A. Finka, R. U. Mattoo, P. Goloubinoff, Experimental milestones in the discovery of molecular chaperones as polypeptide unfolding enzymes. *Annu. Rev. Biochem.* **85**, 715–742 (2016).
12. Y. E. Kim, M. S. Hipp, A. Bracher, M. Hayer-Hartl, F. U. Hartl, Molecular chaperone functions in protein folding and proteostasis. *Annu. Rev. Biochem.* **82**, 323–355 (2013).
13. A. J. Caplan, What is a co-chaperone? *Cell Stress Chaperones* **8**, 105–107 (2003).
14. F. H. Schopf, M. M. Biebl, J. Buchner, The HSP90 chaperone machinery. *Nat. Rev. Mol. Cell Biol.* **18**, 345–360 (2017).
15. E. J. Duncan, M. E. Cheetham, J. P. Chapple, J. van der Spuy, The role of HSP70 and its co-chaperones in protein misfolding, aggregation and disease. *Subcell. Biochem.* **78**, 243–273 (2015).
16. N. Tokuriki, D. S. Tawfik, Protein dynamism and evolvability. *Science* **324**, 203–207 (2009).
17. D. Bogumil, T. Dagan, Cumulative impact of chaperone-mediated folding on genome evolution. *Biochemistry* **51**, 9941–9953 (2012).
18. D. Alvarez-Ponce, J. Aguilar-Rodríguez, M. A. Fares, Molecular chaperones accelerate the evolution of their protein clients in yeast. *Genome Biol. Evol.* **11**, 2360–2375 (2019).
19. S. B. Hedges, J. Marin, M. Suleski, M. Paymer, S. Kumar, Tree of life reveals clock-like speciation and diversification. *Mol. Biol. Evol.* **32**, 835–845 (2015).
20. L. Margulis, M. Chapman, R. Guerrero, J. Hall, The last eukaryotic common ancestor (LECA): Acquisition of cytoskeletal motility from aerotolerant spirochetes in the Proterozoic Eon. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 13080–13085 (2006).
21. C. de Duve, The origin of eukaryotes: A reappraisal. *Nat. Rev. Genet.* **8**, 395–403 (2007).
22. W. Martin *et al*., Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* **393**, 162–165 (1998).
23. N. Dombrowski, J. H. Lee, T. A. Williams, P. Offre, A. Spang, Genomic diversity, lifestyles and evolutionary origins of DPANN archaea. *FEMS Microbiol. Lett.* **366**, fnz008 (2019).
24. F. U. Battistuzzi, A. Feijao, S. B. Hedges, A genomic timescale of prokaryote evolution: Insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol. Biol.* **4**, 44 (2004).
25. F. Burki, A. J. Roger, M. W. Brown, A. G. B. Simpson, The new tree of eukaryotes. *Trends Ecol. Evol.* **35**, 43–55 (2020).
26. A. G. Simpson *et al*., Evolutionary history of "early-diverging" eukaryotes: The excavate taxon *Carpediemonas* is a close relative of *Giardia*. *Mol. Biol. Evol.* **19**, 1782–1791 (2002).
27. M. P. Mayer, Gymnastics of molecular chaperones. *Mol. Cell* **39**, 321–331 (2010).
28. K. Henzler-Wildman, D. Kern, Dynamic personalities of proteins. *Nature* **450**, 964–972 (2007).
29. A. Lafita, P. Tian, R. B. Best, A. Bateman, Tandem domain swapping: Determinants of multidomain protein misfolding. *Curr. Opin. Struct. Biol.* **58**, 97–104 (2019).
30. K. Liu, K. Maciuba, C. M. Kaiser, The ribosome cooperates with a chaperone to guide multi-domain protein folding. *Mol. Cell* **74**, 310–319.e7 (2019).
31. Y. Gong *et al*., An atlas of chaperone-protein interactions in *Saccharomyces cerevisiae*: Implications to protein folding pathways in the cell. *Mol. Syst. Biol.* **5**, 275 (2009).
32. R. D. Finn *et al*., Pfam: The protein families database. *Nucleic Acids Res.* **42**, D222–D230 (2014).
33. R. D. Schaeffer, Y. Liao, H. Cheng, N. V. Grishin, ECOD: New developments in the evolutionary classification of domains. *Nucleic Acids Res.* **45**, D296–D302 (2017).
34. A. P. Pawar *et al*., Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. *J. Mol. Biol.* **350**, 379–392 (2005).
35. D. M. Halaby, J. P. Mornon, The immunoglobulin superfamily: An insight on its tissular, species, and functional diversity. *J. Mol. Evol.* **46**, 389–400 (1998).
36. M. A. Andrade, C. Perez-Iratxeta, C. P. Ponting, Protein repeats: Structures, functions, and evolution. *J. Struct. Biol.* **134**, 117–131 (2001).
37. P. Sormanni, F. A. Aprile, M. Vendruscolo, The CamSol method of rational design of protein mutants with enhanced solubility. *J. Mol. Biol.* **427**, 478–490 (2015).
38. C. Wang, V. N. Uversky, L. Kurgan, Disordered nucleiome: Abundance of intrinsic disorder in the DNA- and RNA-binding proteins in 1121 species from Eukaryota, Bacteria and Archaea. *Proteomics* **16**, 1486–1498 (2016).

39. B. Xue, A. K. Dunker, V. N. Uversky, Orderly order in protein intrinsic disorder distribution: Disorder in 3500 proteomes from viruses and the three domains of life. *J. Biomol. Struct. Dyn.* **30**, 137–149 (2012).
40. Y. Van de Peer, S. Maere, A. Meyer, The evolutionary significance of ancient genome duplications. *Nat. Rev. Genet.* **10**, 725–732 (2009).
41. S. J. Berkemer, S. E. McGlynn, A new analysis of Archaea-Bacteria domain separation: Variable phylogenetic distance and the tempo of early evolution. *Mol. Biol. Evol.* **37**, 2332–2340 (2020).
42. H. Saibil, Chaperone machines for protein folding, unfolding and disaggregation. *Nat. Rev. Mol. Cell Biol.* **14**, 630–642 (2013).
43. S. M. Techtmann, F. T. Robb, Archaeal-like chaperonins in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 20269–20274 (2010).
44. F. L. Sousa, S. Nelson-Sathi, W. F. Martin, One step beyond a ribosome: The ancient anaerobic core. *Biochim. Biophys. Acta* **1857**, 1027–1038 (2016).
45. M. C. Weiss *et al*., The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* **1**, 16116 (2016).
46. T. Kriehuber *et al*., Independent evolution of the core domain and its flanking sequences in small heat shock proteins. *FASEB J.* **24**, 3633–3642 (2010).
47. M. R. Fernández-Fernández, J. M. Valpuesta, Hsp70 chaperone: A master player in protein homeostasis. *F1000 Res.* **7** 1497 (2018).
48. R. Rosenzweig, N. B. Nillegoda, M. P. Mayer, B. Bukau, The Hsp70 chaperone network. *Nat. Rev. Mol. Cell Biol.* **20**, 665–680 (2019).
49. F. U. Battistuzzi, S. B. Hedges, A major clade of prokaryotes with ancient adaptations to life on land. *Mol. Biol. Evol.* **26**, 335–343 (2009).
50. S. M. Hemmingsen *et al*., Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature* **333**, 330–334 (1988).
51. W. Martin *et al*., Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12246–12251 (2002).
52. D. Bogumil, D. Alvarez-Ponce, G. Landan, J. O. McInerney, T. Dagan, Integration of two ancestral chaperone systems into one: The evolution of eukaryotic molecular chaperones in light of eukaryogenesis. *Mol. Biol. Evol.* **31**, 410–418 (2014).
53. H. H. Kampinga *et al*., Function, evolution, and structure of J-domain proteins. *Cell Stress Chaperones* **24**, 7–15 (2019).
54. J. Day, A. Passecker, H. P. Beck, I. Vakonakis, The *Plasmodium falciparum* Hsp70-x chaperone assists the heat stress response of the malaria parasite. *FASEB J.* **33**, 14611–14624 (2019).
55. A. Shonhai, A. G. Maier, J. M. Przyborski, G. L. Blatch, Intracellular protozoan parasites of humans: The role of molecular chaperones in development and pathogenesis. *Protein Pept. Lett.* **18**, 143–157 (2011).
56. J. U. Dahl, M. J. Gray, U. Jakob, Protein quality control under oxidative stress conditions. *J. Mol. Biol.* **427**, 1549–1563 (2015).
57. K. Niforou, C. Cheimonidou, I. P. Trougakos, Molecular chaperones and proteostasis regulation during redox imbalance. *Redox Biol.* **2**, 323–332 (2014).
58. N. K. Sarkar, P. Kundnani, A. Grover, Functional analysis of Hsp70 superfamily proteins of rice (*Oryza sativa*). *Cell Stress Chaperones* **18**, 427–437 (2013).
59. R. G. Smock, I. Yadid, O. Dym, J. Clarke, D. S. Tawfik, De novo evolutionary emergence of a symmetrical protein is shaped by folding constraints. *Cell* **164**, 476–486 (2016).
60. M. Lynch, Evolution of the mutation rate. *Trends Genet.* **26**, 345–352 (2010).
61. M. Lynch, L. M. Bobay, F. Catania, J. F. Gout, M. Rho, The repatterning of eukaryotic genomes by random genetic drift. *Annu. Rev. Genomics Hum. Genet.* **12**, 347–366 (2011).
62. S. Light, R. Sagit, O. Sachenkova, D. Ekman, A. Elofsson, Protein expansion is primarily due to indels in intrinsically disordered regions. *Mol. Biol. Evol.* **30**, 2645–2653 (2013).
63. S. Maisnier-Patin *et al*., Genomic buffering mitigates the effects of deleterious mutations in bacteria. *Nat. Genet.* **37**, 1376–1379 (2005).
64. B. Bukau, A. L. Horwich, The Hsp70 and Hsp60 chaperone machines. *Cell* **92**, 351–366 (1998).
65. T. Scheibel, T. Weikl, J. Buchner, Two chaperone sites in Hsp90 differing in substrate specificity and ATP dependence. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 1495–1499 (1998).
66. S. Rüdiger, L. Germeroth, J. Schneider-Mergener, B. Bukau, Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.* **16**, 1501–1507 (1997).
67. D. S. Tawfik, I. Gruic-Sovulj, How evolution shapes enzyme selectivity—Lessons from aminoacyl-tRNA synthetases and other amino acid utilizing enzymes. *FEBS J.* **287**, 1284–1305 (2020).
68. H. H. Kampinga *et al*., Guidelines for the nomenclature of the human heat shock proteins. *Cell Stress Chaperones* **14**, 105–111 (2009).
69. M. S. Hipp, P. Kasturi, F. U. Hartl, The proteostasis network and its decline in ageing. *Nat. Rev. Mol. Cell Biol.* **20**, 421–435 (2019).
70. G. G. Jayaraj, M. S. Hipp, F. U. Hartl, Functional modules of the proteostasis network. *Cold Spring Harb. Perspect. Biol.* **12**, a033951 (2020).
71. M. Taipale *et al*., A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. *Cell* **158**, 434–448 (2014).
72. T. Haslberger *et al*., Protein disaggregation by the AAA+ chaperone ClpB involves partial threading of looped polypeptide segments. *Nat. Struct. Mol. Biol.* **15**, 641–650 (2008).
73. R. Rosenzweig, S. Moradi, A. Zarrine-Afsar, J. R. Glover, L. E. Kay, Unraveling the mechanism of protein disaggregation through a ClpB-DnaK interaction. *Science* **339**, 1080–1083 (2013).
74. J. R. Glover, S. Lindquist, Hsp104, Hsp70, and Hsp40: A novel chaperone system that rescues previously aggregated proteins. *Cell* **94**, 73–82 (1998).
75. L. Veinger, S. Diamant, J. Buchner, P. Goloubinoff, The small heat-shock protein IbpB from *Escherichia coli* stabilizes stress-denatured proteins for subsequent refolding by a multichaperone network. *J. Biol. Chem.* **273**, 11032–11037 (1998).
76. T. Morán Luengo, R. Kityk, M. P. Mayer, S. G. D. Rüdiger, Hsp90 breaks the deadlock of the Hsp70 chaperone system. *Mol. Cell* **70**, 545–552.e9 (2018).
77. O. Genest, J. R. Hoskins, J. L. Camberg, S. M. Doyle, S. Wickner, Heat shock protein 90 from *Escherichia coli* collaborates with the DnaK chaperone system in client protein remodeling. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 8206–8211 (2011).
78. S. L. Wohlgemuth, B. D. Crawford, D. B. Pilgrim, The myosin co-chaperone UNC-45 is required for skeletal and cardiac muscle function in zebrafish. *Dev. Biol.* **303**, 483–492 (2007).
79. M. R. Woodford *et al*., Tumor suppressor Tsc1 is a new Hsp90 co-chaperone that facilitates folding of kinase and non-kinase clients. *EMBO J.* **36**, 3650–3665 (2017).
80. F. J. Echtenkamp *et al*., Global functional map of the p23 molecular chaperone reveals an extensive cellular network. *Mol. Cell* **43**, 229–241 (2011).
81. A. Melnyk, H. Rieger, R. Zimmermann, Co-chaperones of the mammalian endoplasmic reticulum. *Subcell. Biochem.* **78**, 179–200 (2015).
82. E. A. Craig, J. Marszalek, How do J-proteins get Hsp70 to do so many different things? *Trends Biochem. Sci.* **42**, 355–368 (2017).
83. D. A. Benson *et al*., GenBank. *Nucleic Acids Res.* **41**, D36–D42 (2013).
84. J. Jorda, A. V. Kajava, T-REKS: Identification of tandem repeats in sequences with a K-means based algorithm. *Bioinformatics* **25**, 2632–2638 (2009).
85. P. Sormanni, L. Amery, S. Ekizoglou, M. Vendruscolo, B. Popovic, Rapid and accurate in silico solubility screening of a monoclonal antibody library. *Sci. Rep.* **7**, 8200 (2017).
86. Z. Dosztányi, V. Csizmók, P. Tompa, I. Simon, The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J. Mol. Biol.* **347**, 827–839 (2005).
87. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
88. R. C. Edgar, MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
89. S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
90. M. Wang, C. J. Herrmann, M. Simonovic, D. Szklarczyk, C. von Mering, Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. *Proteomics* **15**, 3163–3168 (2015).

PNAS
www.pnas.org

Supplementary Information for

On the evolution of chaperones and co-chaperones and the expansion of proteomes across the Tree of Life

Mathieu E. Rebeaud [a,1], Saurav Mallik [b,1], Pierre Goloubinoff [a,2] and Dan S. Tawfik [b,2]

[a] Department of Plant Molecular Biology, Faculty of Biology and Medicine, University of Lausanne, CH-1015 Lausanne, Switzerland.

[b] Department of Biomolecular Sciences, The Weizmann Institute of Science, Rehovot 7610001, Israel.

[1] Co-first authors

[2] Corresponding Authors, Email: pierre.Goloubinoff@unil.ch, dan.tawfik@weizmann.ac.il

This PDF file includes

Supplementary Methods

Constructing the Tree of Life

To construct a representative Tree of life, we used the TimeTree database (1) and the NCBI taxonomy database (2). For all major clades of bacterial, archaeal, and eukaryotic domains, non-redundant representative species were selected, ensuring that (i) the minimum splitting time for any pair of taxa is ≥ 65 million years, (ii) their annotated proteome sequences are available in NCBI genome database (3) and (iii) proteome-scale domain assignment data are available in Pfam (4). This analysis rendered 188 representative organisms covering 56 major bacterial, archaeal, and eukaryotic clades (**Data S1**). Note that phylogenetic analyses often assign parasitic and symbiotic organisms that have experienced reductive evolution as the earliest diverging clades of their corresponding kingdoms of life. Examples include *Nanoarchaeum equitans*, an obligate symbiont, assigned as the earliest diverging archaea (5-7), and parasitic Excavates assigned as one of the earliest diverging eukaryotes (8, 9). While these organisms were included in our representative set of organisms, the expansion of proteomes and chaperones has been analyzed for free-living organisms only.

The TimeTree Database comprises phylogenetic relationships, as well as literature-based annotations of predicted emergence times (Million Years Ago, MYA, from now) for > 50,000 species. Using this database, a Tree of Life (ToL) was constructed in which leaves represent the extant clades (*e.g.*, Mammalia, which comprises three representative organisms: human, cat, and mouse). The tree's root comprises the Last Universal Common Ancestor (LUCA), and the nodes represent the hypothetical ancestors. Branch lengths represent the divergence time of the different clades. One representative species from each of our 56 phylogenetic clades was chosen, and this species set was submitted to the TimeTree Database to extract a 'core-tree' for their original tree. The obtained tree topology was manually adjusted to depict the emergence of eukaryotes from Asgard archaea and *Alphaproteobacteria* by an endosymbiosis event. The branch lengths represent the evolutionary divergence times as documented in TimeTree, and were used as proxies of the order of divergence to plot the proteome parameters.

Capturing proteome expansion

Proteome size and median protein length in the proteome. Annotated proteome sequences of the 188 representative organisms were obtained as FASTA-format files from the NCBI genome database (3). For each organism, we used Biopython v1.75 package to compute the total number of proteins in the respective proteome, and the length of each protein. These protein lengths were used to derive the median protein length for each organism (**Data S2**).

Multi-domain proteins in the proteome. Proteome-scale domain annotations of each of the 188 representative organisms were obtained from Pfam (4). In Pfam, the established profile HMMs of known domain families (Hidden Markov Models, probabilistic models used for the statistical inference of homology) are searched against the protein sequences, to find all instances of that domain. A statistical significance score (the probability that the prediction is a random hit) is assigned to each predicted instance based on the sequence similarity with the profile HMM. Any domain assigned with $p < 10^{-5}$ significance were considered for further analysis. Overall, in these 188 representative organisms, we identified 7694 domain families, which amounts to roughly 43% of all annotated domain families (17836 families) in Pfam 32.0, September 2018 release. For each species, the number of proteins comprising <3 and those comprising ≥3 Pfam-annotated domains were counted (**Data S2**).

Number of unique fold types and fold-combinations. In general, sequence homologies between different domain families show that they can be clustered into independently evolved lineages (meaning, members of different lineages do not exhibit any detectable sequence homology). Two databases, Pfam (10) and ECOD (11) perform this clustering considering sequence and structural similarities among domain families as the benchmark. In Pfam, these independently evolved super-families are remarked as Clans (10). In ECOD, this clustering is hierarchical (from bottom to top: F-group, T-group, X-group, and top hierarchy) and the independently evolved lineages are termed as the X-groups (11). For each of the 188 proteomes, we mapped the Pfam-assigned domains to the Pfam- Clans as well as to the ECOD X-groups. The 7694 domain families were mapped to 554 unique Clans, and to 976 X-groups, covering 88% of all annotated Clans (629 Clans in Pfam 32.0), and 42% of all annotated ECOD X-groups (2316 X-groups, ECOD v20191115). The numbers of unique Clans, or X-groups, identified in a given proteome were considered as a measure of the number of unique fold types present in a given organism (**Data S3**).

This analysis further allowed us to count the vectorial combinations (AB ≠ BA) of Clans / X-groups, from N- to C-terminal, present in a given protein. The total number of Clan / X-group combinations present in the proteome represents the total number of unique fold-combinations present in the respective organism. The 554 Clans and 976 X-groups included in our work yielded to 15463 X-group combinations and 16538 Clan combinations respectively (**Data S4, S5**).

Proportion of beta-superfolds in the proteome. To capture the proportion of beta-structures in the proteome, for each representative organism, we identified how many proteins in the proteome are annotated as belonging to all-beta folds, and what fraction of the entire proteome they represent. All beta structures assigned under ECOD top hierarchies beta-barrel, beta meander, beta-sandwich, beta duplicates or obligate multimers, and beta complex topology were considered in the analysis (**Data S6**).

Repeated sequences in the proteome. To capture the abundance of repeated sequences in the proteome, the 188 representative proteomes were scanned by T-REKS repeat-identifier program (12). All repeats that are larger than a 'foldon' unit (~20 aa) (13), and exhibit ≥ 90% sequence similarity were considered for further analysis (**Data S7**). Summing the lengths of all the identified repeats, we derived the total length of repeated sequences in each proteome, and what fraction of the total proteome length it covers (total repeat length normalized by the sum of all protein lengths, **Data S2**).

Expansion of domain lengths. To examine the expansion of domain lengths, we first identified 38 distinct folds that are conserved across the ToL, including parasites and symbionts (**Data S3**). Proteins harboring these X-groups were detected in the representative organisms and the respective domain lengths, as annotated in Pfam, were computed. For each organism, these lengths were pulled together and the average domain length was derived (**Data S2**).

Expansion of domain-flanking regions. For each protein harboring an X-group conserved across the ToL, we used the Pfam-annotated location of the protein domain in the primary sequence to measure the lengths of C- and N-termini segments, and inter-domain linkers. For each organism, the lengths of these domain-flanking regions were pulled together to obtain the average (**Data S2**).

Predicted proportion of aggregation-prone proteins in the proteome. To capture the proportion of aggregation-prone proteins in the proteome, for each representative organism, we identified how many proteins in the proteome comprise ≥ n 'aggregation hotspots', where n is an integer. An 'aggregation hotspot' was defined as a 'poorly soluble' protein segment of ≥5 aa length, with solubility predicted from the protein's sequence using CamSol v2.1 (14). The CamSol method yields the solubility profile of a protein: a solubility score is assigned to each amino acid in a way that regions with scores >1 denote 'highly soluble' regions, while scores <−1 reflect 'poorly soluble' ones. For each of the 188 representative organisms, we computed the percent of proteins in the proteome that comprises ≥ n 'aggregation hotspots', for $2 \leq n \leq 20$ (**Data S8**). Note that the CamSol method does not take structural information into account, which means at least for globular proteins, the predicted poorly soluble regions might reside in the hydrophobic core of globular proteins (14).

Intrinsically disordered regions in the proteome. To capture the expansion of intrinsically disordered segments, we identified the disordered segments of each protein in the proteomes of our species set by IuPred2A that exploits the idea that in disordered regions amino acid residues form fewer energetically favorable contacts than residues in ordered regions (15). IUPred2A does not rely on any information besides the amino acid sequence and is therefore suitable for predicting disorder in large protein datasets. The disorder status of each amino acid site is represented by a

probability score; any amino acid site associated with > 0.5 probability score was assigned as intrinsically disordered. Disorder is, however, a property of segments rather than residues, and thus segments of ≥20 aa long were analyzed. Presented here are the results for ≥100 aa segments (**Data S9**) yet other thresholds yielded similar results. The percent of proteome length that is disordered was subsequently derived (sum of the lengths of all ≥ 100 aa disordered segments multiplied by 100, divided by the sum of lengths of all proteins).

Capturing chaperones in the proteome

To determine the evolutionary appearance and expansion of the core-chaperone (HSP20, HSP60, HSP70, HSP90, and HSP100), and co-chaperone families (HSP10/GroES, JDP/HSP40, HSP110, GrpE, Bag, Fes, Hip, Hop, Chip, Tom70, Cyp40, FKBP52/51, PP5, Unc45, Cdc37, P23, Aha1, Sgt1, Pih1, Tah1, and Tsc1/2 heterodimer) we identified their occurrences in the 188 representative organisms, using two complementary methods.

Identifying chaperones by BLAST-search. Chaperone family proteins were manually curated from the UniProt (16) annotated proteomes of model organisms (*Escherichia coli* and *Saccharomyces cerevisiae*). These protein sequences were then used as queries to find orthologous sequences in the other organisms by a comprehensive protein-protein BLAST (17). BLAST hits associated with 50% sequence coverage and ≤ 10^{-5} e-values (18) were manually inspected to extract the 'true-positive' chaperone family members.

Identifying chaperones by characteristic domain combinations. The second method involved manually looking into the Pfam-assigned domain combinations of annotated chaperones. For instance, bacterial HSP20s (*e.g.*, IbpA protein in *E. coli*) predominantly comprise a single HSP20 domain, whereas chordate HSP20s (*e.g.*, HSPB6 protein in *H. sapiens*) often comprise an additional N-terminal alpha-crystallin domain. For each chaperone family, we investigated their domain organization in Pfam and constructed a library of domain combinations (occurrence of the domain of interest with other domains in annotated proteomes). This library comprised any combination that Pfam reports in at least 10 different sequences. We searched these combinations in the 188 representative proteomes. Any protein comprising any of these domain combinations was assigned to be a member of that core-chaperone family. This analysis excluded eukaryote specific HSP110 co-chaperone family that is composed of two HSP70 domains, and thus cannot be distinguished from HSP70s that comprise two HSP70 domains. In this case, specific hallmark sequences in the linker and the more variable C-terminal domain of HSP110 that differ from HSP70 were used in BLAST searches.

The two complementary methods rendered an identical number of chaperone gene copy-numbers in each organism, reflecting the robustness of the overall approach (**Data S10**).

Phylogenetic analysis

Multiple sequence alignments (MSAs) were generated using Muscle v3.8.31 (19). The alignments were manually curated and gap-majority columns were trimmed using trimAl v1.330 (20). Only for HSP20, alignments were collected from EggNOG database (21), and were manually curated. These alignments were filtered by removing all sequences that are more than 80% identical (usually redundant sequences) and those that are less than 30% identical (usually interferes with rooting by creating extremely long branches) were removed using T-coffee (22). Maximum likelihood phylogenetic trees were generated by MEGA X (23), using JTT distance matrix and NJ/BioNJ initial tree. Phylogenetics trees are provided in **Table S1**. To date the emergence of individual chaperone families, the protein trees were manually compared with the ToL to assign the node of emergence and possible HGT events.

Relative mRNA abundance of core- and co-chaperones

To quantify the variation in chaperone gene-expression across the ToL, genome-scale RNA-seq data of prokaryotic/eukaryotic cells/tissues were collected from various resources. Only samples not subjected to stress or genetic modifications, and processed expression data, were considered (where the abundance of mRNA transcripts is provided as TPMs, transcripts per million). Data for metazoans, chordates, plants, and green alga were collected from Bgee (24) and Expression Atlas (25). Data for free-living prokaryotes and pathogenic bacteria were collected from a comparative transcriptomic resource (26) and Pathogenex database (27). Data for *S. pombe* were collected from YeasTSS database (28). For *S. cerevisiae*, single-cell RNA-seq expression data were collected (29). In total, data from 3287 non-redundant RNA-seq experiments were analyzed in our work, comprising 14 eukaryotes, 13 free-living bacteria, 3 free-living archaea, and 18 pathogenic bacteria. For each experiment, the sum of the provided TPM values of the core- and the co-chaperones were divided by the sum of TPMs of ribosomal proteins. The average and standard deviation over all experiments per given organism were computed (**Data S11**) and were plotted as columns in **Fig. 3C**. Note that in **Fig. 3C**, the mean chaperone expression for all bacterial pathogens is plotted alongside free-living bacteria, and the numerical values for individual species are provided as **Data S11**.

Relative protein abundance of core- and co-chaperones

The abundance of a protein simply means the number of copies of a protein molecule in a cell. Non-redundant genome-scale abundance data was collected for 11 non-extremophilic model organisms from PAXdb (30), measured under normal conditions. This database comprises whole genome protein abundance information across organisms and tissues. Dataset quality was measured by its proteome coverage and interaction consistency score (ICS). Proteome coverage

represents what fraction of the entire proteome is included in the abundance dataset. ICS is based on the assumption that proteins that contribute jointly to a shared function (such as members of a protein complex) should tend to have roughly similar protein abundance levels. For unicellular organisms, datasets that cover ≥40% of the respective proteomes with ≥3.0 ICS and include ≥30% of all core- and co-chaperones identified in the respective organism, were selected for analysis. For multicellular organisms, tissue-specific data was collected. Datasets that cover ≥10% of the respective proteomes with ≥3.0 ICS and includes ≥30% of all core- and co-chaperones identified in the respective organism, were selected for analysis. All datasets related to stress or disease-conditions were removed and the remaining 121 datasets were considered for further analysis (**Data S12**). For each abundance dataset, we classified the list of proteins into three non-overlapping groups: (i) core-chaperones, (ii) co-chaperones, and (iii) proteins that are not chaperones. For each group, the abundance values of all the proteins were summed to measure the 'total abundance'. The 'total abundance of core-chaperones', normalized by the 'total abundance of proteins that are not chaperones' represents the relative abundance of core-chaperones. Similarly, we measured the 'relative abundance of co-chaperones'. For each species, the average and standard deviations among different datasets were measured.

Visualizing the expansion of proteomes and chaperones

For each clade in our core-tree, numerical values of the proteome parameters and chaperone-copy number values obtained for its representative organisms were pulled together to derive the clade average, and the clade standard deviation. The clade average values were subsequently plotted against their predicted divergence time and were fitted into the standard exponential growth curve. The fitted exponential curves are depicted as solid black lines.

Data biases associated with the choice of representative organisms

To what extent the choice of the major clades of life and the set of representative organisms could bias the obtained results? The former is unlikely to bias the results because we considered all clades of life currently annotated in NCBI Taxonomy and TimeTree databases. However, in oppose to analyzing all reference genomes in the NCBI dataset, the set of organisms used here is 'normalized'. Consequently, clades where thousands of genomes are available and clades with only a few genomes are represented to the same extent. Nonetheless, we reexamined the data to rule out a possible bias due to the choice of the organisms that represent each clade. To this end, we re-computed the expansion of proteome size and median protein lengths for an alternate set of 181 organisms (**Data S1**) that represent the same 56 major clades shown in **Fig. 1A**. As plotted in ***SI Appendix,* Fig. S5**, the number of proteins in the proteome and the median protein length expanded 150-fold and 2-fold in the alternative species set (as compared to 200-fold and 2-fold in

the original set). These results indicate the robustness of our analysis in terms of the choice of organisms.

Graphics

The Tree of Life was generated using the Interactive Tree Of Life (iTOL) v4.0 (31). All the plots were generated using OriginPro software v9.1.0. Figures were compiled using Adobe Photoshop CS6 v13.0.1.

Quantification and statistical analysis

All the statistical analyses mentioned in the main text were performed using in-house Python scripts.

Supplementary Figures

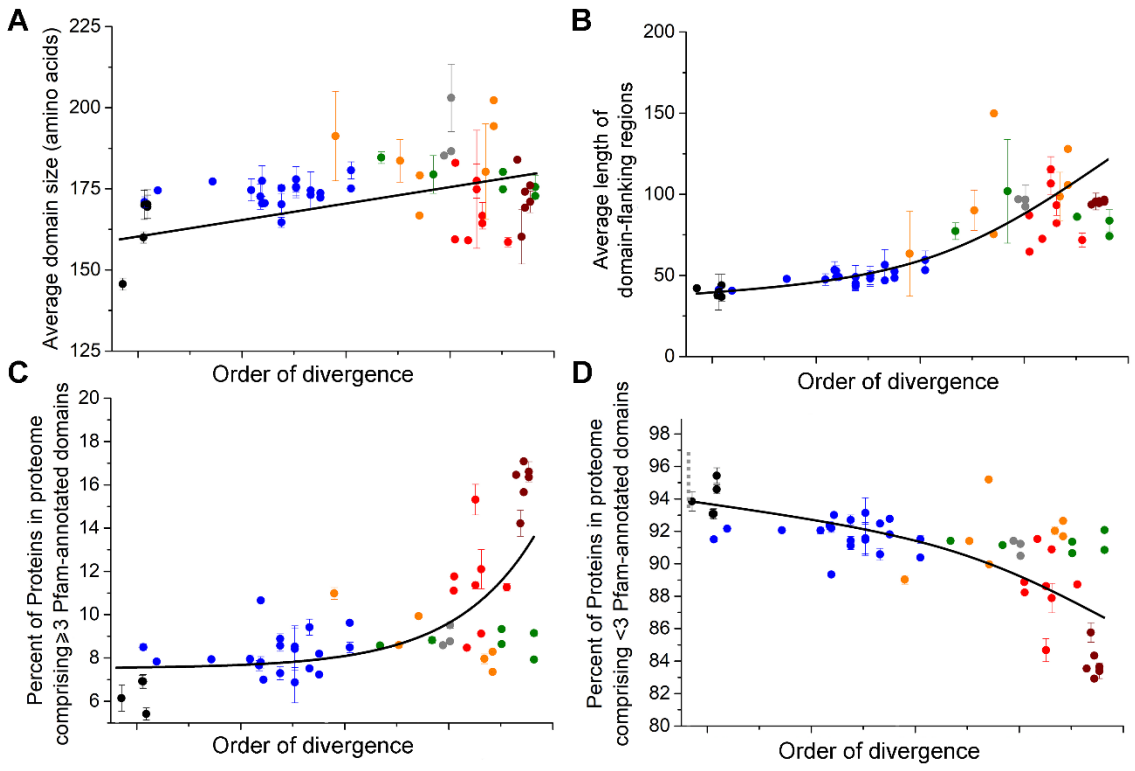


Fig. S1. Figure features follow those of **Fig. 1**. **(A)** Shown on the Y-axis (linear-scale) are the average lengths of the protein domains that belong to 38 ECOD X-groups present throughout the ToL.
(B) Same, for the length of the domain-flanking segments.
(C) Same, for the percent of proteins in the proteome comprising ≥ 3 Pfam-annotated domains.
(D) Same, for the percent of proteins in the proteome comprising < 3 Pfam-annotated domains.

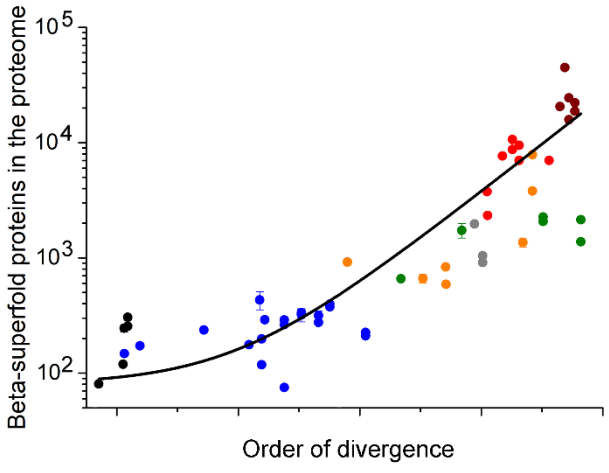


Fig. S2. Figure features follow those of **Fig. 1**. **(A)** Shown on the Y-axis (log-scale) is the number of beta-superfold proteins in the proteome.

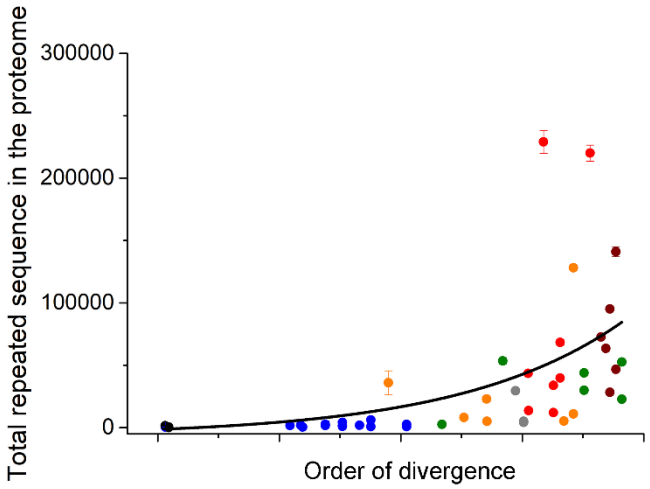


Fig. S3. Figure features follow those of **Fig. 1**. **(A)** Shown on the Y-axis (log-scale) is the total length (number of amino acids) of repeated sequences per proteome.

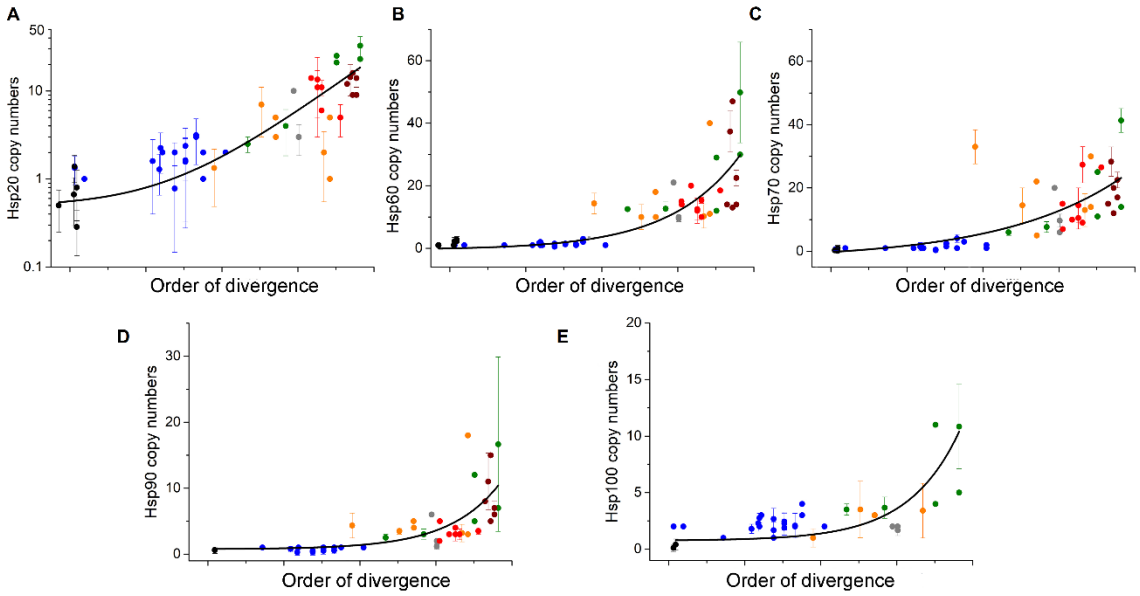


Fig. S4. Figure features follow those of **Fig. 1**. Shown on the Y-axis is the average number, per genome, of genes encoding **(A)** HSP20 (note the log-scale), **(B)** HSP60, **(C)** HSP70, **(D)** HSP90, and **(E)** HSP100.

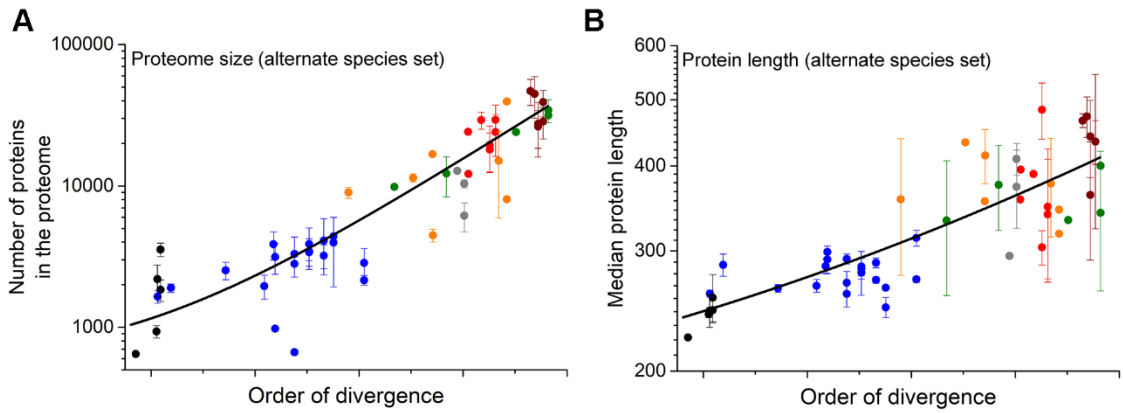


Fig. S5. Figure features follow those of **Fig. 1**. Shown on the Y-axis (linear-scale) is the #proteins per proteome **(A)**, and the median protein length **(B)**, for the alternative species set.

Supplementary Tables

Table S1. Protein trees (newick format) of core and co-chaperone families. These trees were inferred by using the Maximum Likelihood method. Initial trees for the heuristic search were obtained by NJ/BioNJ algorithm, using JTT distance model.

HSP20

((((((((((((((((((((((((Q3IJ58_Gammaproteobacteria:0.11444058,Q3IJ57_Gammaproteobacteria:0.06923611):0.15248496[0.9700],Q47Y72_Gammaproteobacteria_:0.32911559):0.10694500[0.7900],(Q9KVX0_Gammaproteobacteria:0.31956262,(Q8EET9_Gammaproteobacteria:0.25902522,Q7CPF1_Gammaproteobacteria_:0.33057154):0.04154999[0.4100]):0.09539706[0.6400]):0.07878871[0.5800],A4GAF6_Betaproteobacteria:0.38875517):0.08094470[0.6200],Q5LMS8_Alphaproteobacteria:0.39798764):0.05989039[0.1800],(Q9A637_Alphaproteobacteria:0.34958255,(Q5FRB5_Alphaproteobacteria:0.31269259,B3PU08_Alphaproteobacteria:0.38761297):0.28689580[0.9500]):0.08742432[0.0700]):0.06720282[0.0900],Q2RVH4_Alphaproteobacteria:0.37789248):0.10272163[0.1800],(B3PZW3_Alphaproteobacteria:0.39084084,(Q8ZL03_Gammaproteobacteria_:0.58827936,Q5FNB9_Alphaproteobacteria:0.50689915):0.08836353[0.2100]):0.09615326[0.0500]):0.03782656[0.0200],Q5QUR5_Gammaproteobacteria:0.45842963):0.08883975[0.0400],(Q9A2G9_Alphaproteobacteria:0.39650066,(B3PZA6_Alphaproteobacteria:0.35783508,(Q2RMU6_Alphaproteobacteria:0.44668625,(Q5FPX0_Alphaproteobacteria:0.31354848,Q5LV07_Alphaproteobacteria:0.29241684):0.13405965[0.6500]):0.15743591[0.7500]):0.09907613[0.4300]):0.44309075[0.9900]):0.06956799[0.0500],(Q9HZ98_Gammaproteobacteria_:0.43267433,Q5ZTH1_Gammaproteobacteria:0.45722728):0.19711590[0.5400]):0.11019233[0.1100],(Q4FP87_Alphaproteobacteria:0.77623686,(Q2NAA8_Alphaproteobacteria:0.45869360,Q2NBM0_Alphaproteobacteria:0.33927561):0.25416882[0.8000]):0.04300368[0.1200]):0.19307478[0.6000],Q8ZPY6_Gammaproteobacteria_:0.91841419):0.52263206[0.6800],A0LHQ8_Deltaproteobacteria:1.00175896):0.21788683[0.0200],((((Q9RTR5_Deinococci:1.27188290,Q5KXZ2_Bacilli:1.10250256):0.10383775[0.1200],E1QLZ6_Deltaproteobacteria:1.18273167):0.13433264[0.0100],F3YW99_Deltaproteobacteria:0.89428543):0.17271674[0.0000],((Q47QL7_Actinobacteria:0.90693241,Q47PQ5_Actinobacteria:0.61825356):0.20433293[0.4400],(B9LTV5_Euryarchaeota:1.07186334,(B9LSK9_Euryarchaeota:0.42103464,A0A0F7PAY5_Euryarchaeota:0.44227821):0.19666682[0.7300]):0.67172691[0.7900]):0.16488224[0.0800]):0.15846829[0.0000]):0.07877238[0.0000],((P96698_Bacilli:1.36240858,O34321_Bacilli:0.98397267):0.37444495[0.2500],(C0ZG78_Bacilli:1.55364587,(Q3JBB5_Gammaproteobacteria:1.06278297,F2NFQ0_Deltaproteobacteria:0.59473917):0.31449180[0.5000]):0.05148944[0.0000]):0.31539361[0.0000]):0.07607706[0.0000],(Q74MA9_DPANN:0.89219579,(B3E7K3_Deltaproteobacteria:0.70763751,A1AS89_Deltaproteobacteria:0.80190647):0.28821559[0.4500]):0.38201366[0.1100]):0.21558576[0.0000],((((Q8DKI6_Cyanobacteria:0.14187677,B0JSS2_Cyanobacteria:0.71730820):0.14197780[0.4800],B0JUE5_Cyanobacteria:0.40790540):0.18519460[0.7700],Q7NH34_Cyanobacteria:1.11932970):0.41964818[0.3500],((((Q3SIX7_Betaproteobacteria:0.78786390,Q607M7_Gammaproteobacteria:0.42957511):0.11622165[0.2100],Q3J7L2_Gammaproteobacteria:0.49717397):0.51948233[0.9400],B3EA90_Deltaproteobacteria:1.03680843):0.28912446[0.1100],(((((Q3ADV9_Clostridia:0.15605677,Q3A9H1_Clostridia:0.08593047):0.69319574[1.0000],Q2FLP2_Euryarchaeota:1.22691303):0.19542822[0.0000],(Q2IDH5_Deltaproteobacteria:0.96504543,Q67NK6_Clostridia:0.64127035):0.14606051[0.2400]):0.09520148[0.0000],O67316_Aquificae:0.58875516):0.08411968[0.0000],((C1DVH9_Aquificae:0.48236973,F0S317_Aquificae:0.86357364):0.17344310[0.1100],((Q24UE1_Clostridia:0.66201514,Q5KY05_Bacilli:0.49832306):0.25284304[0.2800],(Q3A9G8_Clostridia:0.53063102,(Q2RFK2_Clostridia:0.44954648,A5D2M5_Clostridia:0.40618185):0.16536404[0.6100]):0.13984203[0.2400]):0.10155737[0.1500]):0.11064718[0.0200]):0.07791197[0.0000]):0.05836239[0.00

00]):0.12855959[0.0000]):0.05018037[0.0000],(((Q5ZSM4_Gammaproteobacteria:0.77176411,A0A454Y2I5_Gammaproteobacteria:0.60648074):0.13225789[0.1100],B3PU09_Alphaproteobacteria:0.81640280):0.22863514[0.0400],(Q2NBM1_Alphaproteobacteria:0.68534345,(Q2SF12_Gammaproteobacteria:0.64332404,(Q5ZTH0_Gammaproteobacteria:0.43301608,(A0LM85_Deltaproteobacteria:0.31462919,A0A1Q9MV28_Asgard:1.10672098):0.06687053[0.0700]):0.29350598[0.3900]):0.14850749[0.1200]):0.13993414[0.0600]):0.25589633[0.0000]):0.05678150[0.0000],((((((B3EA88_Deltaproteobacteria:0.13663716,A1AKS9_Deltaproteobacteria:0.11649759):0.68513560[1.0000],(E1QFD8_Deltaproteobacteria:0.54712716,A8ZX67_Deltaproteobacteria:0.49379977):0.17204099[0.6200]):0.06063313[0.1100],F2NE20_Deltaproteobacteria:0.61210765):0.10961453[0.1800],(A0LJ27_Deltaproteobacteria:0.54632298,(A0LKU5_Deltaproteobacteria:0.44272514,F3Z176_Deltaproteobacteria:0.86159792):0.09129255[0.1500]):0.08719991[0.1500]):0.10524493[0.1300],(A8ZX51_Deltaproteobacteria:1.36413643,(Q607M9_Gammaproteobacteria:0.47145034,Q0KBR3_Betaproteobacteria:0.42864486):0.28318449[0.9800]):0.28274629[0.0400]):0.12826099[0.0200],((((Q6AM30_Deltaproteobacteria:0.49476772,Q2SQW1_Gammaproteobacteria:0.40741012):0.19987599[0.8600],(A5D0B6_Clostridia:0.57227808,F2LWB5_Deltaproteobacteria:0.48581395):0.06634968[0.0900]):0.11083877[0.0600],A8ZW92_Deltaproteobacteria:0.89654257):0.02349461[0.0000],(((((A1ATK1_Deltaproteobacteria:0.25567629,B3E7I6_Deltaproteobacteria:0.13397407):0.23389518[0.9700],F3YVM2_Deltaproteobacteria:0.66440104):0.17196044[0.1600],Q2IHQ7_Deltaproteobacteria:0.50500596):0.05794307[0.0800],(F2NGK0_Deltaproteobacteria:0.43329808,Q9WYK7_Thermotogae:0.55382354):0.16340503[0.3200]):0.07932068[0.0200],((Q2IP34_Deltaproteobacteria:0.57333815,Q1DDC1_Deltaproteobacteria:0.34647595):0.47847677[0.9000],((((Q1D4I0_Deltaproteobacteria:0.51647750,A9FKI0_Deltaproteobacteria:0.63785607):0.08425027[0.1800],D0LQ99_Deltaproteobacteria:0.62745376):0.13535230[0.1600],A9GVB5_Deltaproteobacteria:0.84346373):0.25732056[0.3200],(Q83CE9_Gammaproteobacteria:0.55839162,(Q3J7L3_Gammaproteobacteria:0.41331473,(Q8PBS2_Gammaproteobacteria:0.62058653,(Q82T55_Betaproteobacteria:0.32457648,Q604K7_Gammaproteobacteria:0.24673370):0.13166678[0.8300]):0.09004416[0.2500]):0.13877724[0.5900]):0.24854943[0.6600]):0.07707733[0.0600]):0.08046990[0.0200]):0.04757602[0.0000]):0.11481576[0.0000]):0.08175461[0.0000]):0.11796270[0.0000],(((Q2FLI0_Euryarchaeota:1.60739499,E0UTR0_Epsilonproteobacteria:0.45972246):0.32378397[0.0600],Q2FQN2_Euryarchaeota:0.73340792):0.23148519[0.0400],((((A1SVR8_Gammaproteobacteria:0.00694370,A1SV28_Gammaproteobacteria:0.17558577):0.85552885[1.0000],(Q891B1_Clostridia:0.53898296,Q45058_Bacilli:1.44546655):0.27473919[0.1700]):0.06691470[0.0100],(Q24R86_Clostridia:0.67920869,(Q49ZX5_Bacilli:0.52895987,Q49ZX4_Bacilli:0.61863030):0.31384015[0.7400]):0.15829979[0.1400]):0.12968549[0.0100],(((((((A4G7H8_Betaproteobacteria:0.15545474,A4G7K5_Betaproteobacteria:0.28426556):0.31516019[0.9800],Q0KA57_Betaproteobacteria:0.68857266):0.09779057[0.1300],Q3SGI5_Betaproteobacteria:0.22987508):0.19330627[0.2800],Q47IE4_Betaproteobacteria:0.33245439):0.14540269[0.5700],Q3SJ47_Betaproteobacteria:0.60772000):0.37054100[0.6800],Q6MJP8_Oligoflexia:0.70586842):0.12138987[0.0300],((((D1AZG1_Epsilonproteobacteria:0.45355141,E0UTR2_Epsilonproteobacteria:0.41205860):0.16243448[0.5100],(E6X2Y0_Epsilonproteobacteria:0.32553941,B9L6U2_Epsilonproteobacteria:0.56677954):0.05580236[0.1200]):0.10445825[0.5400],(E0UTR1_Epsilonbacteria:0.45757050,E6X2Y1_Epsilonproteobacteria:0.38046809):0.50690501[0.9800]):0.14590817[0.3300],(((B4SEH4_Chlorobia:0.17870218,A1BIT0_Chlorobia:0.18433775):0.08163182[0.7200],B4S543_Chlorobia:0.13669290):0.26294930[1.0000],(B4S6P8_Chlorobia:0.16428618,(A1BET9_Chlorobia:0.25231236,B4SD99_Chlorobia:0.18674679):0.09410733[0.6000]):0.30275446[1.0000]):0.22931923[0.7900]):0.07392737[0.1000]):0.08798254[0.0000]):0.06527090[0.0000]):0.15964789[0.0000]):0.09955893[0.0000],(Q6MPD8_Oligoflexia:0.69989998,(Q01QQ8_Acidobacteria:0.80409650,A9FA72_Deltaproteobacteria:0.82055531):0.18801228[0.1800]):0.36661968[0.0500]):0.00000000[0.0000],((O53673_Actinobacteria:1.03332797,Q67QW2_Clostridia:0.75892013):0.12285868[0.0800],((Q82QP6_Actinobacteria:0.21499914,Q82Q29_Actinobacteria:0.15115

834):0.37536927[1.0000],(Q5Z1K6_Actinobacteria:0.41274012,(Q6ACF4_Actinobacteria:0.000
00000):0.30013986[1.0000]):0.12342038[0.5800]):0.37403790[0.9800]):0.25138217[0.0400]):0.
00000000[0.0000],P9WMK1_Actinobacteria:1.19905526):0.13837647[0.0000],(C1DXW7_Aqui
ficae:0.79252203,(B9LSA4_Euryarchaeota:0.68377244,(B9LSA3_Euryarchaeota:0.60554268,A
0A0F7PCS3_Euryarchaeota:0.43106473):0.12424997[0.3900]):0.15848105[0.3800]):0.1576711
7[0.0700],(((O28308_Euryarchaeota:0.52572009,B1L7F8_TACK:0.78155996):0.35791610[0.27
00],(Q60A86_Gammaproteobacteria:0.75492366,Q3JAU4_Gammaproteobacteria:0.41823864):0
.48362382[0.8100]):0.16824943[0.0100],(A0A1Q9N6X1_Asgard:1.17525957,(((((((((F2NI12_D
eltaproteobacteria:0.49890487,A1SUB1_Gammaproteobacteria:0.31474917):0.64174762[0.9900
],Q1D8N2_Deltaproteobacteria:0.80150271):0.19775796[0.6000],B4SBI9_Chlorobia:0.7279902
8):0.61946532[0.8600],Q8ZTN4_TACK:0.62419459):0.19940802[0.0400],(B1L3F9_TACK:0.4
7280720,(Q9YAJ0_TACK:0.42574035,(Q97W19_TACK:0.37877848,Q974V6_TACK:0.13808
390):0.16429930[0.9500]):0.19247596[0.9800]):0.14113364[0.6300]):0.08868491[0.0400],A0A
1Q9N5R8_Asgard:0.83551802):0.05527982[0.0000],((Q9HKX2_Euryarchaeota:0.58432373,Q6
L1Z6_Euryarchaeota:0.33336247):0.61752216[0.9800],(O28973_Euryarchaeota_:0.27964992,(B
9LSJ8_Euryarchaeota:0.91835692,((Q8PXG4_Euryarchaeota:0.09073786,Q8THE4_Euryarchaeo
ta_:0.03607055):0.50167673[1.0000],(Q12TD1_Euryarchaeota_:0.77431704,(Q8PXG3_Euryarc
haeota:0.07855516,Q8THE3_Euryarchaeota:0.02321810):0.33485945[1.0000]):0.08321486[0.33
00]):0.19925202[0.5500]):0.31598196[0.2800]):0.13484992[0.1100]):0.09544725[0.0100]):0.41
167684[0.0400],(Q6LZE5_Euryarchaeota:0.57123722,(F6BBN6_Euryarchaeota_:0.08777067,Q
57733_Euryarchaeota:0.23349603):0.07298206[0.7400]):0.76439242[1.0000]):0.11006033[0.00
00],((((Q8TUD5_Euryarchaeota:0.04333917,Q8PX03_Euryarchaeota:0.08179882):0.28366742[0
.9900],(Q8TK35_Euryarchaeota_:0.14647739,Q8PZK9_Euryarchaeota:0.06536104):0.48389129
[1.0000]):0.14430664[0.5000],Q12UL4_Euryarchaeota_:0.41978804):0.35998608[0.9300],((R1
G331_DPANN:0.73528606,A0A218NM22_DPANN:0.70517563):0.29448645[0.1500],(E3GX5
6_Euryarchaeota_:0.79751762,(((((E7FI46_Euryarchaeota:0.07501140,O59514_Euryarchaeota_:
0.09721779):0.38576590[0.9800],B1L3V3_TACK:1.06726381):0.06654094[0.0800],(Q8PYA3_
Euryarchaeota:0.43297027,(Q8TK30_Euryarchaeota_:0.12073386,Q8PZK5_Euryarchaeota:0.09
715556):0.42430531[0.9900]):0.84680093[0.9900]):0.11936280[0.0200],(Q8ZVI0_TACK:1.373
61190,A0A1Q9N902_Asgard:0.44268412):0.25759370[0.0300]):0.11585170[0.0000],(((Q9YFZ
9_TACK:1.32727639,I0A078_TACK:1.40026812):0.12787204[0.2000],Q8ZTS4_TACK:0.7456
5761):0.43645778[0.0700],(Q2NH97_Euryarchaeota:1.27541697,(Q8TZC8_Euryarchaeota_:1.2
4064414,((Q970D9_TACK:0.44691272,Q97VL9_TACK:0.24389817):0.28055764[0.9400],(Q9
HJU9_Euryarchaeota:0.77065536,Q6L128_Euryarchaeota:0.54827338):0.25389493[0.4400]):0.3
8250698[0.7900]):0.14389841[0.0600]):0.17344983[0.0000]):0.14167364[0.0000]):0.14031666[
0.0000]):0.13455951[0.0000]):0.12133851[0.0000]):0.12129636[0.0000]):0.08685243[0.0000]):
0.06498201[0.0000]):0.17132212[0.0000])OROOT;

HSP60

((((((((((((((((((P0A1D3_Gammaproteobacteria:0.00408410,P0A6F5_Gammaproteobacteria:0.00
000000):0.04561213[1.0000],Q59687_Gammaproteobacteria:0.04524143):0.03331623[0.8600],(
Q9XAU7_Gammaproteobacteria:0.07154011,Q8CX48_Gammaproteobacteria:0.06006211):0.02
981155[0.3600]):0.01034257[0.0800],Q9KNR7_Gammaproteobacteria:0.04864753):0.00999068
[0.0800],((Q487Q9_Gammaproteobacteria:0.11077307,Q5QVT4_Gammaproteobacteria:0.07298
282):0.02825613[0.2000],(Q9KLC6_Gammaproteobacteria:0.18760641,(A1ST72_Gammaproteo
bacteria:0.00000000,A1SXK4_Gammaproteobacteria:0.00405317):0.11033418[1.0000]):0.02866
845[0.3600]):0.01884868[0.1000]):0.05739394[0.7200],((P19421_Gammaproteobacteria:0.1361
9674,Q5ZXP3_Gammaproteobacteria:0.13630194):0.02151469[0.1000],(Q2SDG0_Gammaprote
obacteria:0.07740545,P30718_Gammaproteobacteria:0.08902615):0.03274745[0.8200]):0.01454
603[0.1600]):0.02353324[0.1000],(((Q3J729_Gammaproteobacteria:0.09434654,Q8PD23_Gam

maproteobacteria:0.10457598):0.02755966[0.4600],Q60AY0_Gammaproteobacteria:0.08315316):0.01556536[0.2000],(Q607Q3_Gammaproteobacteria:0.10585536,(Q82Y60_Betaproteobacteria:0.07173446,(Q3SMK1_Betaproteobacteria:0.07963452,(Q47IZ8_Betaproteobacteria:0.06882175,((Q0KDR7_Betaproteobacteria:0.04743207,A4G837_Betaproteobacteria:0.07416956):0.02543099[0.9200],(Q5F541_Betaproteobacteria:0.09692249,(C1D8S1_Betaproteobacteria:0.03958181,C1DD18_Betaproteobacteria:0.02234271):0.06397113[1.0000]):0.03316005[0.7000]):0.01437601[0.3000]):0.02534230[0.4200]):0.01457229[0.3400]):0.06189908[1.0000]):0.01840906[0.1800]):0.02637930[0.2200]):0.03122099[0.1400],(Q5NEE1_Gammaproteobacteria:0.17674633,Q4FU94_Gammaproteobacteria:0.16848075):0.03244818[0.1000]):0.02837473[0.0600],A5EX17_Gammaproteobacteria:0.10458871):0.07047577[0.9600],(((Q4FPA5_Alphaproteobacteria:0.16304251,Q5LV15_Alphaproteobacteria:0.13222507):0.04124966[0.5600],Q2N5R9_Alphaproteobacteria:0.14555428):0.02069739[0.0400],((Q2RWV4_Alphaproteobacteria:0.06186009,Q2RY28_Alphaproteobacteria:0.04804867):0.06245735[1.0000],((Q5FPQ6_Alphaproteobacteria:0.11527398,P0CAT9_Alphaproteobacteria:0.12307623):0.02161190[0.1400],(B3PRB6_Alphaproteobacteria:0.09146754,(Q2NBL8_Alphaproteobacteria:0.13306597,(B3PU03_Alphaproteobacteria:0.03341306,B3Q351_Alphaproteobacteria:0.04409860):0.04716909[1.0000]):0.02618178[0.6000]):0.02643169[0.3400]):0.03036526[0.1200]):0.01079815[0.0200]):0.09869900[0.9200]):0.03725831[0.3600],(((A1AST1_Deltaproteobacteria:0.05916579,B3E8G0_Deltaproteobacteria:0.05687183):0.05985600[1.0000],(A9GB11_Deltaproteobacteria:0.21132548,Q6MRI1_Oligoflexia:0.17049777):0.06787408[0.4400]):0.02811102[0.1600],((D0LRR3_Deltaproteobacteria:0.09246574,D0LYI5_Deltaproteobacteria:0.17093560):0.05386257[0.5200],(A9GBQ2_Deltaproteobacteria:0.15293368,(((Q2IFK0_Deltaproteobacteria:0.05181650,Q1D2S1_Deltaproteobacteria:0.07713715):0.01288879[0.4800],(Q2IKI2_Deltaproteobacteria:0.10341274,Q1D3Y5_Deltaproteobacteria:0.11571180):0.07569456[1.0000]):0.03400360[0.8000],((A8ZU48_Deltaproteobacteria:0.11596416,Q6ARV6_Deltaproteobacteria:0.27595597):0.02020483[0.1200],(((F3YTT3_Deltaproteobacteria:0.09244111,Q1MQP8_Deltaproteobacteria:0.10114057):0.05967274[0.9600],E1QE82_Deltaproteobacteria:0.08128274):0.02111652[0.3200],(A0LKS4_Deltaproteobacteria:0.11639792,(F2ND31_Deltaproteobacteria:0.09663187,A0LEH2_Deltaproteobacteria:0.10182119):0.02585160[0.5600]):0.03685660[0.5600]):0.01427590[0.2800]):0.02538140[0.6200]):0.02578438[0.5800]):0.01679456[0.2000]):0.01927238[0.3000]):0.06008837[0.8200]):0.02347164[0.6600],P0C0Z7_Chlamydia:0.26598693):0.02148043[0.2000],((((O67943_Aquificae:0.12701934,C1DX20_Aquificae:0.09878925):0.09433556[1.0000],F0S3Q6_Aquificae:0.15068288):0.06259634[0.9200],F2LXB3_Deltaproteobacteria:0.21523512):0.03343721[0.2000],(B9L698_Epsilonproteobacteria:0.10257388,(D1B118_Epsilonproteobacteria:0.08504090,(D3UGP1_Epsilonproteobacteria:0.11697633,(E0UQ31_Epsilonproteobacteria:0.13618173,E6X267_Epsilonproteobacteria:0.10674176):0.04123098[0.6800]):0.03307506[0.6400]):0.04025518[0.9000]):0.10696550[1.0000]):0.05563949[0.4400]):0.00772571[0.0600],(P42375_Bacteroidia:0.22513246,(B4S6H2_Chlorobia:0.02745060,(A1BHS5_Chlorobia:0.01735924,B4SEN1_Chlorobia:0.02463768):0.03803237[0.9600]):0.10884796[1.0000]):0.06496963[0.7800]):0.01528204[0.1800],(((H9UFL5_Spirochaetia:0.10084114,E1R5L0_Spirochaetia:0.07980561):0.05162803[0.9400],Q73NH7_Spirochaetia:0.17068546):0.11196108[1.0000],(((Q8PW06_Euryarchaeota:0.01019520,Q8TGX7_Euryarchaeota:0.03810349):0.17384977[1.0000],Q2FPN5_Euryarchaeota:0.16852479):0.22619766[1.0000],(Q9WYX6_Thermotogae:0.20361358,((Q9RWQ9_Deinococci:0.22277362,Q2RL13_Clostridia:0.17138666):0.03452670[0.2200],(((((((Q9AFA6_Actinobacteria:0.03988183,P9WPE7_Actinobacteria:0.03984894):0.03179162[0.5000],Q6NF74_Actinobacteria:0.09988275):0.03222443[0.9600],(Q6ADK2_Actinobacteria:0.08622537,P69205_Actinobacteria:0.12740980):0.06465843[1.0000]):0.02427517[0.8400],Q82GG6_Actinobacteria:0.05545686):0.01848571[0.5600],Q47TE8_Actinobacteria:0.11915137):0.09712874[1.0000],((Q47LP1_Actinobacteria:0.12896991,Q82DI5_Actinobacteria:0.09717849):0.04770010[0.7600],(Q6NJ37_Actinobacteria:0.26034580,(Q5Z1F9_Actinobacteria:0.07572252,P9WPE9_Actinobacteria:0.19245627):0.07181242[0.8800]):0.17130695[1.0000]):0.07714442

[0.9800]):0.07136682[1.0000],((Q7MBB4_Gloeobacteria:0.07215210,B0JMZ9_Cyanobacteria:0.21919106):0.02616054[0.7400],(Q7MBC7_Gloeobacteria:0.11985192,B0JUI2_Cyanobacteria:0.20529197):0.05177751[0.9000]):0.10133321[1.0000]):0.02721414[0.2800],(Q2RGL8_Clostridia:0.19885199,(Q67KB8_Clostridia:0.18129327,(((A5CZ03_Clostridia:0.11472969,Q3ADX3_Clostridia:0.06993073):0.03745395[0.7800],Q24QE3_Clostridia:0.12229118):0.01594985[0.3600],((Q891G4_Clostridia:0.17166014,Q8R5X7_Fusobacteria:0.25322225):0.04979142[0.2000],(C0ZK52_Bacilli:0.08264749,(B1YEP6_Bacilli:0.07565751,(Q5L3E6_Bacilli:0.06565092,((P28598_Bacilli:0.04391295,A0A127VUF5_Bacilli:0.08669500):0.01666786[0.5800],(Q49YY5_Bacilli:0.15942645,(F2I7V2_Bacilli:0.18691668,(Q93EU6_Bacilli:0.05153995,P37282_Bacilli:0.14712380):0.04815375[0.7400]):0.05425494[0.9200]):0.04882190[0.6200]):0.01446481[0.1400]):0.02104471[0.2800]):0.03032574[0.5400]):0.06491146[0.9400]):0.05354378[0.4400]):0.01726790[0.0600]):0.03199027[0.3000]):0.01959230[0.1400]):0.02000206[0.1200]):0.04050980[0.2200]):0.03042084[0.3800]):0.03874532[0.2800]):0.03283454[0.1000]):0.05662375[0.2600],((A1SYD3_Gammaproteobacteria:0.43007691,P78012_Mollicutes:0.49880966):0.07174034[0.5000],(Q0KA74_Betaproteobacteria:0.36842666,(O84609_Chlamydia:1.05799008,O84760_Chlamydia:1.24151304):0.35915822[0.8800]):0.15530266[0.1600]):0.09940938[0.0600]):0.04748337[0.0400],Q7WZ32_Gammaproteobacteria:0.86804485):0.03984417[0.2000],Q3YRH4_Alphaproteobacteria:0.49485754,(((A5D3X9_Clostridia:0.30224740,Q3AF10_Clostridia:0.62015725):0.17366463[0.9400],Q7NEX9_Gloeobacteria:0.58742848):0.37695744[1.0000],(((Q8THX2_Euryarchaeota:0.04884721,Q8TSE1_Euryarchaeota:0.02168419):0.82947815[1.0000],A0A1Q9MSA6_Asgard:0.96781703):0.23761628[0.7400],(((((B9LU39_Euryarchaeota:0.00271495,A0A0F8CLJ9_Euryarchaeota:0.02464329):0.09562009[1.0000],A0A0F7PBF4_Euryarchaeota:0.08939892):0.11824626[1.0000],(B9LRY6_Euryarchaeota:0.88002292,A0A0F7PDM5_Euryarchaeota:0.45845737):0.12248572[0.2200]):0.05760703[0.0000],(A0A0F7PE45_Euryarchaeota:0.13729536,(B9LSV4_Euryarchaeota:0.01401534,A0A0F8AY03_Euryarchaeota:0.01281373):0.13994856[1.0000]):0.24889391[1.0000]):0.09500664[0.0400],(((((((E3GWZ0_Euryarchaeota:0.12751623,E3GWZ6_Euryarchaeota:0.03358360):0.09657315[1.0000],Q2NHV0_Euryarchaeota:0.16173669):0.04692348[0.9400],(F6BAU4_Euryarchaeota:0.05642699,(Q6LX38_Euryarchaeota:0.13998415,Q58405_Euryarchaeota:0.09143117):0.02375150[0.3800]):0.11191758[1.0000]):0.03760974[0.6200],P50016_Euryarchaeota:0.11315266):0.02506163[0.2400],((Q74N99_Nanoarchaeota:0.10912424,R1E4Z3_DPANN:0.29108124):0.13360382[0.7200],(O57762_Euryarchaeota:0.03701211,Q8TZL6_Euryarchaeota:0.04998157):0.11330996[1.0000]):0.04906707[0.1400]):0.02718634[0.2000],(O28045_Euryarchaeota:0.13497996,O28821_Euryarchaeota:0.13695567):0.06989951[0.9800]):0.01655233[0.3000],((((((Q8PX43_Euryarchaeota:0.01240889,Q8TUI1_Euryarchaeota:0.02573627):0.09445429[1.0000],Q12U60_Euryarchaeota:0.11508726):0.06597595[1.0000],(Q12UN6_Euryarchaeota:0.21243183,(Q8PXX0_Euryarchaeota:0.04153231,Q8THU8_Euryarchaeota:0.02385642):0.11544536[1.0000]):0.09405790[1.0000]):0.05404082[0.9800],Q2FTL9_Euryarchaeota:0.19662174):0.07713254[0.9000],((A0A2D6K5L1_Nanoarchaeota:0.24694327,A0A2D6LPH3_DPANN:0.27658917):0.06862853[0.4800],((Q6KZS2_Euryarchaeota:0.10393662,P48425_Euryarchaeota:0.11216095):0.21366551[1.0000],(Q6L132_Euryarchaeota:0.11291070,P48424_Euryarchaeota:0.07316794):0.15674883[1.0000]):0.08407783[0.7000]):0.02782640[0.1000]):0.02255479[0.0000],(((((Q8Q0R4_Euryarchaeota:0.05675425,Q8TQ70_Euryarchaeota:0.05639981):0.31936373[1.0000],Q12XB3_Euryarchaeota:0.29984813):0.19422787[1.0000],Q2FPE0_Euryarchaeota:0.45104625):0.06577951[0.3000],Q2NHT5_Euryarchaeota:0.69720307):0.06425538[0.1200],((((A0A0F8VDG3_Asgard:0.29247472,A0A0F8VQ95_Asgard:0.58817982):0.10382298[0.1600],A0A1Q9NRV4_Asgard:0.37253987):0.02979765[0.0400],A0A218NP95_DPANN:0.36811868):0.03881078[0.0400],(((((A0RYP4_TACK:0.24249893,A0RZ38_TACK:0.19550867):0.15964700[1.0000],A0A1Q9NRK4_Asgard:0.31527882):0.05883653[0.2400],A0A1Q9N564_Asgard:0.19172480):0.02061826[0.0200],(A0A1Q9N4V4_Asgard:0.25214917,(A0A0F8W5M5_Asgard:0.24684077,A0A1Q9P8V1_Asgard:0.23627647):0.05775551[0.3400]):0.03304208[0.0400]):0.03018537[0.1000],(

B1L6H6_TACK:0.34636165,(((((O24735_TACK:0.04156222,Q9V2T8_TACK:0.10036874):0.13744609[1.0000],I0A243_TACK:0.16949493):0.04473125[0.7000],Q9YA66_TACK:0.21331627):0.05631958[0.5400],Q8ZVU7_TACK:0.27336805):0.08076615[0.9200],(B1L720_TACK:0.36446315,(Q8ZTF8_TACK:0.25510787,(I0A1X5_TACK:0.18812232,(Q9YDK6_TACK:0.20496912,((F9VN97_TACK:0.28268570,Q9V2T7_TACK:0.39037966):0.12065823[0.9000],(O24734_TACK:0.07845424,Q9V2S9_TACK:0.07919724):0.07628575[1.0000]):0.05790841[0.6200]):0.04579251[0.3000]):0.05917007[0.5600]):0.09486952[0.8200]):0.07287596[0.3800]):0.02134249[0.0200]):0.06822151[0.1000]):0.02902848[0.0200]):0.02602922[0.0200]):0.03022876[0.0000]):0.01942117[0.0000]):0.06310628[0.1000]):0.13019390[0.3000]):0.27117567[0.9800]):0.79574325[1.0000])OROOT;

HSP70

((((((((((((((((((((Q486F9_Gammaproteobacteria:0.12628484,Q607A5_Gammaproteobacteria:0.08674305):0.03477286[0.7000],(Q5NFG7_Gammaproteobacteria:0.09214408,Q4FPS9_Gammaproteobacteria:0.15879738):0.02813907[0.6000]):0.02365830[0.2700],(Q5QXL1_Gammaproteobacteria:0.11502507,Q8PAK9_Gammaproteobacteria:0.09633197):0.03595355[0.7100]):0.02664858[0.2500],Q5F6W5_Betaproteobacteria:0.09770327):0.07901616[0.9800],((Q5FSL5_Alphaproteobacteria:0.10942575,Q3YRR6_Alphaproteobacteria:0.15408386):0.04619397[0.7500],(Q4FNP9_Alphaproteobacteria:0.10882343,(P20442_Alphaproteobacteria:0.12784286,Q2NAI8_Alphaproteobacteria:0.09602403):0.02876835[0.6500]):0.02563914[0.4800]):0.02865586[0.6900]):0.04405947[0.8000],(((D0LKN8_Deltaproteobacteria:0.20380710,A9GHU1_Deltaproteobacteria:0.17963067):0.04300466[0.7700],Q2IHN2_Deltaproteobacteria:0.13403388):0.06504408[0.7400],((E1QIB0_Deltaproteobacteria:0.18538353,A0LH28_Deltaproteobacteria:0.07913727):0.03267962[0.4200],(A1ANV0_Deltaproteobacteria:0.11167931,(Q6AMQ3_Deltaproteobacteria:0.10945528,F3Z405_Deltaproteobacteria:0.13604713):0.03219054[0.4600]):0.01143626[0.1900]):0.01715356[0.3300]):0.02707887[0.4800]):0.02718682[0.5000],F2LUD1_Deltaproteobacteria:0.17687369):0.02699900[0.2600],((Q73Q16_Spirochaetia:0.14039678,H9UKG8_Spirochaetia:0.10303039):0.03690795[0.6700],(P17821_Chlamydia:0.15258528,(P0C937_Bacteroidia:0.13469039,B4S6P7_Chlorobia:0.13164995):0.04251830[0.5900]):0.03945027[0.4200]):0.02347589[0.2600]):0.01096555[0.1700],(Q6MNF8_Oligoflexia:0.16540696,(E1QDN0_Deltaproteobacteria:0.12994894,(O67118_Aquificae:0.25753056,(Q2II73_Deltaproteobacteria:0.10934283,(D0LSR1_Deltaproteobacteria:0.14369112,A9G9K4_Deltaproteobacteria:0.13518467):0.03515026[0.5600]):0.04583562[0.8500]):0.03246764[0.3100]):0.02761335[0.2200]):0.04639177[0.3900]):0.03298249[0.2200],(B9L8Z0_Epsilonproteobacteria:0.11969199,(D1B2Q8_Epsilonproteobacteria:0.10658948,E0UPC8_Epsilonproteobacteria:0.11073588):0.04402233[0.7100]):0.17725974[1.0000]):0.03062325[0.1000],(B0JW24_Cyanobacteria:0.21450373,(((Q5KWZ7_Bacilli:0.12787507,A5D3Y1_Clostridia:0.09183830):0.02218356[0.3100],P0CW13_Euryarchaeota:0.14724192):0.02748161[0.1000],(Q9WYK6_Thermotogae:0.27586311,((A0A1Q9PB66_Asgard:0.23404912,Q2FQP2_Euryarchaeota:0.13406151):0.06444787[0.4300],((A0A0F8W6S9_Asgard:0.19933750,A0A1Q9NM92_Asgard:0.22202859):0.03102973[0.3900],(A0A218NM16_DPANN:0.21788127,A0RZ01_TACK:0.21671774):0.03494979[0.3400]):0.13371004[1.0000]):0.01676500[0.0500]):0.03221865[0.0300]):0.01756503[0.0100]):0.02784985[0.0100]):0.00000001[0.0000],(((Q6NEY9_Actinobacteria:0.11180311,Q6AC76_Actinobacteria:0.11467530):0.04468022[0.5300],Q82EX9_Actinobacteria:0.07538018):0.14908309[1.0000],(B0JV83_Cyanobacteria:0.19751065,(Q7NDH1_Gloeobacteria:0.07979620,(Q9RY23_Deinococci:0.25922779,(Q826F6_Actinobacteria:0.14519997,A0LEV9_Deltaproteobacteria:0.10828211):0.07954747[0.9900]):0.03912928[0.5300]):0.05039200[0.5000]):0.08833540[0.9900]):0.04094588[0.1900]):0.03320602[0.0900],(Q8PX42_Euryarchaeota:0.37201073,P75344_Mollicutes:0.29659748):0.08256862[0.2800]):0.06881899[0.6000],((((D0LJ04_Deltaproteobacteria:0.14268495,A9G252_Deltaproteobacteria:0.14422207):0.08066937[0.7900],Q2ILC9_Deltaproteobacteria:0.16917627):0.03672798[0.3300],Q1D7Z8_Deltaproteobacteria:0.19

363640):0.13218446[0.9800],((Q2IL30_Deltaproteobacteria:0.09979948,Q1D8Q9_Deltaproteobacteria:0.11071093):0.27762761[1.0000],((Q2II43_Deltaproteobacteria:0.18901787,Q1D547_Deltaproteobacteria:0.07389268):0.30930111[1.0000],(((Q2IIV2_Deltaproteobacteria:0.11904770,Q1D6Q3_Deltaproteobacteria:0.24402073):0.19680290[1.0000],Q1D211_Deltaproteobacteria:0.75243575):0.09699762[0.3200],(Q1CY00_Deltaproteobacteria:0.71800555,(Q2II28_Deltaproteobacteria:0.30754792,(Q2IKD7_Deltaproteobacteria:0.25405401,Q1D5W3_Deltaproteobacteria:0.18666290):0.07100397[0.8000]):0.05858881[0.5800]):0.10353744[0.7900]):0.11217492[0.3800]):0.16001255[0.5900]):0.05751386[0.1500]):0.10206156[0.1600]):0.03437393[0.0700],E1R5V2_Spirochaetia:0.60709763):0.07096276[0.2700],(Q2FPF2_Euryarchaeota:0.63591377,Q1CWT5_Deltaproteobacteria:0.44630402):0.13042132[0.2800]):0.04927911[0.0500],(Q8ZQY6_Gammaproteobacteria:0.71382291,(Q3JDB0_Gammaproteobacteria:0.76502333,(D0LHM1_Deltaproteobacteria:0.18880624,A9F5L9_Deltaproteobacteria:0.20852539):0.34858677[1.0000]):0.27303410[0.9900]):0.15649259[0.1900]):0.04640956[0.0000],D0LK31_Deltaproteobacteria:0.49692358):0.05708877[0.0200],A0A1Q9P181_Asgard:0.85176277):0.03620439[0.0700],(Q3YS56_Alphaproteobacteria:0.56565123,(((Q2INJ1_Deltaproteobacteria:0.22171583,A9EY14_Deltaproteobacteria:0.24028326):0.05560876[0.4600],Q1D2G2_Deltaproteobacteria:0.23198996):0.07976082[0.7700],(Q4FRM9_Gammaproteobacteria:0.40805066,((((((Q0KCG7_Betaproteobacteria:0.14693558,A4G782_Betaproteobacteria:0.11058878):0.04870677[0.8000],Q47EN1_Betaproteobacteria:0.13763045):0.03382564[0.5800],(C1D4P9_Betaproteobacteria:0.15269212,Q5F8E8_Betaproteobacteria:0.23346886):0.04980646[0.3000]):0.02946583[0.2200],Q3SEX4_Betaproteobacteria:0.12152434):0.05203990[0.4600],Q60C59_Gammaproteobacteria:0.17276458):0.05248651[0.6200],(Q51382_Gammaproteobacteria:0.16709360,((Q486Y6_Gammaproteobacteria:0.19195507,A1SUI8_Gammaproteobacteria:0.19041187):0.07325840[0.6900],((Q8ZN42_Gammaproteobacteria:0.12014311,Q9KTX8_Gammaproteobacteria:0.14811897):0.05150894[0.4500],(Q3IFG5_Gammaproteobacteria:0.22712633,(Q9CNV2_Gammaproteobacteria:0.17003664,Q8EEU5_Gammaproteobacteria:0.15041929):0.03967111[0.3700]):0.02871244[0.1200]):0.04097528[0.2500]):0.055348596[0.6900]):0.04364782[0.5100]):0.07269452[0.8500]):0.06568526[0.4600]):0.06356505[0.4500]):0.10971881[0.7400]):0.14880876[0.5500],A9EQI4_Deltaproteobacteria:1.14896302,(D0LXA2_Deltaproteobacteria:0.63649542,(((Q1D1K0_Deltaproteobacteria:0.39415563,A9GKS7_Deltaproteobacteria:0.39010168):0.05972409[0.4700],Q609C9_Gammaproteobacteria:0.39057744):0.09904663[0.6800],(D0LGP7_Deltaproteobacteria:0.40147868,(Q1D1K1_Deltaproteobacteria:0.25874747,((Q2IDZ4_Deltaproteobacteria:0.28509840,D0LGP6_Deltaproteobacteria:0.27268833):0.10190713[0.7900],(Q609D0_Gammaproteobacteria:0.28193925,(Q0KB65_Betaproteobacteria:0.17927984,A9GKQ6_Deltaproteobacteria:0.28646151):0.06711669[0.3000]):0.04549391[0.1700]):0.09291829[0.4300]):0.26718566[1.0000]):0.10893195[0.7500]):0.80559103[1.0000]):0.12130659[0.3600])OROOT;

HSP90

((((((((((((((Q0K8D7_Betaproteobacteria:0.24020977,A4G388_Betaproteobacteria:0.13384985):0.08224963[0.5400],Q47GZ8_Betaproteobacteria:0.13593974):0.05162089[0.5400],Q82TV8_Betaproteobacteria:0.20102094):0.03528014[0.1000],(C1DC60_Betaproteobacteria:0.15977505,Q3SJW8_Betaproteobacteria:0.15237867):0.04318149[0.6200]):0.03914411[0.2200],Q3JAF6_Gammaproteobacteria:0.21997064):0.01545335[0.1000],Q60AK3_Gammaproteobacteria:0.29227070):0.06146970[0.5200],Q83EL0_Gammaproteobacteria:0.27297235):0.04337826[0.2800],(Q5ZVS1_Gammaproteobacteria:0.30534138,(Q2SKD0_Gammaproteobacteria:0.17480389,(Q9I3C5_Gammaproteobacteria:0.06586996,Q8P855_Gammaproteobacteria:0.13483325):0.16593981[1.0000]):0.11743377[1.0000]):0.03398060[0.2600]):0.04209403[0.2400],((((Q47XA7_Gammaproteobacteria:0.17534663,Q3IKQ2_Gammaproteobacteria:0.18008103):0.03219450[0.6000],Q5QWR2_Gammaproteobacteria:0.21767972):0.04898007[0.5800],A1STI2_Gammaproteobacteria:0.24846308):0.02298072[0.4800],(Q8EFF7_Gammaproteobacteria:0.19664261,(P22359_Gammaproteo

bacteria:0.20893872,(Q9CM20_Gammaproteobacteria:0.15417850,(P58480_Gammaproteobacteria:0.03429536,P0A6Z3_Gammaproteobacteria:0.02109758):0.09708175[1.0000]):0.05507428[0.8600]):0.07749593[1.0000]):0.05339411[0.6000]):0.07807216[1.0000]):0.07945529[0.6800],Q5NHT8_Gammaproteobacteria:0.39796840):0.05440426[0.4200],Q4FQZ1_Gammaproteobacteria:0.39299933):0.11091432[0.9800],((((B3Q3C5_Alphaproteobacteria:0.49147539,Q2RYB8_Alphaproteobacteria:0.31107385):0.06760929[0.3000],Q5FS51_Alphaproteobacteria:0.44185598):0.08086864[0.4400],Q3YSL9_Alphaproteobacteria:0.56234223):0.17549263[1.0000],(((D1B2V5_Epsilonproteobacteria:0.28598497,E0UTY3_Epsilonproteobacteria:0.34797126):0.06117241[0.4000],D3UH07_Epsilonproteobacteria:0.41126879):0.06884740[0.6800],(E6X0D4_Epsilonproteobacteria:0.40011154,(P61188_Spirochaetia:0.27361425,(H9UMB7_Spirochaetia:0.17682104,E1RAF0_Spirochaetia:0.19990318):0.05617327[0.9600]):0.10049339[1.0000]):0.07053906[0.5800]):0.17451768[1.0000]):0.06180512[0.4200]):0.10255859[0.5600],((Q1CZI7_Deltaproteobacteria:0.44902521,A9GDV2_Deltaproteobacteria:0.48428625):0.10950025[0.4000],(A8ZUI6_Deltaproteobacteria:0.39699631,(((Q5Z3N4_Actinobacteria:0.17778642,P9WMJ7_Actinobacteria:0.15212580):0.22159331[1.0000],D0LYJ5_Deltaproteobacteria:0.36423103):0.10764673[0.9400],(P61184_Oligoflexia:0.47618673,(A1ANS1_Deltaproteobacteria:0.11700783,B3E438_Deltaproteobacteria:0.10642882):0.19758534[1.0000]):0.06982348[0.4400]):0.05478285[0.3800]):0.09867752[0.7400]):0.06562625[0.2600]):0.06050549[0.3800],(((A1BFP7_Chlorobia:0.18153749,B4SAD1_Chlorobia:0.11997133):0.09783354[0.9600],B4S7B7_Chlorobia:0.14678316):0.39349990[1.0000],(((A0A0F8VG15_Asgard:0.31907583,A0A1Q9N225_Asgard:0.33179790):0.29350692[1.0000],A0A1Q9P0M4_Asgard:0.42273132):0.19559655[1.0000],(F3YV29_Deltaproteobacteria:0.48214949,(Q24VT7_Clostridia:0.29004829,(A5D630_Clostridia:0.24391675,Q6ARM0_Deltaproteobacteria:0.38840911):0.08940400[0.8600]):0.22934474[1.0000]):0.09841673[0.8200]):0.08982817[0.5200]):0.14985155[0.8000]):0.08845893[0.4000],(Q8RGH4_Fusobacteria:0.35753818,(Q894P6_Clostridia:0.18459830,(B1YG43_Bacilli:0.24703828,(P46208_Bacilli:0.12122093,C0Z8X7_Bacilli:0.13982396):0.04962165[0.6600]):0.11459662[0.9400]):0.34752554[1.0000]):0.23505273[1.0000],(Q2SLM3_Gammaproteobacteria:1.44088058,(P0C938_Bacteroidia:0.52534488,(Q7NJL8_Gloeobacteria:0.26705230,B0JQV2_Cyanobacteria:0.32501593):0.27985305[1.0000]):0.41353926[1.0000]):0.36319680[1.0000])OROOT;

HSP100

((((((((((((((((((((Q7CQ01_Gammaproteobacteria:0.00573105,P63284_Gammaproteobacteria:0.01082935):0.06931612[1.0000],Q9CKC0_Gammaproteobacteria:0.07239693):0.02842290[0.6400],(Q8EBE6_Gammaproteobacteria:0.07402385,Q9KU18_Gammaproteobacteria:0.08304122):0.02037376[0.5800]):0.01165063[0.1000],(Q5QTU7_Gammaproteobacteria:0.09079291,(Q3IEB9_Gammaproteobacteria:0.08788386,A1SZN5_Gammaproteobacteria:0.11197873):0.02002769[0.2200]):0.01662109[0.0800]):0.01721229[0.3000],Q47X99_Gammaproteobacteria:0.12557639):0.03142447[0.6000],(Q5NE82_Gammaproteobacteria:0.16662461,Q4FTI2_Gammaproteobacteria:0.14629299):0.04202317[0.2800]):0.01953009[0.0400],((Q3J8K7_Gammaproteobacteria:0.08955015,Q602F7_Gammaproteobacteria:0.08816926):0.02669614[0.7600],(Q83F55_Gammaproteobacteria:0.11697727,Q5ZUP3_Gammaproteobacteria:0.08141811):0.02051712[0.2000]):0.01729494[0.2200]):0.00922194[0.0000],(Q8P6A0_Gammaproteobacteria:0.11914246,(Q2S9V5_Gammaproteobacteria:0.07669923,Q9HVN5_Gammaproteobacteria:0.07386403):0.03338618[0.7200]):0.01412005[0.0400]):0.02774057[0.4000],(A5EWR7_Gammaproteobacteria:0.19611575,((C1DBN1_Betaproteobacteria:0.04087506,Q5F7W9_Betaproteobacteria:0.09423692):0.04522226[0.9800],(A4G6Z4_Betaproteobacteria:0.10339478,(Q0K9H5_Betaproteobacteria:0.05321563,(Q3SKL1_Betaproteobacteria:0.09425827,(Q47CH4_Betaproteobacteria:0.05653255,Q82SD8_Betaproteobacteria:0.05776050):0.01744451[0.6000]):0.02170774[0.4000]):0.01522149[0.4000]):0.02043878[0.4600]):0.05633407[1.0000]):0.03048389[0.7000]):0.11285944[1.0000],(Q9RVI3_Deinococci:0.19291904,((Q6AC89_Actinobacteria:0.12735471,Q83FI1_Actinobacteria:0.27325220):0.

06205127[0.6200],((Q47TH6_Actinobacteria:0.09000678,Q82EU9_Actinobacteria:0.10639257):0.04491070[0.6400],(Q6NF05_Actinobacteria:0.09270861,(Q5YNI7_Actinobacteria:0.06153180,P9WPD1_Actinobacteria:0.08907721):0.02714655[0.5400]):0.06370308[0.9600]):0.03934798[0.5000]):0.10864844[0.8800]):0.02513329[0.3200]):0.03645877[0.2200],(((((((Q5FQY4_Alphaproteobacteria:0.13235656,Q5LND5_Alphaproteobacteria:0.11918282):0.02409360[0.2800],Q2RWD8_Alphaproteobacteria:0.10680055):0.01442588[0.2400],Q2ND62_Alphaproteobacteria:0.15168622):0.02591599[0.3800],(Q9A9T4_Alphaproteobacteria:0.11674906,B3PP77_Alphaproteobacteria:0.08960840):0.03947969[0.7800]):0.04478930[0.6600],Q3YRH6_Alphaproteobacteria:0.24940690):0.07812032[0.9400],(A0A1Q9NZT2_Asgard:0.22931858,O84115_Chlamydia:0.29084174):0.05815952[0.3200]):0.01893914[0.0200],(((((E1QM11_Deltaproteobacteria:0.12639733,F2NCS7_Deltaproteobacteria:0.14017419):0.04185589[0.3200],A8ZXC0_Deltaproteobacteria:0.13452966):0.03059529[0.3600],(Q2IMX7_Deltaproteobacteria:0.13853653,Q1D277_Deltaproteobacteria:0.11625859):0.05453041[0.9600]):0.01896319[0.3000],(Q24WY5_Clostridia:0.17502499,Q6ALY6_Deltaproteobacteria:0.23393708):0.03824357[0.0400]):0.02240809[0.0400],((A0A1Q9MTN0_Asgard:0.19483833,Q6MIV0_Oligoflexia:0.19497160):0.05062330[0.2000],((A5D0B4_Clostridia:0.14098002,D0LIP3_Deltaproteobacteria:0.19894541):0.04012551[0.3800],(A9FMK6_Deltaproteobacteria:0.19275250,(Q73K92_Spirochaetia:0.13583622,(H9UKL9_Spirochaetia:0.12201809,E1RCI3_Spirochaetia:0.10575296):0.05038159[0.8800]):0.08780430[1.0000]):0.02960230[0.4200]):0.01904957[0.0800]):0.02543092[0.1000]):0.02003184[0.0400]):0.02125299[0.0200]):0.01905482[0.0800],(((Q7NFE9_Gloeobacteria:0.09001274,B0JKF2_Cyanobacteria:0.14294276):0.04356378[0.8200],B0JQ24_Cyanobacteria:0.17569755):0.06861741[0.9600],(F2LWB6_Deltaproteobacteria:0.19753205,(Q12U12_Euryarchaeota:0.22396495,(Q8RHQ8_Fusobacteria:0.25589360,(Q898C7_Clostridia:0.21610358,((F3YX85_Deltaproteobacteria:0.12535203,Q1MS43_Deltaproteobacteria:0.15655536):0.05105381[0.9600],(((((A1APS9_Deltaproteobacteria:0.05942294,B3E2Q5_Deltaproteobacteria:0.05989424):0.05272945[1.0000],Q826F2_Actinobacteria:0.14756702):0.03190854[0.6000],A0LEW2_Deltaproteobacteria:0.09282126):0.03450696[0.6800],Q5L1U6_Bacilli:0.12428680):0.03764182[0.4000],(B1YJ24_Bacilli:0.17069150,(Q49WB4_Bacilli:0.17804868,((Q6F146_Mollicutes:0.20868859,P75247_Mollicutes:0.36578565):0.10660719[0.5600],((F2I689_Bacilli:0.11430984,F2I690_Bacilli:0.36984686):0.04900949[0.2600],(Q831Y7_Bacilli:0.05937356,Q9CFF3_Bacilli:0.09081617):0.06969538[0.9600]):0.06050669[0.2800]):0.03799847[0.1000]):0.04517181[0.2200]):0.03910246[0.1800]):0.03870732[0.1600]):0.02147153[0.1000]):0.04435878[0.1600]):0.02461235[0.3200]):0.02584219[0.2000]):0.02833715[0.2200]):0.02250613[0.1200]):0.03592471[0.0800],((Q2FKR4_Euryarchaeota:0.14024506,Q7MVE7_Bacteroidia:0.15202717):0.11188597[1.0000],(D3UI70_Epsilonproteobacteria:0.16182662,(E0UPI8_Epsilonproteobacteria:0.13560493,(D1B2U8_Epsilonproteobacteria:0.12736787,E6X033_Epsilonproteobacteria:0.10600802):0.01700908[0.3600]):0.02851189[0.5200]):0.18301270[1.0000]):0.04183504[0.3200]):0.02931951[0.3800],A9FG71_Deltaproteobacteria:0.55011874):0.05385166[0.2800],(Q1D2Y0_Deltaproteobacteria:0.19876862,(Q0K1F3_Betaproteobacteria:0.12653421,Q9I659_Gammaproteobacteria:0.10307283):0.16465792[1.0000]):0.09998840[0.9200]):0.04640624[0.2000],(((((((A1BDB6_Chlorobia:0.01206616,B4SFV1_Chlorobia:0.02027243):0.01732050[0.9600],B4S4D9_Chlorobia:0.04413477):0.16853403[1.0000],Q7MXY4_Bacteroidia:0.33208031):0.07984547[0.9400],O84288_Chlamydia:0.36458098):0.03021484[0.2200],(Q73L29_Spirochaetia:0.38762751,(H9UI08_Spirochaetia:0.11309744,E1R1B3_Spirochaetia:0.13459774):0.11484500[1.0000]):0.08596550[0.7400]):0.06415619[0.1400],(F2I7B8_Bacilli:0.30401860,(Q82YZ7_Bacilli:0.14370338,Q9CHS9_Bacilli:0.28804274):0.07303733[0.7400]):0.12269790[1.0000]):0.03533929[0.0200],(((Q7NIW8_Gloeobacteria:0.05937138,B0JJ69_Cyanobacteria:0.04915417):0.06361359[1.0000],(Q7NKD6_Gloeobacteria:0.24548503,B0JIB9_Cyanobacteria:0.20171581):0.08329687[0.9400]):0.08556635[0.8000],((((Q250R5_Clostridia:0.17209781,Q3A9N1_Clostridia:0.07820760):0.03107487[0.2600],Q67JN5_Clostridia:0.14746944):0.02839174[0.2200],(A5D5K8_Clostridia:0.07819387,Q2RM39_Clostridia:0.13378472):0.02541633[0.5800]):0.023

83691[0.2200],((((((B1YGS2_Bacilli:0.10293732,C0ZIE7_Bacilli:0.08846637):0.03174829[0.2400],Q5L436_Bacilli:0.03939788):0.00967876[0.1200],(P37571_Bacilli:0.05613562,A0A127VTU5_Bacilli:0.08695787):0.02476921[0.2800]):0.04467314[0.2600],Q49V34_Bacilli:0.14169326):0.05773176[0.8600],Q890L5_Clostridia:0.20132961):0.03474447[0.1400],(((Q5Z2T4_Actinobacteria:0.01646690,P9WPC9_Actinobacteria:0.03072607):0.01543023[0.5600],Q6NFB1_Actinobacteria:0.08378391):0.05846318[1.0000],((Q6ACR4_Actinobacteria:0.05990043,Q83H06_Actinobacteria:0.16001473):0.08648012[0.9200],(Q47KW3_Actinobacteria:0.06447902,(Q82EB8_Actinobacteria:0.00219985,Q82K04_Actinobacteria:0.01980852):0.03806311[0.9600]):0.02962707[0.7400]):0.02924291[0.5800]):0.18198624[1.0000]):0.02719472[0.0000]):0.04178204[0.1800]):0.03389700[0.1200]):0.03810510[0.0000]):0.03067937[0.0000],(Q9WY41_Thermotogae:0.23618009,(Q9X1B1_Thermotogae:0.27766983,(F0S2V8_Aquificae:0.17293282,(O67588_Aquificae:0.12807284,C1DUW1_Aquificae:0.17808929):0.11309990[1.0000]):0.07111204[1.0000]):0.03577325[0.2800]):0.02902050[0.0400]):0.04085108[0.0000],(F0S121_Aquificae:0.17958925,(O67325_Aquificae:0.10198609,C1DT67_Aquificae:0.11426306):0.10394569[1.0000]):0.10357485[0.9800]):0.01829911[0.0400],(A1STR4_Gammaproteobacteria:0.38267723,(Q899V4_Clostridia:0.33306261,(((Q837W9_Bacilli:0.07765981,Q9CI09_Bacilli:0.11351921):0.09928382[0.9800],F2I5V6_Bacilli:0.21001803):0.09901659[1.0000],(A0A127W0Q0_Bacilli:0.18929591,(Q5L1C7_Bacilli:0.10133515,(O31673_Bacilli:0.13622672,C0ZJ35_Bacilli:0.17129359):0.05729994[0.6200]):0.03333899[0.2800]):0.10860302[1.0000]):0.14059607[1.0000]):0.13173655[0.9200]):0.04546349[0.1000]):0.04841099[0.0800],(Q2IKK0_Deltaproteobacteria:0.20230948,Q1D4S0_Deltaproteobacteria:0.19768735):0.36863196[1.0000]):0.03763243[0.2200],(F2I7G8_Bacilli:0.70119660,(Q9RVB2_Deinococci:0.42708472,(((F3YX74_Deltaproteobacteria:0.25039781,Q1D2Z9_Deltaproteobacteria:0.25985985):0.08487744[0.7600],(Q9I363_Gammaproteobacteria:0.24582641,Q9KN49_Gammaproteobacteria:0.33727242):0.08504248[0.8600]):0.04593564[0.5600],(Q9I1A8_Gammaproteobacteria:0.25989487,(Q47E21_Betaproteobacteria:0.21306279,(Q2SEG4_Gammaproteobacteria:0.29099544,(Q0KDX2_Betaproteobacteria:0.30842002,(B3Q4A0_Alphaproteobacteria:0.18744507,Q9I742_Gammaproteobacteria:0.17122376):0.05095950[0.6800]):0.07778454[0.9000]):0.04813273[0.4600]):0.04544520[0.6000]):0.08917405[0.9600]):0.35538499[1.0000]):0.04200291[0.0600]):0.07169139[0.4000],(Q9RWS7_Deinococci:0.36087740,(((H9UKD7_Spirochaetia:0.23307731,E1R5D7_Spirochaetia:0.28456257):0.07690776[0.8400],Q73KU3_Spirochaetia:0.34370887):0.07919552[0.9800],(((((D3UJ95_Epsilonproteobacteria:0.51089097,E6WZH7_Epsilonproteobacteria:0.22369730):0.07753449[0.4600],E0UPY5_Epsilonproteobacteria:0.26854610):0.06487524[0.7000],D1B3H1_Epsilonproteobacteria:0.28425261):0.05302876[0.5200],B9L8R6_Epsilonproteobacteria:0.30517456):0.11370463[0.6800],(Q1MQQ5_Deltaproteobacteria:0.34580086,((((F3Z1S3_Deltaproteobacteria:0.22102943,A0LE94_Deltaproteobacteria:0.18550572):0.05540222[0.6400],A8ZW51_Deltaproteobacteria:0.24233501):0.03840985[0.5200],Q6APV8_Deltaproteobacteria:0.28304500):0.03854603[0.3200],(((Q2IP33_Deltaproteobacteria:0.18820743,Q1CZL2_Deltaproteobacteria:0.16516163):0.04621905[0.6800],D0LVD0_Deltaproteobacteria:0.24033767):0.04287414[0.6400],((Q3YS04_Alphaproteobacteria:0.37946683,Q6MNG8_Oligoflexia:0.28961312):0.09590500[0.4800],(A9F365_Deltaproteobacteria:0.30185957,((((Q5FTA9_Alphaproteobacteria:0.11135771,Q2N6G4_Alphaproteobacteria:0.17871646):0.03740196[0.5800],Q2RSI6_Alphaproteobacteria:0.09793318):0.04341758[0.7000],(Q5LNN4_Alphaproteobacteria:0.13918890,(Q9A5H9_Alphaproteobacteria:0.12764280,B3PYN2_Alphaproteobacteria:0.15750165):0.02479985[0.3800]):0.04058727[0.6800]):0.09234061[1.0000],(((((((Q0K794_Betaproteobacteria:0.05438026,A4G7T8_Betaproteobacteria:0.06410464):0.05173790[1.0000],Q82TY2_Betaproteobacteria:0.12307751):0.02975269[0.5200],Q47BF6_Betaproteobacteria:0.10273540):0.01901530[0.3400],Q3SJH1_Betaproteobacteria:0.08753867):0.02743485[0.4600],C1DBG3_Betaproteobacteria:0.07360267):0.06945023[0.9800],Q5F9I6_Betaproteobacteria:0.22107091):0.10051258[0.9800],((A5EV83_Gammaproteobacteria:0.26872581,Q4FRL2_Gammaproteobacteria:0.25530887):0.06858617[0.5000],(((((Q3J8G5_Gammaproteobacteria:0.15258761,Q5ZXB4_Gam

maproteobacteria:0.18456694):0.04522821[0.3200],Q607H1_Gammaproteobacteria:0.10834921):0.03139507[0.2000],(Q2SJL2_Gammaproteobacteria:0.09855295,Q9I0L8_Gammaproteobacteria:0.13383226):0.07845626[1.0000]):0.03093522[0.1800],Q8P998_Gammaproteobacteria:0.22771883):0.02139840[0.2200],(Q83CD0_Gammaproteobacteria:0.29459071,(((Q480C5_Gammaproteobacteria:0.12225610,Q8EDW5_Gammaproteobacteria:0.16067561):0.02205046[0.3200],(Q5R0C4_Gammaproteobacteria:0.14127863,Q3IH33_Gammaproteobacteria:0.08270076):0.04281622[0.6400]):0.03561622[0.3800],(A1SSY0_Gammaproteobacteria:0.22336753,(Q9KSW2_Gammaproteobacteria:0.08117855,(Q8ZQE3_Gammaproteobacteria:0.00475935,P0ABH9_Gammaproteobacteria:0.01148430):0.09814721[1.0000]):0.04213259[0.9600]):0.03474224[0.4000]):0.08277973[0.9800]):0.03664497[0.2400]):0.03146202[0.3600]):0.03556241[0.3800]):0.02443392[0.4600]):0.04296371[0.6400]):0.02333574[0.3200]):0.03092794[0.2000]):0.05496737[0.2600]):0.04407384[0.1600]):0.03819533[0.1000]):0.05253885[0.3800]):0.07688311[0.4400]):0.20655553[1.0000])OROOT;

JDPs

((((((((((((((((((((Q47EC2_Betaproteobacteria:0.09671686,A1STR3_Gammaproteobacteria:0.18644890):0.03934143[0.3100],(A4G524_Betaproteobacteria:0.14355552,A4G5R4_Betaproteobacteria:0.34375942):0.04205365[0.2500]):0.04741062[0.0400],Q3SIZ1_Betaproteobacteria:0.18134757):0.04726030[0.0500],(Q4FVP5_Gammaproteobacteria:0.44543038,Q8PBS1_Gammaproteobacteria:0.26470874):0.08684107[0.0900]):0.02536191[0.0000],(Q2SPJ5_Gammaproteobacteria:0.11429019,(Q5ZT99_Gammaproteobacteria:0.08469378,(P63263_Gammaproteobacteria:0.25585606,Q3IJR2_Gammaproteobacteria:0.38950730):0.10078220[0.0100]):0.07242161[0.0000]):0.08044488[0.0100]):0.03528840[0.0000],Q83CJ2_Gammaproteobacteria:0.14747821):0.06071481[0.0300],(A0LEW0_Deltaproteobacteria:0.18084547,(Q826F4_Actinobacteria:0.26388166,B3EA91_Deltaproteobacteria:0.18446141):0.07477966[0.1600]):0.05964554[0.1100]):0.05304651[0.0000],Q5NEU6_Gammaproteobacteria:0.37526436):0.03414217[0.0000],(F3YX83_Deltaproteobacteria:0.18134365,Q1MS45_Deltaproteobacteria:0.17501092):0.07097843[0.4500]):0.01273411[0.0000],Q24XN6_Clostridia:0.23425234):0.07178926[0.0000],(O84345_Chlamydia:0.32693524,(((Q7NGQ4_Gloeobacteria:0.28838688,A5D0B5_Clostridia:0.12317851):0.07853206[0.1100],B0JXG3_Cyanobacteria:0.21644000):0.04048761[0.0600],((Q6ALT1_Deltaproteobacteria:0.22633971,Q6MGQ2_Oligoflexia:0.19814692):0.13077598[0.1800],((A8ZYX1_Deltaproteobacteria:0.20268379,F2NE43_Deltaproteobacteria:0.26548866):0.05584983[0.0700],(((Q2IQI5_Deltaproteobacteria:0.19024470,Q1D782_Deltaproteobacteria:0.25803207):0.12287556[0.3900],Q47TH8_Actinobacteria:0.19966474):0.05525493[0.0000],((Q6AC78_Actinobacteria:0.11107117,Q83FI0_Actinobacteria:0.36221601):0.10637145[0.3100],(Q82EX7_Actinobacteria:0.13648443,(Q6NEZ1_Actinobacteria:0.22423516,Q5YNI2_Actinobacteria:0.19836592):0.12950955[0.6200]):0.04612133[0.2300]):0.08174302[0.0500]):0.07575837[0.0000]):0.00000015[0.0000]):0.02946589[0.0000]):0.03214372[0.0000]):0.04444092[0.0000]):0.06877090[0.0000],(Q6L0S6_Euryarchaeota:0.20602516,((A0A1Q9N5F1_Asgard:0.15377958,A0A218NM31_DPANN:0.18908242):0.09975707[0.1800],(A0A0F8W923_Asgard:0.26747599,(Q892R1_Clostridia:0.24867073,(E1QDM3_Deltaproteobacteria:0.16001997,(A1AJY4_Deltaproteobacteria:0.14798902,B3E335_Deltaproteobacteria:0.10805133):0.13312661[0.7100]):0.06746363[0.0600]):0.03054764[0.0100]):0.05338914[0.0000]):0.04512046[0.0000]):0.00000000[0.0000]):0.03096526[0.0000],(((((A0A1Q9PB78_Asgard:0.18885849,B0JV84_Cyanobacteria:0.26958666):0.09954469[0.3900],O67623_Aquificae:0.26562561):0.06996359[0.0400],(C1DW67_Aquificae:0.10815611,(O66921_Aquificae:0.10139705,Q9WZV3_Thermotogae:0.26136758):0.08478033[0.2500]):0.03789231[0.1100]):0.05547637[0.0000],(B9L955_Epsilonproteobacteria:0.12487340,(E6X0L4_Epsilonproteobacteria:0.18155353,(E0UU84_Epsilonproteobacteria:0.32404420,(D1AZY1_Epsilonproteobacteria:0.20041149,D3UJ13_Epsilonproteobacteria:0.28063081):0.06254152[0.1800]):0.05182089[0.1800]):0.04655794[0.1900]):0.08670291[0.1000]):0.04712969[0.0000],((((A0A1Q9NZP1_Asgard:0.307

01910,A0A1Q9P052_Asgard:0.23974894):0.05987895[0.3000],A0A1Q9NQU0_Asgard:0.14918708):0.13395594[0.2900],Q5NFG8_Gammaproteobacteria:0.29380770):0.05634174[0.0100],(((Q5F5M1_Betaproteobacteria:0.19223088,Q5F952_Betaproteobacteria:0.39344741):0.06991427[0.0900],A5EYE5_Gammaproteobacteria:0.15002200):0.07072381[0.1100],((F2LWB4_Deltaproteobacteria:0.14818642,F2NBX3_Deltaproteobacteria:0.21658891):0.10252893[0.1800],(B9L5J4_Epsilonproteobacteria:0.18160011,(E0UV60_Epsilonproteobacteria:0.22608856,(D3UI47_Epsilonproteobacteria:0.15399274,(D1B2T2_Epsilonproteobacteria:0.15127145,E6X1Y2_Epsilonproteobacteria:0.12548430):0.00000000[0.0900]):0.03263668[0.3500]):0.02362761[0.4100]):0.11638406[0.7500]):0.04933490[0.0900]):0.03120459[0.0000]):0.03032669[0.0000]):0.01436180[0.0000]):0.01852944[0.0000],(Q486D3_Gammaproteobacteria:0.55189966,(Q6AMJ3_Deltaproteobacteria:0.49215554,(F2LUD2_Deltaproteobacteria:0.09370941,Q73KY1_Spirochaetia:0.58116405):0.04469718[0.0000]):0.05417636[0.0000]):0.04868642[0.0000]):0.01433388[0.0000],(((((((A9FI75_Deltaproteobacteria:0.31369943,Q2SMM7_Gammaproteobacteria:0.12745794):0.08366557[0.0900],(Q9HV44_Gammaproteobacteria:0.10564250,Q8PAK8_Gammaproteobacteria:0.25593774):0.06025756[0.0700]):0.02062658[0.0100],(C1DD87_Betaproteobacteria:0.07648766,A4G8D1_Betaproteobacteria:0.22732907):0.04006541[0.1700]):0.02288084[0.0000],(Q5ZTY4_Gammaproteobacteria:0.17439576,(P42381_Gammaproteobacteria:0.22585932,(Q3J7D9_Gammaproteobacteria:0.14291507,Q607A6_Gammaproteobacteria:0.27110849):0.02338927[0.0800]):0.03322774[0.0000]):0.04522252[0.0100]):0.02810758[0.0000],(((Q47XI7_Gammaproteobacteria:0.11599861,A1STE5_Gammaproteobacteria:0.15414482):0.03090688[0.5200],Q9CMS2_Gammaproteobacteria:0.21438800):0.03301813[0.0100],(Q8EHT6_Gammaproteobacteria:0.12743004,(P0A1G7_Gammaproteobacteria:0.11367227,(Q5QXL2_Gammaproteobacteria:0.25268713,Q3IC07_Gammaproteobacteria:0.02971074):0.08898862[0.4500]):0.01143347[0.0900]):0.04549406[0.1400]):0.04903906[0.1100]):0.04447145[0.0000],(((P22305_Alphaproteobacteria:0.27787141,B3PXH2_Alphaproteobacteria:0.11656483):0.06671771[0.3500],Q5FSL4_Alphaproteobacteria:0.24599453):0.07676643[0.1500],(Q5LWJ5_Alphaproteobacteria:0.26045527,(Q2NAI7_Alphaproteobacteria:0.32214594,Q2RNE7_Alphaproteobacteria:0.14312061):0.02313666[0.1300]):0.06815558[0.0600]):0.06198820[0.0300]):0.03439840[0.0000],(O06431_Betaproteobacteria:0.18434122,Q2IHN3_Deltaproteobacteria:0.17925562):0.09815966[0.2000]):0.01450426[0.0000],(((((((A9GVL6_Deltaproteobacteria:0.36573856,P75354_Mollicutes:0.27274262):0.10697211[0.2100],P78004_Mollicutes:0.24812498):0.14198963[0.0200],(A0A0F8VHI1_Asgard:0.31700118,A0A1Q9N0V3_Asgard:0.30550820):0.07515122[0.1300]):0.00000000[0.0000],(Q9RUG2_Deinococci:0.25996984,Q6F150_Mollicutes:0.20762100):0.11278289[0.2200]):0.04610999[0.0000],((Q5KWZ8_Bacilli:0.10927340,Q67S53_Clostridia:0.14798936):0.08536322[0.2700],((B1YKT0_Bacilli:0.18803405,C0ZB49_Bacilli:0.16325426):0.03474025[0.1400],(Q835R5_Bacilli:0.04109853,(A0A127VZ09_Bacilli:0.27414043,P35514_Bacilli:0.20067711):0.04377670[0.0900]):0.05588251[0.1900]):0.04684985[0.0400]):0.03005134[0.0000]):0.02105705[0.0000],(Q24SS4_Clostridia:0.14609403,Q2RKX3_Clostridia:0.12655180):0.03380423[0.1300]):0.06460490[0.0000],((((((P0CW07_Euryarchaeota:0.11256536,A0RZ02_TACK:0.40124405):0.03621116[0.0200],Q12WE7_Euryarchaeota:0.09637094):0.08576035[0.0000],A1ANU9_Deltaproteobacteria:0.19200441):0.01745265[0.0000],(A8ZZE3_Deltaproteobacteria:0.24124901,Q1DEA7_Deltaproteobacteria:0.18688104):0.08762416[0.2400]):0.04111475[0.0000],(Q9XCA6_Bacteroidia:0.21213551,Q3YT99_Alphaproteobacteria:0.20073191):0.08043513[0.1100]):0.04274168[0.0000],((Q8RH03_Fusobacteria:0.18722043,A1SXX5_Gammaproteobacteria:0.29885786):0.07656672[0.0300],((((Q2NE68_Euryarchaeota:0.16518586,Q6MNG0_Oligoflexia:0.22710833):0.06565493[0.1200],B4S9D0_Chlorobia:0.17565832):0.06923902[0.0400],D0LH40_Deltaproteobacteria:0.35727387):0.04047511[0.0000],((Q4FNQ0_Alphaproteobacteria:0.22767436,Q4FVQ7_Gammaproteobacteria:0.23023207):0.11352596[0.2100],(H9UKG7_Spirochaetia:0.19761245,(Q73Q15_Spirochaetia:0.19373566,E1R5G7_Spirochaetia:0.14077156):0.00000010[0.3700]):0.04951304[0.0900]):0.01710583[0.0000]):0.06890436[0.0000]):0.03053845[0.0000]):0.00000015[0.0000]):0.01822806[0.0000]

):0.05230124[0.0000]):0.03737612[0.0000],((F3Z0J0_Deltaproteobacteria:0.25916458,Q1MQI8_Deltaproteobacteria:0.17918191):0.02417502[0.0400],((E1QE45_Deltaproteobacteria:0.14888045,Q6AN63_Deltaproteobacteria:0.27363991):0.09874932[0.3200],(A0LJ41_Deltaproteobacteria:0.14643024,(F2NEH0_Deltaproteobacteria:0.33466371,F2NFR0_Deltaproteobacteria:0.16389110):0.06774893[0.1100]):0.05633731[0.0300]):0.03584388[0.0000]):0.08458888[0.0000]):0.01714234[0.0000],(((Q47D79_Betaproteobacteria:0.61873318,F2LUF1_Deltaproteobacteria:0.60912738):0.05940065[0.0500],(E1R344_Spirochaetia:0.58634864,Q9X2C3_Thermotogae:0.36214061):0.18950212[0.2400]):0.13538485[0.0000],((Q2FSF7_Euryarchaeota:0.63767464,A0LNZ5_Deltaproteobacteria:0.33004823):0.09313684[0.0600],((Q2FQP6_Euryarchaeota:0.30400222,F2I4Z3_Bacilli:0.15871295):0.11171197[0.0600],(D0LKN9_Deltaproteobacteria:0.34118906,(Q7NDG8_Gloeobacteria:0.25559902,B0JW23_Cyanobacteria:0.13392086):0.17939824[0.5300]):0.03760950[0.0100]):0.00000009[0.0000]):0.04246337[0.0000]):0.02319705[0.0000]):0.03226243[0.0000],((((A0A2D6LP47_DPANN:0.33271575,D0LIP4_Deltaproteobacteria:0.48121588):0.00846846[0.0000],B1YJ23_Bacilli:0.35941853):0.10839285[0.0000],(B9LUC6_Euryarchaeota:0.04863403,A0A0F7PCA8_Euryarchaeota:0.25591861):0.35466602[0.9700]):0.12283589[0.0000],(A9G8N1_Deltaproteobacteria:0.30807089,(((Q2N6B6_Alphaproteobacteria:0.36963161,Q2RR22_Alphaproteobacteria:0.20604521):0.07570114[0.1900],B3Q520_Alphaproteobacteria:0.26602010):0.04398090[0.0800],((Q5FR31_Alphaproteobacteria:0.25037076,Q9A4Q8_Alphaproteobacteria:0.28931623):0.11400819[0.3000],(B3PS99_Alphaproteobacteria:0.45416680,Q2RST0_Alphaproteobacteria:0.28738372):0.13409174[0.1700]):0.06190562[0.0100]):0.08173711[0.0300]):0.17855399[0.0300]):0.02204591[0.0000]):0.03407131[0.0000],((C0ZJ82_Bacilli:0.45594901,B0JNV1_Cyanobacteria:0.44735440):0.14007001[0.1400],(F2LVX6_Deltaproteobacteria:0.64289479,A0LKU3_Deltaproteobacteria:0.60991428):0.11503493[0.0400]):0.07647433[0.0000]):0.01542956[0.0000],(Q82BY4_Actinobacteria:0.17589964,(Q47RP0_Actinobacteria:0.02933796,((Q6NG14_Actinobacteria:0.26439289,Q5YZX0_Actinobacteria:0.12759436):0.11677366[0.4700],(Q6AEC0_Actinobacteria:0.06113006,Q83MZ4_Actinobacteria:0.37496400):0.22246594[0.4800]):0.07785648[0.0400]):0.04754918[0.0300]):0.19524180[0.1400]):0.02040870[0.0000],(((((Q7NDI6_Gloeobacteria:0.42438061,B0JN47_Cyanobacteria:0.29632689):0.49680309[0.9700],B9L9F1_Epsilonproteobacteria:0.53836625):0.32865612[0.4100],(Q6AE35_Actinobacteria:0.58835321,H9UJJ3_Spirochaetia:0.72220448):0.24457397[0.0900]):0.11343339[0.0000],(A0LNJ4_Deltaproteobacteria:0.49102176,(F2NI39_Deltaproteobacteria:0.57819439,A0LLI9_Deltaproteobacteria:0.40455726):0.21584529[0.2100]):0.15899649[0.0700]):0.00000010[0.0000],((((F3Z1T4_Deltaproteobacteria:0.51993462,Q1MRK4_Deltaproteobacteria:0.46175699):0.24789301[0.7600],Q2RX13_Alphaproteobacteria:0.61849215):0.09029256[0.1000],A9G249_Deltaproteobacteria:1.02030867):0.26049451[0.0300],(((((Q0KB69_Betaproteobacteria:0.58957571,Q50312_Mollicutes:0.50630745):0.14875186[0.2800],C1DV28_Aquificae:0.67436479):0.09056615[0.0000],Q8PX41_Euryarchaeota:0.56987751):0.11098873[0.0000],A0LIC0_Deltaproteobacteria:0.80639022):0.15214637[0.0000],((Q7NFL3_Gloeobacteria:0.27764080,B0JRS8_Cyanobacteria:0.42004225):0.30281122[0.6200],((Q8EFW5_Gammaproteobacteria:0.52235046,E1R5V3_Spirochaetia:0.56245688):0.13919603[0.0800],((Q7NHY7_Gloeobacteria:0.38613837,B0JN11_Cyanobacteria:0.44108024):0.25087636[0.4300],(Q3JC07_Gammaproteobacteria:0.68410639,(Q3SHD4_Betaproteobacteria:0.47189566,Q6MQ10_Oligoflexia:0.65023312):0.20890484[0.3600]):0.11856432[0.0500]):0.11672596[0.0300]):0.07226464[0.0000]):0.04997858[0.0000]):0.00267096[0.0000]):0.06311855[0.0000]):0.05480564[0.0000],((((Q2IL31_Deltaproteobacteria:0.35243118,Q1D8R1_Deltaproteobacteria:0.26626320):0.79482144[1.0000],Q1DAP5_Deltaproteobacteria:0.55162661):0.34081225[0.0900],Q479W6_Betaproteobacteria:0.45090778):0.06529684[0.0000],((Q47EC9_Betaproteobacteria:0.59190994,Q73LY0_Spirochaetia:1.00319621):0.35870420[0.1500],((((F2NCJ3_Deltaproteobacteria:0.48519134,E6X0E0_Epsilonproteobacteria:0.37321887):0.13694644[0.3300],D1B490_Epsilonproteobacteria:0.68049559):0.00000000[0.2400],Q2RGP0_Clostridia:0.70147919):0.41172537[0.1800],(A0RUN6_TACK:0.95423631,(A0A0F8BHT4_Euryarchaeota:0.394

07683,(B3Q164_Alphaproteobacteria:0.54036759,(A0A0F8D488_Euryarchaeota:0.43672140,Q9A9Q7_Alphaproteobacteria:0.67415592):0.29766139[0.3400]):0.24821633[0.2200]):0.12289653[0.0800]):0.07530627[0.0000]):0.05231867[0.0100]):0.08833334[0.0000]):0.16481883[0.0000])OROOT;

GroES/HSP10

((((((((((((((((('B3PU02_Alphaproteobacteria':0.01703840,'B3Q350_Alphaproteobacteria':0.07621206):0.10618545[1.0000],'B3PRB7_Alphaproteobacteria':0.11542040):0.02359526[0.5100],'P0CAU0_Alphaproteobacteria':0.15495266):0.04339415[0.3400],'Q2NBL7_Alphaproteobacteria':0.09916726):0.06220492[0.2500],'Q2RWV5_Alphaproteobacteria':0.07026984):0.04252048[0.2400],('Q2N5R8_Alphaproteobacteria':0.15896466,'Q5FPQ7_Alphaproteobacteria':0.12272190):0.04627829[0.1700]):0.02196128[0.0900],'Q5LV16_Alphaproteobacteria':0.26493294):0.04490789[0.3400],'Q4FPA6_Alphaproteobacteria':0.26353196):0.04397860[0.2600],'Q2RY29_Alphaproteobacteria':0.23316191):0.12518340[0.5000],'A8ZU47_Deltaproteobacteria':0.13044694):0.04344571[0.0800],(('F2ND30_Deltaproteobacteria':0.28919809,'D0LRR4_Deltaproteobacteria':0.23627159):0.09437981[0.1500],(('A9GB09_Deltaproteobacteria':0.20111925,'A9GBP9_Deltaproteobacteria':0.12884563):0.01294628[0.1400],('Q2IFK1_Deltaproteobacteria':0.05270698,'Q1D2S2_Deltaproteobacteria':0.15309217):0.03586046[0.6100]):0.04351747[0.4000]):0.02370906[0.0400]):0.06080049[0.0400],('Q6MRI2_Oligoflexia':0.25430931,('A1AST2_Deltaproteobacteria':0.10643623,('B3E8F9_Deltaproteobacteria':0.13044400,('D0LYI4_Deltaproteobacteria':0.25787900,('Q1MQP9_Deltaproteobacteria':0.15070739,'F3YTT4_Deltaproteobacteria':0.26402229):0.08331137[0.3700]):0.11829283[0.0600]):0.01933749[0.0200]):0.04867978[0.0100]):0.10963544[0.0300]):0.05836612[0.0000],('A0LKS5_Deltaproteobacteria':0.25228438,('E1QE83_Deltaproteobacteria':0.24638780,'Q6ARV5_Deltaproteobacteria':0.21368771):0.06903249[0.0700]):0.07381972[0.0200]):0.00000009[0.0000],('Q01X67_Solibacter_usitatus':0.25934305,'Q01RQ3_Solibacter_usitatus':0.31098920):0.15524934[0.2600]):0.03510609[0.0000],(('F0S3Q5_Aquificae':0.15905700,'C1DX21_Aquificae':0.13938023):0.09557699[0.9100],('O67942_Aquificae':0.26580851,('F2LXB2_Deltaproteobacteria':0.36052410,('D3UGP0_Epsilonproteobacteria':0.37386127,('B9L697_Epsilonproteobacteria':0.23578089,('E6X266_Epsilonproteobacteria':0.22283919,'E0UQ32_Epsilonproteobacteria':0.43009565):0.06238362[0.2600]):0.04724290[0.2300]):0.23455457[0.7800]):0.10430327[0.1600]):0.14507252[0.0500]):0.08178869[0.0000]):0.01334592[0.0000],('A0LEH1_Deltaproteobacteria':0.12978765,((((((('A4G836_Betaproteobacteria':0.19844627,'Q0KDR8_Betaproteobacteria':0.07128168):0.07170894[0.4200],'Q47IZ9_Betaproteobacteria':0.12333684):0.04736250[0.3100],'Q3SMK0_Betaproteobacteria':0.08833229):0.03909279[0.1800],'Q82Y61_Betaproteobacteria':0.13320568):0.09297606[0.7300],('Q5F542_Betaproteobacteria':0.15168786,('C1D8S2_Betaproteobacteria':0.03828942,'C1DD17_Betaproteobacteria':0.04905471):0.02861969[0.8700]):0.13097121[0.9900]):0.07764391[0.4400],((((('Q60AY1_Gammaproteobacteria':0.13355579,'P19422_Gammaproteobacteria':0.15803916):0.01496035[0.0300],'Q5ZXP4_Gammaproteobacteria':0.09821410):0.06015964[0.2100],('Q3J728_Gammaproteobacteria':0.12169911,'Q607Q2_Gammaproteobacteria':0.23135998):0.00000000[0.0600]):0.04934674[0.3400],(('Q5NEE2_Gammaproteobacteria':0.22467940,'A5EX18_Gammaproteobacteria':0.28112400):0.01455690[0.1300],(((('Q2SDF9_Gammaproteobacteria':0.11237505,'P30720_Gammaproteobacteria':0.18436567):0.04716847[0.5700],'Q4FU95_Gammaproteobacteria':0.17122489):0.01297216[0.5700],('Q9AKT2_Gammaproteobacteria':0.16541571,('Q5QVT3_Gammaproteobacteria':0.08031715,('Q9KLC7_Gammaproteobacteria':0.38425468,(('A1SXK5_Gammaproteobacteria':0.01028800,'A1ST71_Gammaproteobacteria':0.00000000):0.10537137[1.0000],('Q487R0_Gammaproteobacteria':0.05895304,('Q8CX49_Gammaproteobacteria':0.10531033,('Q9KNR6_Gammaproteobacteria':0.11074606,('Q59686_Gammaproteobacteria':0.09559757,('P0A6F9_Gammaproteobacteria':0.03713677,'P0A1D5_Gammaproteobacteria':0.05413543):0.03472565[0.5700]):0.03686810[0.4800]):0.01418635[0.0400]):0.01190796[0.0500]):0.02896591[0.2500]):0.04352990[0.6600]):0.03041741[0.2800

]):0.07879209[0.6500]):0.18897894[0.9200]):0.08472461[0.3300]):0.07972287[0.1200]):0.0400
6033[0.0600]):0.10489350[0.1800]):0.03579687[0.0300]):0.06377354[0.0000],('P0A0R6_Gamm
aproteobacteria':0.33328793,'Q6MI28_Oligoflexia':0.96434840):0.11955518[0.1100],(((((((('Q8
TGX8_Euryarchaeota':0.02520530,'Q8PW07_Euryarchaeota':0.03235242):0.30944667[1.0000],'
Q2FPN6_Euryarchaeota':0.28762133):0.23134440[0.9900],'Q8R5X6_Fusobacteria':0.42781278)
:0.17081527[0.4800],'Q9WYX5_Thermotogae':0.22237401):0.11182641[0.4200],'Q3YRH3_Alp
haproteobacteria':0.75016693):0.10855574[0.0500],'P42376_Bacteroidia':0.22229656):0.086542
44[0.0100],('P0C0Z8_Chlamydia':0.63383298,('A0A3B6VH36_Spirochaetia':0.45484050,('H9U
GQ6_Spirochaetia':0.21413806,('Q73P65_Spirochaetia':0.24822755,'E1R675_Spirochaetia':0.16
348037):0.08873011[0.3200]):0.08315713[0.7500]):0.12729660[0.2600]):0.07699735[0.0900]):0
.05501637[0.0100],('A1BHS6_Chlorobia':0.02125031,('B4SEN0_Chlorobia':0.06777572,'B4S6H
1_Chlorobia':0.06554294):0.01272320[0.4000]):0.26632741[0.9900]):0.04294151[0.0100],(('Q3
ADX4_Clostridia':0.11796165,'A5CZ05_Clostridia':0.21364068):0.05292674[0.5100],(((((((('Q5L
3E7_Bacilli':0.14422918,'A0A127VUH1_Bacilli':0.13952597):0.02982397[0.1700],'P28599_Bac
illi':0.07112957):0.03385674[0.3400],('C0ZK51_Bacilli':0.11758498,'B1YEP7_Bacilli':0.256447
00):0.04783640[0.4600]):0.07893525[0.6500],'Q49YY6_Bacilli':0.44516808):0.06030694[0.180
0],'Q67KB7_Clostridia':0.21287163):0.05593312[0.0600],'Q24QE2_Clostridia':0.18500713):0.0
3311217[0.0100],((((((('Q5Z1G0_Actinobacteria':0.02428166,'P9WPE5_Actinobacteria':0.018890
39):0.07083935[0.9200],'Q6NJ38_Actinobacteria':0.19456955):0.11020634[0.8100],('Q820G1_
Actinobacteria':0.10754187,'A0A147KKR7_Actinobacteria':0.14081932):0.04640915[0.4300]):0
.10632974[0.5700],'Q6AD41_Actinobacteria':0.11793915):0.04533494[0.3600],'Q83N38_Actino
bacteria':0.33853252):0.20489956[0.8000],('Q2RGL7_Clostridia':0.22158804,(((('P0A347_Cyano
bacteria':0.15749904,'B0JUI1_Cyanobacteria':0.17373607):0.13755802[0.8900],('Q7MBB5_Glo
eobacteria':0.17837841,'Q7MBC6_Gloeobacteria':0.20699994):0.04879937[0.4200]):0.19692991
[0.8100],('Q2RL14_Clostridia':0.21697207,('Q9RWR0_Deinococci':0.37553162,('Q93EU7_Baci
lli':0.19060704,('P37283_Bacilli':0.32111232,'F2I7V3_Bacilli':0.36992133):0.09597244[0.5500])
:0.08804380[0.4700]):0.02655362[0.0700]):0.06372940[0.0400]):0.08582540[0.0500]):0.010328
88[0.0000]):0.10426213[0.0100]):0.07962292[0.0300]):0.14150780[0.0400]):0.08217352[0.0100
])OROOT;

GrpE

((((((((((((((((((('Q47XI4_Gammaproteobacteria':0.30372255,'Q5QXM7_Gammaproteobacteria'
:0.39281608):0.11594380[0.6100],'Q8EGS0_Gammaproteobacteria':0.39391971):0.13620156[0.
5200],'O30862_Gammaproteobacteria':0.37487977):0.08709246[0.3100],('Q3IKR2_Gammaprot
eobacteria':0.33446462,'A1STE3_Gammaproteobacteria':0.40678421):0.08080558[0.2800]):0.08
313777[0.4700],('Q9CNU1_Gammaproteobacteria':0.51833597,('Q7CPZ4_Gammaproteobacteri
a':0.04443024,'P09372_Gammaproteobacteria':0.03197714):0.49261583[1.0000]):0.12521681[0.
2100]):0.08780554[0.2000],('Q9HV42_Gammaproteobacteria':0.37953061,('Q2SMM9_Gammap
roteobacteria':0.64137725,('Q3J7D7_Gammaproteobacteria':0.49904300,'Q607A4_Gammaproteo
bacteria':0.40104234):0.16437414[0.5200]):0.10628018[0.1300]):0.13580130[0.1200]):0.147812
77[0.1200],(('A5EYG2_Gammaproteobacteria':0.57578154,'Q5NFG6_Gammaproteobacteria':0.5
5997899):0.12267387[0.1600],('Q83C41_Gammaproteobacteria':0.61366738,'Q5ZTY2_Gamma
proteobacteria':0.54341013):0.16724589[0.6700]):0.09413483[0.0700]):0.10248842[0.0100],('Q4
FPS8_Gammaproteobacteria':0.77259351,('Q6AMQ4_Deltaproteobacteria':0.66727193,('F3Z406
_Deltaproteobacteria':0.35254620,'Q1MPH5_Deltaproteobacteria':0.77952921):0.22670531[0.73
00]):0.13552978[0.1600]):0.15004112[0.0400]):0.09755826[0.0000],((((((('O08384_Betaproteoba
cteria':0.45142673,'Q3SIN5_Betaproteobacteria':0.20996321):0.14519593[0.6300],'Q47HK1_Bet
aproteobacteria':0.24455394):0.14355877[0.6700],'A4G8D3_Betaproteobacteria':0.36985924):0.
08458594[0.4100],'Q0KCJ3_Betaproteobacteria':0.33532684):0.22860773[0.8600],('C1DD89_B
etaproteobacteria':0.33834057,'Q5F6X1_Betaproteobacteria':0.44921076):0.12317200[0.8300]):0

.23800232[0.8600],((('Q2NAJ5_Alphaproteobacteria':0.35938135,'Q4FNP8_Alphaproteobacteria':0.66331864):0.27678028[0.6400],'Q3YSZ3_Alphaproteobacteria':0.90459644):0.08800063[0.1100],('Q2RN58_Alphaproteobacteria':0.85112117,(('Q5FSQ2_Alphaproteobacteria':0.71236975,'P0CAV1_Alphaproteobacteria':0.46803331):0.09008517[0.3100],('B3PZA4_Alphaproteobacteria':0.44858765,'Q5LWF4_Alphaproteobacteria':0.75465013):0.15282058[0.1500]):0.08352392[0.1100]):0.12044393[0.0600]):0.14416143[0.0500]):0.12256833[0.0200]):0.06636738[0.0000],'Q8PAL0_Gammaproteobacteria':0.77080548):0.10501721[0.0100],('E0UPC7_Deltaproteobacteria':0.40794355,('E6X0M1_Epsilonproteobacteria':0.50601687,('B9L8Y9_Epsilonproteobacteria':0.56101566,('D1B2Q9_Deltaproteobacteria':0.38256094,'D3UHE4_Epsilonproteobacteria':0.7321454 3):0.15021508[0.4800]):0.12575540[0.2400]):0.04448674[0.1500]):0.22235693[0.5800]):0.11656512[0.0000],(((((('Q2IIR7_Deltaproteobacteria':0.65972823,'Q1D4B8_Deltaproteobacteria':0.69266705):0.30506487[0.9400],'D0LUJ9_Deltaproteobacteria':1.03119263):0.07908336[0.4200],'Q826F5_Actinobacteria':1.02744391):0.19257267[0.2100],('Q2FQP1_Euryarchaeota':0.66776038,'A0RZ00_TACK':0.99308814):0.14036494[0.1200]):0.23304720[0.0300],('Q9RY24_Deinococci':1.09190784,(('B9LNE5_Euryarchaeota':0.52084523,'A0A0F7P7Y2_Euryarchaeota':0.64050797):0.37913227[0.9600],('Q83MQ1_Actinobacteria':1.04878570,(('Q5YNI1_Actinobacteria':0.30675442,'P9WMT5_Actinobacteria':0.54875248):0.26216638[0.9400],'Q6NEZ0_Actinobacteria':0.67455141):0.23357320[0.8200],('Q6AC77_Actinobacteria':0.78266350,('Q47TH9_Actinobacteria':0.52211913,'Q82EX8_Actinobacteria':0.39503185):0.21472729[0.9300]):0.08961606[0.3600]):0.07928770[0.6200]):0.21711423[0.5800]):0.08606477[0.1300]):0.14454306[0.0600]):0.15955631[0.0200]):0.04181387[0.0000],(((((((('Q2IHN1_Deltaproteobacteria':0.45896688,'P95333_Deltaproteobacteria':0.56690763):0.31993298[0.9400],('D0LKN7_Deltaproteobacteria':0.55751116,'A9GHU4_Deltaproteobacteria':0.46126202):0.15857210[0.4700]):0.09808570[0.1200],('A1ANV1_Deltaproteobacteria':0.24868368,'B3E7X0_Deltaproteobacteria':0.33054719):0.34747246[1.0000]):0.09382342[0.1200],'A0LH27_Deltaproteobacteria':0.45311846):0.12269249[0.1200],('A8ZRW2_Deltaproteobacteria':0.63096659,'F2NCB0_Deltaproteobacteria':0.51268063):0.19741747[0.5700]):0.15684312[0.0900],'E1QIB1_Deltaproteobacteria':0.57126439):0.12817254[0.0500],('Q3AF09_Clostridia':0.79310825,(((((('Q24SS2_Clostridia':0.35975243,'A5D3Y0_Clostridia':0.55934857):0.17826296[0.7100],'C0ZB47_Bacilli':0.67079305):0.11850742[0.1500],('B0JN66_Cyanobacteria':0.84894071,'Q7NDP1_Gloeobacteria':0.41039277):0.32778109[0.8300]):0.06154385[0.0100],('Q67S55_Clostridia':0.54812000,'Q2RKX5_Clostridia':0.64458255):0.22400673[0.2700]):0.09251798[0.0200],((((('Q835R8_Bacilli':0.33334912,'Q9CGY9_Bacilli':0.45723064):0.10317646[0.6000],'F2I4Z5_Bacilli':0.55055371):0.36107429[0.9100],'Q49Y23_Bacilli':0.57451343):0.14852936[0.2400],(('P15874_Bacilli':0.23156409,'Q5KWZ6_Bacilli':0.35130900):0.13516166[0.7400],('B1YKS8_Bacilli':0.49624276,'A0A127VXQ2_Bacilli':0.60374402):0.10197859[0.1500]):0.07407933[0.3100]):0.23052235[0.2400]):0.05201819[0.0000]):0.04496197[0.0100]):0.07551549[0.0000]):0.05778076[0.0000],('Q9WZV4_Thermotogae':0.62855771,('H9UKG9_Spirochaetia':0.49247039,('Q73Q17_Spirochaetia':0.84120957,'E1R5G9_Spirochaetia':0.35938814):0.10626358[0.2600]):0.21835911[0.8100]):0.20253476[0.1700]):0.06896284[0.0000],(('Q7MU00_Bacteroidia':0.87066233,'C0QX60_Spirochaetia':0.62077350):0.10655828[0.0500],('P36424_Chlamydia':0.84289187,'Q892Q9_Chlorobia':0.66390971):0.16317832[0.1500]):0.12265430[0.0200]):0.12206753[0.0000],(((('A0A1Q9N5G9_Asgard':0.76140254,'Q2NE66_Euryarchaeota':0.56492431):0.11732693[0.1400],'Q8RH07_Fusobacteria':0.69601307):0.14972880[0.0600],((('Q6L0S8_Euryarchaeota':0.46139391,'Q9HJ84_Euryarchaeota':0.67933606):0.55860677[0.8400],'F2LUD0_Deltaproteobacteria':0.77702315):0.02146147[0.0100],((('A1BHL2_Chlorobia':0.37424375,'B4SG55_Chlorobia':0.25728314):0.14647929[0.9400],'B4S9D1_Chlorobia':0.33869569):0.61645980[1.0000],('A0A2D6LP76_DPANN':0.46477466,('A0A218NM27_DPANN':0.77790878,('A0A1Q9MTN9_Asgard':0.28087326,'A0A1Q9MTN2_Asgard':0.38594651):0.90697671[1.0000]):0.11752833[0.0400]):0.12188868[0.0400]):0.12630186[0.0200]):0.08881761[0.0000]):0.11031223[0.0000]):0.04464540[0.0000],('A0A1Q9PB65_Asgard':1.19670010,('Q12WE5_Euryarchaeota':0.521

59791,('P0CW11_Euryarchaeota':0.11336972,'Q8TQR3_Euryarchaeota':0.14782896):0.4774569
5[1.0000]):0.19947939[0.8700]):0.17478673[0.0900]):0.13539504[0.0000],('A0A1Q9NMA9_As
gard':1.41129631,'Q6F148_Mollicutes':0.96836347):0.22422619[0.1100]):0.09796083[0.0300],'
Q6MGQ3_Oligoflexia':0.71872903):0.18147678[0.0400],('O66745_Aquificae':0.39575667,'P780
17_Mollicutes':1.11127004):0.25220612[0.4400]):0.10046243[0.0200],'C1DV27_Aquificae':0.55
601476,('A0A1Q9MVP0_Asgard':1.94916008,'Q2FPF1_Euryarchaeota':1.32104290):0.8670734
7[0.4500])OROOT;

Cyp40

((((((((((((((('A0A0F7PFT4_Euryarchaeota':0.28781698,'A9GR29_Deltaproteobacteria':0.50158
847):0.13723195[0.3200],('D0LQF2_Deltaproteobacteria':0.31347110,'D0LX61_Deltaproteobact
eria':0.42434507):0.30706069[0.9700]):0.11027197[0.2200],('B9LPG5_Euryarchaeota':0.000000
00,'A0A081EW71_Euryarchaeota':0.02138152):0.40145798[0.9900]):0.12543475[0.1000],('Q1D
5X7_Deltaproteobacteria':0.33495857,('Q1DD37_Deltaproteobacteria':0.32172735,'Q1D5X8_De
ltaproteobacteria':0.31565633):0.06927418[0.5900]):0.30702858[0.9500]):0.10830284[0.0200],((
'Q47KC0_Actinobacteria':0.24477466,'Q47L19_Actinobacteria':0.34316910):0.11231979[0.5700
],(('Q6NKJ3_Actinobacteria':0.17672619,'P9WHW3_Actinobacteria':0.11000063):0.23001168[0.
9700],('Q6AHM5_Actinobacteria':0.36190335,'Q82FC6_Actinobacteria':0.20197629):0.1083564
3[0.5100]):0.12424152[0.5300]):0.11263242[0.2600]):0.05811747[0.0100],('Q6MRB4_Oligofle
xia':0.40820546,('Q2IIE8_Deltaproteobacteria':0.58298291,'Q2IIE7_Deltaproteobacteria':0.2607
7096):0.12776766[0.3400]):0.07864280[0.1000]):0.11537862[0.0200],'Q5ZU46_Gammaproteob
acteria':0.49490913):0.04219123[0.0000],((('Q82ZZ9_Bacilli':0.19318585,'Q9CH46_Bacilli':0.3
3896422):0.21065885[0.9700],'F2I5E9_Bacilli':0.33322203):0.15882297[0.1000],('E6X070_Epsi
lonproteobacteria':0.32732246,('E0UPT2_Deltaproteobacteria':0.29437649,('A0A1Q9MZU3_As
gard':0.58531697,'D1AZ05_Deltaproteobacteria':0.33596823):0.19810657[0.7700]):0.18057414[
0.4600]):0.27215167[0.4900]):0.12545865[0.0100]):0.09815704[0.0000],('A0A1Q9NIH8_Asgar
d':0.22446681,'Q5YSZ1_Actinobacteria':0.37432436):0.54915387[1.0000]):0.02997412[0.0000],
((((('A0A1Q9MZH7_Asgard':0.33494157,'Q8RGF6_Fusobacteria':0.29201074):0.15964101[0.58
00],'Q8PC26_Gammaproteobacteria':0.35886985):0.16573397[0.1700],'Q8RI36_Fusobacteria':0.
99166487):0.06651207[0.0200],(('Q67KZ4_Clostridia':0.58787546,'Q67L36_Clostridia':0.2587
6098):0.13264375[0.3700],'C0ZEP1_Bacilli':0.36639987):0.26860147[0.5000],('Q486E3_Gamm
aproteobacteria':0.71118116,('Q9RXR9_Deinococci':0.53353024,('Q9RT72_Deinococci':0.35826
980,'Q9RRF0_Deinococci':0.40904975):0.26078775[0.9600]):0.07394550[0.1600]):0.07900603[
0.0500]):0.10959856[0.0300]):0.11140629[0.0100]):0.06858966[0.0000],('E1R4X8_Spirochaetia
':0.61038426,((((('Q5YTH8_Actinobacteria':0.53312741,'P9WHW1_Actinobacteria':0.37663518):
0.13917270[0.6500],'Q6NGY5_Actinobacteria':0.71420379):0.15796714[0.5800],('Q6AFA6_Ac
tinobacteria':0.79039052,'Q827T2_Actinobacteria':0.68843171):0.09601384[0.2200]):0.5448769
3[0.8100],('Q49W93_Bacilli':0.34971228,('A0A127W4C9_Bacilli':0.19168317,('Q834U8_Bacill
i':0.46378728,('C0ZAF3_Bacilli':0.61625826,'Q9CIJ2_Bacilli':0.57396050):0.13296324[0.3100])
:0.14398535[0.1900]):0.06643175[0.2000]):0.25278199[0.2000]):0.07767073[0.0000]):0.054033
35[0.0000]):0.15604931[0.0100],(('A0A2D6K5Z2_Nanoarchaeota':0.40602138,'D0LUK1_Delta
proteobacteria':0.51959596):0.14511228[0.4600],('Q6L1D0_Euryarchaeota':0.18013014,('Q2FP4
6_Euryarchaeota':0.23757207,('Q8PT88_Euryarchaeota':0.03810379,'Q8TQ90_Euryarchaeota':0.
01433887):0.14705916[1.0000]):0.09631448[0.9100]):0.11706131[0.8500]):0.11985828[0.5000]
):0.00754636[0.0000],'A9F2I1_Deltaproteobacteria':1.09687630):0.16348186[0.0300],((((('A1BE
D2_Chlorobia':0.09178568,'B4SD18_Chlorobia':0.08017986):0.09932501[0.8500],'B4S3D7_Chl
orobia':0.13488251):0.22182461[0.9800],'Q7MV65_Bacteroidia':1.00191395):0.12716669[0.510
0],(('B0JTK1_Cyanobacteria':0.62722384,'Q7NG65_Gloeobacteria':1.44195204):0.07710165[0.1
300],('Q7NLZ6_Gloeobacteria':0.41147410,('Q7NHC7_Gloeobacteria':0.51270469,('B0JTU3_C
yanobacteria':0.75909202,'Q7NKH8_Gloeobacteria':0.57116804):0.40426734[0.9600]):0.174745

20[0.3500]):0.19144436[0.3000]):0.42760672[0.5300]):0.21842733[0.1500]):0.01969594[0.0000],((((('Q2N9X2_Alphaproteobacteria':0.68614658,'Q2IFL3_Deltaproteobacteria':0.55092836):0.07887289[0.1100],'Q9A8L6_Alphaproteobacteria':0.61306704):0.11799882[0.2300],'Q9A9K1_Alphaproteobacteria':0.54675630):0.54265882[1.0000],(((((((('Q488Q5_Gammaproteobacteria':0.35292912,'Q8EAT0_Gammaproteobacteria':0.35899718):0.09131997[0.7200],'Q3IHQ4_Gammaproteobacteria':0.69128870):0.12509103[0.3000],'Q5QWT2_Gammaproteobacteria':0.36827273):0.30910757[0.9100],'Q5F959_Betaproteobacteria':0.51500232):0.10567722[0.1400],'A5EWG7_Gammaproteobacteria':0.31365324):0.07098724[0.0400],('Q488X1_Gammaproteobacteria':0.66006607,'Q3IC17_Gammaproteobacteria':0.72320412):0.58309754[0.9500]):0.06365629[0.0300],(('Q6LY62_Euryarchaeota':0.15397856,'Q6LY63_Euryarchaeota':0.20002228):0.31485764[1.0000],'Q2SD51_Gammaproteobacteria':0.50742258):0.14821693[0.1300],((((((('A0LNA2_Deltaproteobacteria':0.26632196,'A0LHX7_Deltaproteobacteria':0.25926695):0.05016798[0.4600],'F3YVE6_Deltaproteobacteria':0.37755856):0.09159216[0.2700],('A1ATD4_Deltaproteobacteria':0.13981432,'B3E5N9_Deltaproteobacteria':0.16222472):0.17568430[1.0000]):0.09605142[0.1500],(('Q1MS04_Deltaproteobacteria':0.20126380,'A1STV4_Gammaproteobacteria':0.61152617):0.15159781[0.3100],('Q5ZRZ4_Gammaproteobacteria':0.21285755,'Q8EGU9_Gammaproteobacteria':0.54134297):0.12204945[0.4300]):0.07411540[0.0200]):0.02868569[0.0000],('Q3SH89_Betaproteobacteria':0.46292340,('Q59641_Gammaproteobacteria':0.29415102,('P20753_Salmonella':0.01356354,'P0AFL3_Gammaproteobacteria':0.02794496):0.26967622[1.0000]):0.12032903[0.5700]):0.09472799[0.0900]):0.05137552[0.0000],('E1QL41_Deltaproteobacteria':0.41886341,((((('Q47HK9_Betaproteobacteria':0.35997255,'Q9KPR7_Gammaproteobacteria':0.48311028):0.08576890[0.2700],('C1D6W5_Betaproteobacteria':0.23181927,'Q82Y46_Betaproteobacteria':0.42745514):0.08749395[0.5300]):0.13111784[0.0800],('Q60BG1_Gammaproteobacteria':0.26027205,('Q0KCB1_Betaproteobacteria':0.42616719,'A1ASL0_Deltaproteobacteria':0.41793965):0.09251338[0.0800]):0.05926353[0.0700]):0.07413088[0.0000],(((('Q5F9L7_Betaproteobacteria':0.24134523,'A5EWG8_Gammaproteobacteria':0.22584425):0.08488371[0.2000],'Q8EE18_Gammaproteobacteria':0.32843707):0.09693118[0.0500],('Q3J8Y7_Gammaproteobacteria':0.36772553,((((('A8ZZB1_Deltaproteobacteria':0.17954492,'Q4FTV6_Gammaproteobacteria':0.24188832):0.07537761[0.4700],'F2NF28_Deltaproteobacteria':0.31044329):0.06324553[0.2100],'F3Z1A4_Deltaproteobacteria':0.22768150):0.08504161[0.2500],((('Q5Z3F8_Actinobacteria':0.18119109,'Q1MS03_Deltaproteobacteria':0.49325123):0.09309434[0.2200],'A4G4M0_Betaproteobacteria':0.33988746):0.10129671[0.1000],('Q82Y47_Betaproteobacteria':0.24774470,('Q3SHM9_Betaproteobacteria':0.13129149,(('Q0KCB2_Betaproteobacteria':0.15577126,'A4G4L9_Betaproteobacteria':0.27607791):0.03516716[0.0400],(('Q47HL0_Betaproteobacteria':0.18370629,'C1D6W4_Betaproteobacteria':0.20007856):0.04702567[0.1600],('Q9I2U9_Gammaproteobacteria':0.13246604,(((((('Q8XFG8_Salmonella':0.01853500,'P23869_Gammaproteobacteria':0.00538122):0.12295006[1.0000],'Q9CM92_Gammaproteobacteria':0.34823146):0.06035740[0.3000],'Q3IF50_Gammaproteobacteria':0.12172642):0.06503920[0.2400],'Q5QYP4_Gammaproteobacteria':0.24456823):0.10449846[0.4000],(('Q47XL4_Gammaproteobacteria':0.28211958,'Q2SK42_Gammaproteobacteria':0.16772511):0.09688367[0.3100],('A1SWR1_Gammaproteobacteria':0.17328963,('Q8EG23_Gammaproteobacteria':0.14914905,('Q9CM91_Gammaproteobacteria':0.22590593,'Q9KQZ8_Gammaproteobacteria':0.14608564):0.06921142[0.3700]):0.10007882[0.3300]):0.07254581[0.2300]):0.03757322[0.0300]):0.05906014[0.0500]):0.05114859[0.0100]):0.07234645[0.0400]):0.01088943[0.0000]):0.02878456[0.0100]):0.06825957[0.0000]):0.04998095[0.0000]):0.05246671[0.0000]):0.01840850[0.0000]):0.03548457[0.0000]):0.09962905[0.0000]):0.09238257[0.0000]):0.05895251[0.0100]):0.00000021[0.0200]):0.12668237[0.7500]):0.21998974[0.7700]):0.13032443[0.0300],(('Q897Q3_Clostridia':0.38621008,'Q24WG2_Clostridia':0.14637038):0.21579305[0.9300],(((('A0A3B6VF87_Spirochaetia':0.20864787,'A0A3B6VBW1_Spirochaetia':0.11094294):0.30168945[1.0000],'A0RXJ3_TACK':0.29397358):0.15328905[0.6700],('A0RYN7_TACK':0.81648297,(('Q1D3D1_Deltaproteobacteria':0.29588820,'A9ETL1_Deltaproteobacteria':1.04604194):0.34354752[0.7900],(((

(((('E0UUL5_Deltaproteobacteria':0.22398647,'B9LA78_Epsilonproteobacteria':0.25340931):0.05802159[0.6000],'D1B1F2_Deltaproteobacteria':0.30210887):0.11445634[0.4700],'E6WXU5_Epsilonproteobacteria':0.21321501):0.09728782[0.6200],'D3UH51_Helicobacter_mustelae':0.37397013):0.22919250[0.9500],'A0A1Q9P5P9_Asgard':0.48317520):0.16160159[0.4400],(('Q2NI82_Euryarchaeota':0.12417125,'Q12VY4_Euryarchaeota':0.21177156):0.07803929[0.2900],('C0ZC74_Bacilli':0.24159855,('P35137_Bacilli':0.20325342,('Q5KXK3_Bacilli':0.14857460,'B1YM99_Bacilli':0.28618853):0.04720429[0.5300]):0.09514110[0.6600]):0.11465594[0.8700]):0.08596158[0.5000]):0.09521752[0.4900]):0.30446872[0.5900]):0.08579725[0.0200]):0.12089761[0.0300]):0.10658170[0.0600],('Q4FL03_Alphaproteobacteria':1.15302974,('Q5FSI5_Alphaproteobacteria':0.39763272,((((('B3PZQ9_Alphaproteobacteria':0.41446175,'Q5LR96_Alphaproteobacteria':0.29859086):0.15732753[0.5300],'Q5LR95_Alphaproteobacteria':0.23203787):0.09331866[0.2700],'Q4FLZ3_Alphaproteobacteria':0.23515312):0.07035308[0.1800],'B3PZR0_Alphaproteobacteria':0.21939293):0.12566756[0.2400],('Q9A7Y5_Alphaproteobacteria':0.14388247,'Q2N943_Alphaproteobacteria':0.16049063):0.16687928[0.9700]):0.05877875[0.0400],('Q2RTK3_Alphaproteobacteria':0.26527472,(('B4S5R9_Chlorobia':0.26937520,'B4SHG7_Chlorobia':0.29346744):0.13871651[0.9400],('Q2NAT2_Alphaproteobacteria':0.63841669,(('Q9A7Y7_Alphaproteobacteria':0.40127546,'Q9A7Y6_Alphaproteobacteria':0.30323108):0.40416394[0.9800],('Q47UB6_Gammaproteobacteria':0.81410098,('Q3IIS9_Gammaproteobacteria':0.77365537,('Q2N8M0_Alphaproteobacteria':0.67787104,('D0LL82_Deltaproteobacteria':0.48966567,'Q8P7I6_Gammaproteobacteria':0.38715507):0.13150351[0.5400]):0.35275663[1.0000]):0.25235876[0.7000]):0.40812813[0.9600]):0.14040861[0.4800]):0.14554367[0.3900]):0.07406060[0.0000]):0.08874286[0.1000]):0.13213842[0.1300]):0.21608804[0.5400]):0.19084772[0.4500])OROOT;

Legends for Datasets

Data S1. The 188 representative organisms and their phylogenetic classifications. This file also lists the 181 alternative representative organisms.

Data S2. Numerical values of the proteome parameters.

Data S3. A list of all the unique ECOD X-groups, and Pfam clans, detected in representative proteomes of our ToL, and a list of ECOD X-groups conserved across the ToL.

Data S4. A list of all the unique ECOD X-group combinations detected in the representative proteomes of our ToL.

Data S5. Unique Pfam domain combinations, and clan combinations, detected in the representative proteomes; 'no clan' relates to Pfam families that are not assigned to a clan, and thus reflect an independent lineage.

Data S6. Expansion of beta-superfolds beyond the expansion of proteome size: for representative organisms of our ToL, shown are the percent of proteins in the proteomes that contain a domain related to one of the ECOD top hierarchies.

Data S7. Repeat segments identified in the representative proteomes.

Data S8. The percent of aggregation-prone proteins per proteome.

Data S9. The predicted disordered segments in the representative proteomes.

Data S10. Pfam-annotated domain combinations, and the UniProt identifiers, of the core and co-chaperones identified in the representative proteomes.

Data S11. mRNA abundance of core- and co-chaperones.

Data S12. Protein abundance of core- and co-chaperones.

SI References

1. S. Kumar, G. Stecher, M. Suleski, S. B. Hedges, TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol Biol Evol* **34**, 1812-1819 (2017).
2. S. Federhen, The NCBI Taxonomy database. *Nucleic Acids Res* **40**, D136-143 (2012).
3. D. A. Benson *et al.*, GenBank. *Nucleic Acids Res* **41**, D36-42 (2013).
4. R. D. Finn *et al.*, Pfam: the protein families database. *Nucleic Acids Res* **42**, D222-230 (2014).
5. H. Huber *et al.*, A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417**, 63-67 (2002).
6. E. Waters *et al.*, The genome of Nanoarchaeum equitans: insights into early archaeal evolution and derived parasitism. *Proc Natl Acad Sci U S A* **100**, 12984-12988 (2003).
7. S. B. Hedges, J. Marin, M. Suleski, M. Paymer, S. Kumar, Tree of life reveals clock-like speciation and diversification. *Mol Biol Evol* **32**, 835-845 (2015).
8. A. G. Simpson *et al.*, Evolutionary history of "early-diverging" eukaryotes: the excavate taxon Carpediemonas is a close relative of Giardia. *Mol Biol Evol* **19**, 1782-1791 (2002).
9. F. Burki, A. J. Roger, M. W. Brown, A. G. B. Simpson, The New Tree of Eukaryotes. *Trends Ecol Evol* **35**, 43-55 (2020).
10. R. D. Finn *et al.*, Pfam: clans, web tools and services. *Nucleic Acids Res* **34**, D247-251 (2006).
11. R. D. Schaeffer, Y. Liao, H. Cheng, N. V. Grishin, ECOD: new developments in the evolutionary classification of domains. *Nucleic Acids Res* **45**, D296-D302 (2017).
12. J. Jorda, A. V. Kajava, T-REKS: identification of Tandem REpeats in sequences with a K-meanS based algorithm. *Bioinformatics* **25**, 2632-2638 (2009).
13. S. W. Englander, L. Mayne, The nature of protein folding pathways. *Proc Natl Acad Sci U S A* **111**, 15873-15880 (2014).
14. P. Sormanni, L. Amery, S. Ekizoglou, M. Vendruscolo, B. Popovic, Rapid and accurate in silico solubility screening of a monoclonal antibody library. *Sci Rep* **7**, 8200 (2017).
15. Z. Dosztanyi, V. Csizmok, P. Tompa, I. Simon, The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J Mol Biol* **347**, 827-839 (2005).
16. C. UniProt, UniProt: a hub for protein information. *Nucleic Acids Res* **43**, D204-212 (2015).
17. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J Mol Biol* **215**, 403-410 (1990).
18. M. Marcet-Houben, T. Gabaldon, Beyond the Whole-Genome Duplication: Phylogenetic Evidence for an Ancient Interspecies Hybridization in the Baker's Yeast Lineage. *PLoS Biol* **13**, e1002220 (2015).
19. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797 (2004).
20. S. Capella-Gutierrez, J. M. Silla-Martinez, T. Gabaldon, trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973 (2009).
21. J. Huerta-Cepas *et al.*, eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* **47**, D309-D314 (2019).
22. P. Di Tommaso *et al.*, T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res* **39**, W13-17 (2011).
23. S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35**, 1547-1549 (2018).
24. F. B. Bastian *et al.*, The Bgee suite: integrated curated expression atlas and comparative transcriptomics in animals. *Nucleic Acids Res* **49**, D831-D847 (2021).
25. I. Papatheodorou *et al.*, Expression Atlas update: from tissues to single cells. *Nucleic Acids Res* **48**, D77-D83 (2020).

26. O. Cohen *et al.*, Comparative transcriptomics across the prokaryotic tree of life. *Nucleic Acids Res* **44**, W46-53 (2016).
27. K. Avican *et al.*, RNA Atlas of Human Bacterial Pathogens Uncovers Stress Dynamics Linked to Infection. *bioRxiv* 10.1101/2020.06.29.177147, 2020.2006.2029.177147 (2020).
28. J. McMillan, Z. Lu, J. S. Rodriguez, T. H. Ahn, Z. Lin, YeasTSS: an integrative web database of yeast transcription start sites. *Database (Oxford)* **2019** (2019).
29. M. Nadal-Ribelles *et al.*, Sensitive high-throughput single-cell RNA-seq reveals within-clonal transcript correlations in yeast populations. *Nat Microbiol* **4**, 683-692 (2019).
30. M. Wang, C. J. Herrmann, M. Simonovic, D. Szklarczyk, C. von Mering, Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. *Proteomics* **15**, 3163-3168 (2015).
31. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**, W256-W259 (2019).

CHAPTER 3

Interdomain communication suppressing high intrinsic ATPase activity of Sse1 is essential for its co-disaggregase activity with Ssa1

Interdomain communication suppressing high intrinsic ATPase activity of Sse1 is essential for its co-disaggregase activity with Ssa1

Vignesh Kumar, Joshua Jebakumar Peter, Amin Sagar, Arjun Ray, Mainak Pratim Jha, Mathieu E Rebeaud, Satyam Tiwari, Pierre Goloubinoff, Fnu Ashish, Koyeli Mapa

I contributed to this work conceptually and in the manuscript writing.

Author contributions

KM conceived the project, supervised the work, analyzed the biophysical and biochemical data, and wrote the manuscript. VK made all the constructs, purified the proteins, and performed the biochemical and yeast experiments. MPJ performed the SPR measurements along with VK. JJP performed the sm-FRET experiments, JJP and KM analyzed the sm-FRET data. JJP and AS obtained the SAXS data and JJP, AS, and FA analyzed the SAXS data. AR performed the MD simulation. ST purified the Sse1 domains and chimera for luciferase refolding and ATPase assays. MR and PG designed and performed the luciferase refolding and ATPase assay and analyzed the rates of refolding experiments and ATPase rates.

The FEBS Journal

FEBS PRESS science publishing by scientists

Interdomain communication suppressing high intrinsic ATPase activity of Sse1 is essential for its co-disaggregase activity with Ssa1

Vignesh Kumar[1,2], Joshua Jebakumar Peter[1], Amin Sagar[3], Arjun Ray[1,2,*], Mainak Pratim Jha[4], Mathieu E. Rebeaud[5], Satyam Tiwari[5], Pierre Goloubinoff[5], Fnu Ashish[3] and Koyeli Mapa[2,4]

1 Proteomics and Structural Biology Unit, CSIR-Institute of Genomics and Integrative Biology, New Delhi, India
2 Academy of Scientific and Innovative Research (AcSir), CSIR-HRDC, Ghaziabad, Uttar Pradesh, India
3 CSIR-Institute of Microbial Technology, Chandigarh, India
4 Department of Life Sciences, School of Natural Sciences, Shiv Nadar University, Greater Noida, India
5 Department of Plant Molecular Biology, University of Lausanne, Switzerland

Keywords
Hsp110; Hsp70; molecular chaperones; single-molecule FRET; small angle X-ray scattering

Correspondence
Fnu Ashish, CSIR-Institute of Microbial Technology, Sector 39 A, Chandigarh 160036, India
Tel: +91 172 6665472
E-mail: ashgang@imtech.res.in
and
K. Mapa, Department of Life Sciences, School of Natural Sciences, Shiv Nadar University, Greater Noida, India
Tel: +91 120 3819 138
E-mail: koyeli.mapa@snu.edu.in

***Present address**
Centre for Computational Biology, Indraprastha Institute of Information Technology, New Delhi,110020,India

Vignesh Kumar and Joshua Jebakumar Peter contributed equally for this work

(Received 27 February 2019, revised 8 July 2019, accepted 16 August 2019)

doi:10.1111/febs.15045

In eukaryotes, Hsp110s are unambiguous cognates of the Hsp70 chaperones, in primary sequence, domain organization, and structure. Hsp110s function as nucleotide exchange factors (NEFs) for the Hsp70s although their apparent loss of Hsp70-like chaperone activity, nature of interdomain communication, and breadth of domain functions are still puzzling. Here, by combining single-molecule FRET, small angle X-ray scattering measurements (SAXS), and MD simulation, we show that yeast Hsp110, Sse1 lacks canonical Hsp70-like interdomain allostery. However, the protein exhibits unique noncanonical conformational changes within its domains. Sse1 maintains an open-lid substrate-binding domain (SBD) in close contact with its nucleotide-binding domain (NBD), irrespective of its ATP hydrolysis status. To further appreciate such ATP-hydrolysis-independent exhaustive interaction between two domains of Hsp110s, NBD-SBD chimera was constructed between Hsp110 (Sse1) and Hsp70 (Ssa1). In Sse1/Ssa1 chimera, we observed undocking of two domains leading to complete loss of NEF activity of Sse1. Interestingly, chimeric proteins exhibited significantly enhanced ATPase rate of Sse1-NBD compared to wild-type protein, implying that intrinsic ATPase activity of the protein remains mostly repressed. Apart from repressing the high ATPase activity of its NBD, interactions between two domains confer thermal stability to Sse1 and play critical role in the (co)chaperoning function of Sse1 in Ssa1-mediated disaggregation activity. Altogether, Sse1 exhibits a unique interdomain interaction, which is essential for its NEF activity, suppression of high intrinsic ATPase activity, co-chaperoning activity in disaggregase machinery, and stability of the protein.

Abbreviations
MD simulation, molecular dynamic simulation; NEF, nucleotide exchange factor; SAXS, small angle X-ray scattering; SD, standard deviation; SEM, standard error of the mean; smFRET, single-molecule Förster Resonance Energy Transfer.

Introduction

Heat shock protein 70s (Hsp70s) belong to an extremely ancient, conserved group of molecular chaperones that are ubiquitously present in all ATP-containing subcellular compartments. They use ATP to fuel structural changes in stress-damaged proteins and are in general involved in protein quality control and a multitude of cellular regulatory processes. Hsp70s are assisted and regulated by two types of co-chaperones to accomplish their functions in physiological time scales: J-domain proteins (such as the Hsp40s or DNAJA/Bs) and nucleotide exchange factors (NEFs) [1–3]. While Hsp40s bind and target protein substrates onto Hsp70s and accelerate their weak ATPase rate, the NEFs accelerate ADP and protein release from Hsp70s, thereby promoting ATP rebinding and commencement of a new chaperone cycle. Hsp110 proteins are known to act as NEFs for cytosolic and ER (endoplasmic reticulum)-resident Hsp70s. They are exclusively found in eukaryotes along with several other structurally different NEFs, such as HspBP1 (Fes1 in yeast) and Bag domain NEFs [4]. While the other groups of NEFs that lack ATPase activity, albeit interact with overlapping surfaces on the NBD of Hsp70s, strangely, Hsp110s possess highly homologous primary sequences, starkly similar domain structure as well as domain organization with the Hsp70s, in particular a highly homologous NBD or ATPase domain. The structural intricacies along with molecular mechanism of NEF activity of Hsp110s were captured by a series of studies on yeast Hsp110, Sse1 [4–8]. Sse1 structure was solved in the ATP or ATP-analog-bound states either in isolation or in complex with various Hsp70 partners (with Hsp70 NBDs) [6–8]. In these structures, Sse1 has a lid-open structure, with extensive contacts between the nucleotide-binding domain (NBD) and substrate/peptide-binding domain (SBD/PBD) [6–8], closely mimicking the lid-open structure of ATP-bound bacterial Hsp70 (DnaK) and human BiP (Hsp70 of ER) [9,10]. Despite several ATP-bound structures, there is a caveat in alternate states (e.g., ATP-hydrolyzed state) of the Hsp110 protein so far, except a study where the authors had employed HD-X coupled to mass spectrometry on Sse1 in both ADP and ATP states [5]. The authors demonstrated that the signature of tryptic fragments of Sse1 in ATP and ADP-bound states was nearly identical leading to the conclusion that, in contrast to its Hsp70 partners (Ssa/Ssbs), Sse1 lacks ATP-hydrolysis-driven domain allostery [5]. It is therefore interesting to address further whether Hsp110s exhibit any noncanonical domain movements and the reason for such rigid static structure that is so contrasting with the ATP-fueled domain dynamics of their distant Hsp70s cognates.

Here, by employing single-molecule Förster Resonance Energy Transfer (FRET) measurements (sm-FRET) on yeast *Saccharomyces cerevisiae*, Hsp110, Sse1 in solution, we report that ATP hydrolysis-driven docking–undocking of NBD-SBD and opening–closing of alpha helical lid of SBD are absent in Sse1, in contrast to Hsp70s. Importantly, we were able to capture unique changes in conformation in the ADP and ATP states of the Sse1 using small angle X-ray scattering (SAXS) measurements. Using MD simulation, we captured subtle changes within the SBD of the Sse1 protein. Furthermore, to understand the importance of close NBD–SBD interactions, we constructed chimeric proteins of Sse1 and Ssa1 by domain shuffling. In chimeric proteins of Sse1/Ssa1, we observed Hsp70-like undocking of NBD and SBD resulting in abolished NEF activity of Sse1. This result indicated the importance of interaction of Sse1 NBD with self-SBD, which was disturbed despite the presence of structurally similar Hsp70 SBD. This data reiterated the importance of unique contacts between two domains of Sse1 for its NEF function, as was indicated by previous structural studies on the protein. Importantly, we found that isolated NBD of Sse1 contained an extremely high ATPase activity, which was also exhibited by the chimeric proteins of Sse1/Ssa1 (Hsp110/70) harboring Sse1-NBD, probably due to the disruption of NBD–SBD contacts, as revealed by SAXS measurements. This data implied that apart from maintaining the NEF activity of Sse1, unique interdomain contacts are critical for the repression of the high intrinsic ATPase rate of its NBD. Altogether, here we show that in spite of the unambiguous similarity in sequence, domain organization, and domain structure between Hsp110s and Hsp70s, a unique SBD–NBD interaction has been evolved in Hsp110s of yeast and possibly in other higher eukaryotes. The extensive two domain contacts are essential to suppress the otherwise high intrinsic ATPase activity of Sse1 NBD to facilitate its interaction as NEF with Hsp70 partners, which predominantly takes place in the ATP-bound state of the Sse1.

Results

Yeast Hsp110, Sse1, lacks prominent nucleotide-dependent canonical domain movements like Hsp70s

Several Sse1 structures are available in ATP or ATP-analog-bound states although there is scarce

information on the ADP state or other functional states of the protein. Previously, using different Hsp70 proteins, we and others have shown that the ADP-bound states of Hsp70s are inherently heterogeneous, compared to the ATP-bound states which remain masked in ensemble measurements [11–13]. Thus, to capture conformational changes and the presence of any heterogeneity in structural states of Hsp110s, we took yeast Hsp110, Sse1, as a model protein for single-molecule FRET (smFRET) measurements. For position-specific fluorophore labeling, we substituted the native cysteine residues of Sse1 to alanine and the cysteine-substituted Sse1 protein remained functionally unaffected (Fig. 1A). On the cysteine-less version of Sse1, we engineered the cysteine pairs at desired location for maleimide reactive fluorophore labeling. We made three double-cysteine mutants to monitor the canonical Hsp70-like domain movements by following the changes in distance by sm-FRET due to possible conformational alterations, between (a) NBD and SBD (E319C-D412C), (b) lid and base of SBD (D412C-D600C), and (c) SBD-lid and NBD (K45C-K600C) (Fig. 1B). We took single-molecule FRET measurements of donor–acceptor (Alexa 488-maleimide and Alexa 647-maleimide) labeled Sse1 proteins in solution in extremely low concentration (~ 50 pM) to achieve the single molecule resolution [11,14]. We used pulsed overlaid laser excitation to obtain FRET-ratio distribution exclusively from molecules which are labeled with an active donor and acceptor fluorophores. The sm-FRET methodology also termed as POLEX [modified version of Pulsed Interleaved Excitation (PIE), ALEX, and PAX [15–17], warranted that only molecules with active donor and acceptor fluorophores are considered for FRET ratio (FR) histogram] has been described in details before [11,14] and this technique enabled us to investigate structural transitions around zero FRET as the canonical donor-only zero FRET peak can be negated by this methodology and analogous smFRET techniques like PIE and ALEX [16,17].

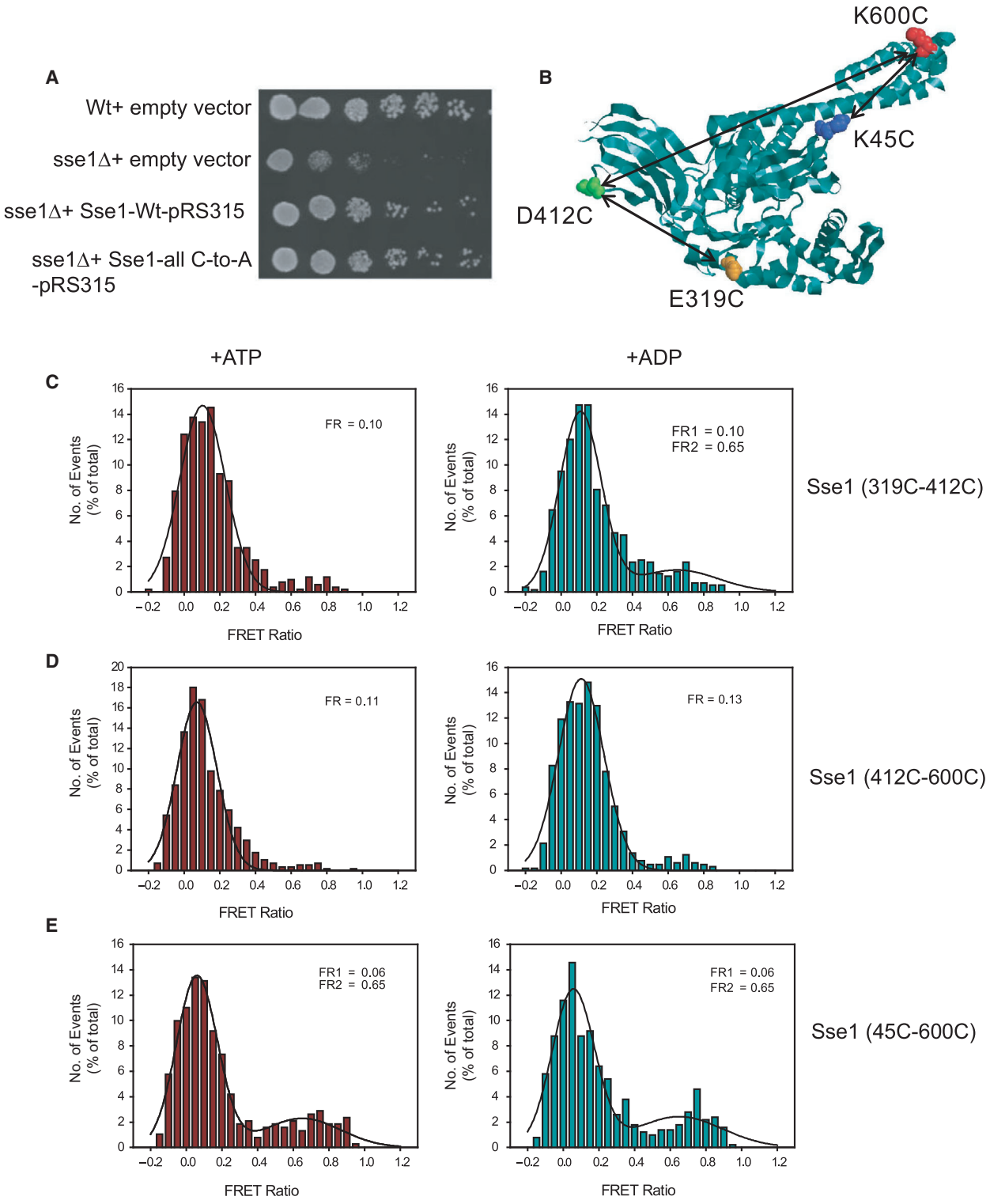
To monitor interdomain communication between NBD and SBD, we subjected Sse1 (319C-412C) for sm-FRET measurements. In the ADP-bound state, we observed a heterogeneous distribution of molecules. Majority of molecules populated at low FRET Ratio region with a peak centering at FR of 0.1. Additionally, a small population of molecules were also present at high FR region (FR = 0.65) (Fig. 1C, right panel). In the ATP-bound state, the FR distribution of molecules remained essentially the same (Fig. 1C, left panel). This finding is in stark contrast to Hsp70s where a prominent change in FR or FE (FRET efficiency) distribution was reported due to change in NBD–SBD interaction with ATP-hydrolysis due to domain undocking [12,13]. In case of Sse1, in spite of ATP-hydrolysis, we did not observe any alteration in the FR histograms, indicating an insignificant alteration in the interdomain distance in Hsp110s, as compared to Hsp70s.

Next, we probed the movements of SBD-lid with the mutant (D412C-D600C) and observed a single prominent peak at low FR region (FR = 0.1) in the ATP state, indicating an open-lid conformation of SBD, in complete agreement with the reported Sse1-ATP crystal structure [6]. In the ADP state, the FR distribution remained unchanged indicating the absence of prominent SBD-lid movements owing to ATP-hydrolysis (Fig. 1D).

To probe the relative position of open SBD-lid with NBD, we made a third double-cysteine mutant to report for any distance change between SBD-lid and NBD (K45C-K600C). This mutant revealed prominent bimodal distribution with FR peaks at 0.1 and 0.65, indicating heterogeneity in SBD-lid-NBD distance due to wide opening of lid and its contact with NBD (Fig. 1D). This conformational distribution remained unchanged with bound nucleotides and was essentially identical in both ATP and ADP states (Fig. 1E). These data indicate that indeed the lid of SBD mostly remains wide open and comes near NBD although opening and closing of lid do not dependent on ATP hydrolysis as has been shown with many Hsp70s. Altogether, the sm-FRET data with Sse1 show that the domain communication in case of Hs110s is unique and apparently does not recapitulate the nucleotide-dependent domain communication of Hsp70s. As we could not detect any noticeable ATP-hydrolysis-mediated changes in interdomain or intra-domain (SBD) distances, even in single molecule resolution, we reasoned that the nucleotide-induced conformational changes in Hsp110, if any, are noncanonical and are significantly different from the ones exhibited by their Hsp70 cognate proteins.

Small angle X-ray scattering measurements revealed distinctive nucleotide-dependent conformational changes of Hsp110

To detect any noncanonical conformational alteration during chaperone cycle of Hsp110s, we subjected Sse1 for SAXS measurements to determine large-scale changes in domain orientation. To investigate the effect of nucleotides on the solution conformation of Sse1, we acquired and analyzed SAXS data in nucleotide-bound states under saturating concentrations of



ATP and ADP (see Materials and methods and Tables 1 and 2) and without nucleotide. This without added nucleotide state is considered as an apo state of the molecule as there is instantaneous binding of added nucleotides as observed with fluorescently tagged ATP or ADP to the purified Sse1 (shown later). The SAXS intensity profiles showed no upward or downward profile at low angles indicating that the

Fig. 1. Sse1 lacks canonical domain movements as probed by single-molecule FRET experiments. (A) To assess any loss of function of Sse1 due to cysteine to alanine substitution, growth rescue potential of sse1Δ cells by mutant Sse1 (Sse1-allC-A mutant) was checked by transforming the mutant gene from centromeric plasmid pRS315 under endogenous promoter of sse1. Growth phenotype was assessed by drop dilution assay of sse1Δ cells harboring the mutant sse1 in pRS315 in comparison to empty pRS315 (vector control), and Wt-Sse1-pRS315 as positive control. (B) Ribbon structure of Sse1 (PDB ID: 2QXL) showing residues substituted with cysteine for donor–acceptor (Alexa 488-maleimide/Alexa 647-maleimide) fluorophore incorporation. Arrows indicate FRET pairs used to monitor domain movements. Following an initial incubation of individual dual-labeled Sse1 mutant (final 1 μM) with either ATP or ADP (final 2 mM), the protein was serially diluted in buffer A (as mentioned in the Materials and methods) containing respective nucleotides (final 2 mM), to a final concentration of ~ 50 pM, to perform the sm-FRET measurements. Representative histograms at two different nucleotide-bound states are shown for each mutant. Mean values of Gaussian fits for the FRET ratio distributions (FR1/FR2) are indicated. For plotting each sm-FRET ratio histogram, at least 500 single particles were taken. (C–E) smFRET ratio histograms showing population of Sse1 in various conformational states during ATP (Left panel) and ADP (right panel) bound condition. (C) smFRET ratio histograms of Sse1 (319C-412C) and Sse1 (45C-600C) (Panel E) containing FRET pairs between PBD and NBD show a low FRET ratio population in both ATP- and ADP-bound condition indicating lack of domain docking–undocking movements as observed in canonical Hsp70s. (D) smFRET measurements on Sse1 (412C–600C) with FRET pairs between β-sheet subdomain and α helical lid of PBD show significant low FRET population during ATP- and ADP-bound states indicating a lid-open conformation of the PBD in both the states.

Table 1. SAXS data collection and scattering-derived parameters.

Data collection parameters	
Instrument	SAXSpace (Anton Paar)
Beam geometry	10 mm slit
Wavelength (Å)	1.5418
q range ($Å^{-1}$)	0.010–0.300
Temperature (K)	283
Molecular mass determination	
Partial-specific volume ($cm^3 \cdot g^{-1}$)	0.724
Calculated monomeric M_r from sequence	73 470
Dry volume calculated from sequence ($Å^3$)	83 220
Software employed	
Primary data reduction	SAXSQUANT
Data processing	PRIMUS QT
Ab initio analysis	DAMMIF/DAMMIN
Validation and averaging	DAMAVER
Rigid-body modeling	N/A
Computation of model intensities	CRYSOL
Three-dimensional graphics representations	PYMOL

samples were free of aggregation and there was no interparticle interference in any of the conditions (Fig. 2A). The molecular weight of Sse1 was calculated to be in the range of 75–86 kDa using the consensus Bayesian assessment of different concentration-independent measures of molecular weight including Porod Invariant, volume of correlation, size and shape of *ab initio* reconstructed models [18]. Linear profile of the Guinier analysis confirmed monodispersity of the scattering species profile (Fig. 2A inset). The slope of the linear fits provided radius of gyration (R_g) of Apo-, ATP-, and ADP-enriched states to be 3.71 ± 0.14, 3.85 ± 0.11, and 3.75 ± 0.12 nm, respectively, corroborating nicely with the sm-FRET results eliminating the possibility of major conformational changes of Sse1 with the nature of the bound nucleotide. Computed $P(R)$ curves for Apo-, ATP-, and ADP-enriched states showed a maximum linear dimension (D_{max}) values of 13, 12, and 12.2 nm and R_g values of 4.0 ± 0.07, 3.8 ± 0.16, and 3.8 ± 0.07 nm, respectively (Fig. 2B). Normalized Kratky plots for the Apo- and ATP-bound states had maxima close to 1.73 supporting a compact, globular, and well-folded conformation (Fig. 2B inset). The ADP-bound state had the maxima shifted slightly toward a higher value (~ 2.0) possibly indicating an increase in the flexibility in the ADP-bound state. Overall, the shape parameters indicated no large-scale shape change in the Sse1 protein as a function of bound nucleotides except a possible increase in flexibility in the ADP-bound form.

To visualize the global shape of the Sse1 protein with or without bound nucleotides, SAXS data profile was used to restore the scattering shape of the molecule under Apo-, ATP-, or ADP-bound state. Dummy atom modeling calculations were performed to restore 10 different models for each state and were averaged after alignment of their inertial axes. The values of average Normalized Spatial Discrepancy with the corresponding standard deviations between the 10 models were 0.55 ± 0.017, 0.76 ± 0.012, and 0.86 ± 0.02 for Apo-Sse1, ATP-Sse1, and ADP-Sse1, respectively, indicating that the modeling procedure was stable and reproducible (Fig. 3A). Upon superimposition, it was clear that the model for Apo state was best comparable with crystal structure of Sse1 in the ATP state (PDB ID: 2QXL) with NSD of 1.3 (Fig. 3B, columns 1–4) [6]. ATP- or ADP-bound states were best comparable to Sse1 crystal structure in ADP-BeF_3 (non-hydrolysable ATP analog) in complex with bovine Hsc70 NBD (PDB ID: 3C7N) with NSD of 1.2 in contrast to ATP-bound structure (PDB ID: 2QXL) with NSD of

Table 2. SAXS data collection and scattering-derived parameters.

Data collection parameters	Apo-Sse1	ATP-Sse1	ADP-Sse1	Chimera AEE
Exposure time (min)	45	45	45	45
Concentration range (mg·mL^{-1})	6	6	6	3
Structural parameters				
$I(0)$ (AU) [from $P(R)$]	29 960 ± 418	31 330 ± 1487	33 840 ± 493	7414 ± 748
R_g (Å) [from $P(R)$]	4.0 ± 0.07	3.8 ± 0.16	3.8 ± 0.07	4.25 ± 0.91
$I(0)$ (AU) (from Guinier)	29 384 ± 682	34 680 ± 644	34 727 ± 566	7335 ± 46
R_g(Å) (from Guinier)	3.71 ± 0.14	3.85 ± 0.11	3.75 ± 0.12	4.13 ± 0.82
D_{max} (nm)	13.3	12.0	12.2	15.2
Molecular mass determination				
Molecular mass M_r (from Consensus Bayesian Estimation)	79.2	76.8	85.5	92.6

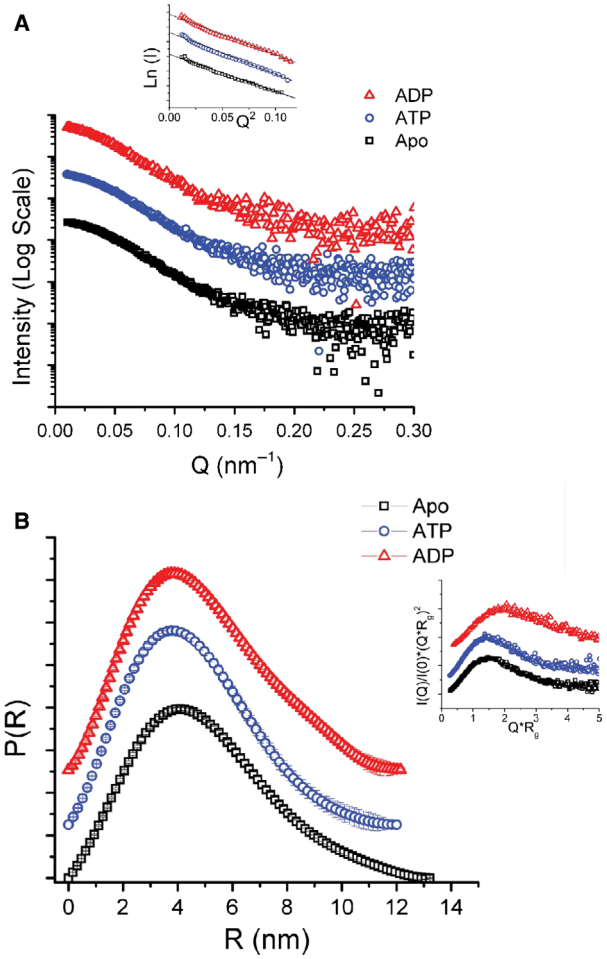


Fig. 2. SAXS profile of Apo- and ATP/ADP-bound forms of Sse1. (A) SAXS intensity profiles of Sse1 in Apo form (Apo-Sse1) and bound to ATP (ATP-Sse1) or ADP (ADP-Sse1) are plotted as $I(Q)$ (log scale) vs. Q (linear scale) with the Guinier Plots [ln $I(Q)$ vs. Q] in the inset. The black solid lines in the Guinier plot show the linear fits. (B) The distance distribution curves [$P(R)$ vs. R] derived by Indirect Fourier transformation of the intensity profiles with the normalized Kratky plots [$I(Q)/I(0)*(Q*R_g)^2$ vs. $Q*R_g$] in the inset.

2.5 (Fig. 3C,D columns 1–4 and 3E) [8]. Technically, the two crystal structures are grossly the same, their RMSD being only 0.5 Å over 453 residues (residues 385–389, 466–473, and 501–525 are missing in PDB: 2QXL due to poor resolution). We compared the theoretical SAXS profiles of the models based on the structures used for the inertial axis alignment with the corresponding experimental SAXS profiles. The theoretical SAXS profiles of the model of Sse1 based on the crystal structure of ATP-bound state fit well with the experimental SAXS profile in Apo state with a χ^2 value of 1.4 (Fig. 3B, column 5). In addition, the theoretical SAXS profiles of the model based on ADP-BeF_3 bound crystal structure agrees with the experimental SAXS profiles in the ATP- and ADP-bound states with χ^2 values of 1.5 and 1.8, respectively (Fig. 3C,D, column 5). It is to be noted here that the theoretical SAXS profiles of the two models are very similar owing to the high structural similarity. Altogether, comparison of the SAXS-based models and SAXS profiles (theoretical and experimental) confirmed that Sse1 protein remains monomeric and its shape largely remains unchanged except for few regions of SBD, regardless of the bound nucleotide status.

Thus, the SAXS data very nicely correlate with sm-FRET data with three different FRET-based reporters described in the last section. Altogether, sm-FRET spectroscopy and SAXS measurements collectively suggest that Hsp110s do not exhibit canonical ATP-hydrolysis-driven domain movements as seen in Hsp70s.

To further realize the apo state of Sse1 at a higher resolution, we generated it using all-atomistic MD (Molecular Dynamics) simulation. MD simulation was performed on the nucleotide-removed form of Sse1 (PDB ID: 3C7N) using GROMACS 4.6.1 (see Materials and methods for details). The simulation revealed that the protein undergoes ~ 5 Å deviation from the

starting structure within the first 100 ns, we confirmed the stability of the protein for a longer duration (~ 900 ns) (Fig. 4A, pink vs blue cartoons and Fig. 4B). MD stabilized final structure (Fig. 4C) was found to be very similar to the nucleotide bound structure, but SBDβ was stabilized and was folded within and was not extended as seen in the ADP-BeF_3-bound crystal structure of Sse1 (PDB ID: 3C7N). We also found that few segments of SBDβ and SBDα were highly dynamic (Fig. 4E, left and right panels). Finally, comparison of MD simulations (pink cartoon in Fig. 4A) with SAXS data (golden cartoon in Fig. 4A,D) suggest that nucleotide-driven conformational changes occur prominently at localized regions of SBDβ, especially at residues 510–540. The SAXS-fitted model also shows an extension of the last ~110 residues at the C terminus of SBDα (Fig. 4A). All these results collectively demonstrate that nucleotide binding imparts atypical conformational changes, which might be characteristic of Sse1 or Hsp110s in general, the significance of such conformational changes for chaperoning function of the protein remains to be explored in details.

Hsp70-like domain conformation is recapitulated in chimeric constructs of Ssa1 and Sse1

As the previous high-resolution structures of Sse1 (either in isolation or in complex with Hsp70 partners) have shown extensive interdomain contacts between its NBD and SBD, we were curious to know the outcome of artificial fusion of the N-terminal domain of Hsp70 (NBD) and the C-terminal domain (SBD) of Hsp110 protein and vice versa. We asked the question whether the chimeric molecule would be stable as a two-domain protein. If it is stable, will the protein regain Hsp70-like domain movements, or remain static like Hsp110s? Is there a dominant effect of one of the domains in determining the nature of interdomain allostery? There are reports in literature on chimeric yet functional Hsp70s [19–22] and Hsp110s [23], resulting from domain shuffling between different Hsp70s or Hsp110s. These chimeric proteins proved that individual domain function as well as allosteric communication can be maintained in a synthetic protein with similarly folded domains from foreign proteins. Considering the overall similar domain architecture of Hsp70s and Hsp110s, we made chimeric proteins by shuffling the domains of Sse1 and its Hsp70 partner, Ssa1. Taking into consideration the importance of the interdomain linker for solubility of NBDs of some Hsp70s [24], we shuffled the NBDs with and without the respective linker sequences resulting in four different chimeric proteins. For simplicity, we have named them as chimera AAE, AEE, EEA, and EAA, where the three letters denote the source of NBD, linker, and SBD, respectively, from either Ssa1 (A) or Sse1 (E) (Fig. 5A). All chimeras were expressed with cleavable N-terminal hexa-histidine tags in *Escherichia coli* cells and we found that after induction at 30 °C, chimera AAE and AEE were mostly found in the soluble fraction, although chimera EEA and EAA were majorly found in the inclusion bodies (Fig. 5B,C). Upon expression at lower temperature (18 °C overnight induction), solubility of chimera EEA and EAA was improved significantly and we were able to purify these proteins from the soluble fraction (Fig. 5D). *In vivo* expression of full-length chimeric proteins was checked in yeast cells by expressing the chimera in *sse1*Δ cells under the endogenous promoter followed by immunoblotting with polyclonal anti-Sse1 antibody raised against the full-length Sse1 protein. Indeed, all four chimeric proteins were expressed as full-length proteins in yeast cells and could be detected by anti-Sse1 antibody (Fig. 5E).

After we established that the chimeric proteins are stable, we were interested to study the domain arrangement of the chimeric proteins. To achieve that, we subjected chimera AEE to SAXS measurements (Fig. 6A). SAXS data analysis and modeling were done as described previously. The Guinier analyses revealed larger R_g of 4.13 ± 0.8 [4.25 ± 0.9 based on $P(R)$] for chimera AEE in its apo state (Fig. 6A inset) and $P(R)$ distribution profile showed larger D_{max} of 15.2 nm (Fig. 6B) compared to the wild-type Sse1 protein. Normalized Kratky plot of chimera AEE had the peak significantly shifted from the value of 1.73 indicating an increase in the flexibility of the protein (Fig. 6B, inset). This shift is much more prominent than the shift seen in case of ADP-bound Sse1. Notably, SAXS model of apo-AEE could not be fitted to neither the Sse1-ATP structure (PDB code 2QXL [6]; NSD 3.7) nor to the Sse1-ADP*BeF_3 structure (PDB code 3C7N with NSD 4.8) [8]. On the contrary, its overall shape is most similar to the shape of ADP- and peptide-bound state of DnaK (PDB code: 2KHO; NSD 1.1) [25] (Fig. 6C,D). We explicitly state here that the ideal scenario would be to model an ensemble of structures for chimera AEE because of the flexibility indicated by the SAXS data. However, this analysis is not included in the current study because of the absence of an all-atom structure. An overlay of the theoretical and experimental SAXS intensity profile is also not included for the same reason.

Overall, our data of chimera AEE suggest that the domain-undocked, lid-closed state of Hsp70s typically

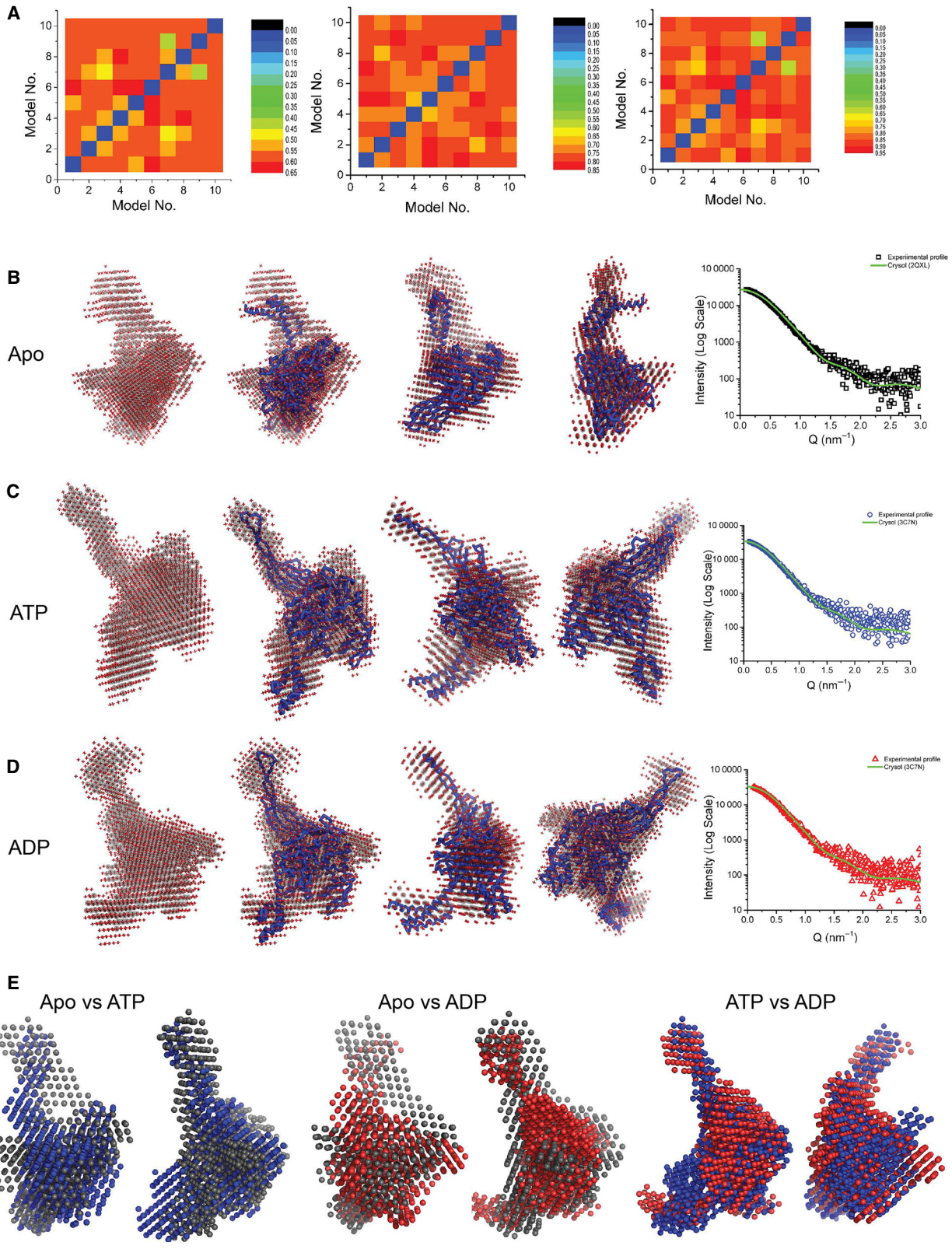

A
Model No.
Model No.
B
Apo
Experiimental profile
Crysol (2QXL)
Intensity (Log Scale)
Q (nm^{-1})
C
ATP
Experimental profile
Crysol (3C7N)
Intensity (Log Scale)
Q (nm^{-1})
D
ADP
Experimental profile
Crysol (3C7N)
Intensity (Log Scale)
Q (nm^{-1})
E
Apo vs ATP
Apo vs ADP
ATP vs ADP

Fig. 3. SAXS models of Sse1 in apo and ATP/ADP states. (A) From left to right, color maps of the cross-NSD values of all DAMMIF generated models for Sse1-Apo, Sse1-ATP, and Sse1-ADP. All the models satisfied the inclusion criteria of Mean NSD + 2*Standard Deviation and were included in the averaging procedure. (B–D) Starting from top, the first column shows the damfilt (Black) and damaver (Red) models derived after the averaging of 10 dummy residue models for Apo-Sse1 (Panel B), ATP-Sse1 (Panel C), and ADP-Sse1 (Panel D), respectively. The following three columns show three orthogonal views of the SAXS-derived models of Apo-Sse1 overlaid with the crystal structure of ATP bound form (PDB ID: 2QXL) using inertial axis alignment in the top row and the SAXS-derived models of ATP-Sse1 and ADP-Sse1 overlaid with the ADP-BeF_4-bound crystal structures (PDB ID: 3C7N) in the C and D panel rows, respectively. The last column shows the comparison of the theoretical SAXS profile of the model of Sse1 based on the crystal structure used in the inertial axis alignment and the experimental SAXS profile. (E) Starting from left, two orthogonal views each of the overlay of the SAXS models of Apo-Sse1 (Black) with ATP-Sse1 (Blue), Apo-Sse1 with ADP-Sse1 (Red), and ATP-Sse1 with ADP-Sse1.

found in peptide- and ADP-bound states of Hsp70s, which seems to have been lost in Hsp110s, however, can be artificially recapitulated in Hsp110, especially the lid-closed state of its SBD. We tried to capture the structural features of other type of chimera where the NBD is obtained from Sse1 and SBD from Ssa1 (chimera EEA and EAA). Unfortunately, at the high protein concentrations required for SAXS measurements, both of these chimeras were extremely aggregation-prone and we were unable to produce the data required to build a SAXS model.

Ssa1-Sse1 chimeric proteins lack the NEF activity but can rescue the growth defects due to deletion of sse1

As the chimera AEE exhibited undocking of NBD and SBD, and specifically the apparent closure of the lid onto its base which was never visible for wild-type Sse1, it was interesting to check whether the chimera retains the activities of the Sse1 protein and possibly gain some activities of Ssa1. For that purpose, we addressed the NEF activity of chimera *in vivo* by checking the rescue capacity of synthetic lethal phenotype of *sse1*/*sse2* double deletion by the chimera. None of the chimera could rescue the synthetic lethal phenotype upon expression under Sse1 promoter (Fig. 7A). Importantly, overexpressing chimeras also failed to rescue the synthetic lethal phenotype of double deletion of *sse1* and *sse2*, whereas, in contrast, overexpressing Fes1, a cytosolic NEF, could rescue the lethal phenotype as described earlier (Fig. 7B) [26]. This result led to the conclusion that although the chimeric proteins are stable two-domain proteins and despite the fact that chimera EEA and EAA presented a *bona fide*, properly folded NBD from Sse1, which should in principle harbor all of Sse1's NEF activity by interacting with the Hsp70 NBD, they are not able to perform the NEF function of holo Sse1 (EEE). This finding reiterated that NBD–SBD contacts are playing an essential role for the NEF activity of the Hsp110 proteins, as shown by previous mutational studies based on the crystal structure of the protein [7]. Next, we checked whether the chimeric proteins harbor any residual exchange factor function that might be insufficient to rescue the synthetic lethal phenotype of double deletion of *sse1* and *sse2*. All four recombinant chimeras were checked for nucleotide exchange activity by following the change in fluorescence signal of MANT-ATP upon release from Hsp70s after being exchanged with nonfluorescent nucleotides. Wild-type Sse1 showed prominent exchange activity while none of the four chimeras showed any exchange activity (Fig. 7C). This result points toward the fact that important communication or interactions between SBD and NBD are critical for the NEF activity of Sse1 (HSP110) proteins.

*sse1*Δ cells show a prominent growth phenotype even at permissive temperature despite the presence of paralogous *sse2,* another Hsp110 NEF for Hsp70s, as well as other NEFs like Fes1, it indicates other important cellular functions Sse1 that are in addition to its NEF function. To check whether growth phenotype of *sse1*Δ strain can be rescued by these chimeric proteins, all chimeric proteins were expressed from endogenous promoter of Sse1 retaining the 3′UTR of Sse1 in centromeric plasmids and the plasmids were transformed in *sse1*Δ cells. All the resulting *sse1*Δ yeast strains harboring chimera were grown under permissive (30 °C) as well as heat shock condition (37 °C) and in the presence of proteotoxic stressor 2-D-G (an inhibitor of glycolysis) [27]. When the growth of *sse1*Δ cells were compared with the wild-type cells, prominent growth defect of *sse1*Δ cells was evident both in the presence and the absence of stressor molecules (Fig. 7D). This growth phenotype was efficiently rescued by wild-type Sse1 from plasmid. Interestingly, chimera AAE and AEE showed efficient growth rescue under all conditions, chimera EEA and EAA were less efficient in rescuing the growth phenotype like AAE and AEE, especially at heat shock condition (Fig. 7D). This result indicate that the conserved Sse1-SBD, which is

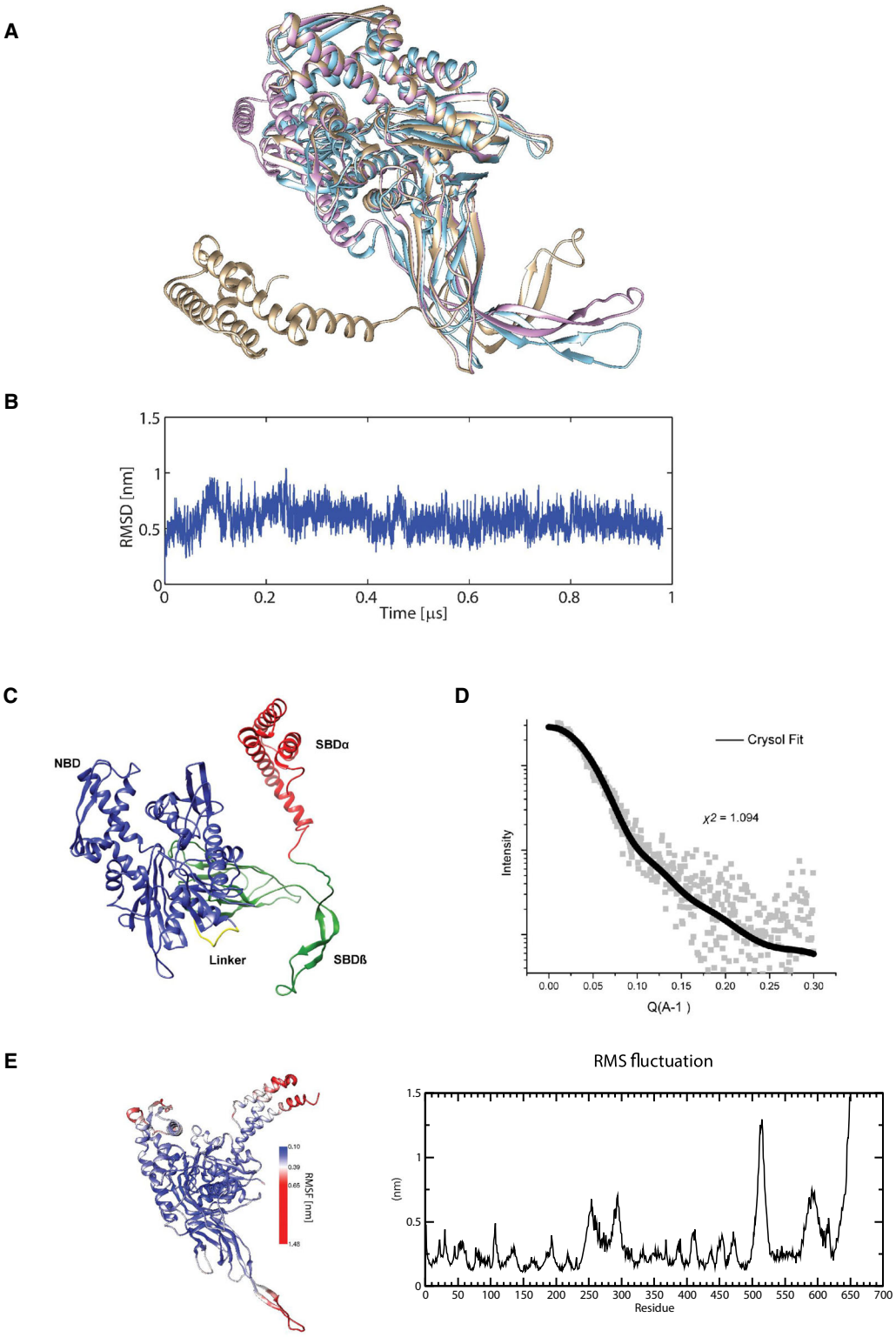

A
B
RMSD [nm]
Time [μs]
C
NBD
SBDα
Linker
SBDß
D
Crysol Fit
χ2 = 1.094
Intensity
Q(A-1)
E
RMSF [nm]
RMS fluctuation
(nm)
Residue

Fig. 4. Molecular dynamics simulation of Sse1 reveals highly dynamic regions within the protein. (A) Structural alignment of (a) Sse1 crystal structure (PDB ID: 3C7N) (blue), (b) molecular dynamics resolved final structure (pink) and SAXS-fitted structure (golden) reveals the maximum motion at the SBD domain. The SAXS-fitted model also shows an extension of the last ~ 110 residues at the C terminus. (B) Root mean square deviation (RMSD) was calculated for the all-atomistic one microsecond (1 μs) molecular dynamics simulation of native Sse1 and the structure was found to be stable post ~ 100 ns. (C) Cartoon representation of MD stabilized final structure of Sse1 in Apo form. NBD has been depicted in blue, interdomain linker in yellow, SBD-β in green and SBD-α in red. (D) CRYSOL fitting of SREFLEX refined MD stabilized structure of Apo form of Sse1 with the experimental SAXS intensity profile gave χ^2 value of 1.094 suggesting a good fit with the experimental SAXS profile. (E) Left panel: Root mean square fluctuation of the protein reveals C-terminal helices and part of the beta-sheet region to be highly dynamical. Right panel: Residue-wise RMS fluctuation.

structurally distal to the interaction surfaces involved in the NEF function of the Sse1-NBD, is nonetheless necessary to carry additional chaperone functions, which can be accomplished *in vivo* by Sse1/Ssa1 chimeric proteins containing the Sse1-SBD.

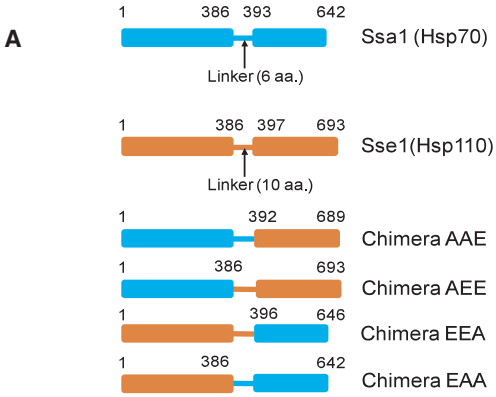
The SBD of Sse1 exerts chaperoning activity on self-NBD and structurally similar foreign NBDs

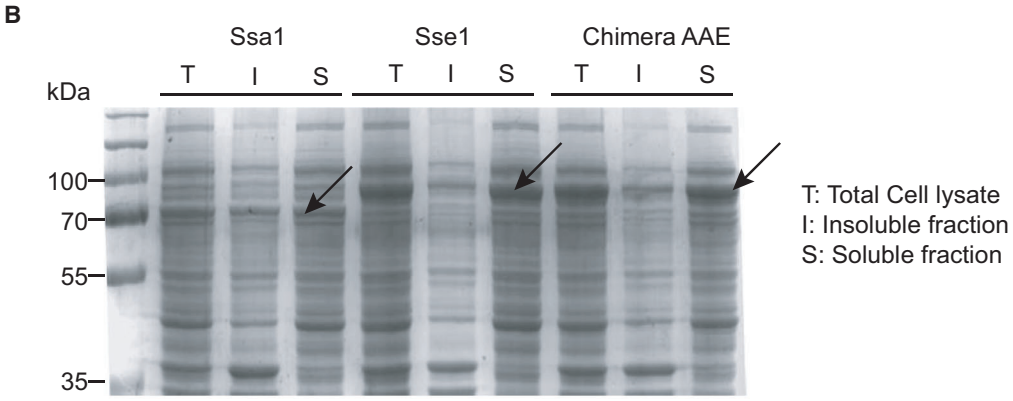
To corroborate our *in vivo* observation that chimera EEA and EAA were less efficient compared to chimera AAE and AEE in rescuing the growth defects of *sse1*Δ cells under high temperature plausibly due to low stability, we compared *in vitro* thermal stability of all chimeric proteins at 42 °C. Wild-type Sse1 and Ssa1 were kept as controls. We monitored the intensity of light scattering at 320 nm as an indicator of aggregate formation (Fig. 7E). The light scattering of chimera AAE, AEE, and wild-type Sse1 remained in the baseline level indicating high intrinsic stability. In contrast, within the first five minutes of incubation at 42 °C chimera EEA and EAA showed a significant increase in scattering indicating aggregation. Interestingly, wild-type Ssa1 also demonstrated significant scattering albeit appreciably less compared to chimera EEA and EAA. This experiment hinted that chimera EEA and EAA are unstable and aggregate at 42 °C due to thermal denaturation, whereas chimera AAE and AEE are more stable at this temperature. We further subjected all protein to thermal melting by gradually increasing the temperature from 25 °C to 95 °C and by monitoring the scattering at 320 nm wavelength. As shown in Fig. 7F, chimera AAE and AEE start aggregating only above 65 °C and are thus as thermally stable as wild-type Sse1. In contrast, chimera EEA and EAA are much less stable and started aggregating at 45 °C and above. Interestingly, in accordance with Ssa1 undergoing much more dynamic structural changes than Sse1 during the chaperone cycle, we found that wild-type Ssa1 was much less thermally stable than Sse1. In summary, we observed that chimeric proteins with Ssa1 NBD and Sse1 SBD are thermally stable like wild-type Sse1, and are significantly more stable than Ssa1 and the chimera containing the Sse1 NBD and the Ssa1 SBD. Our data demonstrate that Sse1 SBD has a strong chaperoning capacity which helps to stabilize not only its own NBD, as exhibited by great thermal stability of wild-type Sse1 protein, but also the NBDs of similar architecture of foreign proteins, like Ssa1 (as shown by substantial thermal stability of chimera AAE and AEE). This chaperoning activity of Sse1-SBD on NBDs (of Hsp110 or Hsp70) is independent of wild-type NBD-SBD contacts found in Sse1 as evidenced by domain-undocked SAXS profile of chimera AEE and loss of NEF activity of chimera AAE and AEE as shown previously by *in vitro* and *in vivo* assays.

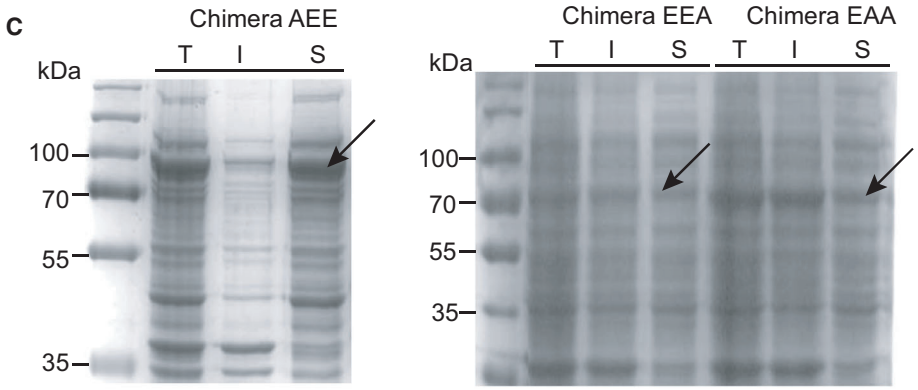
Sse1-Ssa1 chimera reveal the importance of substrate-binding domains of Sse1 and Ssa1 as effective disaggregating co-chaperone and chaperone respectively

As the Sse1-Ssa1 chimera could complement the growth phenotype of *sse1*Δ strain, at least partially, we wanted to check the chaperoning activity of the chimera *in vitro*. For that purpose, we performed refolding assay of heat-inactivated firefly luciferase that were pre-aggregated in the absence of chaperones. We tested the activity of the four chimeras (AAE, AEE, EEA, and EAA) compared to Sse1 (EEE), as potential co-chaperones of the ATP-dependent-disaggregase Ssa1, or as potential disaggregating chaperones by themselves.

First, we chose near physiological concentrations of Ssa1 and Sis1 (J-domain co-chaperone), to address the ability of near physiological concentrations of Sse1 to promote the disaggregation of pre-aggregated luciferase, and test the possible Sse1-like properties of the Sse1/Ssa1 chimera. As expected, WT-Sse1 strongly activated the ATP-fuelled Ssa1-Sis1-mediated disaggregation and refolding reaction. In contrast, addition of the same amount of any of the four chimeras did not compensate for the lack of WT-Sse1, and any


A
1
386
393
642
Ssa1 (Hsp70)
Linker (6 aa.)
1
386
397
693
Sse1(Hsp110)
Linker (10 aa.)
1
392
689
Chimera AAE
1
386
693
Chimera AEE
1
396
646
Chimera EEA
1
386
642
Chimera EAA


B
Ssa1
Sse1
Chimera AAE
T I S T I S T I S
kDa
100
70
55
35
T: Total Cell lysate
I: Insoluble fraction
S: Soluble fraction


C
Chimera AEE
T I S
kDa
100
70
55
35
Chimera EEA
Chimera EAA
T I S T I S
kDa
100
70
55
35

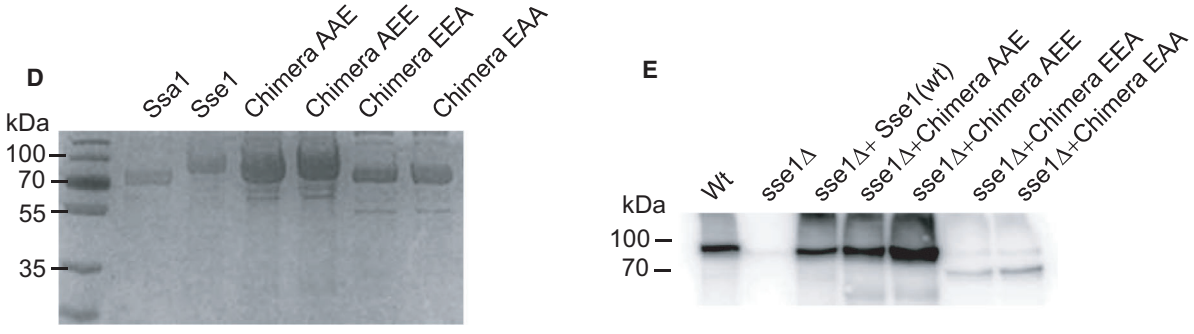

D
Ssa1
Sse1
Chimera AAE
Chimera AEE
Chimera EEA
Chimera EAA
kDa
100
70
55
35
E
Wt
sse1Δ
sse1Δ+ Sse1(wt)
sse1Δ+Chimera AAE
sse1Δ+Chimera AEE
sse1Δ+Chimera EEA
sse1Δ+Chimera EAA
kDa
100
70

Fig. 5. Construction of Sse1–Ssa1 chimeric proteins. (A) Schematic representation of chimeric constructs of Sse1 and Ssa1. The number denotes the amino acid numbers. (B, C) Solubility assay for wild-type Ssa1, Sse1, and Hsp110-70 chimeric proteins: The expression and the solubility of the wild-type Sse1 and all the chimeras were assayed by fractionating the whole cell lysate of BL21-DE3 cells into soluble (S) and insoluble fractions (I) by centrifugation. The total cell lysate before (T) and the soluble (S) and insoluble fractions (I) were run in SDS/PAGE and the gels were stained with Coomassie brilliant blue (CBB). The overexpressed protein bands have been shown with an arrow in the respective soluble fraction lanes. (D) The wild-type and chimeric proteins of Sse1 and Ssa1 were purified as described by in the Materials and methods. The CBB-stained SDS/PAGE shows the purified proteins. (E) The expression of wild-type Sse1 and four chimera were checked by expressing the proteins from overexpression vector (pJV340) in the *sse1*Δ cells and detected with anti-Sse1 polyclonal antibody by western blot.

significant increase over the basal rate of Ssa1 (Sis1)-mediated unfolding/refolding was not observed (Fig. 8A). This data confirmed that WT-Sse1 is an effective disaggregating co-chaperone of Ssa1 (Fig. 5A). In contrast, none of the Sse1-Ssa1 chimera, even EEA and EAA presenting a bona fide Sse1 NBD for NEF activity, could assist Ssa1(Sis1) in disaggregation of pre-aggregated luciferase, implying loss of co-chaperoning activity due to abolished NEF function. The current proposed disaggregation mechanism by Ssa1 assisted by NEF, Sse1, assign no functional role either to the linker or to the SBD of Sse1, although both structural elements are highly conserved in evolution since the birth of eukaryotes where Hsp110 diverged from the Hsp70s. Our result indicates that the presence of a *bona fide* Sse1 linker and SBD is essential for an effective disaggregation/refolding by the Ssa1(Sis1)-Sse1 chaperone system, and that the transient binding of only of an Sse1-NBD, acting as a mere nucleotide exchange factor of Ssa1 apparently does not suffice to mediate an effective disaggregation reaction.

Furthermore, transient-specific binding of Ssa1, through its NDB, to the M-domain of Hsp104, is known to cause the transient activation of Hsp104's disaggregating activity, leading to a cooperative conversion of stable pre-aggregated luciferase into native species by both chaperones. In the presence of constant near physiological amounts of Hsp104 and Sis1, subphysiological limiting amounts of Ssa1 led to a rather effective disaggregation and refolding of luciferase and further addition of 1 μM of Ssa1 expectedly increased the refolding yield by 36% (Fig. 8B). Addition of 1 μM of Sse1, or of chimera AAE and AEE both harboring Ssa1-NBD, did not affect the basal refolding yield. In contrast, addition of 1 μM chimera EEA and EAA harboring Sse1-NBDs strongly inhibited the basal refolding yields of the reaction (Fig. 8B). The fact that Sse1 did not inhibit Ssa1-Sis1-Hsp104-mediated disaggregation and refolding reaction and in contrast EAA or EEA, both strongly inhibited the reaction, further indicates these two chimeras, because of the presence of Ssa1-SBD possibly capture some of the misfolded luciferase which, however, cannot be subsequently released to refold to the native state, like the case with Wt-Ssa1.

It is well established that upon binding to the M-domain of Hsp104, NBD of Ssa1 activates the disaggregation activity of Hsp104 and that Ssa1 further assists Hp104 in the ATP-fuelled conversion of the disaggregation products into native proteins [5,28,29]. Near-physiological amounts of Hsp104 and Sis1 without addition of Ssa1 expectedly remained ineffective as disaggregase, while addition of a limiting amount of Ssa1 activated well the ATP-fuelled disaggregation and refolding reaction (Fig. 8C). In contrast, neither Sse1 nor any of the chimera was able to replace the activity and aggregate processing functions of the wild-type Ssa1 molecule.

Our results suggest that although Sse1's NBD is structurally highly homologous to Ssa1's NBD and that both can bind and hydrolyze ATP, Sse1-NBD may neither interact with the M-domain of Hsp104 to activate it, nor act alike Ssa1, as an ATP-fuelled disaggregase by itself both up- and downstream the Hsp104 machinery. Interestingly, chimera AAE and AEE harbor Ssa1-NBD and which were expected to bind Hsp104's M-domain and activate its disaggregation activity, remained ineffective. This suggests that in addition to the NBD of Ssa1, the presence of a functional Ssa1-SDB is central for activation of Hsp104 as a disaggregase and for their effective collaboration at converting stably misfolded proteins into native proteins.

Sse1-NBD harbors a high intrinsic ATPase activity which remains suppressed by intimate NBD–SBD contacts of the protein

As the NEF activity was completely abrogated in chimeric proteins of Sse1/Ssa1 and (co)chaperoning activity of parent proteins was not observed for chimera, we were interested to check the conservation of individual domain functions of chimera. Domain functions were determined by measuring the ATP-

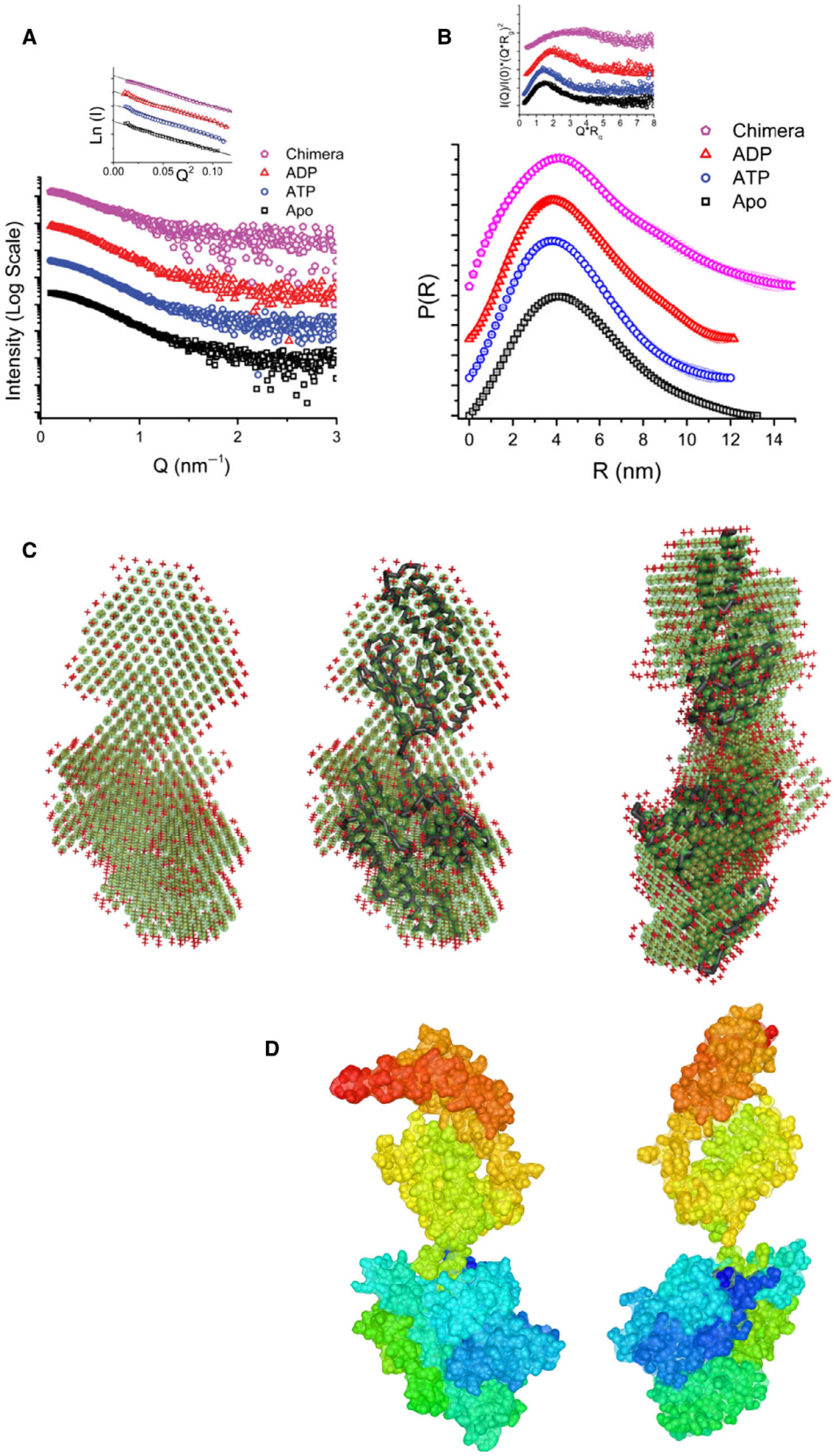

A
Ln (I)
0.00
0.05 Q²
0.10
Chimera
ADP
ATP
Apo
Intensity (Log Scale)
0
1
2
3
Q (nm⁻¹)
B
I(Q)/I(0)*(Q*R_g)²
0 1 2 3 4 5 6 7 8
Q⁴R_g
Chimera
ADP
ATP
Apo
P(R)
0 2 4 6 8 10 12 14
R (nm)
C
D

Fig. 6. SAXS profile and modeling of the chimeric protein AEE. (A) SAXS intensity profiles of the Ssa1-Sse1 chimera (Chimera AEE) along with the SAXS profiles of Apo-Sse1, ATP-Sse1 and ADP-Sse1 shown for comparison. The corresponding Guinier plot for the four datasets is shown in inset with the black solid lines showing the linear fits. (B) The distance distribution curves [$P(R)$ vs. R] for Chimera AEE, Apo-Sse1, ATP-Sse1, and ADP-Sse1 along with the normalized Kratky [$I(Q)/I(0)*(Q*R_g)^2$ vs. $Q*R_g$] plots in the inset. (C) Starting from left, the damfilt (green) and damaver (red) models for Chimera AEE followed by two orthogonal views of their overlay with the crystal structure of peptide-bound DnaK (PDB ID: 2KHO). (D) Comparison of Sse1 3D structure generated by FATCAT (Flexible structure AlignmenT by Chaining Aligned fragment pairs allowing Twists) [48] alignment with the DAMFILT (green) and DAMAVER (red) models for Chimera AEE shown in Panel C. Sse1 3D structure was generated by structural alignment to the crystal structure of ADP-bound DnaK (PDB ID: 2KHO) by FATCAT and is displayed in two orientations (left and right) best matching with two orthogonal views of the SAXS models (middle and right column of Panel C).

binding and substrate-binding activity of NBD and SBDs, respectively. All chimeras exhibited ATP binding and (Fig. 9A,B) comparable affinity to the LIC peptide [30], as measured by SPR measurements (Fig. 9C and Table 3). Thereafter, we subjected all four chimeras for measurement of steady-state ATPase activity in the presence of J-domain protein Sis1 and compared the ATPase rates with the wild-type parent proteins, Ssa1 and Sse1. As expected, Ssa1 had an appreciable ATPase activity, whereas Sse1 had a negligible ATPase activity (Fig. 9D). Chimera AAE and AEE exhibited an ATPase activity comparable to that of Ssa1, as the NBD in these two chimeras was taken from Ssa1. Isolated Ssa1-NBD also exhibited very similar ATPase rates as the Wt-Ssa1, as has been shown by other Hsp70s. Surprisingly, for chimera EEA and EAA, we observed very high ATPase activity, which were ~ 4-fold higher than the ATPase rates of Ssa1 and ~ 150-fold higher than that of Wt Sse1. This result was indeed surprising as NBDs in these two chimeras (EEA and EAA) were from Sse1 and the wild-type Sse1 harbored only a negligible ATPase activity. To confirm that the high ATPase activity of EEA and EAA was due to Sse1 NBD only, we purified the isolated Sse1-NBD and measured its ATPase assay. Interestingly, Sse1-NBD exhibited even higher ATPase rate than chimera EEA and EAA. The ATPase rate of isolated Sse1-NBD was approximately fourfold higher than these chimeras and approximately 600-fold higher than the wild-type Sse1 protein. This result indicates that Sse1-NBD remained a very potent ATPase since approximately two billion years ago when it diverged from its Hsp70-like ancestor by gene duplication in the first eukaryotes. However, over the course of evolution, Hsp110s have apparently developed intimate contacts between its SBD and NBD that has evolved to suppress such a high ATPase activity during the chaperone cycle. This repression of intrinsic ATPase activity has happened possibly to prevent futile ATP hydrolysis, as well as to keep the protein in an ATP state to facilitate its interaction and subsequent NEF activity with its Hsp70s partners.

Discussion

In the current study, we present unconventional domain communication and domain functions of Hsp110 proteins taking yeast Hsp110, Sse1 as a model. Although the high-resolution structure of Sse1 was solved more than a decade ago, the dynamic nature of the domains that may be crucial for cellular functions of this group of molecular chaperones remained enigmatic. By the state-of-the-art technique single-molecule FRET spectroscopy, we have completely ruled out the existence of Hsp70-like domain allostery in Hsp110 in single molecule resolution which might have remained hidden in previous ensemble experiments. Simultaneously, we have captured unique conformational changes of the Hsp110 protein using another state-of-the-art technique SAXS. Furthermore, we have performed MD simulation to observe finer changes that were difficult to capture by SAXS. Indeed, using a combination of various techniques, we conclusively demonstrate that despite significant similarity in domain organization and individual domain structures, the nature of domain communication in Hsp110s is unique and it has deviated significantly from its structural homolog, Hsp70. The most prominent and well-described conformational changes in Hsp70s are the opening–closing of the SBD α-helical lid and the docking–undocking of SBD with NBD, in the ATP and ADP states of the chaperones, respectively. In contrast, only ATP or ATP-analog bound states of Hsp110 were captured in high resolution thus far, leaving behind a caveat in understanding of the structural features of the protein in its other functional states. Here, we show that indeed, Hsp70-like lid movements or interdomain communication are lacking in Hsp110. Yet, SAXS models also showed that when the protein binding domain of Sse1 was artificially disconnected from its own NBD in the AEE chimera, the α-helical

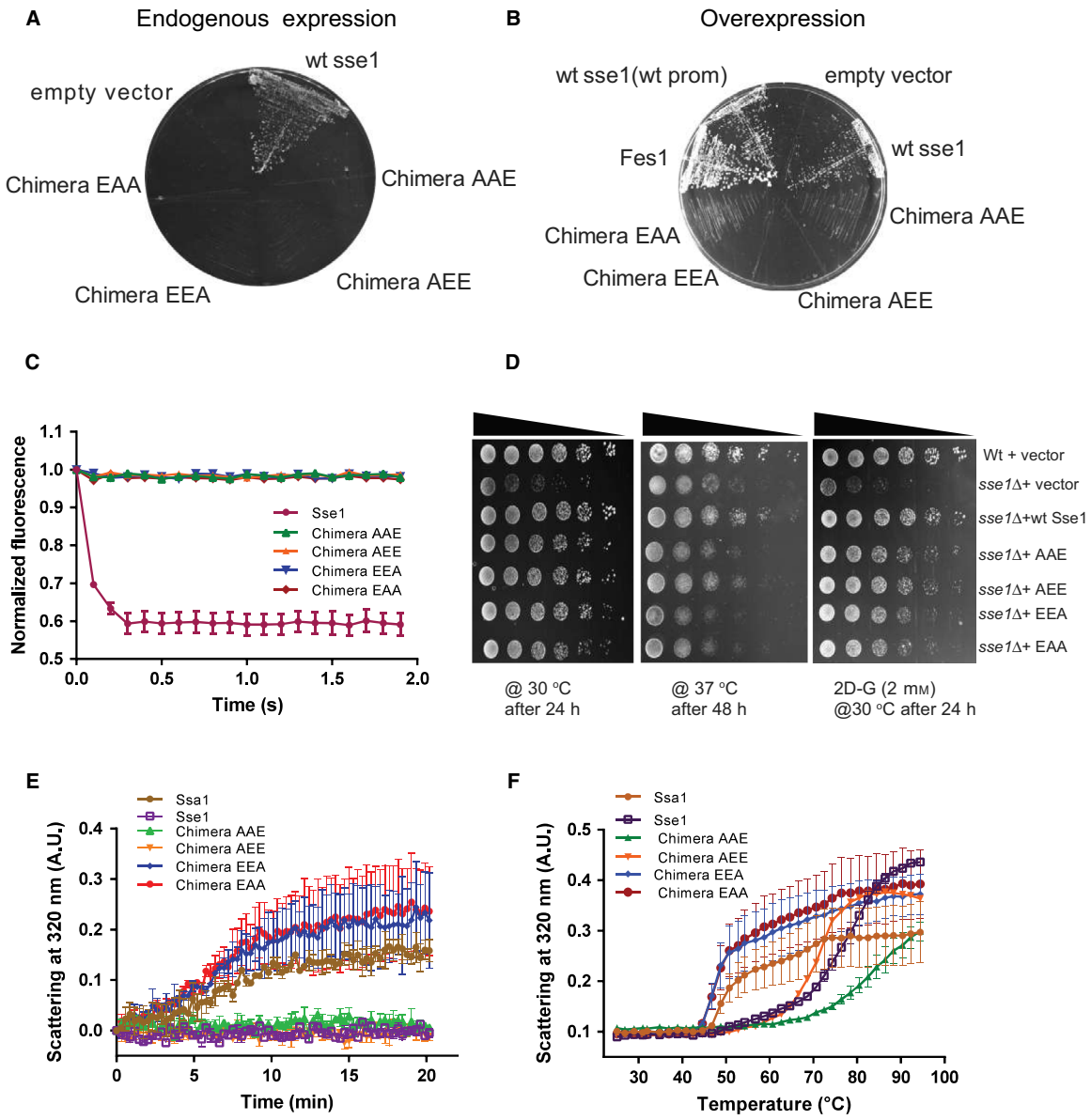


Fig. 7. Nucleotide exchange factor activity is lost in Ssa1-Sse1 chimeric proteins, however; Sse1-SBD containing chimeric proteins reveal its chaperoning activity on self and structurally similar NBDs. The endogenous (A) and overexpression (B) of wild-type Sse1 and chimera in sse1Δ-sse2Δ strain (sse1::His3 sse2::KanMX4 pCM189-SSE1-URA3) show that chimera is not capable of rescuing the lethal phenotype of sse1Δ-sse2Δ whereas the overexpression of another cytosolic nucleotide exchange factor, Fes1 can rescue sse1Δ-sse2Δ phenotype. Panel (C) shows the nucleotide exchange assay for the wild-type and chimera by the measuring the decrease in the fluorescence of Mant-ATP upon release due to exchange, which reveals that wild-type Sse1 efficiently exchange the nucleotide whereas all four chimeras lack the nucleotide exchange factor activity. The error bars represent the SD between replicates (n = 3). (D) Growth phenotype of sse1Δ cells is rescued by all four chimeras like wild-type *sse1* under permissive temperature although when the cells are subjected to proteotoxic stress 2-Deoxy glucose) or heat shock condition (37 °C) chimera AAE and AEE exhibit appreciable growth rescue compared to wild-type sse1 whereas rescue by chimera EEA and EAA is much less efficient. (E) Aggregation propensity of chimeric proteins was measured by subjecting the wild-type proteins and chimeras to high temperature (42 °C) and by following the light scattering at 320 nm. Chimera EEA and EAA along with Ssa1 exhibited increased light scattering with time (20 min) indicating aggregation whereas chimera AAE and AEE and wild-type Sse1 did not exhibit any increase in light scattering at 42 °C indicating no thermal aggregation at this temperature. The error bars represent the SD between replicates (n = 3). (F) Chimera and wild-type Sse1 and Ssa1 proteins were subjected to thermal melting from 25 °C to 95 °C by following the light scattering at 320 nm. Ssa1 and Chimera EEA and EAA exhibit appreciable light scattering at ~ 45 °C indicating aggregation whereas Sse1 and chimera AAE and AEE are much more thermally stable and shows aggregation beyond 65 °C only. The error bars represent the SEM between replicates (n = 3).

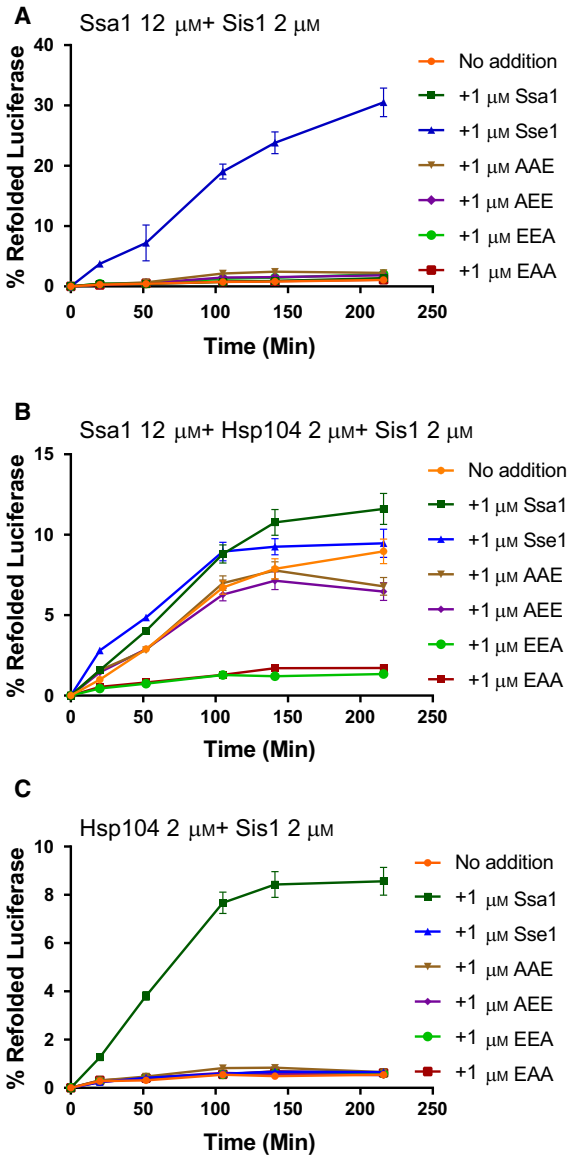


Fig. 8. Hsp110-70 chimera reveal important role of substrate-binding domains (SBD) of Ssa1 and Sse1 as a disaggregating chaperone and co-chaperone. Time-dependent refolding of pre-aggregated luciferase by yeast chaperones and Ssa1-Sse1 chimera: 50 nM of stable inactive luciferase aggregates was incubated up to 220 min at 25 °C in the presence of 5% glycerol, 5 mM ATP, 2 μM Sis1, with or without 2 μM Hsp104, without or with WT Ssa1, Sse1, or with 1 or 2 μM Ssa1-Sse1 chimera, as indicated. (A) Luciferase reactivation by a constant excess of Ssa1 (12 μM) and 2 μM Sis1, without (no addition) or with 1 μM Ssa1, or Sse1, or chimera AAE, AEE, EEA, and EAA. (B) Luciferase reactivation by a constant limiting amount of Ssa1 (2 μM), 2 μM Sis1 and 2 μM Hsp104 (protomers), without (no addition) or with 1 μM Sse1, or Ssa1, or chimera AAE, AEE, EEA, and EAA. (C) Luciferase reactivation by 2 μM Sis1, 2 μM Hsp104 (protomers), without (no addition) or with 1 μM Sse1, or Ssa1, or chimera AAE, AEE, EEA, and EAA. The error bars in all panels represent the SD between replicates (n = 3).

lid of Sse1-SBD was found tightly associated with its β-stranded base precisely reproducing the closed SBD state of ADP-bound Hsp70 (DnaK), which is a central intermediate state of the Hsp70-mediated polypeptide unfoldase cycle. The conserved Hsp70-like ability of the Sse1 lid to tightly close onto its SBD-base in Sse1, over one billion years since the Hsp110 parted from Hsp70 partition, suggests that an essential transient lid closure step may still occur, albeit on an extremely short time scale compared to Hsp70.

Yet, within the limits of the methods used here, we only show that α-helical lid remains mostly wide open both in ATP-bound and ADP-bound states, although there are additional changes in the β-sheet base of SBD and the α-helical lid as captured by SAXS models and further recapitulated by MD simulation. By constructing chimeric proteins between Hsp70 and Hsp110, we subsequently showed that Hsp70-like domain movements are recapitulated in chimeric proteins, although such domain movements appear to be detrimental for the NEF activity of the Hsp110, Sse1. *In vivo*, the chimeric constructs can rescue the growth phenotypes of *sse1* deletion to some extent although *in vitro*, the chimeric proteins are not capable to assist the potent disaggregases like Ssa1 and Hsp104 or lack the disaggregase activity of their own. Since the chimera lack the NEF activity, they are deficient in assisting Ssa1 in the *in vitro* disaggregation of luciferase. It has been shown that for assisting Ssa1 as a disaggregating chaperone, the NEF activity resulting from the direct binding of the NBD of Sse1 to the NBD of Ssa1 [7,8] is crucial. Although Hsp110s, including Sse1, can also bind misfolded proteins (by the holding activity), thereby possibly preventing some protein aggregations, the role of the conserved SBD of Sse1 in assisting the active mechanisms of Ssa1-Sse1-mediated protein disaggregation was not clear. Recently, Garcia *et al.* [31] reported that an Sse1 mutant, $Sse1_{sbd}$, which was severely deficient in holding activity, yet fully retained its nucleotide-binding and NEF activity on Ssa1, had apparently no effect of these mutations in reconstituted disaggregation or refolding reactions *in vitro*. This result implied that substrate binding by Sse1 is not necessary for Ssa1-mediated disaggregation [31]. We also show that none of the chimera can assist Ssa1 in its disaggregase activity *in vitro* due to loss of NEF function when added simultaneously in the reconstituted disaggregase system. Rather interestingly, upon pre-incubation with denatured substrate, all four chimeras are able to hold the non-native substrate in a state that subsequent disaggregation by Ssa1/Sis1 machinery becomes more effective (Fig. 10). This result is in nice agreement with our substrate-binding data using SPR where we have shown that all the chimeric

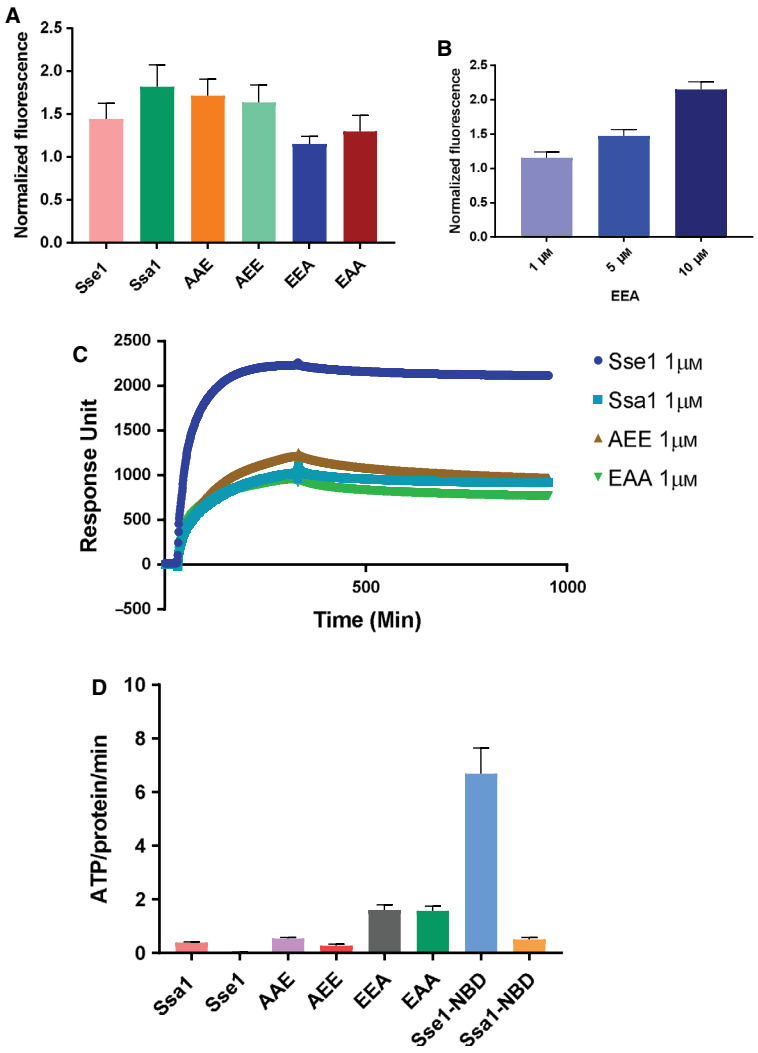


Fig. 9. Sse1-NBD harbors a very high intrinsic ATPase activity which remains suppressed by exhaustive NBD-SBD contacts. (A, B) ATP-binding assay for wild-type Ssa1, Sse1, and Sse1-Ssa1 Chimeric proteins. (A) The nucleotide binding affinity for wild-type Ssa1, Sse1, and Sse1-Ssa1 Chimeric proteins was assayed using MANT-ATP (2′-(or-3′)-*O*-(*N*-Methylanthraniloyl) Adenosine 5′-Triphosphate). The ATP-binding activity of wild-type Sse1 and Chimeras was measured following the increase in fluorescence (Excitation/Emission: 355 nm/428 nm) after binding to the proteins by incubating 100 nm of MANT-ATP with 1 μM of protein (wild-type Sse1, Ssa1, and Chimeras) in binding buffer containing 25 mM HEPES, pH 7.4, 150 mM KCl, and 5 mM $MgCl_2$. The relative fluorescence was calculated as the ratio of fluorescence emission of MANT-ATP bound to protein of interest and the emission of free Mant-ATP. Panel A shows that chimera AAE and AEE show ATP-binding affinity is equivalent to Ssa1 (NBD of AAE and AEE are obtained from Ssa1) and similarly, chimera EAA shows similar binding affinity like Sse1 (NBD of chimera EEA and EAA are obtained from Sse1). In contrast, chimera EEA exhibited less affinity toward ATP. The error bars represent the SD between replicates (n = 3). (B) The binding affinity of Chimera EEA with Mant-ATP with increasing concentrations of the proteins was checked and we found consistent increase of binding to MANT-ATP with increasing concentrations of the protein. The error bars represent the SD between replicates (n = 3). (C) The dissociation rate constants for wild-type Sse1 and chimera's interaction with LIC peptide were measured by SPR measurements by BIAcore3000. LIC peptide (100 mM) was immobilized on CM5 amine coupling chip (GE Healthcare) according to manufacturer's protocol using immobilization buffer (10 mM sodium acetate, pH 4.5). Before the run, the peptide-immobilized lane and the reference lane (blank) were equilibrated with BIAcore assay buffer (25 mM HEPES, 150 mM KCl, 5 mM $MgCl_2$, pH 7.4). Wild-type Sse1, Ssa1, and chimera AEE and EAA were passed over the peptide-immobilized chip at five different concentrations (10 nM, 50 nM, 100 nM, 500 nM, 1000 nM) in BIAcore assay buffer and response was recorded. A representative sensogram of Ssa1, Sse1, and chimera AEE and EAA is shown. (D) Steady-state ATPase rates of WT-Ssa1, WT-Sse1, and their NBDs along with four chimeric proteins AAE, AEE, EEA, and EAA were measured as described in Materials and methods. The ATPase assay was performed in the presence of J-domain co-chaperone, Sis1. The initial NADH absorbance at 340 nm was ~ 0.8–1.0 and the decrease in NADH absorbance due to conversion in NAD^+ by the ATPase activity of proteins was monitored for 20 min and ATPase rates were calculated and plotted as bar graphs. The error bars represent the SD between replicates (n = 3).

Table 3. Table with equilibrium dissociation constant (K_D) for interaction of LIC peptide. The equilibrium dissociation constant (K_D) for interaction of LIC peptide with wild-type Sse1, Ssa1 and four chimeras of Ssa1-Sse1 were calculated as described in the Materials and methods section and have been tabulated.

Protein	(K_D) M^{-1} for LIC peptide binding
Sse1	1.88×10^{-9}
Ssa1	4.37×10^{-9}
Chimera AAE	6.80×10^{-8}
Chimera AEE	4.55×10^{-8}
Chimera EEA	$6.86E \times 10^{-8}$
Chimera EAA	1.69×10^{-9}
Sse1 SBD	1.88×10^{-7}
Ssa1 SBD	5.08×10^{-8}

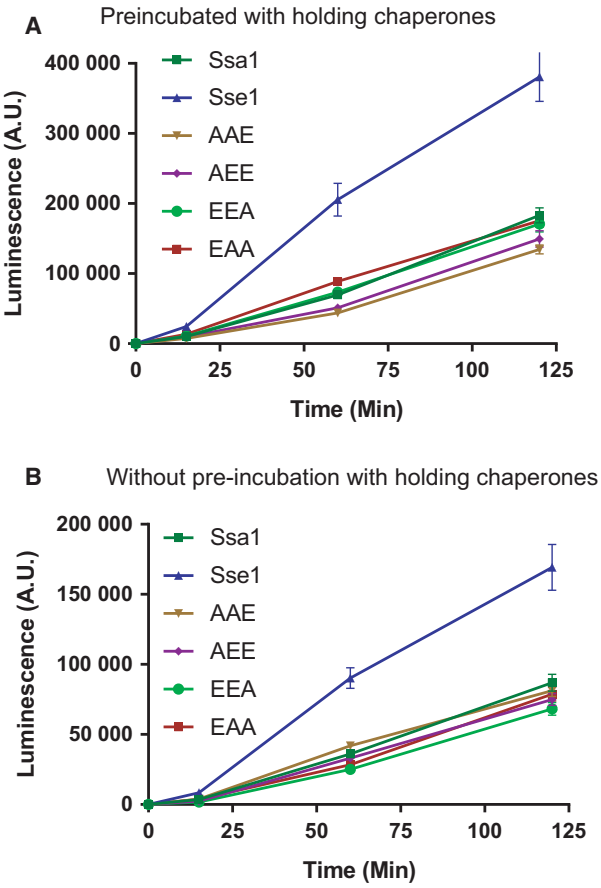


Fig. 10. Time-dependent reactivation of urea-pre-denatured Luciferase that was pre-incubated with (A) or without (B) holding chaperones (Sse1/Ssa1/chimera). The reaction mix was then supplemented with the fully active chaperone mix (6 μM Ssa1, 2 μM Sis1), to reactivate the chimera-bound or unbound aggregated luciferase at 25 °C. 50 μM Luciferase was first inactivated for 2 min in 6M Urea, then diluted 100 folds in buffer (25 mM HEPES-KOH pH 7.5, 150 mM KCl, 10 mM $MgAc_2$, 1 mM DTT, and 5 mM ATP) at 30 °C in the presence of 1 μM chimera AAE/AEE/EEA/EAA, or of 1 μM Ssa1/Sse1 (A), or without any putative holding chaperone (B). 30 min of incubation at 30 °C was done to allow the misfolding luciferase to bind to the holding chaperones (A), or to aggregate (B). Following this, full disaggregation chaperone mix (6 μM Ssa1, 2 μM Sis1) was added to all samples and luciferase reactivation was followed in real time at 25 °C. The error bars represent the SD between replicates (n = 3).

proteins are able to efficiently bind substrate peptides with comparable affinity (Table 3). This substrate binding and holding activity of chimera play role in other cellular activities of Sse1, as we found that chimera, especially AAE and AEE were able to efficiently rescue the growth phenotype of *sse1* deletion.

With the exception of human Hsp105, for which some remodeling activity has been reported with a particular type of misfolded luciferase monomers [32], purified Sse1 failed to show typical ATP- and Hsp40-dependent Ssa1-like protein remodeling activity of its own [3,5,33]. We show that chimera like Sse1 do not possess disaggregase activity *in vitro* although they can rescue the growth phenotype of s*se1* deletion *in vivo* due to its holdase function. We speculate that the frequent ATP-fueled domain movements that occur in Ssa1 might enable it to actively disaggregate and refold inactive aggregates *in vitro*. In contrast, the very rare domain movements that occur in Sse1 can explain the inability of Sse1 to actively disaggregate by itself and refold inactive aggregates *in vitro*. Although using SAXS, we show chimera AEE SBD can close, yet it not able to disaggregate misfolded luciferase, probably due to a yet unknown co-chaperone, that would have helped the Sse1-SBD to reopen and enabling it for disaggregation.

Importantly, we found that chimera EEA and EAA both harboring Sse1-NBD exhibited very high ATPase activity compared to wild-type Sse1 and even prominently higher than Ssa1. This led us to determine the ATPase activity of the isolated Sse1-NBD, which revealed a significantly higher ATPase rate than the wild-type Sse1 (~ 600 fold higher) and Ssa1. These data indicate that despite such high ATPase activity of Sse1-NBD, the wild-type protein exhibits negligible ATPase activity likely due to a tight suppression of the ATPase activity exerted by the SBD. From our data, we can conclude that extensive NBD–SBD interaction is responsible for suppression of high intrinsic ATPase activity of the Sse1-NBD and whenever the interdomain contacts are compromised (like in chimera EEA and EAA) or completely lost (in isolated Sse1-NBD), high ATPase activity is exhibited by the protein. This unexpected finding about high intrinsic ATPase activity of Sse1-NBD also explained the strong inhibition of luciferase disaggregation by chimera EEA and EAA

(Fig. 8B) as rapid ATP hydrolysis by these two chimeras would lead to strong capture and sequestering of aggregated luciferase by the Ssa1-SBD present in these proteins. This in turn would be prevented from being efficiently unfolded and released by the Hsp104-Ssa1-Sis1 machinery to spontaneously refold in solution. This unusually high ATPase activity could also capture unfolded/misfolded proteins *in vivo* which would explain the lesser ability by these two chimeras to rescue *sse1Δ* growth defects. This is a completely new aspect of evolution of loss of canonical domain allostery of Hsp110 chaperones, despite extremely similar domain architecture with Hsp70s.

Altogether, in the current study, we show that yeast Hsp110 chaperone Sse1 has lost domain movements that characterize their Hsp70 cognates although Sse1 undergoes unique intra-domain movements especially within its SBD. The loss of interdomain movements have evolved to maintain extensive contacts between NBD and SBD apparently to tightly suppress a high intrinsic ATPase activity (summarized in Table 4). Evolution of atypical interdomain contacts for repression of high intrinsic ATPase activity of Sse1 rather than mutation at the active site indicates that ATPase activity of Sse1 or Hsp110s in general is essential for its cellular functions, the molecular details of which is yet to be elucidated.

Materials and methods

Strains and plasmids

Wild-type *sse1*, *ssa1,* and chimeras were cloned in the pET-Duet1, pRS313, and pJV340 vector for propagation in *E. coli* and *S. cerevisiae*. The chimeras were constructed by fusing the domains of sse1 and ssa1 by overlap PCR. These constructs and wild-type sse1, ssa1 were inserted into the pETDuet1 using the restriction sites BamH1/Sal1 or Sac1 for sse1 and ssa1, respectively. The yeast centromeric plasmids (pRS313 and pRS316) were inserted with the Sse1 Promoter and 3′ UTR using the restrictions site Sac1/BamH1 and Xho1/Kpn1, respectively. Furthermore, we introduced wild-type sse1, ssa1 and chimeras in between Promoter and 3′ UTR using the restriction sites BamH1/Xho1. And also the wild-type Sse1 and chimeras were cloned into the Yeast 2μ plasmid (pJV340) using the restriction sites BamH1/Xho1. All the constructs were confirmed by sequencing. Yeast cells were cultured in the selective synthetic medium at 30 °C for growth and maintenance of plasmids. For drop dilution assay, yeast cells were serially diluted from 1 OD and equal volume were spotted on selective synthetic medium with or without proteotoxic stress.

Purification of proteins

For purification of the His_6 tagged wild-type Sse1, Ssa1 and chimeras, we expressed them in the *E. coli* BL21(DE3) cells with IPTG induction (final 0.5 mM) at 18 °C, overnight. The overnight grown cells were harvested and the cells were lysed in the Buffer A with 1 mM PMSF (final) [Buffer A: 25 mM HEPES (pH 7.4), 250 mM KCl, 1 mM β–Mercaptoethanol, 10 mM Imidazole]. The lysate was incubated with Ni^{2+} NTA resin (Sigma-Aldrich). The bound proteins were eluted with Buffer A + 250 mM Imidazole. The Eluted proteins were desalted using PD10 desalting column from GE using storage buffer [25 mM HEPES (pH 7.4), 150 mM KCl, 5 mM $MgCl_2$, 1 mM β–Mercaptoethanol, 10% Glycerol] and concentrated using amicon ultra-centrifugal devices from GE. The concentrated proteins were purified further in HiLoad 16/600 Superdex 200 pg (GE Healthcare Life Sciences) size exclusion column in AKTA explorer system from GE. The Superdex 200 column was pre-equilibrated with storage buffer as per manufacturer's instruction and proteins were loaded and different protein fractions were eluted and check for purity. The purified proteins were collected, concentrated, and stored with 15% glycerol at −80 °C for further use.

Table 4. Summary table of different activities used by different chaperones or chimera of chaperones used in this study. ND, not done; NA, not applicable.

Protein	Solubility	Thermostability	Expression in yeast	Rescue potential of sse1/sse2 double deletion	NEF activity	Rescue potential of heat sensitivity of sse1Δ strain	Assists Ssa1 in disaggregation	Inhibition of disaggregation by Hsp104	ATPase activity
Ssa1	Good	Less than Sse1	Good	No	NA	ND	NA	NA	Basal
Sse1	Good	Good	Good	Yes	Yes	Yes	Yes	No	Insignificant
AAE	Good	Good	Good	No	No	Yes	No	No	Basal
AEE	Good	Good	Good	No	No	Yes	No	No	Basal
EAA	Poor	Poor	Poor	No	No	Insignificant	No	Yes	High
EEA	Poor	Poor	Poor	No	No	Insignificant	No	Yes	High
Fes1	ND	ND	ND	Yes only upon overexpression	Yes	Insignificant even after overexpression	No	ND	NA

Fluorophore labeling of double cysteine mutants of Sse1

Double cysteine mutants of Sse1 were resuspended in buffer containing 25 mM HEPES pH 7.5, 150 mM KCl and 5 mM $MgCl_2$ (Buffer A) were incubated with Alexa 488-C5 maleimide and Alexa 647-C2 maleimide dyes (Molecular Probes, Invitrogen) at a molar ratio of 1 : 1.5, respectively, for 2 h on ice. Following incubation, excess of free dye was removed using NAP-5 desalting columns.

Single-molecule FRET experiments

The methodology used for single-molecule FRET measurements has been described in details earlier [11,14]. Our methodology is a modified derivative of earlier described methods like PIE (Pulsed interleaved excitation) [17], ALEX (Alternate Laser Excitation) [16], and PAX [15]. This methodology enables us to detect only double-labeled molecules with active donor–acceptor fluorophores. The measurements of Sse1 mutants were performed on a Zeiss confocal system [Zeiss Axiobserver Microscope (Carl Zeiss, GmbH, Jena, Germany)] in an inverted microscope. Lab-Tek chambered cover glasses were used as the sample holders. We have simultaneously used one continuous laser source at 488 nm and a pulsed laser source of 630 nm set at a repetition frequency of 20 MHz to excite a small confocal detection volume.

Surface plasmon resonance

The dissociation rate constants for wild-type Sse1 and chimera's interaction with LIC peptide was measured by SPR measurements by BIAcore 3000 (Biacore SPRTM systems, GE Healthcare Life Sciences). LIC peptide (100 mM) was immobilized on CM5 amine coupling chip according to manufacturer's protocol using immobilization buffer (10 mM sodium acetate, pH 4.5). Before the run, the peptide-immobilized lane and the reference lane (blank) were equilibrated with BIAcore assay buffer (25 mM HEPES, 150 mM KCl, 5 mM MgCl2, pH 7.4). Wild-type Sse1 and chimeras were passed over the peptide-immobilized chip at five different concentrations (10 nM, 50 nM, 100 nM, 500 nM, 1000 nM) in BIAcore assay buffer and response was recorded. SPR data were exported to BIAevaluation tool and the response curves were fitted to Langmuir's equation for calculating the equilibrium dissociation constant (K_D) for wild-type Sse1 and chimeras.

Small angle X-ray scattering data collection and analysis

Small angle X-ray scattering measurements were done on SAXSpace system (Anton Paar GmbH, Graz, Austria) with sealed tube (line collimation) X-ray generator ($\lambda = 1.5418$) operated at 40 kV and 50 mA mounted with 1D CMOS MYTHEN (Dectris, Baden-Dättwil, Switzerland) detector. The scattering intensity $I(q)$ was measured for scattering angles ($Q = 4\pi\sin\theta/\lambda$) from 0.01 to 3 nm^{-1}. The samples were placed in a quartz capillary on a thermostatic capillary holder. The data were collected at 25 °C and the integrated exposure time for one acquisition was 45 min containing three frames each of 15 min. Approximately 50 µL of each sample was exposed to X-rays. Unless specified otherwise, the purified concentrated proteins were dialyzed against 20 mM HEPES, 150 mM KCl, 5 mM $MgCl_2$, and 2 mM DTT at a final pH of 7.2. Immediately prior to SAXS experiments, samples were centrifuged in a tabletop centrifuge, and the protein concentrations were determined using a Nanodrop spectrophotometer (Thermo Scientific). SAXSTREAT software (Anton Paar GmbH) was used for data reduction and calibration of primary beam position. Data processing, including buffer subtraction, scaling, and de-smearing, was done using SAXSQUANT software (Anton Paar GmbH, Graz, Austria). Estimation of radius of gyration (R_g), pairwise distribution function ($P(R)$), and D_{max} from the SAXS profile of the protein samples was done using PRIMUSQT available in ATSAS suite [34]. To obtain solution structures, modeling of low resolution structures was carried out using DAMMIF [35]. Ten independent *ab initio* models were generated and averaged and filtered using DAMAVER suite of programs [36]. All high resolution structures were overlaid with SAXS-derived model using SUPCOMP20 [37]. Molecular graphical representations were made using PYMOL (The PyMOL Molecular Graphics System, Version 1.8, Schrödinger, LLC) and UCSF CHIMERA 1.12 [38]. The IUCr guidelines published by Jacques *et al.* [39] were followed for reporting SAXS data parameters.

Establishing the structure of the apo-form of Sse1 protein by MD simulation

To elucidate the high-resolution apo-form of the protein, molecular dynamic simulation was performed on the nucleotide-removed form of Sse1 (PDB ID: 3C7N) using GROMACS 4.6.1 [40]. All atomistic simulations were carried out using the CHARMM36 all-atom force field (November release) [41–43] using periodic boundary condition and TIP3P water model, developed by Jorgensen *et al.*, was used to solvate the protein and ions [44]. The starting model was solvated in a periodic box with 97916 TIP3 water molecules. Twenty-six Na^+ ions were added to the solvent to neutralize electrical net charge of the protein. The system was then minimized for 50 000 steps using a steepest decent algorithm with an emtol of 200 $KJ{\cdot}mol^{-1}$ after a minimization with emtol of 100 $KJ{\cdot}mol^{-1}$. This was followed by an equilibration run of 100 ps in NVT ensemble with restrains on the protein atoms. The NPT ensemble was used for production simulation. Systems were simulated at 310 K, maintained by a Berendsen thermostat with a time constant of 1 ps with the protein and non-protein molecules coupled separately. Pressure coupling was done employing a Parrinello-Rahman barostat using a 1 bar

reference pressure and a time constant of 2.0 ps with compressibility of 4.5e-5 bar using isotropic scaling scheme. Electrostatic interactions were calculated using the Particle Mesh Ewald (PME) summation. The production run was performed for 1 μs.

Molecular dynamic stabilized final structure was used as an input for an iterative SAXS-guided refinement using SREFLEX [45]. The final model was validated using CRYSOL [46], which gave a chi-square fitting score of 1.094 to the apo-SAXS data (Fig. 4D).

Luciferase refolding assay

Chaperone disaggregation and reactivation assays were performed according to Ref. [33] with minor modifications. About 50 μM of firefly luciferase (50 μM) was incubated for 3 min in 7 M urea at 30 °C, then diluted 50-fold in 25 mM HEPES-KOH pH 7.5, 150 mM KCl, 10 mM $MgAc_2$, 1 mM DTT, 10 mM ATP, 4 mM PEP, PK 10 $\mu g \cdot mL^{-1}$ and further incubated for 20 min 30 °C. About 100 μL aliquots containing 100 nM of stable inactive luciferase aggregates were snap frozen and stored at −80 °C until use. For the chaperone activity assays, an equal volume of pre-aggregated luciferase was mixed with chaperones and co-chaperones at final concentrations ranging between 0 and 5 μM, as indicated and incubated at 25 °C. At various incubation times, 3–5 μL aliquots were automatically mixed with 80 μL of luciferase reaction mix (50 mM Tris-acetate pH 7.5, 15 mM $MgCl_2$, 5 mM ATP, 50 mM KCl, 200 μM Coenzyme A, 5 μM D-Luciferin) in a Victor Light 1420 Luminescence Counter from Perkin-Elmer (Turku, Finland). The emitted photons were counted during 20 s. The light units emitted at the end of the 20 s measure, which best correlated with preset amounts of native (active) luciferase, were used as a relative measure of chaperone-mediated disaggregation and refolding activity.

ATPase assay

ATPase assay was performed as described before [47]. The assay was performed in 384-well plates in total 50 μL of reaction volume in each well. For each reaction, 4 μM (final) of Sse1, Ssa1, or chimera was added in the assay buffer (25 mM of HEPES-KOH pH 7.5, 100 mM of KCl, 10 mM of $MgAc_2$) along with 3 μM (final) of Sis1 (J-domain co-chaperone) containing 2 mM of PEP, 30 Units of PK, 20U of LDH, and ~ 0.25 mM NADH. The initial NADH absorbance at 340 nm was 0.8–1.0 and the change in NADH absorbance was monitored for 20 min.

Acknowledgements

We thank Prof Ulrich Hartl from MPI, Biochemistry to generously provide the strain sse1::His3 sse2::KanMX4 pCM189-SSE1-URA3. VK acknowledges CSIR-Senior Research Fellowship. KM acknowledges SERB YSS grant (YSS/2015/000532) for funding, Shiv Nadar University for infrastructural support. PG acknowledges the Grant 31003A_156948 from the Swiss National Fund for funding. We thank Joao Pereira Cordeiro and Joe Dickinson for setting up and repeating some of the chaperone assays. We acknowledge Kausik Chakraborty and Asmita Ghosh for their critical comments.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

KM conceived the project, supervised the work, analyzed the biophysical and biochemical data, and wrote the manuscript. VK made all the constructs, purified the proteins, and performed the biochemical and yeast experiments. MPJ performed the SPR measurements along with VK. JJP performed the sm-FRET experiments, JJP and KM analyzed the sm-FRET data. JJP and AS obtained the SAXS data and JJP, AS, and FA analyzed the SAXS data. AR performed the MD simulation. ST purified the Sse1 domains and chimera for luciferase refolding and ATPase assays. MR and PG designed and performed the luciferase refolding and ATPase assay and analyzed the rates of refolding experiments and ATPase rates.

References

1 Finka A, Sharma SK & Goloubinoff P (2015) Multi-layered molecular mechanisms of polypeptide holding, unfolding and disaggregation by HSP70/HSP110 chaperones. *Front Mol Biosci* **2**, 29.
2 Hartl FU, Bracher A & Hayer-Hartl M (2011) Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324–332.
3 Nillegoda NB, Wentink AS & Bukau B (2018) Protein disaggregation in multicellular organisms. *Trends Biochem Sci* **43**, 285–300.
4 Bracher A & Verghese J (2015) GrpE, Hsp110/Grp170, HspBP1/Sil1 and BAG domain proteins: nucleotide exchange factors for Hsp70 molecular chaperones. *Subcell Biochem* **78**, 1–33.
5 Andreasson C, Fiaux J, Rampelt H, Mayer MP & Bukau B (2008) Hsp110 is a nucleotide-activated exchange factor for Hsp70. *J Biol Chem* **283**, 8877–8884.
6 Liu Q & Hendrickson WA (2007) Insights into Hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1. *Cell* **131**, 106–120.

7 Polier S, Dragovic Z, Hartl FU & Bracher A (2008) Structural basis for the cooperation of Hsp70 and Hsp110 chaperones in protein folding. *Cell* **133**, 1068–1079.

8 Schuermann JP, Jiang J, Cuellar J, Llorca O, Wang L, Gimenez LE, Jin S, Taylor AB, Demeler B, Morano KA *et al.* (2008) Structure of the Hsp110:Hsc70 nucleotide exchange machine. *Mol Cell* **31**, 232–243.

9 Qi R, Sarbeng EB, Liu Q, Le KQ, Xu X, Xu H, Yang J, Wong JL, Vorvis C, Hendrickson WA *et al.* (2013) Allosteric opening of the polypeptide-binding site when an Hsp70 binds ATP. *Nat Struct Mol Biol* **20**, 900–907.

10 Yang J, Nune M, Zong Y, Zhou L & Liu Q (2015) Close and allosteric opening of the polypeptide-binding site in a human Hsp70 chaperone BiP. *Structure* **23**, 2191–2203.

11 Banerjee R, Jayaraj GG, Peter JJ, Kumar V & Mapa K (2016) Monitoring conformational heterogeneity of the lid of DnaK substrate-binding domain during its chaperone cycle. *FEBS J* **283**, 2853–2868.

12 Mapa K, Sikor M, Kudryavtsev V, Waegemann K, Kalinin S, Seidel CA, Neupert W, Lamb DC & Mokranjac D (2010) The conformational dynamics of the mitochondrial Hsp70 chaperone. *Mol Cell* **38**, 89–100.

13 Marcinowski M, Holler M, Feige MJ, Baerend D, Lamb DC & Buchner J (2011) Substrate discrimination of the chaperone BiP by autonomous and cochaperone-regulated conformational transitions. *Nat Struct Mol Biol* **18**, 150–158.

14 Mapa K, Tiwari S, Kumar V, Jayaraj GG & Maiti S (2012) Information encoded in non-native states drives substrate-chaperone pairing. *Structure* **20**, 1562–1573.

15 Doose S, Heilemann M, Michalet X, Weiss S & Kapanidis AN (2007) Periodic acceptor excitation spectroscopy of single molecules. *Eur Biophys J* **36**, 669–674.

16 Kapanidis AN, Lee NK, Laurence TA, Doose S, Margeat E & Weiss S (2004) Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc Natl Acad Sci USA* **101**, 8936–8941.

17 Muller BK, Zaychikov E, Brauchle C & Lamb DC (2005) Pulsed interleaved excitation. *Biophys J* **89**, 3508–3522.

18 Hajizadeh NR, Franke D, Jeffries CM & Svergun DI (2018) Consensus Bayesian assessment of protein molecular mass from solution X-ray scattering data. *Sci Rep* **8**, 7204.

19 Makhoba XH, Burger A, Coertzen D, Zininga T, Birkholtz LM & Shonhai A (2016) Use of a chimeric Hsp70 to enhance the quality of recombinant *Plasmodium falciparum* S-adenosylmethionine decarboxylase protein produced in *Escherichia coli*. *PLoS One* **11**, e0152626.

20 Shonhai A, Boshoff A & Blatch GL (2005) *Plasmodium falciparum* heat shock protein 70 is able to suppress the thermosensitivity of an *Escherichia coli* DnaK mutant strain. *Mol Genet Genomics* **274**, 70–78.

21 Strub A, Rottgers K & Voos W (2002) The Hsp70 peptide-binding domain determines the interaction of the ATPase domain with Tim44 in mitochondria. *EMBO J* **21**, 2626–2635.

22 Suppini JP, Amor M, Alix JH & Ladjimi MM (2004) Complementation of an *Escherichia coli* DnaK defect by Hsc70-DnaK chimeric proteins. *J Bacteriol* **186**, 6248–6253.

23 Polier S, Hartl FU & Bracher A (2010) Interaction of the Hsp110 molecular chaperones from *S. cerevisiae* with substrate protein. *J Mol Biol* **401**, 696–707.

24 Blamowska M, Sichting M, Mapa K, Mokranjac D, Neupert W & Hell K (2010) ATPase domain and interdomain linker play a key role in aggregation of mitochondrial Hsp70 chaperone Ssc1. *J Biol Chem* **285**, 4423–4431.

25 Chang YW, Sun YJ, Wang C & Hsiao CD (2008) Crystal structures of the 70-kDa heat shock proteins in domain disjoining conformation. *J Biol Chem* **283**, 15502–15511.

26 Raviol H, Sadlish H, Rodriguez F, Mayer MP & Bukau B (2006) Chaperone network in the yeast cytosol: Hsp110 is revealed as an Hsp70 nucleotide exchange factor. *EMBO J* **25**, 2510–2518.

27 Liu XD, Morano KA & Thiele DJ (1999) The yeast Hsp110 family member, Sse1, is an Hsp90 cochaperone. *J Biol Chem* **274**, 26654–26660.

28 Glover JR & Lindquist S (1998) Hsp104, Hsp70, and Hsp40: a novel chaperone system that rescues previously aggregated proteins. *Cell* **94**, 73–82.

29 Goloubinoff P, Mogk A, Zvi AP, Tomoyasu T & Bukau B (1999) Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. *Proc Natl Acad Sci USA* **96**, 13732–13737.

30 Goeckeler JL, Petruso AP, Aguirre J, Clement CC, Chiosis G & Brodsky JL (2008) The yeast Hsp110, Sse1p, exhibits high-affinity peptide binding. *FEBS Lett* **582**, 2393–2396.

31 Garcia VM, Nillegoda NB, Bukau B & Morano KA (2017) Substrate binding by the yeast Hsp110 nucleotide exchange factor and molecular chaperone Sse1 is not obligate for its biological activities. *Mol Biol Cell* **28**, 2066–2075.

32 Mattoo RU, Sharma SK, Priya S, Finka A & Goloubinoff P (2013) Hsp110 is a bona fide chaperone using ATP to unfold stable misfolded polypeptides and reciprocally collaborate with Hsp70 to solubilize protein aggregates. *J Biol Chem* **288**, 21399–21411.

33 Shorter J (2011) The mammalian disaggregase machinery: Hsp110 synergizes with Hsp70 and Hsp40

to catalyze protein disaggregation and reactivation in a cell-free system. *PLoS One* **6**, e26319.
34 Franke D, Petoukhov MV, Konarev PV, Panjkovich A, Tuukkanen A, Mertens HDT, Kikhney AG, Hajizadeh NR, Franklin JM, Jeffries CM *et al.* (2017) ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J Appl Crystallogr* **50**, 1212–1225.
35 Franke D & Svergun DI (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr* **42**, 342–346.
36 Volkov A & Svergun D (2003) Uniqueness of ab initio shape determination in small-angle scatterin. *J Appl Crystallogr* **36**, 860–864.
37 Kozin MB & Svergun DI (2001) Automated matching of high- and low-resolution structural models. *J Appl Crystallogr* **36**, 860–864.
38 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004) UCSF Chimera–a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612.
39 Jacques DA, Guss JM, Svergun DI & Trewhella J (2012) Publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution. *Acta Crystallogr D* **68**, 620–626.
40 Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE & Berendsen HJ (2005) GROMACS: fast, flexible, and free. *J Comput Chem* **26**, 1701–1718.
41 Brooks BR, Brooks CL III, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S *et al.* (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* **30**, 1545–1614.
42 Denning EJ, Priyakumar UD, Nilsson L & Mackerell AD Jr (2011) Impact of 2'-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. *J Comput Chem* **32**, 1929–1943.
43 Huang J & MacKerell AD Jr (2013) CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem* **34**, 2135–2145.
44 Lamb ML & Jorgensen WL (1998) Investigations of neurotrophic inhibitors of FK506 binding protein via Monte Carlo simulations. *J Med Chem* **41**, 3928–3939.
45 Panjkovich A & Svergun DI (2016) Deciphering conformational transitions of proteins by small angle X-ray scattering and normal mode analysis. *Phys Chem Chem Phys* **18**, 5707–5719.
46 Svergun DI, Barberato C & Koch MH (1995) CRYSOL – a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* **28**, 768–773.
47 Kiianitsa K, Solinger JA & Heyer WD (2003) NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. *Anal Biochem* **321**, 266–271.
48 Ye Y & Godzik A (2004) FATCAT: a web server for flexible structure comparison and structure similarity searching. *Nucleic Acids Res* **32**, W582–W585.

CHAPTER 4

Increased energy efficiency of eukaryotic Hsp70 Unfoldases by a stop-start mechanism in J-domain cochaperones

Increased energy efficiency of eukaryotic Hsp70 Unfoldases by a stop-start mechanism in J-domain cochaperones

Mathieu E. Rebeaud, Bruno Fauvet, Satyam Tiwari, Paolo De Los Rios and Pierre Goloubinoff

I contributed to this work conceptually and in the manuscript writing.

Conceptualization: PG, PDLR, MER
Methodology: PG, PDLR, MER, BF, ST
Investigation: PG, PDLR, MER, BF, ST
Visualization: PG, PDLR, MER, BF, ST
Supervision: PG, PDLR
Writing—original draft: PG, PDLR, MER, BF
Writing—review & editing: PG, PDLR, MER, BF, ST

Increased energy efficiency of eukaryotic HSP70 unfoldases by a stop-start mechanism in J-domain cochaperones

Mathieu E. Rebeaud[ab], Bruno Fauvet[b], Satyam Tiwari[a], Paolo De Los Rios[b,c] and Pierre Goloubinoff[a,d,1]

[a] Department of Plant Molecular Biology, Faculty of Biology and Medicine, University of Lausanne, CH-1015 Lausanne, Switzerland

[b] Institute of Physics, School of Basic Sciences, École Polytechnique Fédérale de Lausanne – EPFL, 1015 Lausanne, Switzerland

[c] Institute of Bioengineering, School of Life Sciences, École Polytechnique Fédérale de Lausanne – EPFL, 1015 Lausanne, Switzerland

[d] School of Plant Sciences and Food Security, Tel-Aviv University, Ramat Aviv 69978, Israel

[1] Corresponding author: Pierre Goloubinoff; Pierre.Goloubinoff@unil.ch

Abstract

The HSP70 chaperones centrally control protein homeostasis in all the cellular compartments with ATP. J-domain proteins (JDPs) co-evolved with the HSP70s to trigger ATP-hydrolysis and catalytically upload various substrate polypeptides that are in need to be structurally modified by the chaperone. Here, the protein disaggregation and refolding activities of the main yeast cytosolic HSP70, SSA1, were measured in the presence of its natural JDPs, SIS1 (SS), YDJ1 (YY), and two of their J-domain swap mutants, referred to SY and YS. Remarkably, both swap mutants were more effective than their wild-type progenitors at uploading the misfolded polypeptide substrates, triggering SSA1's ATPase, causing polypeptide unfolding and at releasing the unfolded polypeptide products. Limited proteolysis-coupled to mass spectrometry showed that without bound substrates, SS and YY, but not YS and SY, could adopt an auto-repressed conformation with their J-domains reversibly associated with their own substrate-binding domain, thereby minimally binding and triggering SSA1's ATPase. The data suggest that eukaryotic DNAJAs and DNAJBs use this reversible auto-repressed state as a stop-start mechanism to reduce wasteful ATP hydrolysis in the absence of polypeptide substrates, thereby optimizing the energy consumption of the protein unfolding/refolding work by the HSP70 chaperone machinery.

Keywords: DNAJA, DNAJB, autorepression, co-evolution, JDPs, HSP70

Introduction

Life is a non-equilibrium phenomenon. Organisms are composed of highly complex macromolecules with biologically active conformations that can be unstable even under physiological conditions (*1*). Especially under stress, many native proteins may spontaneously lose their functional 3D structure, misfold and form stable aggregates lacking their dedicated biological activity (*2*). Although the primary sequence of a nascent polypeptide chain may, in principle, suffice to drive its spontaneous folding into a specific native conformation (*3*), lengthy multidomain polypeptides enriched in beta structures, tend to form stable inactive aggregates (*4*), which in metazoans can be toxic and cause degenerative disorders, as in the case of Alzheimer's and Parkinson's diseases (*5*). Early in Life's history, prokaryotes evolved protective mechanisms and ATP-consuming chaperones that can specifically target and bind conformationally-compromised proteins and transiently unfold them. Upon chaperone dissociation, the formerly damaged unfolded polypeptides may either refold to their native conformation, even transiently, under denaturing conditions (*6*) or be degraded by chaperone-gated proteases (*1*). The HSP70s, Hsp60s, and Hsp100s are highly conserved families of molecular chaperones acting as ATP-fueled polypeptide-unfolding enzymes (*7*). The HSP70s centrally coordinate the transfer of polypeptides between the Hsp20s, Hsp60s, Hsp100, and Hsp90s chaperone families and control their activity, thereby forming a highly effective collaborative proteostasis network. The triage by HSP70s of polypeptides substrates in need to be structurally modified is performed by obligate J-Domain Protein (JDP) co-chaperones (*8*). The JDPs catalyze the uploading onto HSP70 of substrate polypeptides that are either misfolded and aggregated or active (alter)native proteins, such as HSF1 trimers and clathrin cages, which at a given stage of the cell's life, need to be de-oligomerized and maintained inactive by the chaperone (*9, 10*). Similarly, the organellar polypeptides that are synthesized in the cytosol, need first to translocate while completely unfolded through narrow pores across membranes and then need to be uploaded onto organellar HSP70s through specific organellar JDPs at the import pore's exits (*11*). As a consequence, the translocating preproteins can be effectively pulled by HSP70 and unfolded from within the organelles ,and successfully imported into the mitochondrial matrix, the ER lumen or the chloroplast stroma (*12*).

To carry out their polypeptide-uploading and HSP70's ATPase-triggering action, all JDPs comprise at least two domains: one that must specifically recognize and bind onto misfolded or alternatively-folded polypeptide substrates and the other maned J-domain (JD), that must specifically recognize and bind

HSP70. JD is made of ~70 residues, some of them highly conserved, that form four compact alpha-helices. All classic JDs expose a characteristic "HPD" motive (*13*) that can specifically anchor into a pocket of the HSP70-ATP complex. Once bound, the HPD triad binds a conserved hydrophobic linker, typically VLLL, between HSP70's N- and C-terminal domains (see PDB 5NRO (*14*)). The concomitant interaction of a JD and a polypeptide substrate with HSP70 (*15-17*), greatly accelerates HSP70's ATPase cycle (*14, 18*).

Following denaturing stress, such as heat-shock, conserved homodimeric JDPs from the DNAJAs and DNAJBs classes (traditionally called Hsp40s) predominantly bind to misfolded and aggregated proteins (*19*) through their two C-terminal domains (CTDs) and, at the same time, bind ATP-HSP70 molecules through their two N-terminal JDs. Following JD-induced ATP-hydrolysis and the consequent ultra-affinity-driven locking of HSP70-ATP molecules onto the JDP-delivered misfolded polypeptide substrate (*20*), a force is applied by entropic pulling (*11*), causing the polypeptide unfolding and leading to its gradual extraction/solubilization from compact aggregates (*21*). Concomitantly, the JDPs dissociate from the ADP-HSP70-unfolded polypeptide complex (*22*) and are free to scout the cellular environment for additional misfolded/aggregated substrates, to bind and upload them onto nearby ATP-HSP70s. Next, the transient binding of a nucleotide exchange factor (NEF) accelerates ADP dissociation from the HSP70 chaperone and the release of the transiently unfolded polypeptides (*23*) which may then spontaneously refold to their native state (*24*).

Owing to the high flexibility of their flanking G/F-rich linkers (see Figure 1), the N-terminal JDs of DNAJAs or DNAJBs are expected to freely seek and unrestrictedly bind nearby ATP-HSP70 molecules, which in the cell are about 10-fold more abundant than the JDP co-chaperones (Table S1). Yet, growing evidence indicates that the JD of particular eukaryotic JDPs can adopt an alternative configuration and bind their CTDs, rather than readily seek anchoring to HSP70s. For example, the JDs of human DNAJB1 (a canonical class B JDP) and DNAJB6 and DNAJB8 (two non-canonical class B JDPs) interact with a helix in the nearby G/F regions, forming a structural arrangement that apparently competes with the binding of the J-domain to HSP70 (*25*). In the case of yeast SIS1 (a canonical class B JDP), an interaction between the JD with specific residues in the G/F region forming Helix V has been reported to block partnership with HSP70 (*25, 26*). Moreover, it was found that the ability of individual DNAJA or DNAJB homodimers to activate the ATP-fueled unfolding-refolding activity of HSP70s, is much higher when they are in the presence of one another, as compared to when they are individually presented to HSP70 (*21*). This may be due to an autoinhibited conformation of the DNAJA or DNAJB dimers, which may, by an unclear mechanism, become derepressed by when DNAJA homodimers interact with DNAJB homodimers (*21*).

Here, we investigated the intramolecular regulatory mechanism by which the two main JPDs in the yeast cytosol, YDJ1 (class A) and SIS1 (class B) (henceforth called YY and SS), regulate the binding, unfolding, and disaggregating activities of the main yeast cytosolic HSP70, SSA1, by comparing their action with two J-domain swap chimeras (henceforth called YS and SY), in which the HSP70-binding of the J-domain and the substrate-binding of the CTDs may have been unlocked, because the hypothesized intramolecular interactions between the JDs and their CTDs, that co-evolved in wild type SS and YY (*27*), may have been disrupted. We found that the swap chimeras were more active at lower concentrations than their WT progenitors. The differences in protein disaggregation and refolding, in the prevention of protein aggregation and the stimulation of HSP70's ATPase in the absence and presence of substrates, strongly suggest that intra-molecular self-inhibitory interactions exist in SS and YY and were disrupted in YS and SY, rendering the swap mutant more active albeit also more wasteful in ATP terms of consumption. Limited proteolysis-coupled mass spectrometry and machine-learning structural prediction studies also suggested that residues of the conserved JDs, have co-evolved in SS and YY, to interact with surfaces of their CTDs, which overlap surfaces that bind the misfolded protein substrates, thereby forming a regulated stop-start mechanism to improve the coupling between energy consumption and the polypeptide unfolding work of eukaryotic HSP70s.

RESULTS

Design and analysis of domain swap YS and SY mutants.

When active, the compact N-terminal JD of YY and SS, which are both connected by a long flexible glycine-rich linker to their CTDs, are generally thought to freely seek binding to ATP-HSP70 molecules, which are 10-fold more abundant than the JDPs. In an auto-inhibitory state, the JDs of YY and SS should instead be locked in a restrained conformation preventing optimal binding to ATP-HSP70. This could happen if residues on the surface of the JDs would have co-evolved to reversibly bind residues on the surface of their G/F-rich regions and CTDs. To address this possibility, we reasoned that chimera in which the conserved JD from SIS1 (SS) would have been swapped with that from YDJ1 (YY) should lose their ability to adopt an autoinhibited state, while fully maintaining their HSP70-anchoring ability through their exchanged JD and fully maintaining their substrate-binding ability through their original G/F-rich regions and CTDs. We thus designed two J-domain swap mutants, in which the 80 N-terminal residues of YDJ1 (YY) and SIS1 (SS), containing the entire J-domain and a small 8-10 segment of the G/F rich region (Fig. 1A, underlined residues), were fused to the G/F-rich and the CTD of SIS1 and to the G/F-rich, the

cysteine-rich domain and the CTD of YDJ1 (See Fig. 1B), to form the SY and YS chimeras, respectively. Note that at variance with previously described similar swap mutants (*28-33*), the swap mutants here left the G/F rich regions mostly uninterrupted (Fig 1A, green).

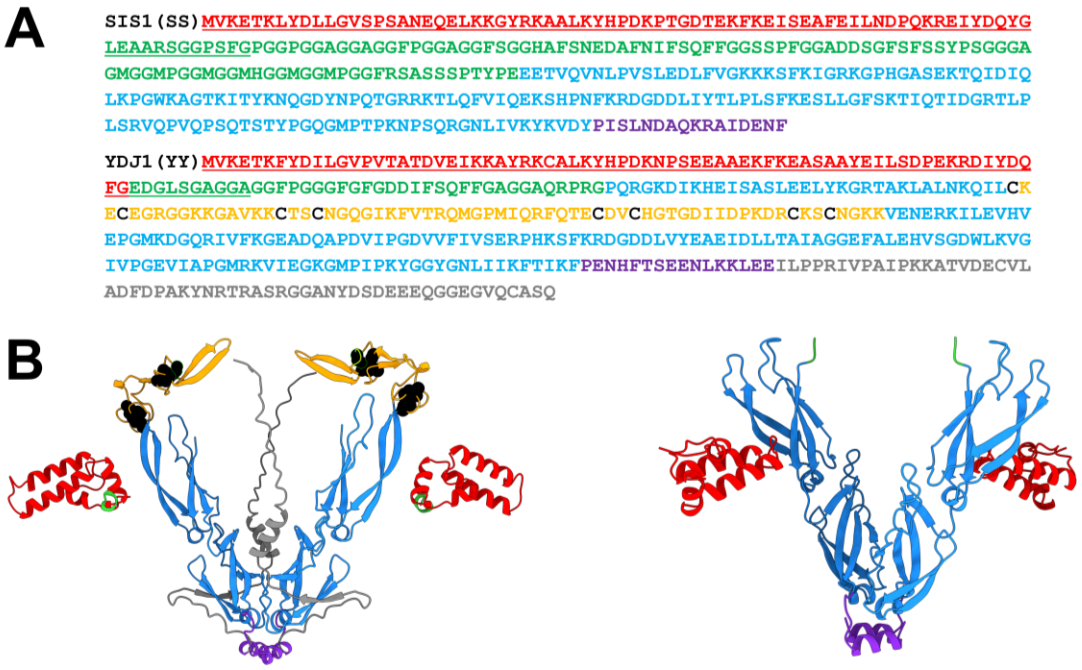


Figure 1: Domain organization of class A (YDJ1) and B (SIS1) JDPs. A: Sequence and domain organization yeast YDJ1, SIS1. J-domain, red; GF, green; CTDI & CTDII, light blue; Dimerization Domain (DD), purple; Zinc-finger-like region (ZFLR), orange; Cysteines pairs, black; DNAJA-specific C-terminal domain, grey. The underlined N-terminal amino acids of SIS1 (SS) and YDJ1 (YY) were exchanged to form the YS and SY chimera. **B**: AlphaFold-generated (*34*) homodimer models of SIS1 and YDJ1, using the domain color code as in A. The variable G/F region (green) is not well-resolved by AlphaFold, thus part of it is missing from the 3D modeling. Cysteine pairs are represented as black spheres.

The YS and SY chimeras better activate protein refolding by SSA1 than the wild-type SS and YY.

We tested the *in-vitro* ability of a constant amount of SSA1, SSE1, and ATP and of increasing concentrations of recombinant YS, SS, SY, and YY, to drive the native refolding of stable, preformed urea- and heat-preformed luciferase aggregates (*35*) (Fig. 2A) and preformed heat-aggregated G6PDH (Fig. 2B). Remarkably, refolding rates revealed that lower concentrations of YS were substantially more effective at refolding the pre-aggregated reporter enzymes, than the wild type SS. Whereas YY was virtually inactive at all concentrations, surprisingly, increasing concentrations of SY, which cumulated both the

apparently less active J-domain from SS and the apparently less active CTD from YY, was a more effective JDP cochaperone at lower concentrations than the wild-type YY.

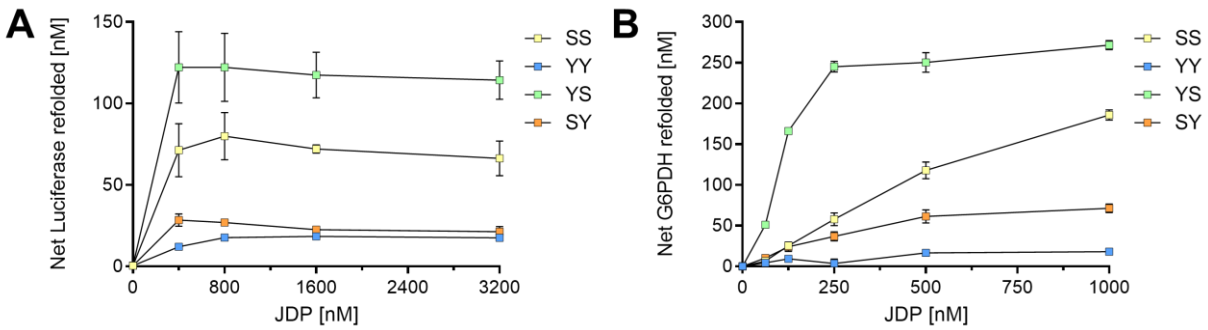
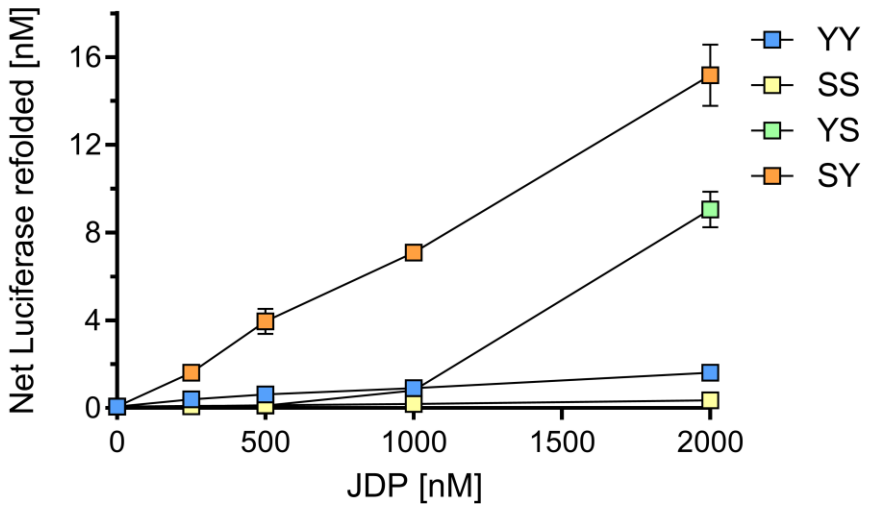


Figure 2: J-domain protein power HSP70-mediated disaggregation in a dose-dependent manner of different substrates. A: Efficiency of different JDPs proteins (SIS1 and YDJ1 and the two swap SY and YS) to power SSA1-mediated disaggregation of heat-urea pre-aggregated Luciferase. 0.2 µM of stable inactive luciferase aggregates were incubated for up to 120 minutes at 25°C in the presence of 5 mM ATP, 6 µM of SSA1, 0.75 µM SSE1, and between 0 to 3.2 µM of either YY, SS, SY or YS. B: Efficiency of four different JDPs proteins to activate the SSA1-mediated disaggregation of heat-denatured G6PDH. 0.5 µM of stable inactive luciferase aggregates were incubated up to 120 minutes at 25°C in the presence of 5 mM ATP, 6 µM of SSA1, 0.75 µM SSE1, and between 0 to 1 µM of either YY, SS, SY or YS. In all panels, error bars represent mean ± SD (n = 3).

The remarkable apparent gain of activity observed in the case of YS, as compared to SS, and of SY, as compared to YY, is strong evidence that in isolation, wild-type SS and YY may adopt an auto-inhibitory state, which was disrupted in the chimera, while both their JDs G/F and CTDs remained fully potent. When in these activity assays, we replaced the yeast SSA1 and SSE1 proteins with *E. coli* DnaK and GrpE (Fig. S1), the wild-type yeast JDPs, YY and SS, were expectedly virtually ineffective JDPs at activating luciferase disaggregation and refolding, owing to the great evolutionary distance between the bacterial DnaK and the yeast JDPs. Remarkably, however, the chimera YS and SY were significantly more effective than their wild yeast-type progenitors at activating DnaK-mediated disaggregation (Fig. S1). Interestingly, here SY was more effective than YS. This indicates that an intrinsic change in the molecular structure of the SY and YS, compared to SS and YY, is responsible for their apparent gain of JDP activity and is independent of the ability of SSA1 and SSE1 to bind misfolded substrates and hydrolyze ATP.



Supplementary Figure S1: J-domain protein power DnaK-mediated disaggregation in a dose-dependent manner of pre-aggregated Luciferase. The efficiency of different yeast JDPs proteins (SIS1 and YDJ1 and the two swap SY and YS) to power DnaK-mediated disaggregation of heat-urea pre-aggregated Luciferase. 200 nM of stable inactive luciferase aggregates were incubated for up to 120 minutes at 25°C in the presence of 5 mM ATP, 6 µM of DnaK, 2 µM GrpE, and between 0 to 2 µM of each JDPs. Error bars represent mean ± SD (n = 3).

The efficiency of YY, SY, SS, and YS correlates with their ability to dissociate the bound substrate protein substrate.

Whereas the role of the conserved N-terminal JD in class A (YY) and B (SS) JDPs, is to bind ATP-HSP70 and dissociate from ADP-HSP70, the role of their CTDs and the cysteine-rich domain of class A JDPs (*36*), is to bind misfolded protein substrates and present them to ATP-HSP70 and thereafter dissociate from both the ADP-HSP70 and the HSP70-unfolded protein product of the unfolding reaction (*25, 37, 38*). It is generally thought that the stronger the initial binding of a misfolded protein substrate to a JDP (the so-called "holdase" activity of JDPs), the more efficient will a JDP be at promoting the unfolding of that polypeptide substrate by HSP70 and ATP and its subsequent native refolding (*39*). Yet, given that in the cytosol, ER, and mitochondria of yeast cells, the JDPs are 10-fold less abundant than the HSP70s, JDP-mediated uploading of misfolded proteins onto the HSP70s needs to be catalytic, where, under optimal conditions, a single JDP molecule should successively upload misfolded substrates onto up to ten surrounding HSP70 molecules. This implies that as a JDP must seek to initially bind a misfolded

polypeptide, it must similarly seek to dissociate from it once it has been uploaded onto ATP-HSP70. We therefore next tested the ability of SS, YY, YS, and SY to bind a misfolded protein by measuring their ability to reduce the light-scattering signal of aggregating luciferase. Whereas without JDPs the time-dependent light scattering signal of urea pre-unfolded luciferase alone, upon dilution and incubation at 30^{o}C, was maximal after 24 min, it was strongly inhibited by the presence of 2 μM SY, less inhibited by the same amount of YY, then by SS and finally, it was virtually not inhibited by YS (Fig. 3A and S2).

Furthermore, measures of light scattering inhibition in the presence of increasing concentrations of the four JDPs confirmed that at all concentrations, SY and YY, which were the least effective at catalyzing the transfer of misfolded substrates onto HSP70 and promoting refolding, were the most potent at binding the misfolding luciferase and preventing its aggregation. In contrast, SS and YS, which were the most effective at catalyzing the transfer of misfolded substrates onto HSP70 and promoting refolding, were the least potent at binding the misfolding luciferase and preventing its aggregation (Fig. 3B).

The different apparent binding affinities of the four JDPs to aggregating protein intermediates were independently confirmed by FRET, using MLucV, which is composed of a urea pre-unfolded luciferase, flanked on both sides by urea-resistant mTFP1 and Venus FRET pairs (*40*). When MLucV was preincubated in 4M urea and then diluted 80 times, it readily formed stable aggregates with a FRET signal that was 142% of native MLucV (*40*) (Fig. 3C). The FRET signal of the aggregate was strongly reduced in the presence of increasing concentrations of SY and YY, whereas it remained virtually unaffected by increasing concentrations of YS and SS, indicating that note the binding but the subsequent release the misfolded protein substrate from onto SSA1, is the rate-limiting step of the reaction.

Noticeably, both swap constructs reduced aggregation more effectively than the wild-type JDPs, although containing the same G/F and CTDs. It is tempting to speculate that this is owing to their autoinhibitory conformation, with their J-domain possibly binding their protein binding regions.

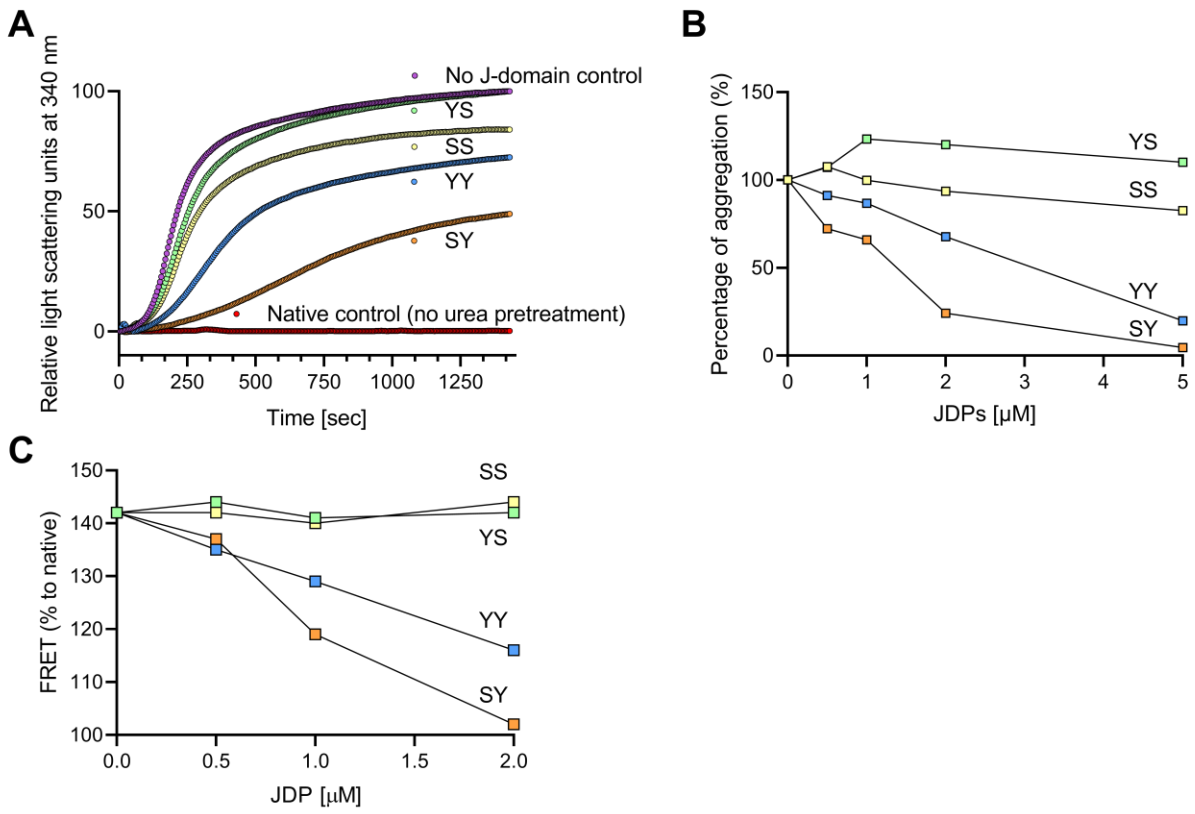
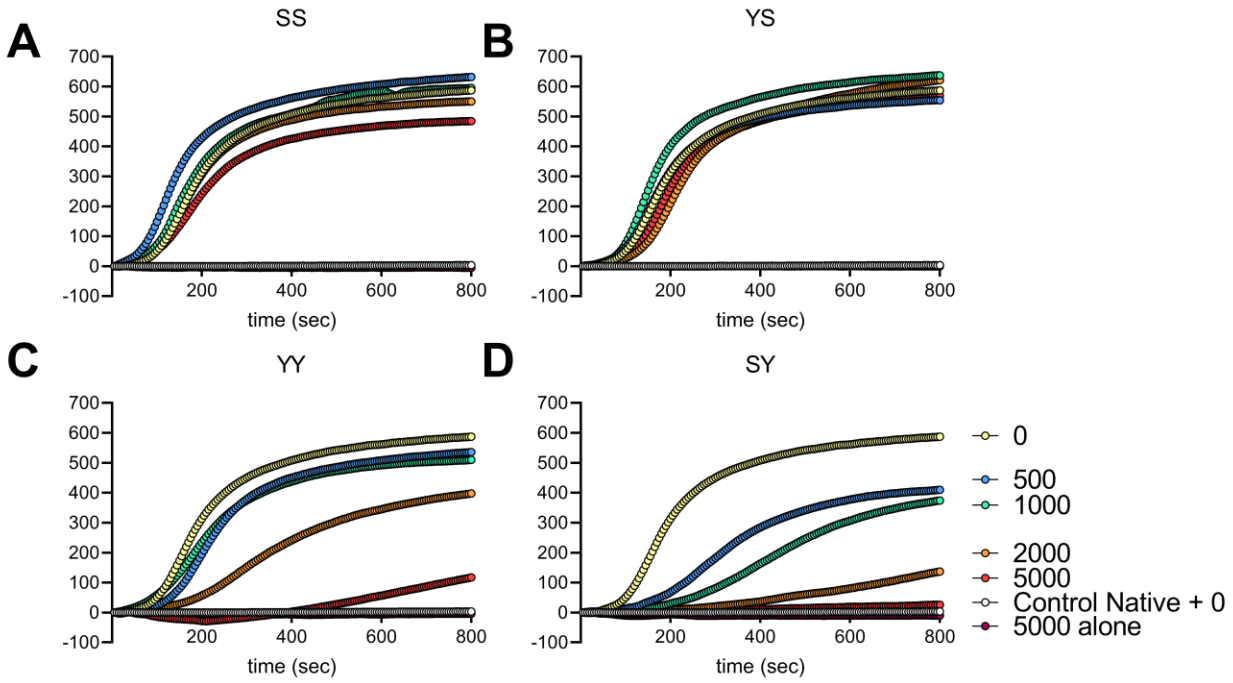


Figure 3: J-domain protein acting as holding chaperone. A: Efficiency of SS, YY, SY, and YS at preventing luciferase aggregation. 20 µM Luciferase was pre-unfolded with 4 M urea at 25℃ for 30 min, then diluted to a final concentration of 0.3 µM in buffer (50mM HEPES-KOH pH 7.5, 150mM KCl, 10mM MgCl2, 2mM DTT), the time-dependent aggregation was monitored by light scattering at 340 nm at 30 in the absence or presence of 2 µM SS, or YY or SY or YS. A native control (without urea-pretreatment) of 0.3 µM of Luciferase was also monitored B: Efficiency of dose-dependent JDPs proteins (SS and YY and the two swap SY and YS) to prevent Luciferase aggregation. 0.3 µM of native Luciferase was incubated and aggregation was monitored by light scattering with or without JDP (between 0 to 5000 nM), at 340 nm at 30°C for 25 min using Perkin Elmer Fluorescence Spectrophotometer. 100% of aggregation was set on the native Luciferase without any addition of JDPs. C: 200 nM of native MLucV (*40*) in the presence of 4 µM BSA and indicated concentrations (0, 500, 1000, and 2000 nM) of JDPs were incubated for 340 seconds at 42°C, following which, only 0.11% of the luciferase was still active. Luciferase activity and FRET efficiency were measured at the indicated time points and expressed as % of the initial luciferase activity and FRET efficiency at t=0. Because time-dependent measures of protein aggregation by light scattering are notoriously variable the presented results are only representative of three replicates.



Supplementary Figure S2: Prevention of urea-pre-unfolded luciferase aggregation by increased YY, SS, SY, and YS concentrations. 20 µM luciferase was preincubated in 4 M urea, then diluted to a final concentration of 0.3 µM in buffer containing 0, 500, 1000, 2000, 5000 nM of SS (**A**), YS (**B**), YY (**C**), SY (**D**). Aggregation was monitored by light scattering at 340 nm at 30°C. The presented results are representative of three replicates.

This data strongly indicates that in the catalytic action of the JDPs, the affinity, to be distinguished from the specificity, for a misfolded substrate is not as central to their function as generally thought (*41*). Rather, the rate-limiting step of the JDP catalytic cycle appears to be its ability, once the misfolded polypeptide is bound and unfolded by HSP70, to dissociate from both the unfolded polypeptide and from the ADP-HSP70.

Given that JDPs are ten times less abundant in the cells than their HSP70 partners, this result strongly indicates that the co-chaperoning activity of JDPs is not obligatorily associated with the ability of some JDPs to act as passive "holdases" by possibly preventing protein aggregation in the cell (*42, 43*).

SY and YS trigger SSA1's ATPase more strongly than SS and YY.

To further address the possibility that, at variance with wild-type YY and SS, the chimera are constitutively activated even in the absence of a bound substrate, we measured the rates of ATP hydrolysis by SSA1 in the presence of either of the four JDPs, without (Fig. 4A) and in the presence of misfolded luciferase (Fig. 4B). 2 μM of YY alone was twice less ATPase activatory than SS, suggesting that the YY's JD has less affinity for SSA1 than SS. Both YS and SY were more activatory than YY and SS. Remarkably, although the J-domain of YY and YS are identical and should have the same affinity for ATP-HSP70, the activation of SSA1's ATPase by YS was 2.7 times higher than by YY. This demonstrates that the YY's lower activity compared to SS is not because YY's JD would have an intrinsically lower affinity for ATP-SSA1, compared to SS's JD. It rather suggests that YY is more tightly auto-repressed and less prone to become de-repressed by the binding of a misfolded substrate, than in the case of SS. Similarly, but less extensively, the activation of SSA1 by SY was higher than by SS, further in line with the hypothesis that the swap mutants are more in a dis-inhibited state than their wild-type progenitors (Fig. 4A).

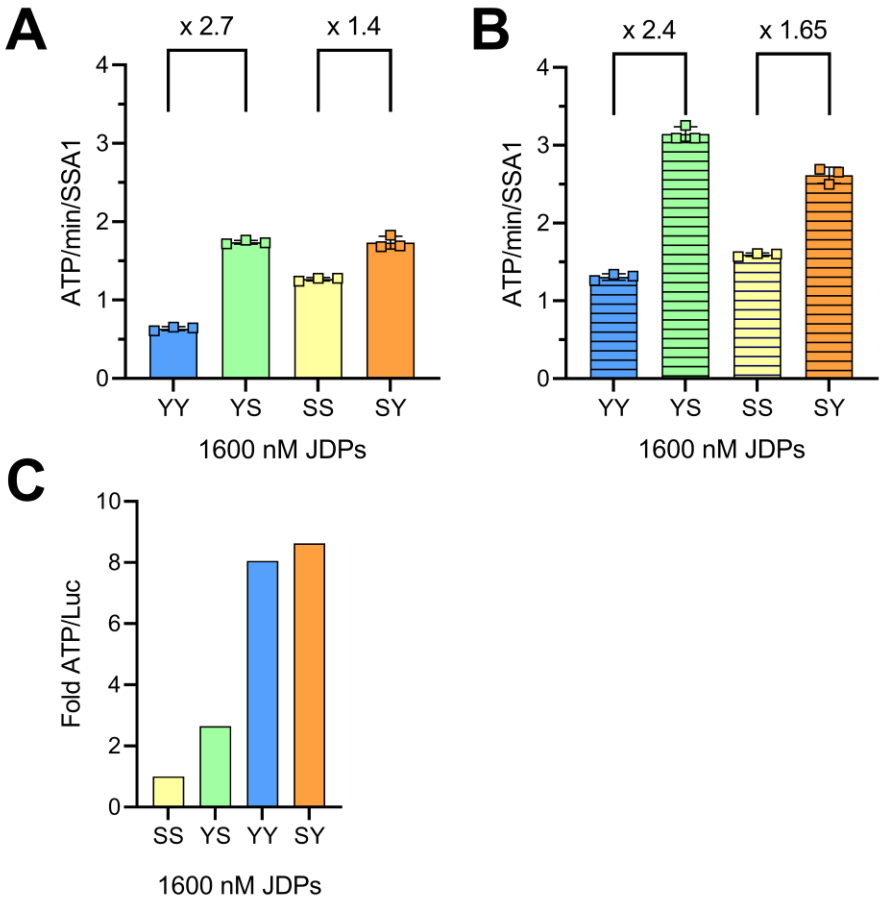
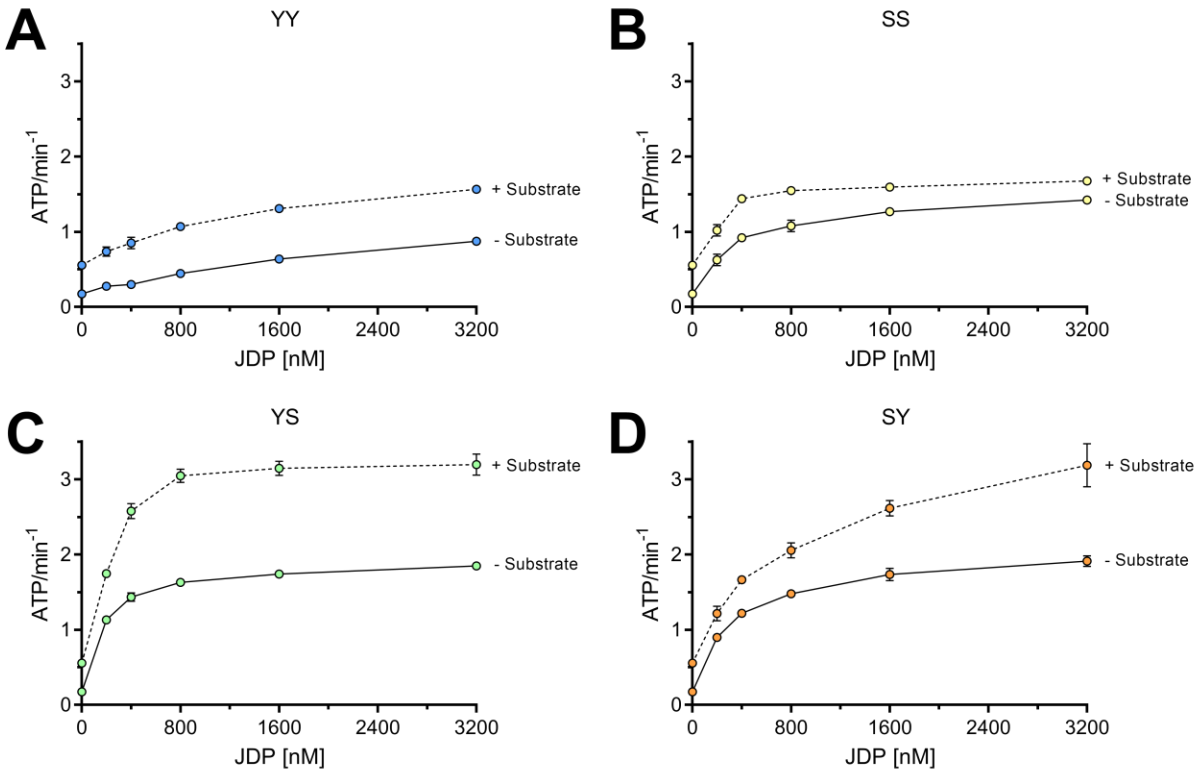


Figure 4: Rates of ATP hydrolysis of the chaperone assays in the presence of JDPs. A) Rates of ATP hydrolysis of 6000 nM SSA1, 750 nM SSE1 in the presence of 1600 nM YY, SS, YS, or SY without pre-aggregated luciferase. Brackets indicate the fold change of hydrolyzed ATP between the WT and Swap. **B)** Rates of ATP hydrolysis of 6000 nM SSA1, 750 nM SSE1 in the presence of 1600 nM YY, SS, YS, or SY in the presence of 200 nM pre-aggregated luciferase. Brackets indicate the fold change of hydrolyzed ATP between the WT and Swap. **C**) Fold-change of ATP hydrolyzed per natively refolded luciferase with SIS1 set as a reference, by 6000 nM SSA1 and 750 nM SSE1, in the presence of either 1600 nM YY, SS, SY, or YS. The error bars in panels A and B represent mean ± SD (n = 3).

The presence of a bound misfolded substrate further nearly doubled the rate of YY-induced ATP hydrolysis by SSA1 and also the rate of SS-induced ATP hydrolysis, albeit to a lesser extent. This indicates that while YY-induced refolding was very ineffective, much more ATP was hydrolyzed than in the case of productive SS-induced refolding. Noticeably, the misfolded substrate also mildly increased the rates of YS- and SY-induced ATP hydrolysis, suggesting not all of the swapped JDP molecules are in a fully disinhibited state, despite the disruption of their presumed co-evolved binding surfaces on the J-domain and corresponding G/F and CTDs (Fig. 4B). The activatory effect of increasing amounts of the four JDPs on SSA1's ATPase was further addressed in the absence or presence of SSE1 and/or of the misfolded substrate (Fig. S3).



Supplementary Figure S3: Substrate and JDPs increase ATPase activity of HSP70 (SSA1). A-D: Rates of ATP hydrolysis of 6000 nM SSA1 and 750 nM SSE1 in the presence of increasing concentrations of YY, SS, YS, and SY, with or without 200 nM of pre-aggregated luciferase. The error bars in all panels represent mean ± SD (n = 3).

Confirming that JDP activity does not directly correlate with the apparent specific affinity of the J-domain for ATP-HSP70, the activation of SSA1's ATPase was half-maximal with 360 nM YS, whereas at the concentration, the activation of G6PDH disaggregation and refolding was already maximal (see Fig. 2B). Conversely, the activation of SSA1's ATPase was already maximal with 230 nM SS, whereas the activation of G6PDH disaggregation and refolding was only a third of the maximal refolding.

Knowing the number of ATP molecules hydrolyzed by one SSA1 molecule per hour and the number of luciferase molecules natively refolded per hour, allowed us to estimate the relative ATP cost of the *in vitro* disaggregation/refolding reaction, carried by SSA1 and SSE1 in the presence of a constant amount (1600 nM) of SS, YY, YS or SY (Fig. 4A, B). The most energy-efficient JDP of the four was SS (whose relative efficiency was set to 1). The second most energy-efficient JDP was YS, which hydrolyzed 2.6-

times more ATPs per refolded Luciferase. YY and SY were equally energy inefficient. They hydrolyzed at least 8 times more ATP per refolded luciferase, than SS (Fig. 4C).

Note that the ATP cost of the chaperone-mediated protein disaggregation reactions is probably much lower in the yeast cytosol than in our poorly performing *in vitro* chaperone assay. Yet, these results show that the YS swap mutant, although being the most effective JDP in terms of activated protein disaggregation and refolding by SSA1-SSE1, does so at an ATP cost that is 2.6-higher than for wild-type SS (Fig. 5C). Thus, when limited in food and energy supply, yeast in the wild, may have favored the evolution of SS-like JDPs that are slightly less efficient but are with an energy-saving auto-repression mechanism that can reduce futile ATP hydrolysis in the absence of misfolded proteins to repair.

Identification of possible interaction surfaces between J-domain and CTDs of wild-type JDPs.

We next attempted to identify by Limited Proteolysis – Mass Spectrometry (LiP-MS) (*44*) and by AlphaFold-based analysis, potential residues on the surface of the JDs of YY and SS that may have co-evolved with residues on the surface of their own G/F and CTDs, to form reversible intramolecular associations in the autoinhibited state. Following limited proteinase K digestion and total trypsin digestion, we used my mass spectrometry to identify semitryptic peptides in YY and SS, whose normalized abundances changed significantly in YS and SY, indicative of surfaces on the JDs, G/F, and CTDs of the Swap JDPs, whose exposure to limited proteinase K digestion changed significantly, as compared to the WT JDPs. In general, we found that YS and SY contained more PK-accessible regions than their corresponding WT counterparts. This was in particular the case for the J-domain of YS, as compared to that of YY (Fig. 5A) and for the G/F and CTD of YY, compared to SY (Fig. 5C). We mapped the most significant PK-protected positions in the J-domain and CTDs of wild-type JDPs, as compared to the swap mutants, on the AlphaFold structures of YY and SS (see Methods). These positions are highlighted in red in Fig. 5E (YY) and 5F (SS). The CTD_{YY} residues that are more protected in YY than in SY and the JD_{YY} residues that are more protected in YY than in YS are highlighted in red in Fig. 5A. On the CTD, protected regions are localized on the Cys-domain (H189 and G190), on CTD_2 close to the dimerization domain (T332 and E337) and close to the substrate-binding region of CTD_1 (E217). On the J-domain, protected surfaces are localized on helix I (V13) and in the middle of helix III (H48).

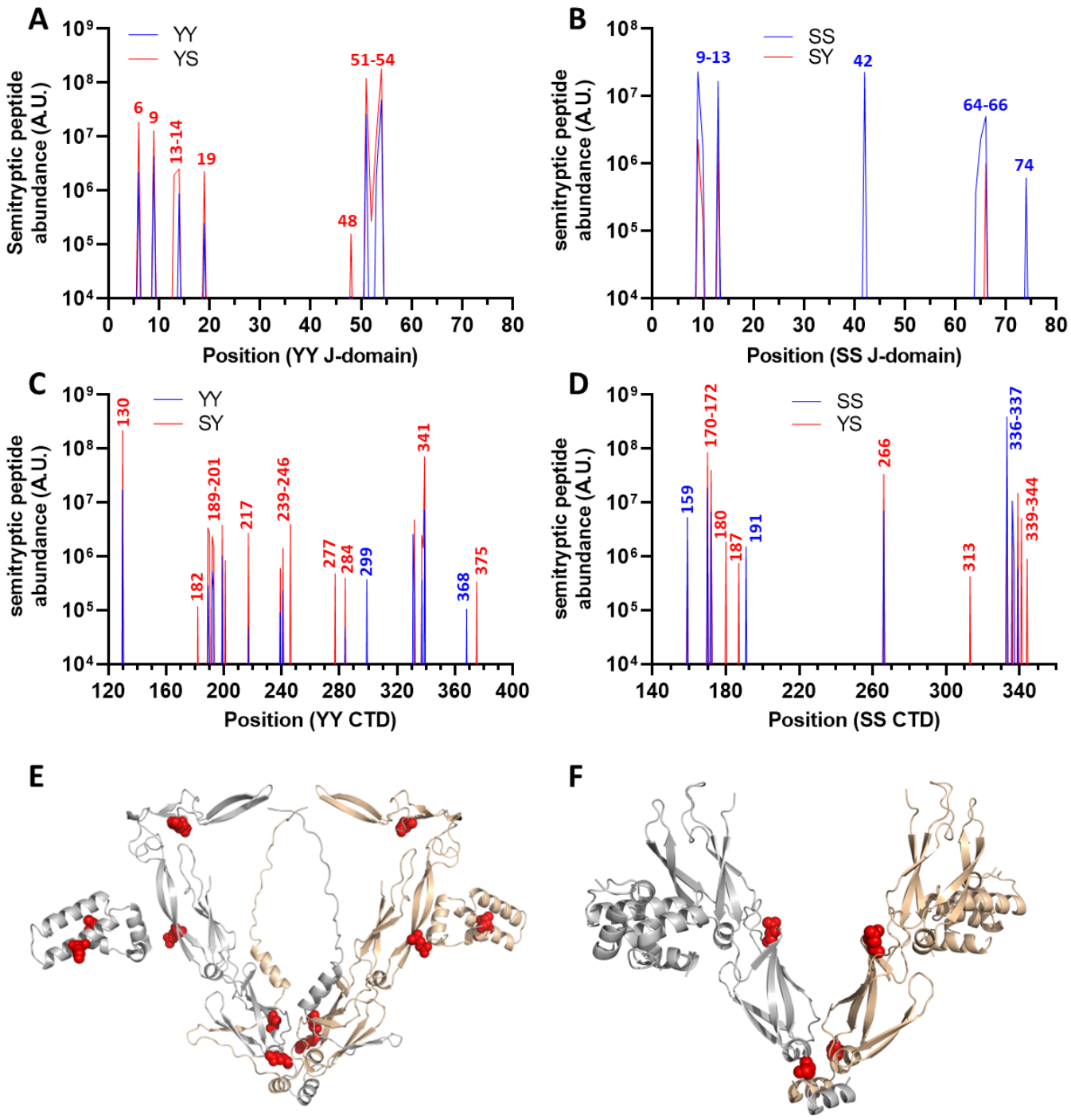
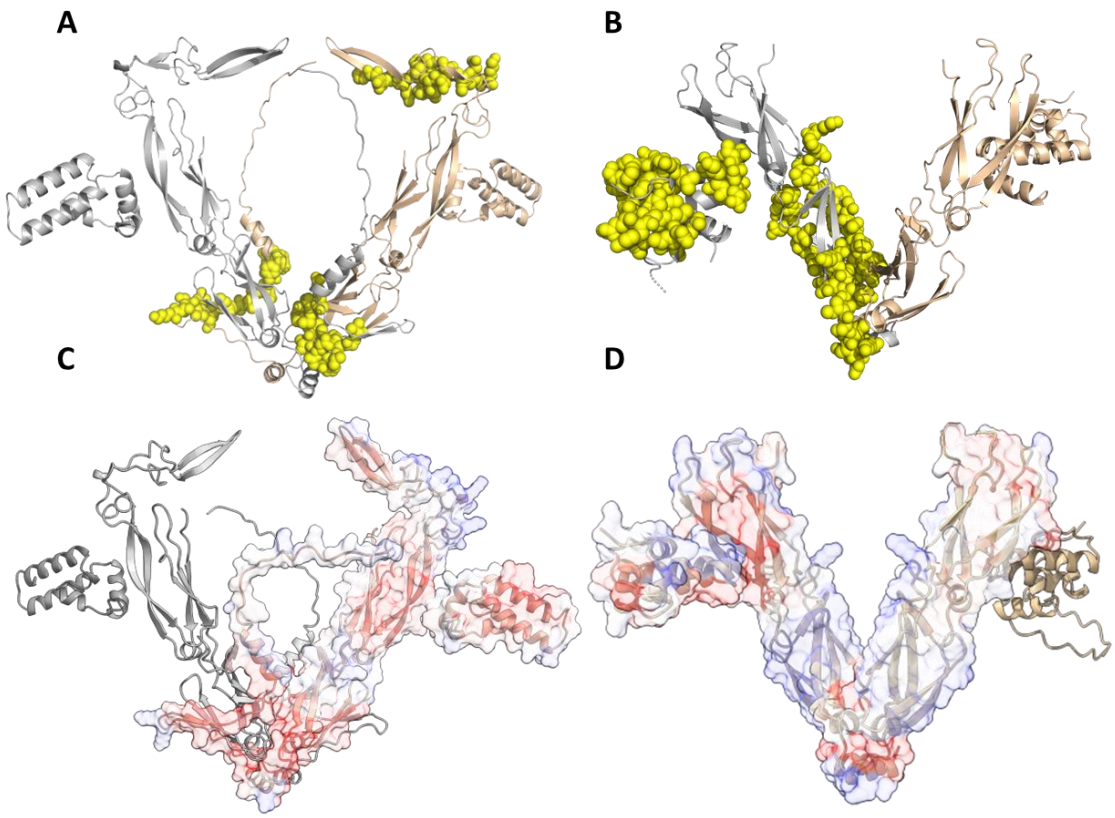


Figure 5: Limited proteolysis – mass spectrometry analysis of the JDPs. A-D: Mapping of proteinase K cutting sites in pairs of compared domains across JDPs, as determined by LiP-MS. **A:** PK sites in the J-domains of YY *vs.* YS. **B:** PK sites in the J-domains of SS *vs.* SY. **C:** PK sites in the CTDs of YY *vs.* SY. **D:** PK sites in the CTDs of SS *vs.* YS. In all panels, numbers above the peaks indicate identified PK cutting sites; red numbers correspond to sites that were more exposed to PK in the Swap compared to the

corresponding WT protein; numbers in blue correspond to sites more protected from PK in the swap than in the WT protein. **E** and **F:** Residues that are more protected in the wild types than in the swap mutants in YY (**E**) and SS (**F**).

We then compared the LiP-MS predicted YY-protected regions to the most temperature-sensitive regions identified by Leuenberger *et al.* (*45*), (yellow in Fig. S4A (YY) and S4B (SS) and to regions predicted to be the most prone to molecular interactions by MaSiF (*46*) (see Methods) (see Fig. S4C for YY and S4D for SS). Interestingly, the regions identified in my MaSIF agreed with the substrate-binding regions previously identified by NMR (*47*). In the case of YY, the JD-CTD complex predicted by AlphaFold 2, placed helix I of the JD near the MaSiF-predicted interaction region on CTD_1, despite AlphaFold 2 not being trained to take surface interaction propensities into account. Taken together, these data gave strong support to the biochemical data indicating that the JD of YY interacts with the substrate-binding region in CTD_1, leading to the protection of some close-by residues (as a matter of fact, the next most protected residue, is not represented in Fig. 5E, is G130, just about 9Å away from E217). In the case of SIS1 CTD, the more PK-protected regions in SS, we localized in the proximity of the dimerization domain (S339) and close to the hinge region connecting CTD_1 to CTD_2 (D266)(Fig. 5D and 5F). A further position was localized on the flexible G/F region (S172), whose structural prediction by AlphaFold 2 is unreliable. No positions were more protected in SS on the JD than SY.



Supplementary Figure S4: Thermally-sensitive and interaction-prone regions in YY and SS. A and **B**: Thermally sensitive regions according to Leuenberger *et al.* (*45*) on YY (**A**) and SS (**B**); **C** and **D**: MaSiF (*46*) predictions for the most interaction-prone regions of YY (**C**) and SS (**D**); red to blue: strong to no interaction propensity.

Looking at the LiP-MS findings of YY and SS sensitivity to a higher temperature (*45*) (Fig. S4A, B) involves most of CTD_2, thus somewhat correlating with our LiP-MS findings with the swap/WT analysis (Fig. 5F) and helices I and IV (and a small portion of helix II) of the JD. MaSiF analysis of SS (Fig. S4D) revealed that the only CTD region with a strong interaction propensity is located on CTD_1, which has been identified earlier as the substrate-binding region of SIS1 (*47*) and also contains the interaction site for the EEVD motif that in the cytosol of eukaryotes is to be found at the C-terminus of a subclass of HSP70s (SSA1-4) and Hsp90s (Hsc82). The same region was also chosen by AlphaFold-2 to be most probable to form interactions with their own JD. On the JD side, the most prominent interaction regions were predicted to be, expectedly, in helix II (that interacts with HSP70) and helix IV. These results were for the most part consistent with each other and with what has been previously reported in the literature and suggest that

the CTD-bound JD of SIS1 may compete with the binding of the misfolded substrate as well as with the binding though HSP70's C-terminal EEVD motif to CTD_1 (*37*).

Thus, limited proteolysis-coupled to mass spectrometry of YY, SS, SY, and YS provided evidence that in WT YY and SS, the J-domains can reversibly interact with their CTDs, thereby reducing the ability of the JD to interact with ATP-HSP70. Together with the ATPase measurements, our results suggest that when JD-CTD interactions are disrupted by a bound misfolded substrate in WT JDP, or near constitutively in a swap mutant, the JDs are mostly in their active conformation, poised to bind and trigger ATP-hydrolysis in SSA1, even in the absence of a misfolded substrate when unfolding energy is not needed.

Individual YY and SS may have an inactive state.

To further address evidence of an auto-repressed state of DNAJAs and DNAJBs, we used *in vitro* protein disaggregation and refolding assays to revisit the known synergism between the two mains cytosolic JDPs of yeast, possibly resulting from their mutual dis-inhibition, by an unknown mechanism: DNAJA type, YDJ1 (YY) and DNAJB type SIS1 (SS). Heat-pre-aggregated luciferase (0.2 μM), SSA1 (6 μM), SSE1 (0.75 μM), and ATP (4 mM) were incubated with different mixes of YY and SS, keeping their sum constant (at 4 μM).

Luciferase refolding yields (Fig. 6A) showed that, whereas up to 4 μM YY alone remained virtually ineffective at activating the luciferase refolding, growing concentrations of SS increasingly activated the chaperone mediate luciferase refolding. Yet, a mixture of YY and SS was even more effective: in the presence of 3 μM SS, 1 μM YY, by itself ineffective, nearly doubled the luciferase refolding yields, in keeping with previous findings for the human HSP70/Hsp40/NEF system (*21*). This strong synergism, which was observed at all ratios between YY and SS (Fig. 6A) was confirmed using a different protein substrate: heat-pre-aggregated G6PDH (*48*). SSA1, SSE1, and ATP with increasing amounts of up to 1 μM of YY alone merely produced 20 nM of active G6PDH in two hours, out of 600 nM heat-preaggregated G6PDH (Fig. 6B). Whereas 0.5 μM SS alone produced up to 40 nM native G6PDH, the addition of 0.25 μM YY, which alone led to the mere production of 11 nM G6PDH, nearly tripled the refolding yields to 130 nM native G6PDH (Fig. 6C). This is in line with the view that when alone, SS and YY can independently adopt an auto-repressed inactive state and that under certain conditions, as when together, change into de-repressed JDPs effectively triggering SSA1's ATPase and polypeptide unfolding activity.

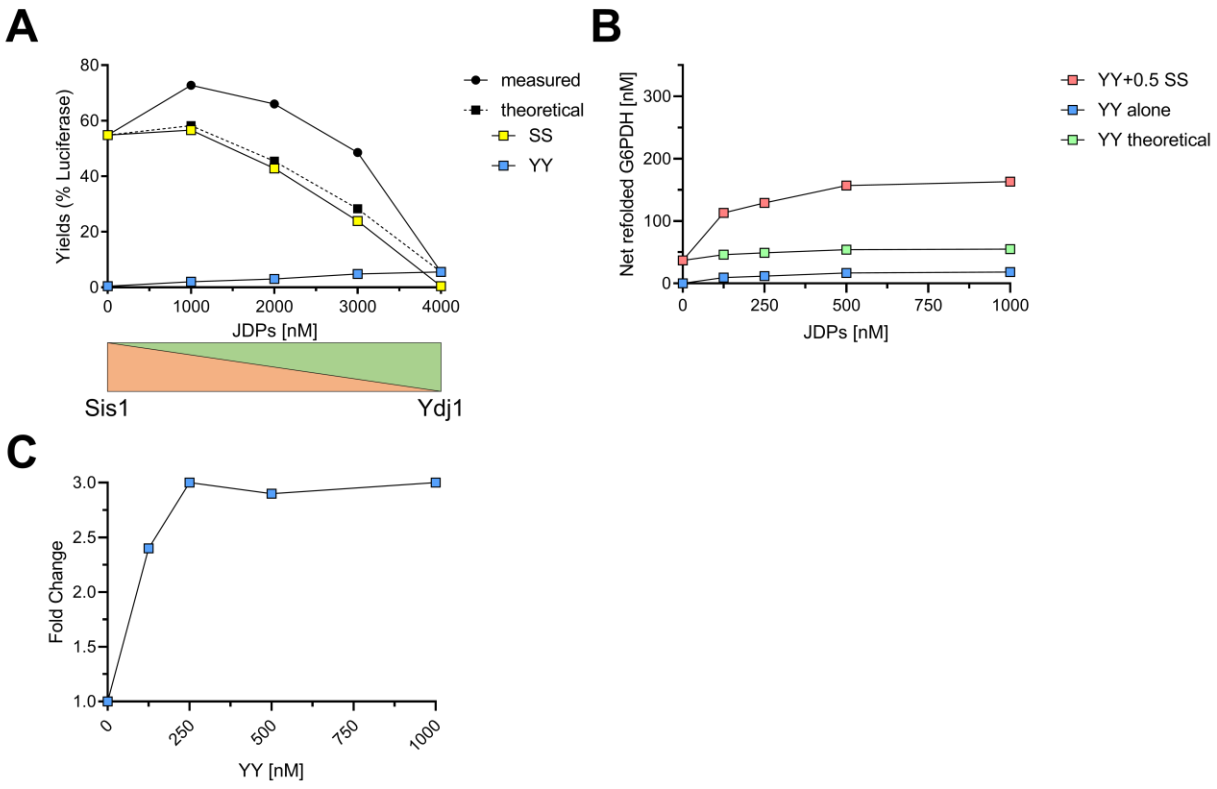
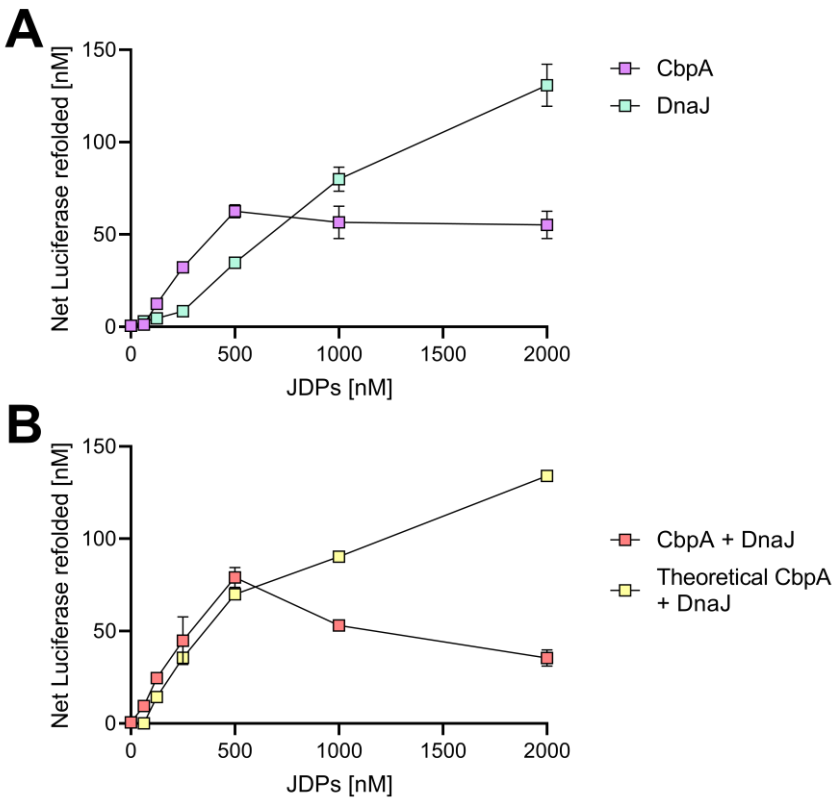


Figure 6: Synergism between YY and SS. (A) Refolding yields of heat-pre-aggregated luciferase (200 nM) in the presence of SSA1 (6 µM), SSE1 (0.75 µM), ATP (4 mM) at indicated SS: YY ratios, between 4 micromolar SS and 0 micromolar YY (left), up to 4 micromolar YY and 0 micromolar SS (right). Hatched lines, theoretical values if SS does not collaborate with YY. (B) one hour of refolding yields heat-pre-aggregated G6PDH (0.2 µM) in the presence of SSA1 (6 µM), SSE1 (0.75 µM), ATP (4 mM), and increasing concentrations of YY, without (Blue squares), or in the presence (red squares) of 500 nM SS. (C) effect in fold change on the basal chaperone activity activated by 500 nM SS of increasing concentrations of YY.

We next asked if bacterial DNAJAs and DNAJBs can also alternate between active and auto-inhibited states. Heat pre-aggregated luciferase (0.2 µM) was incubated for 2 hours with DnaK (4 µM), GrpE (1 µM), and ATP (4 mM), in the presence of increasing amounts of *E. coli* DNAJ (DNAJA) or CBPA (DNAJB). The luciferase refolding yields showed that individually, increasing concentrations of individual bacterial DNAJ or CBPA, up to 0.5 µM, drove equally well a steady increase of chaperone-mediated luciferase renaturation. Whereas an excess of CBPA above 0.5 µM inhibited the refolding reaction, DNAJ was not inhibitory below 2 µM (Fig. S5A). Remarkably, increasing amounts of equimolar DNAJ and CBPA, up to 250 nM each, showed no apparent synergism between the two (Fig. S5B) Higher equimolar

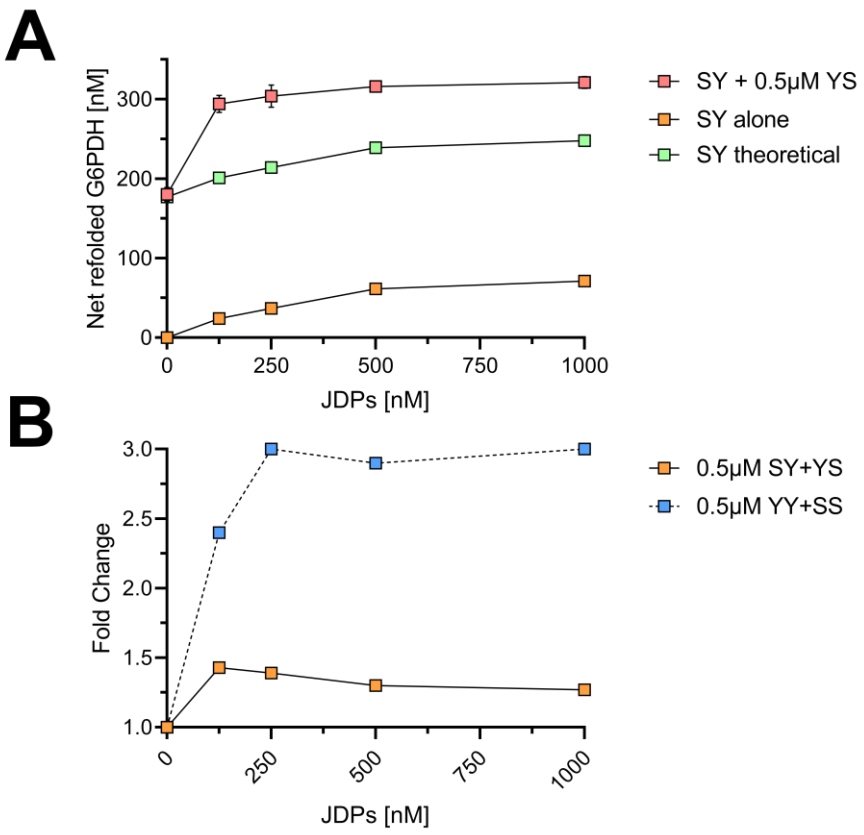
concentrations of the two, led to a remarkable decrease of the refolding yields, whereas in contrast, the calculated theoretical sum of their standalone activities predicted increasing yields. This suggests that contrary to the eukaryotic DNAJAs and DNAJBs that seem to activate each other, the *E. coli* DNAJA and DNAJBs apparently inhibited each other's activity. It should be noted that exponentially growing *E. coli* cells contain ~10 times less CBPA than DNAJ molecules (*49*). CBPA only accumulates in the late log phase, when there is no more cell growth and therefore, there is no massive *de novo* protein synthesis and folding (*50*). Thus, it is unlikely that in exponentially growing *E. coli* cells CBPA significantly inhibits DNAJ's co-chaperoning activity.



Supplementary Figure S5: Refolding yields of 200 nM heat-pre-aggregated Luciferase by 4µM DnaK, 1µM GrpE, and 4 mM ATP, in the presence of increasing amounts of DNAJ and/or CBPA. (A), or increasing equimolar DNAJ+CBPA (B), compared to the theoretical summed-up values expected when DNAJ and CBPA would not have cross-interactions.

As separately, YY and SS are in an auto-inhibited state and their mixing is causing their apparent dis-inhibition by a yet unclear mechanism (Fig. 6B), we next tested the effect of mixing YS with SY which

separately appear to be constitutively more active. The potential synergism of increasing amounts of SY was measured in the presence of pre-aggregated G6PDH, SSA1, SSE1, and ATP in the absence or presence of YS (0.5 μM) (Fig. S6). Increasing concentrations up to 1 μM of the less active SY, in the presence of a constant limiting amount (0.5 μM) of the over-active YS chimera, merely promoted 1.3 folds the chaperone activity. For comparison, the addition of merely 0.125 μM YY, which alone is virtually inactive, tripled the activity of 0.5 μM SS alone (Fig. 6B and S6). Thus, in contrast to YY and SS, which alone seems to be in an auto-repressed state and become derepressed when mixed, YS and SY, appear to be mostly in a constitutively active state that cannot be strongly activated further when mixed (Fig. S6).



Supplementary Figure S6: Effects of mixed-classes of hybrid JDPs on SSA1-mediated disaggregation and refolding activity. A: Refolding yields (2 hours) of 0.6 μM of heat-pre-aggregated G6PDH treated with 6 μM SSA1, 0.75 μM SSE1, 4 mM, ATP, without (orange) or with 0.5 μM YS (green) and increasing amounts of SY. **B:** Fold change of chaperone refolding yields in the presence of 0.5 μM YS and increasing

concentrations of SY (orange). For comparison, the refolding yields in the presence of 0.5 µM SS and increasing concentrations YY are shown (blue).

Discussion

Quantitative proteomics shows that in both prokaryotes and eukaryotes, the number of HSP70 molecules is ~10 times greater than the sum of the DNAJA, B, and C (*13*), indicating that optimal protein homeostasis is achieved when a single DNAJA or B dimer successively delivers misfolded polypeptides to several HSP70s (*48, 51*). Thus, an optimal catalytic unfolding cycle by HSP70 must include not only an initial stage, during which a JDP binds to a misfolded substrate and triggers ATP-hydrolysis in HSP70 but, as importantly, a final stage in which the JDP dissociates, both from the ADP-HSP70the and from in unfolded polypeptide product of the chaperone unfolding reaction.

Whereas yeast HSP70 alone spontaneously hydrolyses ATP at a rate of ~0.1 per min (*52*), the binding of either a misfolded protein substrate, or of a WT JDP (depending on the concentration), triggers an increase of the ATPase rates, respectively by ~3 and 2-6 folds (*16*), yet without producing natively refolded proteins. Only when both misfolded substrate and JDP are synergically triggering HSP70's ATPase by 4-10 folds, some coupling between energy expenditure and the protein unfolding/refolding process is occurring. Proteomic estimate that a yeast cell cytosol contains ~15-20 µM of HSP70s and 2 µM of JDPs (SS and YY) (Table S1). Most yeasts in the wild live a quiescent lifestyle with very slow doubling times exceeding a week. They do not need massive amounts of proteins involved in biosynthetic processes whose misfolding would necessitate ATP-fueled chaperones and in general undergo very occasional *de novo* protein folding (*53*). Unstressed cells lacking an auto-repressive mechanism in their JDPs would be expected to unnecessarily trigger futile ATP-hydrolysis by HSP70 to hydrolyze ATP while unfolding work is not necessary.

Uniprot ID	Protein Name	Class	Function(s)	Compartment-Specific Concentration (μM) Mean and SD
P32589	SSE1	Hsp110	HSP70 nucleotide exchange, substrate binding, constitutively expressed	3.27±1.05
P32590	SSE2	Hsp110	HSP70 nucleotide exchange, substrate binding, stress-inducible	0.29±0.11
P10591	SSA1	HSP70	Protein folding, translocation, constitutively expressed	14.87±6.58
P10592	SSA2	HSP70	Protein folding, translocation, constitutively expressed	4.35±1.77
P09435	SSA3	HSP70	Protein folding, translocation, stress-inducible	0.15±0.19
P22202	SSA4	HSP70	Protein folding, translocation, stress-inducible	0.45±0.40
P25491	YDJ1	JDP	HSP70 ATPase activator, substrate binding	1.10±0.00
P25294	SIS1	JDP	HSP70 ATPase activator, substrate binding	0.89±0.22

Table S1: Protein concentration in *Saccharomyces cerevisiae*. Calculations were made by extracting several results from previously published papers by estimating the concentration of proteins derived from the mass, density, and volume of yeast during its cell cycle and by calculating organelles and cytosol volume deducted and calculated from soft X-ray tomography studies (*54-59*).

A central question is why hasn't evolution favored the formation of a YS-like JDP, which appears to be more efficient at lower concentrations than SS. A slightly different yeast chimera named YYS has been described earlier by (*28*), in which 104 N-terminal residues of Ydj1 were fused to Sis1's CTD, rather than only the first 80 N-terminal residues in the case of YS. Interestingly, YYS was found to rescue the growth of a *ΔSIS1* mutation in a glucose-rich medium (*28*). It would be interesting to further investigate the ability of YYS and YS to rescue the growth of a *ΔSIS1* yeast on media limited in energy source and under heat shock. It should be noted that YYS, at variance with YS, included 24 residues from Ydj1's G/F region and did not include 41 residues from Sis1's G/F region involved in substrate binding and as found here, in the reversible binding of the JD to form the auto repressed state.

We show here that although lower concentrations of YS mediate the unfolding/refolding of more misfolded luciferase in a shorter period and reached higher yields, it is using 2.6 times more ATP to refold a single luciferase molecule, compared to SS. It is not clear whether a similar stop-start system in the JDPs exists in prokaryotes. We find (Fig. S3 & S5) that *in vitro*, there is a great difference between the *E. coli* system (DnaK+GrpE with DNAJ or CpbA) and the yeast eukaryote system (SSA1+SSE1 with wild-type YDJ1 or SIS1). In vitro, we find that the yeast system is 10-100-fold slower and it is also apparently proportionally much more wasteful in ATP than the prokaryotic system. DNAJ and CpbA do not collaborate; actually, they rather inhibit one another (Fig. S5). Here we also showed that yeast class A (YY) and B (SS) JDPs can collaborate to actively power HSP70 disaggregation, as previously shown by Nillegoda *et al (21)*. Interestingly, when comparing the ability of WT JDPs and SWAPs to collaborate and activate substrate disaggregation by HSP70, the fold change of SWAPs is significantly lower than that of WTs, being already in a de-repressed state (Fig. S5). It is tempting to speculate that because eukaryotes are more wasteful, they need their DNAJA and DNAJBs to have an effective Stop-start mechanism, whereas prokaryotes being more ATP effective their DNAJA and DNAJBs less need a Stop-start mechanism (Fig. 7). This, however, needs to be further addressed in future experiments, by performing among others JD swaps between E. coli DNAJ and CBPA.

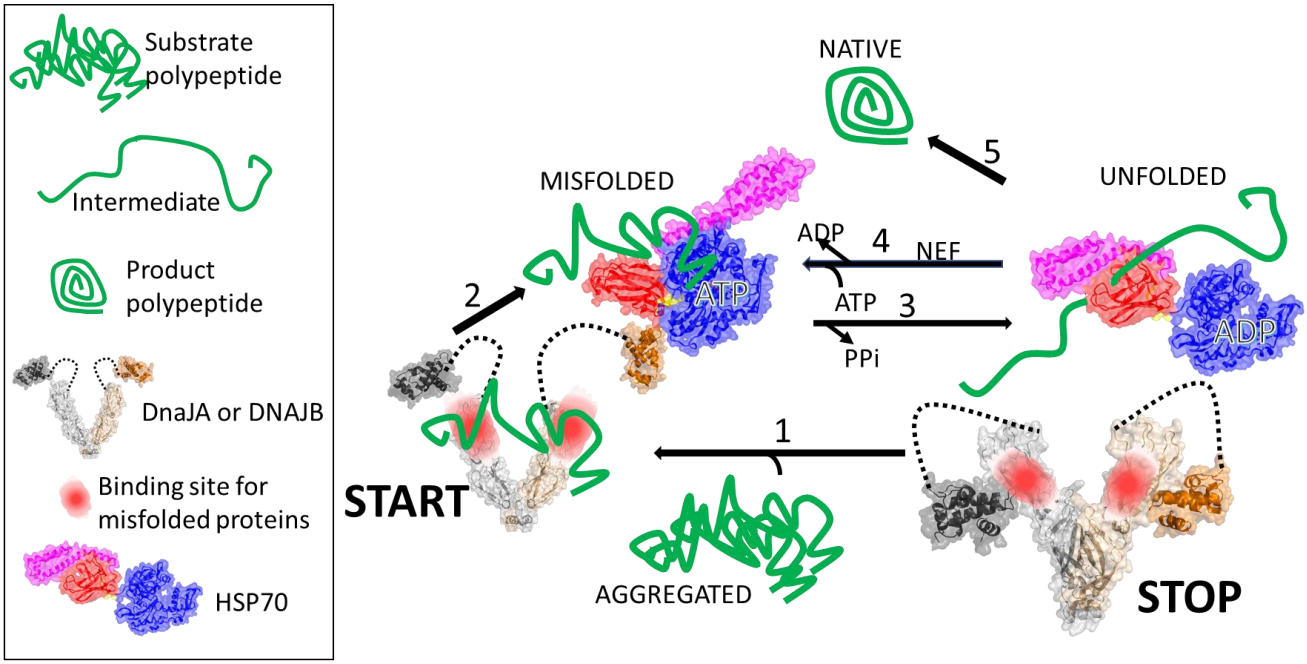


Figure 7: Eukaryotic DNAJA and DNAJB co-chaperones use a stop-start mechanism to optimize the coupling between protein unfolding and ATP hydrolysis by HSP70. In the absence of misfolded protein

substrates, a DNAJ (A or B) dimer has an arrested confirmation (stop), with its J-domains binding adjacent substrate-binding surfaces (red) rather than ATP-bound HSP70s. Binding of misfolded protein (green)(step 1), the DNAJ(A or B) dimer is activated (start), the J-domains bind ATP-HSP70, and the misfolded protein is uploaded (step 2). ATP-hydrolyses and the bound misfolded protein are unfolded (step 3). The nucleotide exchange factor (NEF) releases ADP (step 4) and the unfolded protein intermediate (step 5) that spontaneously folds into a low-affinity native product.

Thus, in the absence of a bound misfolded protein substrate, futile ATP hydrolysis by HSP70 can be reduced in Eukaryotes. LiP-MS experiments suggested that in the absence of a substrate, the J-domains of YY and SS may interact with their own CTDs. The identified regions mostly correlated with positions known to be more thermosensitive. While this does not necessarily imply that they are direct interaction regions, this indicates that the intra-molecular CTD-JD interactions in wild-type JDPs may rigidify and stabilize the overall structure of the molecule. When in the swap mutants these interactions are disrupted, the structure acquires increased flexibility, impacting the regions that are also more sensitive to increased temperatures. Furthermore, AlphaFold 2 predicted JD-CTD complexes that matched the regions predicted by MaSiF to be the most interaction-prone on the CTD_1s and, for YY, on the JD. Taken together, these results suggest a picture whereby in the absence of bound substrates, the wild-type JDs do tend to interact, although possibly only weakly, with their CTDs and by doing so, compete with substrates binding (and with the binding to SS of the EEVD C-terminal motif from SSA1) for the same interaction sites, further supporting the hypothesis that in the absence of substrates, these interactions can partially prevent the interaction of JDPs with ATP-HSP70s, thereby reducing unnecessary futile energy consumption.

The importance of Stop-start mechanisms to save ATP in biology

The mass of the ATP-hydrolyzing chaperones Hsp110s Hsp100 Hsp90s HSP70s and Hsp60s can together approximate 5-8% of the total protein mass of a healthy eukaryotic cell, a quarter of which being members of the HSP70 family (*60*). In addition, the proteasome, other ATPase machines, ATP-transports, pumps, GTPases, kinases, and so on, can sum up to be 10% of the proteome mass. Thus, when a yeast cell in nature is living on extremely low sugar, it may grow and divide only once a week. Given that the yeast cytosol contains about 4 mM ATP, 16 micromolar HSP70 (SSA1), 1 micromolar YDJ1, and 1 micromolar SIS1 (Table S1) and assuming a very conservative rate of futile ATP hydrolysis by JDP-triggered SSA1 of 2 per minute, the absence of a stop-start mechanism would imply complete ATP

consumption in the cell in less than 3 hours. Therefore, evolution is likely needed to favor the apparition of a stop-start regulatory mechanism to tame futile ATP hydrolysis by molecular machines such as HSP70s.

Remarkably, in addition to the described JDP-dependent stop-start mechanism, the coupling of ATP consumption to the protein remodeling (unfolding) work of HSP70 can additionally be regulated by other stop-start mechanisms in two of its ATP-fueled disaggregase co-chaperones: HSP110 and HSP104. Hence, the basal, stand-alone disaggregase activity of JDP-HSP70-NEF machinery, is synergized by the co-disaggregase HSP104 (ClpB in bacteria), whose ATP consumption is regulated by Off-and-On states of its M-domains, reducing ClpB's futile ATP-hydrolysis in the absence of HSP70 and protein substrates (*61, 62*). Similarly, in the cytosol and the endoplasmic reticulum of Eukaryotes, the basal disaggregase activity of JDP-HSP70 machinery is also synergized by the Hsp110 co-disaggregases, whose futile ATPase activity is dramatically reduced by its own protein-binding domain (*17*). Finally, the basal unfoldase activity of the JDP-HSP70-Hsp110 machinery in the cytosol of eukaryotes can also be synergized by the co-unfoldase activity of Hsp90, whose futile ATPase activity is regulated by the Hop and Aha1 co-chaperones (*63, 64*).

Material and Methods

Strains and plasmids

Wild-type *SSA1*, *SSE1*, *SIS1*, *YDJ1, and* swap chimeras were cloned in the pE-SUMO vector for propagation in *E. coli.* The swap chimeras were constructed by PCR by amplifying the 80 first amino acids (J-domain and a small part of the Glycine-Rich region) with the following primers: JYDJ1-F 5'-GCGAACAGATTGGAGGTATGGTTAAAGAA-3', JYDJ1-R 5'-GCCAGGACC ACCAGGAGCGCCACCAGCAC-3', JSIS1-F 5'-GCGAACAGATTGGAGGTATGGTCAAGGAG-3', JSIS1-R 5'-ACCTGGGAATCCGCCACCAAAGCTTGGACCAC-3') of the two JDPs and insert them inside the vector of the other one using these set of primers (YDJ1V-F 5'-GGCGGATTCCCAGGTGGTGGATTC-3', YDJ1V-R 5'-ACCTCCAATCTGTTCGCGGTGAGC-3', SIS1VF 5'-TGGTCCTGGTGGTCCTGGCGGTGC-3', SIS1VR 5'-CCAATCTGTTCGCGGTGAGCCTCA-3') by Gibson Assembly (NEB). All constructs were confirmed by sequencing.

Purification of proteins

For purification of the His10-SUMO tagged wild-type SSE1, SSA1, SIS1, YDJ1, and swap chimeras, were expressed and purified from *E. coli* BL21-CodonPlus (DE3)-RIPL cells with IPTG induction (final 0.5 mM for SSA1 and SSE1 and 0.2 for YDJ1, SIS1, and swap chimeras) at 18 °C, overnight. Briefly, cells were grown in LB medium + ampicillin at 37 °C to OD600 ~0.4-0.5. Protein expression was induced by the addition of 0.5 mM IPTG for 3 hours. Cells were harvested, washed with chilled PBS, and resuspended in buffer A (20 mM Tris-HCl pH 7.5, 150 mM KCl, 5% glycerol, 2mM DTT, 20mM MgCl2) containing 5 mM imidazole, 1mg/ml Lysozyme, 1mM PMSF for 1 h. Cells were lysed by sonication. After high-speed centrifugation (16000 rpm, 30 min/4°C), the supernatant was loaded onto a gravity flow-based Ni-NTA metal affinity column (2 ml beads, cOmplete His-Tag Purification Resin from Merck), equilibrated and washed with 10 column volumes of buffer A containing 5 mM imidazole. After several washes with high salt buffer A (+150 mM KCl, 20mM Imidazole and 5 mM ATP), N-terminal His10-SUMO (small ubiquitin-related modifier) Smt3 tag was cleaved with Ulp1 protease (2mg/ml, 300 µl, added to beads with buffer (20 mM Tris-HCl pH 7.5, 150mM KCl, 10mM MgCl2, 5% glycerol, 2mM DTT). Digestion of His10 Smt3 was performed on the Ni-NTA resin by, His6-Ulp1 protease. Because of dual His tags, His6- Ulp1 and His10-SUMO display a high affinity for Ni-NTA resin and remain bound to it during cleavage reaction. After overnight digestion at 4°C, the unbound fraction is collected (which contains only the native proteins). Proteins were further purified by concentrating to ~3mg/ml and applying to a size exclusion column (Superdex-200 increase, 10/30 GE Healthcare) equilibrated in buffer A containing 5 mM ATP. Pure fractions were pooled, concentrated by ultrafiltration using Amicon Ultra (Millipore), aliquoted, and stored at -80 °C. All protein concentrations were determined spectrophotometrically at 562 nm using BCA Protein Assay Kit− Reducing Agent Compatible (cat no. 23250).

The purified proteins were collected, concentrated, and stored at -80°C for further use.

Luciferase refolding assay

Luciferase activity was measured as described previously (*65, 66*). In the presence of oxygen, luciferase catalyzes the conversion of D-luciferin and ATP into oxyluciferin, CO2, AMP, PPi, and hv. Generated photons were counted with a Victor Light 1420 Luminescence Counter from Perkin-Elmer (Turku, Finland) in a 96-well microtiter plate format.

G6PDH refolding assay

Heat-pre-aggregated G6PDH was refolded by the SSA1 chaperone system as described previously for the DnaK chaperone system (*48*), with the following modifications; 500 nM heat-aggregated G6PDH (final concentration) was reactivated in the presence of 6 µM SSA1, incrementing (0–1 µM) JDPs, 0.75 µM SSE1

(the full SSA1 chaperone system) and 5 mM ATP. G6PDH activity was measured at different times of chaperone-mediated refolding reaction at 25°C.

ATPase assay (Malachite Green)

Colorimetric determination of Pi produced by ATP hydrolysis was performed using the Malachite Green Assay Kit (Sigma-Aldrich, Switzerland) and as described previously (*67*). Several concentrations of HSP70 (SSA1) and JDPs (YY, SS, SY, YS) were mixed with or without substrate (200 nM of pre-aggregated luciferase) and with 1 mM of ATP and incubated for 1 hour at 25°C. 4 µL of each sample was taken and put inside a 96-Well plate with 76 µL of H_2O. A 20-µl volume of Malachite Green reaction buffer was added and the samples were mixed thoroughly and incubated at 25°C for 30 min before measuring at 620nm on a plate reader (HIDEX-Sense 425-301, Finland). The rate of intrinsic ATP hydrolysis was deduced by subtracting the signal from ATP in the absence of a chaperone.

Light Scattering

To monitor the aggregation propensity of urea denatured Luciferase as described previously (*40*), 30 µM Luciferase was denatured with 4M urea at 25°C for 10 min, then diluted to a final concentration of 0.3 µM in buffer A (50 mM HEPES-KOH pH7.5, 150 mM KCl, 10 mM MgCl2, 2 mM DTT), immediately aggregation was monitored by light scattering at 340 nm at 30°C using Perkin Elmer Fluorescence Spectrophotometer. To monitor aggregation propensity of heat-denatured Luciferase or Luciferase (0.3 µM) was kept at 37°C in Buffer A and aggregation was monitored by light scattering at 340 nm for 30 min using Perkin Elmer Fluorescence Spectrophotometer.

FRET measurements and FRET efficiency calculation

All ensemble relative FRET efficiencies were calculated from maximum fluorescence emission intensities of the donor (ED) and acceptor (EA) fluorophore by exiting the donor only at 405 nm wavelength (*68, 69*). Fluorescence emission spectra analysis of MLucV reporter was performed on PerkinElmer LS55 fluorometer. Emission spectra were recorded from 480 to 580 nm wavelength with excitation slit 5 nm and emission slit 10nm. Average intensity values of spectral crosstalk were minimized by excitation donor at 405nm. Spectra were background-subtracted with spectra of buffer only or non-transformed cells in case of in-vivo measurements, samples acquired in the same conditions. The relative FRET efficiencies were calculated using the following equation:

$$FRET_{ensemble} = \frac{E_{acceptor}}{E_{donor} + E_{acceptor}}$$

Normalized FRET efficiencies relative to that of native MLucV were calculated as follows:

$$FRET_{norm} = \frac{FRET_{ensemble} - FRET_{separated}}{FRET_{native} - FRET_{separated}}$$

where $FRET_{ensemble}$ is the measured ensemble FRET efficiency, $FRET_{separated}$ is the calculated ensemble FRET measured in a solution of separated mTFP1 and Venus (0.33) and $FRET_{native}$ is the measured ensemble FRET of native MLucV (0.43). Unless otherwise specified, all ensemble FRET measurements were performed at 400 nM of MLucV. The temperature was maintained at 25°C unless otherwise specified. All experiments were performed in LRB (20mM Hepes-KOH pH7.4,150mM KCl, 10mM MgCl2) refolding buffer containing 5 mM ATP, 2mM DTT, unless otherwise specified. 4µM BSA was used in assays with chaperones to avoid MLucV species sticking to the vessel, it does not affect the fate of the formed aggregates, nor affects the activity of the chaperones. All experiments were repeated at least three times.

Limited proteolysis – Mass spectrometry (LiP-MS)

LiP-MS experiments were carried out on samples containing isolated JDPs at a concentration of 3 µM in LRB, using a procedure adapted from (46). For each JDP, four independent replicates were performed. Proteolysis with proteinase K (PK, Promega, cat. # V3021) was performed in thin-walled PCR microtubes, with a final reaction volume of 75 µL. PK was added at a 1:500 w/w ratio concerning the JDPs, corresponding to a final PK concentration of 240 ng/mL. Upon addition of PK, samples were incubated at 25°C for exactly 2 min, immediately followed by PK inactivation at 100°C for 5 min using a pre-heated PCR block. After cooling down to room temperature, sodium deoxycholate was added to a final concentration of 5% w/v (from a freshly prepared 10% w/w stock solution in 100 mM ammonium bicarbonate pH 8.5). Then reduction/alkylation of cysteine residues by adding 12 mM DTT and incubating at 37°C for 30 min, followed by addition of 40 mM iodoacetamide during 45 min at RT in the dark. Protein digestion was then performed by first incubating with 1:100 w/w LysC (Promega, cat. # VA1170) at 37°C for 4h with 800 rpm orbital agitation. Deoxycholic acid was then diluted to 1% w/w using 100 mM ammonium bicarbonate pH 8.5 and then 1:100 w/w trypsin (Promega, cat. # V5111) and incubated overnight (16h) at 37°C under agitation.

Peptide extraction

Digested peptides were separated from excess deoxycholic acid by liquid-phase extraction: two volumes of ethyl acetate (EtOAc) + 1% trifluoroacetic acid (TFA) were added to each sample, followed by vigorous mixing for 2 min and centrifugation at 5000g for 2 min. The bottom aqueous phase was transferred out and kept for peptide purification using 200 µL pipette tips loaded with strong cation-exchange resin

(Thermo Scientific, StageTips SCX). SCX tips were conditioned with 200 µL methanol and equilibrated with 200 µL SCX buffer A (2% acetonitrile, 1% formic acid in water). Samples were loaded from the back of the SCX tips and binding was carried out by slow centrifugation (1000g) until all samples passed through the resin. SCX tips were then washed twice with 200 µL EtOAc containing 0.5% TFA, then three times with SCX buffer A. Finally, peptides were eluted with 200 µL of freshly prepared SCX buffer B (80% acetonitrile, 19% water, 1% w/w ammonium hydroxide) and slow centrifugation (500g, 4min). Eluates were passed once more through the SCX tips and then dried using a speed vac centrifuge (organic solvent evaporation settings). Dried peptides were stored at -20°C until analysis by LC-ESI-MS/MS.

Mass spectrometry analysis

Dried peptide samples were resuspended in 100 µL aqueous 2% acetonitrile containing 0.05% TFA. After loading onto a trapping microcolumn (Acclaim PepMap100 C18, 20 mm x 100 µm ID, 5 µm, Dionex), peptides were separated on a custom-packed nanocolumn (75 µm ID × 40 cm, 1.8 µm particles, Reprosil Pur, Dr. Maisch), with a flow rate of 250 nL/min and a gradient from 4% to 76% acetonitrile in water + 0.1% formic acid, over 140 min. Eluted peptides were analyzed on an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) operated in data-dependent mode, controlled by Xcalibur software (version 3.0.63). Full survey scans were performed at a 120000 resolution and a top-speed precursor selection strategy was applied to maximize the acquisition of peptide tandem MS spectra with a maximum cycle time of 3s. HCD fragmentation model was used at a normalized collision energy of 32%, with a precursor isolation window of 1.6 m/z and MS/MS spectra were acquired in the ion trap. Peptides selected for MS/MS were excluded from further fragmentation during the 60s. Raw data were analyzed using MaxQuant version 1.6.10.43 (Andromeda search engine) with a protein database consisting of the *E. coli* (strain K12) proteome (UniProt proteome ID: UP000000625), to which the sequences of the four yeast HSP40 variants (YY, SS, YS, SY) and proteinase K were added, as well as common contaminants. Reversed sequences were used as decoys. To identify semitryptic peptides, a semi-specific peptide search was conducted using trypsin cut settings, minimal and maximal peptide lengths of 5 and 30 residues, respectively, and a maximum peptide mass of 6000 Da. The LFQ intensities of identified peptides were used as an abundance measure and were used to determine PK accessibility changes between matching domains of the yeast HSP40 swaps (i.e., the J-domains of YY vs. YS, the J-domains of SS vs SY, the CTD of YY vs SY and the CTD of SS vs YS) using custom Matlab scripts, based on procedures described in (*44*).

Acknowledgments

Funding:

Swiss National Fund grant 31003A_175453

Israel Science Foundation grant 90 3150/2

Author contributions:

Conceptualization: PG, PDLR, MER

Methodology: PG, PDLR, MER, BF, ST

Investigation: PG, PDLR, MER, BF, ST

Visualization: PG, PDLR, MER, BF, ST

Supervision: PG, PDLR

Writing—original draft: PG, PDLR, MER, BF

Writing—review & editing: PG, PDLR, MER, BF, ST

Data and materials availability:

All data are available in the main text or supplementary materials.

Competing interests:

The authors declare that they have no competing interests.

Reference

1. B. Fauvet, M. E. Rebeaud, S. Tiwari, P. De Los Rios, P. Goloubinoff, Repair or Degrade: the Thermodynamic Dilemma of Cellular Protein Quality-Control. *Frontiers in molecular biosciences* **8**, (2021).
2. V. N. Uversky, A. V. Finkelstein, Life in Phases: Intra- and Inter- Molecular Phase Transitions in Protein Solutions. *Biomolecules* **9**, 842 (2019).
3. C. B. Anfinsen, Principles that govern the folding of protein chains. *Science* **181**, 223-230 (1973).
4. M. E. Rebeaud, S. Mallik, P. Goloubinoff, D. S. Tawfik, On the evolution of chaperones and cochaperones and the expansion of proteomes across the Tree of Life. *Proc Natl Acad Sci U S A* **118**, (2021).
5. P. Goloubinoff, Mechanisms of protein homeostasis in health, aging and disease. *Swiss Med Wkly* **146**, w14306 (2016).
6. P. Goloubinoff, A. S. Sassi, B. Fauvet, A. Barducci, P. De Los Rios, Chaperones convert the energy from ATP into the nonequilibrium stabilization of native proteins. *Nat Chem Biol* **14**, 388-395 (2018).
7. A. Finka, R. U. Mattoo, P. Goloubinoff, Experimental Milestones in the Discovery of Molecular Chaperones as Polypeptide Unfolding Enzymes. *Annu Rev Biochem* **85**, 715-742 (2016).
8. R. Zhang, D. Malinverni, D. M. Cyr, P. L. Rios, N. B. Nillegoda, J-domain protein chaperone circuits in proteostasis and disease. *Trends Cell Biol*, (2022).
9. R. Sousa *et al.*, Clathrin-coat disassembly illuminates the mechanisms of Hsp70 force generation. *Nat Struct Mol Biol* **23**, 821-829 (2016).
10. A. Guihur, M. E. Rebeaud, P. Goloubinoff, How do plants feel the heat and survive? *Trends Biochem Sci*, (2022).
11. P. Goloubinoff, P. De Los Rios, The mechanism of Hsp70 chaperones: (entropic) pulling the models together. *Trends Biochem Sci* **32**, 372-380 (2007).
12. P. De Los Rios, P. Goloubinoff, Hsp70 chaperones use ATP to remodel native protein oligomers and stable aggregates by entropic pulling. *Nat Struct Mol Biol* **23**, 766-769 (2016).
13. H. H. Kampinga *et al.*, Function, evolution, and structure of J-domain proteins. *Cell Stress Chaperones* **24**, 7-15 (2019).
14. R. Kityk, J. Kopp, M. P. Mayer, Molecular Mechanism of J-Domain-Triggered ATP Hydrolysis by Hsp70 Chaperones. *Mol Cell* **69**, 227-237 e224 (2018).
15. H. Raviol, H. Sadlish, F. Rodriguez, M. P. Mayer, B. Bukau, Chaperone network in the yeast cytosol: Hsp110 is revealed as an Hsp70 nucleotide exchange factor. *EMBO J* **25**, 2510-2518 (2006).
16. P. G. Needham, D. C. Masison, Prion-impairing mutations in Hsp70 chaperone Ssa1: effects on ATPase and chaperone activities. *Arch Biochem Biophys* **478**, 167-174 (2008).
17. V. Kumar *et al.*, Interdomain communication suppressing high intrinsic ATPase activity of Sse1 is essential for its co-disaggregase activity with Ssa1. *FEBS J* **287**, 671-694 (2020).
18. R. Russell, A. W. Karzai, A. F. Mehl, R. McMacken, DnaJ Dramatically Stimulates ATP Hydrolysis by DnaK: Insight into Targeting of Hsp70 Proteins to Polypeptide Substrates. *Biochemistry* **38**, 4165-4176 (1999).

19. H. H. Kampinga, E. A. Craig, The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature reviews. Molecular cell biology* **11**, 579-592 (2010).
20. P. De Los Rios, A. Barducci, Hsp70 chaperones are non-equilibrium machines that achieve ultra-affinity by energy consumption. *Elife* **3**, e02218 (2014).
21. N. B. Nillegoda *et al.*, Crucial HSP70 co-chaperone complex unlocks metazoan protein disaggregation. *Nature* **524**, 247-251 (2015).
22. M. P. Mayer, The Hsp70-Chaperone Machines in Bacteria. *Frontiers in Molecular Biosciences* **8**, (2021).
23. C. Andreasson, J. Fiaux, H. Rampelt, M. P. Mayer, B. Bukau, Hsp110 is a nucleotide-activated exchange factor for Hsp70. *J Biol Chem* **283**, 8877-8884 (2008).
24. S. K. Sharma, P. De Los Rios, P. Goloubinoff, Probing the different chaperone activities of the bacterial HSP70-HSP40 system using a thermolabile luciferase substrate. *Proteins* **79**, 1991-1998 (2011).
25. O. Faust *et al.*, HSP40 proteins use class-specific regulation to drive HSP70 functional diversity. *Nature* **587**, 489-494 (2020).
26. C. Y. Fan, S. Lee, H. Y. Ren, D. M. Cyr, Exchangeable chaperone modules contribute to specification of type I and type II Hsp40 cellular function. *Mol Biol Cell* **15**, 761-773 (2004).
27. D. Malinverni, S. Zamuner, A. Barducci, N. B. Nillegoda, P. De Los Rios, Data-driven large-scale genomic analysis reveals an intricate phylogenetic and functional landscape in J-domain proteins. *bioRxiv*, 2022.2002.2028.482344 (2022).
28. W. Yan, E. A. Craig, The glycine-phenylalanine-rich region determines the specificity of the yeast Hsp40 Sis1. *Mol Cell Biol* **19**, 7751-7758 (1999).
29. J. C. Silva, J. C. Borges, D. M. Cyr, C. H. Ramos, I. L. Torriani, Central domain deletions affect the SAXS solution structure and function of yeast Hsp40 proteins Sis1 and Ydj1. *BMC Struct Biol* **11**, 40 (2011).
30. B. A. Schilke *et al.*, Broadening the functionality of a J-protein/Hsp70 molecular chaperone system. *PLoS Genet* **13**, e1007084 (2017).
31. H. Y. Yu *et al.*, Roles of intramolecular and intermolecular interactions in functional regulation of the Hsp70 J-protein co-chaperone Sis1. *J Mol Biol* **427**, 1632-1643 (2015).
32. M. Reidy *et al.*, Hsp40s specify functions of Hsp104 and Hsp90 protein chaperone machines. *PLoS Genet* **10**, e1004720 (2014).
33. J. C. Borges *et al.*, Identification of regions involved in substrate binding and dimer stabilization within the central domains of yeast Hsp40 Sis1. *PLoS One* **7**, e50927 (2012).
34. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
35. J. Shorter, The mammalian disaggregase machinery: Hsp110 synergizes with Hsp70 and Hsp40 to catalyze protein disaggregation and reactivation in a cell-free system. *PLoS One* **6**, e26319 (2011).
36. K. Linke, T. Wolfram, J. Bussemer, U. Jakob, The roles of the two zinc binding sites in DnaJ. *J Biol Chem* **278**, 44457-44466 (2003).
37. H. Wyszkowski *et al.*, Class-specific interactions between Sis1 J-domain protein and Hsp70 chaperone potentiate disaggregation of misfolded proteins. *Proceedings of the National Academy of Sciences of the United States of America* **118**, e2108163118 (2021).
38. A. Finka, S. K. Sharma, P. Goloubinoff, Multi-layered molecular mechanisms of polypeptide holding, unfolding and disaggregation by HSP70/HSP110 chaperones. *Front Mol Biosci* **2**, 29 (2015).
39. D. W. Summers, P. M. Douglas, C. H. I. Ramos, D. M. Cyr, Polypeptide transfer from Hsp40 to Hsp70 molecular chaperones. *Trends in biochemical sciences* **34**, 230-233 (2009).
40. S. Tiwari, B. Fauvet, S. Assenza, P. De los Rios, P. Goloubinoff, A novel fluorescent multi-domain protein construct reveals the individual steps of the unfoldase action of Hsp70. *bioRxiv*, 2022.2002.2017.480908 (2022).
41. J. Perales-Calvo, D. Giganti, G. Stirnemann, S. Garcia-Manyes, The force-dependent mechanism of DnaK-mediated mechanical folding. *Science Advances* **4**, eaaq0243 (2018).
42. A. M. R. de Graff, D. E. Mosedale, T. Sharp, K. A. Dill, D. J. Grainger, Proteostasis is adaptive: Balancing chaperone holdases against foldases. *PLOS Computational Biology* **16**, e1008460 (2020).
43. J. M. Nunes, M. Mayer-Hartl, F. U. Hartl, D. J. Müller, Action of the Hsp70 chaperone system observed with single proteins. *Nature Communications* **6**, 6307 (2015).
44. S. Schopper *et al.*, Measuring protein structural changes on a proteome-wide scale using limited proteolysis-coupled mass spectrometry. *Nat Protoc* **12**, 2391-2410 (2017).
45. P. Leuenberger *et al.*, Cell-wide analysis of protein thermal unfolding reveals determinants of thermostability. *Science* **355**, (2017).
46. P. Gainza *et al.*, Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning. *Nature methods* **17**, 184-192 (2020).
47. Y. Jiang, P. Rossi, C. G. Kalodimos, Structural basis for client recognition and activity of Hsp40 chaperones. *Science* **365**, 1313-1319 (2019).
48. R. U. Mattoo *et al.*, Synergism between a foldase and an unfoldase: reciprocal dependence between the thioredoxin-like activity of DnaJ and the polypeptide-unfolding activity of DnaK. *Front Mol Biosci* **1**, 7 (2014).
49. B. Fauvet *et al.*, Bacterial Hsp90 Facilitates the Degradation of Aggregation-Prone Hsp70-Hsp40 Substrates. *Front Mol Biosci* **8**, 653073 (2021).
50. T. Ali Azam, A. Iwata, A. Nishimura, S. Ueda, A. Ishihama, Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *Journal of bacteriology* **181**, 6361-6370 (1999).
51. B. Misselwitz, O. Staeck, T. A. Rapoport, J proteins catalytically activate Hsp70 molecules to trap a wide range of peptide sequences. *Mol Cell* **2**, 593-603 (1998).
52. P. Lopez-Buesa, C. Pfund, E. A. Craig, The biochemical properties of the ATPase activity of a 70-kDa heat shock protein (Hsp70) are governed by the C-terminal domains. *Proc Natl Acad Sci U S A* **95**, 15253-15258 (1998).
53. S. Sun, D. Gresham, Cellular quiescence in budding yeast. *Yeast* **38**, 12-29 (2021).
54. R. J. Mackenzie *et al.*, Absolute protein quantification of the yeast chaperome under conditions of heat shock. *Proteomics* **16**, 2128-2140 (2016).

55. P. Brownridge *et al.*, Quantitative analysis of chaperone network throughput in budding yeast. *Proteomics* **13**, 1276-1291 (2013).
56. S. Ghaemmaghami *et al.*, Global analysis of protein expression in yeast. *Nature* **425**, 737-741 (2003).
57. C. Lawless *et al.*, Direct and Absolute Quantification of over 1800 Yeast Proteins via Selected Reaction Monitoring. *Mol Cell Proteomics* **15**, 1309-1322 (2016).
58. M. Uchida *et al.*, Quantitative analysis of yeast internal architecture using soft X-ray tomography. *Yeast* **28**, 227-236 (2011).
59. A. K. Bryan, A. Goranov, A. Amon, S. R. Manalis, Measurement of mass, density, and volume during the cell cycle of yeast. *Proc Natl Acad Sci U S A* **107**, 999-1004 (2010).
60. A. Finka, V. Sood, M. Quadroni, L. Rios Pde, P. Goloubinoff, Quantitative proteomics of heat-treated human cells show an across-the-board mild depletion of housekeeping proteins to massively accumulate few HSPs. *Cell Stress Chaperones* **20**, 605-620 (2015).
61. P. Goloubinoff, A. Mogk, A. P. Zvi, T. Tomoyasu, B. Bukau, Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. *Proc Natl Acad Sci U S A* **96**, 13732-13737 (1999).
62. H. Mazal *et al.*, Tunable microsecond dynamics of an allosteric switch regulate the activity of a AAA+ disaggregation machine. *Nat Commun* **10**, 1438 (2019).
63. T. Moran Luengo, R. Kityk, M. P. Mayer, S. G. D. Rudiger, Hsp90 Breaks the Deadlock of the Hsp70 Chaperone System. *Mol Cell* **70**, 545-552 e549 (2018).
64. A. Lott, J. Oroz, M. Zweckstetter, Molecular basis of the interaction of Hsp90 with its co-chaperone Hop. *Protein Science* **29**, 2422-2432 (2020).
65. P. Bischofberger, W. Han, B. Feifel, H.-J. Schönfeld, P. Christen, d-Peptides as Inhibitors of the DnaK/DnaJ/GrpE Chaperone System*. *Journal of Biological Chemistry* **278**, 19044-19047 (2003).
66. S. K. Sharma, P. De los Rios, P. Christen, A. Lustig, P. Goloubinoff, The kinetic parameters and energy cost of the Hsp70 chaperone as a polypeptide unfoldase. *Nat Chem Biol* **6**, 914-920 (2010).
67. S. Lee *et al.*, Cryo-EM Structures of the Hsp104 Protein Disaggregase Captured in the ATP Conformation. *Cell reports* **26**, 29-36.e23 (2019).
68. R. D. Fritz *et al.*, A Versatile Toolkit to Produce Sensitive FRET Biosensors to Visualize Signaling in Time and Space. *Science Signaling* **6**, rs12-rs12 (2013).
69. R. J. Wood *et al.*, A biosensor-based framework to measure latent proteostasis capacity. *Nature Communications* **9**, 287 (2018).

CHAPTER 5

How do plants feel the heat and survive?

How do plants feel the heat and survive?

Anthony Guihur, Mathieu E. Rebeaud and Pierre Goloubinoff

I contributed to this work conceptually, in writing the manuscript and preparing the figures.

Trends in Biochemical Sciences

CellPress OPEN ACCESS

Review

How do plants feel the heat and survive?

Anthony Guihur,[1,*] Mathieu E. Rebeaud,[1] and Pierre Goloubinoff[1,2,*]

Climate change is increasingly affecting the quality of life of organisms on Earth. More frequent, extreme, and lengthy heat waves are contributing to the sixth mass extinction of complex life forms in the Earth's history. From an anthropocentric point of view, global warming is a major threat to human health because it also compromises crop yields and food security. Thus, achieving agricultural productivity under climate change calls for closer examination of the molecular mechanisms of heat-stress resistance in model and crop plants. This requires a better understanding of the mechanisms by which plant cells can sense rising temperatures and establish effective molecular defenses, such as molecular chaperones and thermoprotective metabolites, as reviewed here, to survive extreme diurnal variations in temperature and seasonal heat waves.

Plant heat sensing and signaling to build up effective molecular defenses

All organisms on Earth are facing rapid climate change. Indeed, the Intergovernmental Panel on Climatic Change estimates a 0.3°C rise in global mean temperature per decade [1], and a greater and more frequent occurrence of heat waves is expected to compromise human food security [2,3]. This alarming situation is calling for crop breeders to better understand how some wild plants have been able to adapt to harsh environments by ameliorating existing molecular strategies to protect their heat-labile macromolecules during **noxious heat stresses** (HS; see Glossary). HS can disrupt ecosystems, especially when combined with other stresses, such as drought [4]. Excessive heat can alter the structure and function of thermo-labile macromolecular ensembles and accelerate water loss, resulting in heat-aggregated proteins deprived of their dedicated biological activities. Misfolded proteins trigger the production of **reactive oxygen species** (ROS), which can damage lipids in membranes, cause DNA mutations, and result in leaf shedding, apoptosis, and, ultimately, plant death [5]. Here, we review recent advances in our understanding of heat-sensing and heat stress-related responses leading to the onset of molecular defenses against heat damage in land plants. Data from omics approaches are discussed, highlighting the central role of heat-induced proteins, among which **heat shock transcription factors** (HSFs), heat-signaling proteins, heat-induced metabolic enzymes producing thermoprotective and ROS-scavenging metabolites, and heat-induced chaperones, commonly but confusingly called **heat shock proteins** (HSPs), which can prevent and revert protein aggregation. In addition, we discuss the central role of plant heat-sensory calcium channels that can progressively respond to increasing temperatures by way of reacting to changes in the fluidity of the plasma membrane [6].

At sunrise, most land plants face a rapid increase in ambient temperature of up to 20°C over the following 12 h [7]. This can occur more suddenly when winds chase sun-veiling clouds. Thus, a seemingly harmless increase in the ambient temperature at dawn induces an optimal heat signal, leading to the synthesis of metabolic enzymes, which, by noon, can result in the production of dozens of millimolars of thermoprotective metabolites and tens of micromolars of protective HSP chaperones. Metabolites, such as carotenoids and glutathione, can mitigate harmful heat-induced ROS, whereas others, such as proline, glycine betaine, and trehalose, can protect

Highlights

As the sun rises, the temperature rapidly increases and, by noon, heat may damage labile macromolecular complexes and impair the vital biological functions of plants.

Plants have a heat shock response (HSR), which is activated via fluidity changes in the plasma membrane and heat-responsive cyclic nucleotide-gated ion channels (CNGCs), which use Ca^{2+} and reactive oxygen species (ROS) as messengers to mediate a signaling pathway, leading to the upregulation of heat-induced mRNA in minutes, and to the accumulation of protective heat shock proteins (HSPs) and metabolites in hours.

While some studies have been conducted on plant molecular chaperones, their precise role in acquired thermotolerance and the identity of their thermolabile protein substrates remain unknown.

Based on 'omics analyses, plant HSR is a multigenic trait and new thermoresistant crops should harbor complementary mechanisms combining HSP chaperones and enzymes producing thermo- and ROS-protecting metabolites.

[1]Department of Plant Molecular Biology, Faculty of Biology and Medicine, University of Lausanne, CH-1015 Lausanne, Switzerland
[2]School of Plant Sciences and Food Security, Tel-Aviv University, Tel Aviv 69978, Israel

*Correspondence:
anthony.guihur@gmail.com (A. Guihur) and Pierre.Goloubinoff@unil.ch (P. Goloubinoff).

heat-labile proteins and membranes. The massive accumulation of HSP20s by mid-morning may prevent protein misfolding and aggregation occurring by noon, whereas, in the afternoon, heat-accumulated HSP60s, HSP70s, HSP90s, and HSP100s can resolubilize protein aggregates into biologically active proteins [8].

The design of new crops that can withstand more frequent and acute heat waves likely necessitates combined activities of several heat-induced enzymes to produce thermoprotective metabolites, and of chaperones to protect labile proteins during, and repair heat-damaged proteins after, stress exposure [9]. The breeder's challenge is immense, given the expected trade-off between crop yields and the increased expression of expectedly costly molecular defenses.

Heat sensing and signaling

HS is a transient increase in ambient temperature beyond the optimal plant growth temperature, with deleterious effects on the plant physiology. The severity of heat damage is determined by the rate of temperature increase, the intensity of the HS, its duration, and by the presence of other stresses [10,11]. At the molecular level, higher temperatures can damage fragile complexes, mostly proteins and membranes, whereas polysaccharides may be more resistant. Although heat-denatured RNA and DNA may spontaneously revert to their functional native state after heat stress, heat-induced DNA methylations and demethylations may have long-lasting transgenerational phenotypic and epigenetic consequences [12]. It is generally thought, albeit poorly demonstrated experimentally, that severe HS in plants causes the transient unfolding of heat-labile proteins, leading to their aggregation into insoluble inactive species. In metazoans, protein aggregates can be cytotoxic and clog protein quality control machineries, such as **molecular chaperones** and the proteasome [13,14].

By translating an early moderate temperature increment into an effective heat-priming signal, leading to the onset of new molecular defenses (Figure 1, blue), plant cells can prevent heat damage during exposure to a subsequent noxious HS and repair damaged proteins and membranes once the stress has passed [15]. All organisms, including plants, express several conserved families of core chaperones: ATP-independent HSP20s, four ATP-dependent unfoldases, HSP60s, HSP70s, the co-disaggregases HSP100s, and HSP90s. Together with a plethora of co-chaperones, these control the quality of protein structures and cellular functions [8].

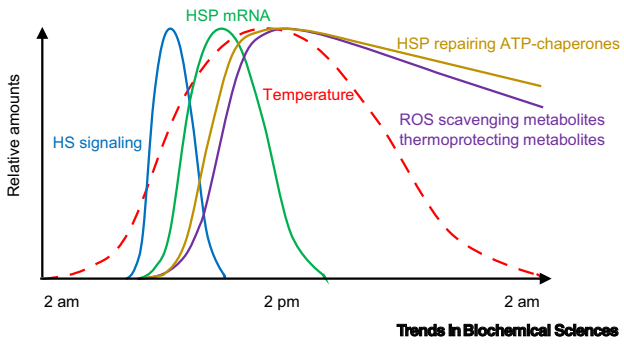


Figure 1. Plant sensing and response to heat stress during a 24-h cycle. Early in the morning of a typical hot summer's day, land plants must sense and evaluate the need for molecular defenses by mid-day to prevent heat damage and repair it by evening. The activity of lipid desaturases at different basal ambient temperatures may set different activation thresholds for the plant heat shock response. By dawn, plants need to detect mild, harmless increments in the ambient temperature in anticipation of an upcoming damaging heat stress by noon (red dashed line). Within minutes, plants need to emit a transient heat shock signal (blue) to produce, within a couple of hours, heat shock protein (HSP) mRNA (green). During the morning, HSP transcription factors and HSP chaperones accumulate (gold) to prevent and repair heat-damaged proteins and produce HSP enzymes accumulating thermoprotecting and reactive oxygen species (ROS)-scavenging metabolites (purple), which can last several days. This thermomemory predisposes plants to better withstand ensuing heat stresses over the following days.

Glossary

Acquired thermotolerance (AT): ability of a plant to accumulate HSPs and thermo- and ROS-protective metabolites, in response to a mild and harmless, prior warming, conferring the ability to survive an upcoming severe harmful HS for a few hours.
Calmodulins (CaMs): calcium-binding proteins that typically bind other proteins, such as cyclic nucleotide gated channels. Upon binding of entrant periplasmic Ca^{+2} ions, CaMs change their conformation and send a specific cellular signal to produce the HSR.
Heat shock proteins (HSPs): proteins that massively accumulate in response to mild HS. Many HSPs are molecular chaperones. Others are heat shock signaling and transcription factors and enzymes that produce thermo- and ROS-protective metabolites. Other HSPs have unknown functions.
Heat shock response (HSR): homeostatic transcriptional program highly conserved in all organisms, where, in response to a mild heat shock, thermoprotective HSPs and metabolites are massively produced, conferring cells acquired thermotolerance.
Heat shock transcription factors (HSFs): family of transcription factors that bind specific sequences in the promoter regions of HSP genes, the concomitant derepression and activation of which lead to the accumulation of HSPs and confer plant AT.
Molecular chaperones: proteins that control the quality of the structure and function of other proteins. Some but not all chaperones are HSPs. Most use ATP to unfold heat-aggregated proteins to be repaired into functional native proteins.
Noxious heat stress (HS): harmful HS causing irreversible damage to the structure and function of thermolabile macromolecules, such as membranes and protein complexes, possibly leading to apoptosis and plant death.
Reactive oxygen species (ROS): highly reactive chemicals or radicals, formed from O_2, such as H_2O_2, ozone, singlet oxygen, superoxide, and hydroxyl radical, that can chemically damage essential lipids, membranes, nucleotides and proteins.
Targetases: substoichiometric co-chaperones that can catalytically recruit an excess of HSP70s onto specific (mostly misfolded) polypeptides

The plant heat shock response

Within minutes of a mild heat-priming treatment (several degrees above ambient temperature without reaching noxious HS) [16,17], ~1% of plant genes, including some encoding chaperones, are massively transcribed (Figure 1, green). The ability of plants to respond to abrupt changes in conditions was recently reviewed by Kollist *et al.* [18]. In the case of HS, genes encoding molecular chaperones, such as HSP101, HSA32, some HSP70s and HSP90s, possibly acting as repressors of HSFA1 activity at ground temperatures [19], and many cytosolic HSP20s, possibly blocking heat-induced apoptosis, are overexpressed and accumulate predominantly in the cytosol [16,20]. Within hours, various HSPs involved in HS defenses accumulate (Figure 1, gold). In animal and moss cells, 100 or so newly heat-accumulated HSPs represent a net 2% mass gain in total cellular proteomes at the expense of an across-the-board net 2% protein mass loss of thousands of proteins, mostly with house-keeping functions, which are outcompeted on the ribosomes by the abundant new HSP mRNAs [16,21].

that need to be structurally and functionally modified by the chaperone into differently active proteins.

When a promoter is tuned by evolution to constantly create just enough new mRNA to replace a degraded polypeptide with a new one, terminally differentiated cells do not necessitate sensors and signaling pathways to maintain a constant cellular level of that protein. By contrast, to express HSPs only on a need-to-defend basis, cells may require thermosensors connected to a specific signaling pathway. Both in animals and plants, the transcription of HSP genes at non-HS temperatures is tightly repressed by bound HSP70-HSP90 chaperones, which maintain HS transcription factors (i.e., HSFA1) inactive in the cytosol, and by histones enwrapping HSP genes in the nucleus [22] (Figure 2, Key figure).

Upon mild warming (priming), thermosensors must first send a signal to derepress transcription of HSP genes by dissociating the inhibitory HSP70-HSP90 chaperones from HSFA1. In addition, a specific heat-induced Ca^{2+}-entry signal from the plasma membrane must activate a kinase, which in turn not only activates the chaperone-liberated HSFA1 to transcribe new HSP genes [23] (Figure 2, Box 1), but also acts on the chromatin remodeling machinery to evict the bound histones from HSP genes [24]. Priming is a required step in **acquired thermotolerance** (AT), which involves molecular modifications that may be maintained longer than the initial heat priming, resulting in thermomemory lasting several hours to a few days. Hence, before this memory vanishes, plants can 'learn' to respond more readily to a new HS (Figure 1). Most of the heat-induced mRNA produced during the first hour of HS is degraded in less than 24 h [21,25]. By contrast, the degradation of heat-induced chaperones, transcription factors, and thermoprotective metabolites may take a few days [12]. Thermomemory can also involve chromatin remodeling, which can last up to a week [12,26]. Experimental data suggest that transgenerational heat stress memory can be inherited via an HSFA2-activated H3K27me3 demethylase [27,28]. Consequently, on subsequent hot days, plants are better prepared to withstand a noxious HS [12].

From the plasma membrane to the production of thermoprotective compounds

During the first minutes of an HS, the phosphoinositide-specific phospholipases PLC3 and PLC9 are rapidly activated, with the concurrent accumulation of cytosolic Ca^{2+} [29,30] (Figure 2). PLC3 and PLC9 hydrolyze phosphatidylinositol 4,5-bisphosphate into diacylglycerol (DAG) and inositol-1,4,5-trisphosphate (IP_3) [30], respectively, which was suggested to activate Ca^{2+} release from intracellular stores [29] and trigger a signaling cascade involving **calmodulins** (CaMs) and kinases [31]. Whereas a chloroplast-specific calcium signal was identified in response to heat, without evidence for an increase in cytosolic calcium [32], ratiometric Ca^{2+} reporter-based analyses provided strong evidence that HS induces cytosolic Ca^{2+} signals in plant leaves [33]. Moreover, electrophysiology showed specific heat-induced transient entry of external Ca^{2+} into the cytosol of protoplasts [34].

Key figure

Heat shock sensing, signaling, and responses in plant

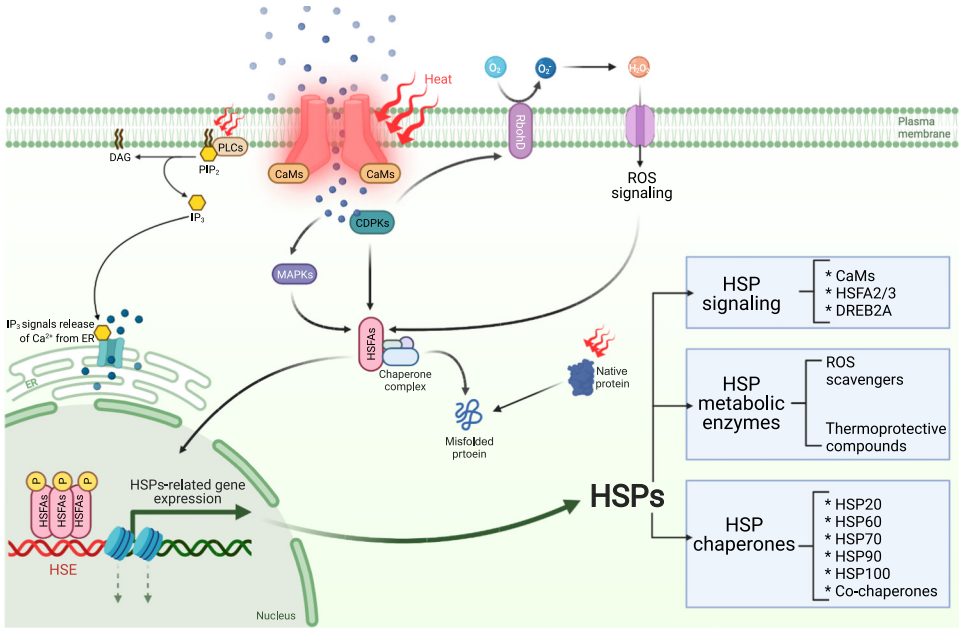


Figure 2. The plant heat shock response (HSR) is initiated by an increase in the fluidity of the plasma membrane (PM) and the possible activation of phospholipases (PLCs) hydrolyzing phosphatidylinositol 4,5-bisphosphate (PIP_2) into inositol-1,4,5-trisphosphate (IP_3) and diacylglycerol (DAG), leading to the controlled entry of Ca^{2+} from the endoplasmic reticulum (ER) and activation of calmodulins (CaMs) and the heat shock signaling pathway. The cyclic nucleotide-gated ion channels (CNGCs) embedded in the PM (red) respond to the temperature-increased fluidity of the PM and mediate the controlled entry of periplasmic Ca^{2+} into the cytosol. Ca^{2+} binding to CNGC-bound CaMs initiates a specific signaling cascade that activates kinases, which then phosphorylate and activate heat shock transcription factor A (HSFA). Reactive oxygen species (ROS) production by NADPH oxidase RbohD at the PM may also lead to co-activation of some heat shock factors (HSFs). The Ca^{2+} entry-dependent heat shock signal phosphorylates HSFA1 and evicts the bound HSP70–HSP90, thereby activating HSFA1 to translocate to the nucleus and bind heat shock elements (HSEs). This leads to the eviction of bound histones from HSP genes and the recruitment of RNA polymerase. The consequent massive synthesis, in minutes, of HSP mRNA and the accumulation, in hours, of heat shock signaling proteins, enzymes producing thermo- and ROS-protective metabolites and of molecular chaperones, ultimately confers acquired thermotolerance to the plant cell, which can last for days. Abbreviation: MAPK, mitogen-activated protein kinase.

Land plants contain ~20 different cyclic nucleotide-gated ion channels (CNGCs) [6,34,35]. In particular, CNGC2/4, which, similar to their distant animal relatives, the heat- and nociceptive TRPV1 channels [36], form homo- and heterotetrameric transmembrane ion channels, with their cytosolic parts interacting with CaMs and cyclic nucleotides (Box 2). In animals and plants, the **heat shock response** (HSR) depends on the degree of fluidity of the plasma membrane in which the thermosensory channels are embedded [37,38]. At low temperatures, CNGC2/4 channels are closed and poised to readily respond to a heat-induced increase in the fluidity of the surrounding plasma membrane. Under HS, the channels quickly open, allowing periplasmic Ca^{2+} to enter and bind CaMs associated with the cytosolic C-terminal domain of the CNGCs [39] (Figure 2). Consequently, kinases phosphorylate HSFs, which translocate to the nucleus and

Box 1. Feel the heat (sensing)

The conductor of an orchestra must first produce a gesture to signal musicians when to execute a musical movement. Similarly, plant cells must sense a mild temperature rise to signal to the protein synthesis machinery when to execute a HSR and build up effective molecular defenses in anticipation of damage from an upcoming noxious HS (see Figure 1 in the main text). Whereas sending a signal from cellular thermosensors to the nucleus may take seconds, and mRNA synthesis minutes, the accumulation of HSPs and thermo- and ROS-protective metabolites, takes hours (see Figure 1 in the main text). Thus, to induce an effective HSR, land plants need thermosensors to detect mild warming, which is unlikely to cause protein aggregation, and yet enables them to respond fully to HS.

The accepted model is that as-yet unidentified thermolabile proteins must first respond to warming upon undergoing thermal unfolding and recruiting HSP70 and HSP90 chaperones from inactive HSF1, thereby initiating a specific HS signal [31,47]. Indeed, an 'unfolded protein response' (UPR) has been identified in the ER and cytosol, suggesting that thermolabile proteins act as cellular thermosensors [113]. A putative heat-responsive retrograde pathway was also reported in chloroplasts, in which the photosynthetic apparatus is a primary target for heat damage. HS in the chloroplast triggers a signal leading to the transcription of HSPs in the nucleus [114]. The dissociation of histone variant H2A.Z at high temperatures from HSP genes has also been suggested to serve as a direct thermosensory mechanism in the nucleus, inducing the HSR in higher plants [50]. Yet, because H2A.Z is associated with many promoters of genes that are not activated by heat, this histone is unlikely to serve as the primary heat sensor of plant cells. It rather stands at the very end of the HS signaling pathway, receiving eviction orders from heat-activated histone-remodeling complexes, such as ARP6 [50], which in turn receive orders from the plasma membrane via a calcium entry-dependent signal [34,35,42].

Moreover, light and temperature signals are intertwined. The photoreceptor phytochrome B and transcription factors called 'phytochrome interaction factors' (PIFs) change conformation and activity under warming [115,116]. When the activity of phytochrome B is reduced by far-red light or by mildly elevated temperatures, PIFs accumulate and promote hypocotyl elongation. Indicative of a crosstalk between light signaling, the circadian clock, and temperature signaling [53,117], plants in the dark may be less responsive to HS than in the light, possibly due to thermosensitive photosensors [25,118,119].

bind conserved *cis*-elements called heat shock elements (HSE) in the promoters of HSP genes [40,41] (Figure 2).

HSFs are classified into three classes with different functions: HSFA, B, and C [42]. HSFA1A is a master regulator of plant AT that triggers the HSR through the induction of HSFA1b and the overexpression of new HSFA2; HSFA2 thus becomes a major heat stress transcription factor amplifying HSPA1a's activatory effect, leading to the HSR and AT [43,44] and, together with HSFA3, is prolonging thermomemory, [45]. HSFB1 acts as a co-regulator enhancing the activity of both HSFA1A and HSFA2 [40,42,46]. The so-called 'titration model' for HSF activation was proposed based on the observation that, at non-HS temperatures, hypophosphorylated HSFAs are maintained inactive in the cytosol by bound HSP70 and HSP90 [47] (Figure 2). Under HS, HSFAs are

Box 2. The role of CNGCs

Growing evidence points at the plasma membrane as a central thermosensing component in land plants [31,120]. Hence, moss plants constantly grown at 22°C, with plasma membranes naturally enriched with unsaturated lipids, produced more HSPs in response to 1 h of HS at 38°C than when constantly grown at 28°C with plasma membranes enriched in saturated lipids [37]. This, together with evidence that a membrane fluidizer can artificially decrease the threshold temperatures at which a strong HSR is produced, suggests that plant thermosensors do not directly respond to a given elevated temperature: it is the basal temperature of growth that dictates the ratio between fluidizing unsaturated fatty acids and rigidifying saturated fatty acids. Thus, plant thermosensors embedded in the plasma membrane can react to a specific net gain of membrane fluidity, enacted by a given ratio of saturated and unsaturated lipids set by the ground temperature of growth [37].

Further demonstrating the key role of the plasma membrane in thermosensing, depletion or chelation of external Ca^{+2} ions completely, albeit reversibly, blocked the heat-induced HSR in both plants and human cells [48], pointing at members of the plasma membrane-embedded CNGCs, as potential membrane fluidity-responsive protein thermosensors in plants. Indeed, disruption of *CNGC2* or *CNGC4* resulted in a hyperthermosensitive phenotype with a HSR occurring at heat-priming temperatures lower by ~4°C. The hyperthermosensitive mutants unnecessarily accumulated HSPs at nonstressful temperatures and grew extremely slowly, but were less in need of heat priming to effectively resist a noxious HS, indicating their increased thermotolerance [34].

found to dissociate from the chaperones and become functional trimers in the nucleus. Although evidence for protein aggregation in the cytosol of plant cells under mild heat-priming temperatures is lacking, it is generally believed that chaperone binding to these heat-aggregated proteins causes chaperone dissociation from the inactive HSFA1 [40]. Noticeably, during HS, when the entry of external Ca^{2+} is prevented by EGTA, the plant HSR does not occur [35]. This implies that, without a Ca^{2+}-entry dependent signal from the plasma membrane, chaperone titration by putative heat-labile protein aggregates does not suffice to elicit full activation of HSFA1. Similarly, blockage by a specific inhibitor, capsazepine, of Ca^{2+} entry under HS through the TRPV1 thermosensory channel in the plasma membrane of vertebrates does not suffice to activate the accumulation of HSPs [48]. By analogy, releasing the handbrake is not enough to start moving a car uphill: one must concomitantly press on the gas.

Upon binding to HSE, the activated phosphorylated HSFAs may also need to instruct the chromatin remodeling complexes to evict bound histones that would otherwise repress the transcription of HSP genes [49,50]. Alongside HSFs, other transcription factors control HSP gene expression, such as dehydration-responsive element-binding protein (DREB 2A), some NAC family transcription factors [51], multiprotein-bridging factor 1c (MBF1c), and some members of the Ethylene-responsive transcription factor (ERF) family [52]. Moreover, the time of the day impacts the translation of HSP transcripts in response to HS [25] and an HSF-independent pathway involving circadian-regulated genes has been described [53]. Thus, evolution may have favored that, for the same extent of HS, a land plant may need to produce a costly optimal HSR at noon, while the high cost of a full HSR may be spared at night, when there is less light and fewer ROS-generating stresses. Whereas an excess of ROS has detrimental effects on plant cells, low ROS can mediate stress responses and affect plant development [54]. In addition to Ca^{2+}, H_2O_2 levels at the plasma membrane also increase quickly in response to severe HS. Suggesting a crosstalk between the ROS and Ca^{2+} signaling cascade, this process, which is catalyzed by the NADPH oxidase RbohD, is activated by Ca^{2+} binding to EF-hand motives [55,56] (Figure 2). H_2O_2 was also shown to activate particular HSFs [57] (Figure 2).

Surviving the heat: thermoprotective metabolite production

Abiotic stresses have a wealth of cumulative adverse effects on plant fitness and survival. Plants can accumulate organic compounds in the tens of millimolar range with thermo- and ROS-protective properties that can mitigate some of the heat-damaging effects [58]. Hydrophobicity is central to the maintaining of membranes and proteins in the native state. However, as the temperature increases, hydrophobicity is not strong enough to compensate for the low entropy of the single native state. Whereas membranes may undergo hyperfluidization [59], proteins may transiently unfold and readily acquire different, more compact misfolded and aggregated conformations that are more entropic than the native state, while still satisfying the requirement of sequestering most of their hydrophobic parts [60].

From a thermodynamic point of view, many protective metabolites, often called osmolytes, can lower the osmotic potential, thereby acting as thermoprotectants of native proteins and membranes. Thermoprotective metabolites can be amino acids (proline), polyamines, quaternary ammonium compounds (glycine-betaine), sugars (trehalose), polyols, sugar alcohols (mannitol, galactinol), or tertiary sulfonium compounds [61]. Amino acids, such as tyrosine, valine, proline, tryptophan, and glutamine, accumulate during plant responses to various abiotic stresses, including HS [62]. Amino acids likely have an essential role in the regulation of osmotic adjustment to keep plant turgor pressure under heat-accelerated evaporation. However, various metabolites do not equally protect plants from HS damage. Whereas proline accumulates in some drought-stressed plants, it can inhibit growth of *Arabidopsis* seedlings upon HS [63]. Under a combination

of drought and HS, proline is replaced by sucrose [64]. Glycine-betaine is known to stabilize macromolecular structures in response to dehydration and HS and protect the cytoplasm and photosystem II in chloroplasts from heat and ion toxicity [65]. It was suggested that, during HS, increased levels of trehalose, *myo*-inositol, and galactinol, which are precursors of oligosaccharides, stabilized membranes by interacting with the phosphates of the phospholipids [66]. Moreover, HSFs also control the increased levels of essential metabolites, such as galactinol and its derivatives, with potential thermoprotective effects [67]. The accumulation and production of thermoprotective metabolites is regulated by several pathways [68], including ROS-mediated mitogen-activated protein kinases (MAPKs), salt overly sensitive (SOS), abscisic acid and calcium signaling, via CaMs.

The term 'chemical chaperones' frequently used to describe these metabolites is misleading. These thermoprotective metabolites act by stabilizing labile macromolecular structures, such as native proteins and membranes, under otherwise heat-denaturing conditions. By contrast, HSP60, HSP70, and HSP100 are sophisticated ATP-fueled molecular machines with strong affinities for already-formed misfolded and aggregated proteins. Thus, chaperones may 'repair' proteins that have been heat damaged into their native, functional, state. By accumulating protective metabolites typically during the morning of a summer's day, heat-primed plants may stabilize thermolabile macromolecules in their native state, despite the denaturing temperatures at noon (Figure 1, purple). Priming leads to a transcriptional upshift and subsequent accumulation of heat-induced proteins, including enzymes that accumulate thermo- and ROS-protecting metabolites, allowing plants to withstand repeated HS [69], and ATPase chaperones to repair heat-damaged proteins during and after HS.

Surviving the heat: heat-accumulated molecular chaperones

The term 'molecular chaperones' was reinstated by John Ellis [70] to describe proteins assisting the native (re)folding and assembly of various proteins complexes, without being part of the final assembled oligomers. Experimental data have since shown that many chaperones act as polypeptide unfolding catalysts, and are major components of the cellular protein homeostasis network [13], involved in controlling both physiological processes and repairing stress-damaged proteins during and following HS. Noticeably, the general designation of chaperone proteins as being HSPs is misleading; although about one-third of chaperones and co-chaperone genes are overexpressed under HS, the remaining two-thirds are not. Nevertheless, chaperone and co-chaperone genes are ~15 times more likely to be heat inducible compared with other gene categories, confirming that heat-induced chaperones are key to the prevention and repair of structural heat damage in labile proteins [16,21]. Chaperones are found in all cellular compartments (Table 1). Whereas protein crowding is generally thought to aggravate protein aggregation, heat-induced plant chaperones mostly accumulate, counter-intuitively, in the cytosol, where protein crowding is significantly lower than in the endoplasmic reticulum (ER), mitochondria and chloroplasts [16] (Figure 3A). In general, viable single T-DNA insertion lines have no detectable HS or AT phenotypes, likely because of the high redundancy of orthologous genes in the various chaperone families (except HSP101) [71].

Interestingly, HS generates stress granules [72–74] containing mRNAs, elongation initiation factors, and several chaperones ,such as HSP20s and HSP101. HSP20s are unable to actively promote the solubilization of already-formed stable aggregates, but can prevent protein aggregation. HSP101 is a co-chaperone that increases the stand-alone disaggregase activity of HSP70s and mediates protein hydrolysis by the 26S proteasome [75,76]. In heat-stressed cells, HSP chaperone levels can persist from hours to several days and, therefore, are central to the onset of plant AT [12,28] (Figure 1, gold).

Table 1. Chaperone families and their representative members in four plant species

Class	Representative members				Subcellular localization	Refs
	Arabidopsis thaliana	*Physcomitrium patens*	*Oryza sativa*	*Populus trichocarpa*		
HSP20s	18	12	23	29		[77]
Subfamily						
I	6	9	9	16	Cytosol	
II	2	2	2	1	Cytosol	
III	1		1	1	Cytosol	
IV	1		1	1	Cytosol	
V	1		1	2	Cytosol	
VI	1			2	Cytosol	
ER	1		2	1	ER	
Mitochondria/plastids	5	1	7	5	Mitochondria/plastids	
HSP60s	17	14	20	28		[82,121]
Subfamily						
Group 1: CPN60	9	6	11	10	Mitochondria/plastids	
Group 2: CCTs	8	8	9	18	Cytosol	
HSP70-110	18	17	32	20		[82,121,122]
Subfamily						
Hsp/Hsc70	11	11	14	13	Cytosol/Mitochondria, plastids	
Bip	3	2	6	4	ER	
Hsp110	4	4	8	3	Cytosol/ER	
HSP90	7	6	9	10	Cytosol/Mitochondria, plastids/ER	[82,100,121]
HSP100-ClpB	4	4	5	5	Cytosol/mitochondria, plastids	[82,121]
HSP40-JDPs	99	100	95	163	Cytosol/mitochondria, plastids	[92]
Subfamily						
Class A	8	14	13	11		
Class B	9	9	7	11		
Class C	82	77	75	141		

HSP20s: the most heat-responsive plant chaperone

HSP20s are a family of ancient ubiquitous proteins with a conserved α-crystallin domain suggested to bind denatured proteins. They are divided into about a dozen subclasses (Table 1) (Figure 3A,B), with diverse N-terminal domains responsible for their assembly into oligomers of 12 or more subunits [77]. It is not known what the state of the HSP20-bound polypeptides is in the cell: unfolded, misfolded, aggregated, or native. Following the suggestion that HSP20s bind misfolded species, they were shown to collaborate *in vitro* with the ATPase unfoldases HSP70-40 and HSP60, and also indirectly with HSP100 in the refolding of heat-predenatured proteins [78].

In plants, HSP20s are the most heat responsive of all chaperone classes (see Outstanding questions). Remarkably, the basal expression of HSP20s at non-HS temperatures is generally null, implying a strong repressive mechanism [15,16] (Figure 3B) and suggesting that their overexpression has a high cost to the plant fitness. Heat-accumulated HSP20s account for ~30% of the overall net mass gain of heat-induced proteins. Although being the most dramatically heat-induced chaperones in the cytosol of higher plants, current knowledge is limited on the

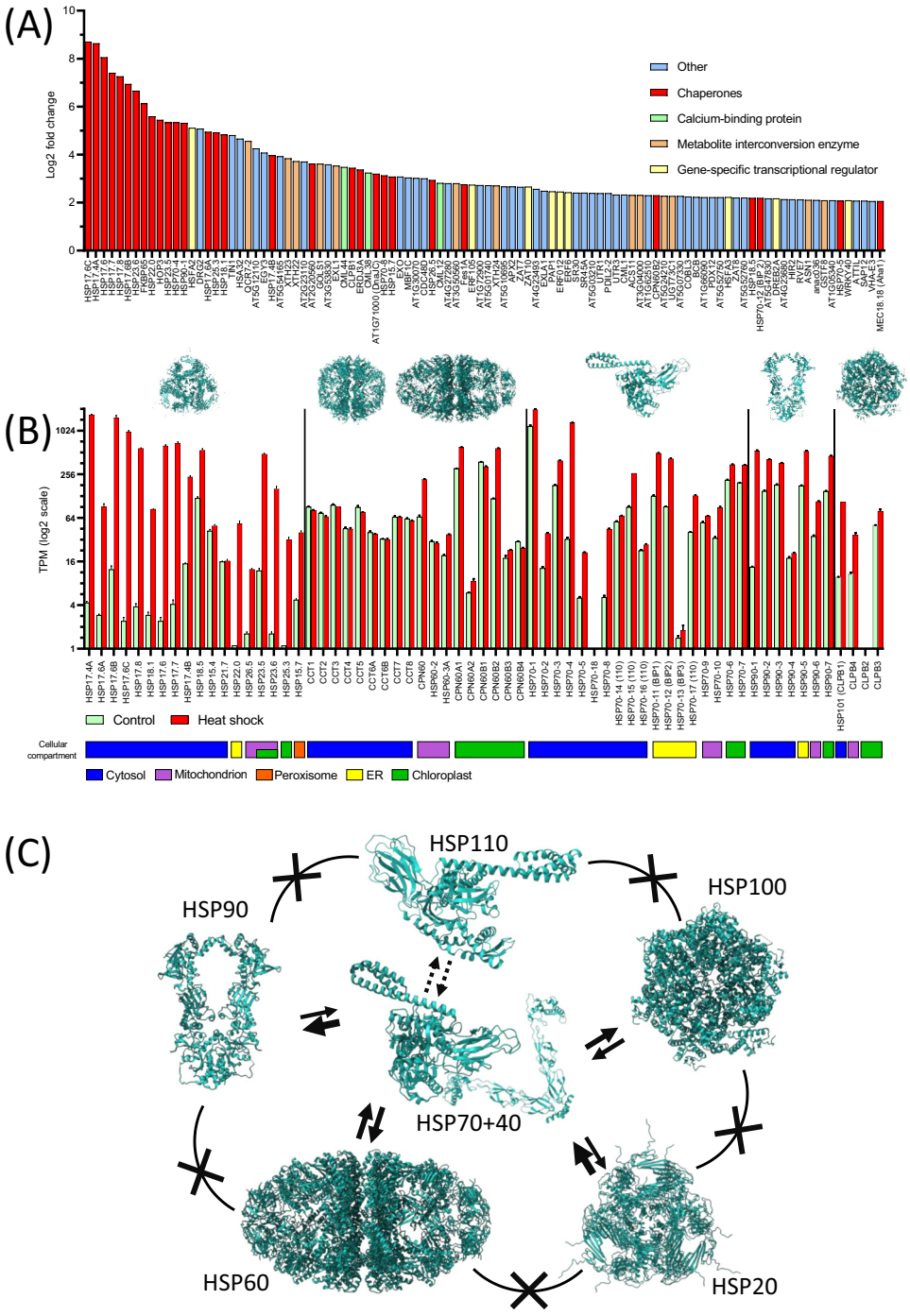


Figure 3. RNA sequencing (RNA-seq) of untreated and mild heat-treated *Arabidopsis thaliana* seedlings and the chaperone collaboration network. (A) Log_2 fold-change of the 100 most heat-induced genes in 10-day-old *A. thaliana* seedlings, treated for 90 min at a priming temperature of 33°C. Chaperones genes are in red, calcium-binding genes are in green, metabolic genes are in orange, transcriptional regulators are in yellow and others are in blue. (B) mRNA transcripts per million (TPM) of mRNAs of the core-chaperone families in untreated (green) and warmed (33°C) (red) *A. thaliana* seedlings. From left to right: HSP20s, HSP60s, HSP70s, HSP90s, and HSP100s. Insert above the

(Figure legend continued at the bottom of the next page.)

mechanism by which HSP20s effectively contribute to plant AT. During HS, HSP20s have been shown to stabilize lipid bilayers and possibly protect membranes from hyperfluidization [79,80]. In addition to their ability to passively prevent aggregation, HSP20s may also carry out specific physiological functions, such as repressing heat-induced apoptosis [16]. The plastid metalloprotease FtsH6 and HSP21 jointly regulate thermomemory in *Arabidopsis* [81]. Noticeably, higher plants are the only eukaryotes that express a HSP20 in the ER [82] (Figure 3A,B).

HSP60s: cage unfolding chaperonins

Together with HSP20s, HSP60s [83] belong to an ancient chaperone family that was already present in the last common ancestor to all organisms [82]. Group 1 HSP60s (CPN60 and GroEL) are found in bacteria, chloroplasts, and mitochondria, and were initially found to mediate ribulose-1,5-bisphosphate carboxylase oxygenase assembly in chloroplasts [70]. Group 2 HSP60s (CCT/TRiC) are found in archaea and the cytosol of eukaryotes, including plants. Remarkably, CCTs, which predominantly fold actin and tubulin, are poorly heat induced in differentiated plant cells [10] (Figure 3B). Yet, enhanced expression of Group 2 HSP60s in root stem cell may maintain proteome integrity and suppress protein aggregation [14].

HSP70s and HSP40s: the central hub of the chaperone network

HSP70s [84] are one of the most highly conserved class of chaperones, accounting for ~1% of the total protein mass of unstressed eukaryotic cells [85]. HSP70s act as ATP-fueled polypeptide-unfolding enzymes that coordinate the activity of the other families of chaperones in the cellular proteostasis network [82] (Figure 3C). They control protein homeostasis in all ATP-containing compartments of eukaryotic cells, both under physiological and HS conditions. Some HSP70s are constitutively expressed, whereas others, such as HSP70-4, are strongly heat upregulated [86] (Figure 3B). Overexpression of heat-inducible HSP70-1 improves plant thermotolerance, whereas its reduced expression is lethal [87]. HSP70s can use the energy of ATP hydrolysis to apply a pulling and unfolding force that remodels high-affinity (alter)natively folded or stress-misfolded proteins into differently folded native proteins, each with a different structure and carrying a specific biological activity [88,89]. To perform their protein-remodeling action, HSP70s rely on J-domain co-chaperones, JDPs (HSP40s and DNAJs), acting as specific obligate HSP70-**targetases**. JDPs, which are about ten times less abundant than HSP70s, act as catalysts that upload misfolded or alternatively folded protein substrates onto the HSP70 machineries to be structurally modified into low-affinity native products [89–91].

Plants are eukaryotes with the largest number of JDP-encoding genes [92], which are divided in three classes [90]. Classes A and B are conserved JDPs families, also called HSP40s, which preferentially act as 'generalists' bringing misfolded protein substrates onto HSP70s for unfolding [93]. By contrast, class C JDPs are diverse with different conserved domains, mostly with

columns are representative structures of members from each chaperone family [Protein Data Bank (PDB): 3J07, 7LUM, 4PKN, 4B9Q, 2IOP and 5VJH, respectively]. Below, members of each chaperone family grouped according to the cellular compartments in which they occur; cytosol (blue), mitochondria (purple), chloroplasts (green), endoplasmic reticulum (yellow), and peroxisome (orange). (C) Plant chaperone network organization. HSP60s use ATP hydrolysis to unfold misfolded polypeptides. HSP20s bind misfolded polypeptides and prevent their aggregation ('holdase'). HSP70-JDPs use ATP hydrolysis to unfold misfolded polypeptides and solubilize proteins aggregates, which can be partially prevented from forming by HSP20s. Polypeptides that are incompletely unfolded by HSP70-JDPs can be further unfolded by HSP60s or HSP90s. HSP100s act as HSP70-dependant co-disaggregases and HSP90s act as HSP70-dependant co-foldases. HSP110s act as additional HSP70-dependant co-disaggregases. Whereas HSP100s, HSP20s, HSP110s HSP90s, and HSP60s do not directly interact, they closely collaborate and exchange unfolding/misfolding intermediates through HSP70-JDPs, which act as integrative hubs for the entire chaperone network. Arrow directions and widths indicate the preferred flow of protein un/folding intermediates between the different members of the network. RNA-seq data adapted from [123]; adapted from [82] (C).

unknown functions [94]. The few known class C JDPs serve as specialized co-chaperones to target unique, alternatively folded protein complexes onto HSP70, for example to import polypeptides into organelles, [95,96]. Less than a dozen out of 80 bioinformatically identified class C JDPs in the genome of *Arabidopsis* have biological functions assigned. The variety of JDPs in plants likely serves to recruit HSP70s onto proteins that control various plant-specific physiological and stress-related processes. Several plant class A and B JDPs are stress upregulated, indicating a role in protein quality control during and following HS. Deletion of AtADJA1 and A2 (orthologs of yeast Ydj1) in *Arabidopsis thaliana* showed impaired thermotolerance in seedlings [97].

HSP90s: mysterious ATP-consuming co-chaperones of HSP70s

HSP90s [98] are a family of abundant chaperones in prokaryotes and in the cytosol, ER, and organelles of eukaryotes [85], accounting for 1% of total cellular protein. Many HSP90s are induced by various stresses, including HS (Figure 3A,B). HSP90s appear to systematically act in tandem with HSP70s (Figure 3C), particularly in the cytosol, where their ATPase cycle is regulated by several conserved co-chaperones [99]. In *A. thaliana*, seven genes encoding HSP90s are expressed in all ATP- and HSP70-containing cellular compartments (Table 1). Overexpression of HSP90s was shown to confer plant resistance to several biotic and abiotic stresses, such as heavy metal, oxidative, and salt stresses and pathogen infection [100]. Bound HSP90s are thought to repress HSFA1 at ambient temperature, while, under HS, they dissociate, with HSFA1 becoming concomitantly active in the transcription of HSP genes [101]. Remarkably, HSP90 inhibitors can cause abnormal morphological phenotypes in seedlings and affect organ growth, supporting a role of HSP90s in buffering protein evolution [82] and in plant development. Unlike HSP70-JDPs, which show stand-alone autonomous unfolding/refolding chaperone activity *in vitro* and in primitive bacteria naturally lacking Hsp90, the reverse is not true, suggesting that HSP90s act as downstream co-chaperones of HSP70s, facilitating the folding of HSP70 substrates that failed to properly fold by themselves [102]. It is unclear how, in close collaboration with HSP70s, energy from ATP hydrolysis is harnessed by HSP90s to remodel the structure of heat-damaged proteins. Thus, there is a vital need to identify the most heat-labile proteins, the heat denaturation of which would be limiting for plant growth and compromise survival when ineffectively repaired by the combined action of HSP90s and HSP70s.

HSP100s and HSP110s: co-disaggregases of HSP70s

The HSP70-JDP chaperone machinery is, by itself, able to target stable preformed protein aggregates and use ATP hydrolysis to unfold, solubilize, and reactivate them back into native proteins [103]. HSP100s [104] are specific co-disaggregases of HSP70s belonging to the AAA+ ATPases superfamily. They form hexameric cylinders in the central cavity from which, once activated by HSP70s, misfolded loops protruding from aggregates are forcefully stretched [105]. The genome of *A. thaliana* encodes four classes of HSP100, which are constitutively expressed at low temperature and may accumulate under HS [106]. In *A. thaliana*, the cytosolic Hot1 HSP101 mutant is more sensitive to HS compared with wild-type, and HSP101 is required for basal thermotolerance, acquired thermotolerance [107–109] and the memory of heat acclimation [110]. Besides their role during HS, HSP100s may perform housekeeping functions in plant growth and chloroplast development.

HSP110s are eukaryote-specific co-disaggregases of HSP70s. Sequence-wise, HSP110s share a common ancestor with HSP70s. Generally, HSP110s are ten times less abundant compared with HSP70s, and are obligate ATP-dependent catalysts that accelerate the exchange of ADP into ATP, thereby ameliorating the basal intrinsic disaggregase activity of cytosolic and ER HSP70s. Similar to fungi, plants co-express both types of HSP70-codisaggregase, HSP101 and HSP110, in their cytosol, suggesting that plants have a powerful HSP70-centrered disaggregation network that is able to process a wide array of protein aggregates. In contrast to aging metazoans that lack

HSP100 co-disaggregases, there is no evidence for the formation of toxic aggregates and amyloid fibrils *in planta*, despite predictions of prions in their proteomes [111]. It has been reported that plant stem cells have an enhanced ability to prevent protein misfolding and aggregation under stress conditions, compared with their differentiated counterparts [14]. It is not clear whether plant cells chose to undergo programed cell death before accumulating toxic aggregates. Possibly owing to their sessile life style, higher plants may have evolved proteomes particularly poor in aggregation-prone proteins compared with metazoans [82].

Concluding remarks

Land plants have evolved powerful but costly protection and repair mechanisms to counteract heat damage, mostly to heat-labile proteins and membranes. A plant cell can be compared to a country, the security of which would necessitate the expensive maintenance, even in times of peace, of a large army in anticipation of rare but possible external aggressions. Plant cells have adopted a cheaper but more risky strategy: to maintain a small core of highly specialized molecular defenses, such as chaperones and enzymes to produce metabolites, that, in the case of an upcoming HS, will benefit within a few hours from the back-up of a large pool of freshly recruited 'militia' in the form of heat-accumulated HSPs, produced in response to a preceding milder heat-priming signal. This strategy demands a sophisticated 'intelligence service' that can identify early warnings and transform a mild harmless warming into a signal to produce high levels of costly defensive HSPs. It also requires that, without HS, the expression of HSP genes be tightly repressed, although they remain poised to be readily de-repressed in response to a signal from the heat sensors in the plasma membrane.

The dramatic increase in atmospheric temperatures due to global warming has become a major concern for human food security. To produce more thermotolerant crops, breeders need detailed understanding of the molecular mechanisms leading to the onset of plant thermotolerance. Some loss-of-function mutants in model plants and crops are more sensitive to various stresses, whereas the overexpression of metabolic enzymes and HSPs result in slightly more heat-resistant crops, albeit mostly under laboratory conditions. Recent transcriptomic, proteomic, and metabolomic studies revealed that temperature tolerance in plants is a polygenic trait. It includes not only HSP chaperones, generally suggested to prevent the heat-induced aggregation of ill-defined thermolabile proteins, but also metabolic enzymes, which can produce ROS-quenching and thermoprotective metabolites, as well as heat-induced signaling proteins and transcription factors. Moreover, plants adapt to warmer temperatures by changing their physiological and morphological characteristics. They may choose to reorient or bluntly shed their leaves and, for plants that can afford it, cause local cooling by increased water evaporation. Many desert plants evade scorching temperatures by choosing to be semi-dry quiescent embryos within dehydrated heat-resistant seeds, deeply buried in the soil, waiting to germinate in a year with a cooler and rainier winter.

Thus, researchers aiming to maintain agricultural productivity in a world undergoing climate change need to investigate the molecular mechanisms that promote stress resilience in model plant species and in crops. Researchers have recently benefited from rapidly accumulating data from various 'omics approaches and from genome editing with CRISPR/Cas9, allowing the engineering of combined new traits in crops [9,112]. Yet, the major challenge remains to produce heat-resistant crops by expressing costly energy-consuming defenses, while maintaining high yields that can still feed the ever-growing human population of the planet.

Outstanding questions

How do HSP20 chaperones prevent the aggregation of heat-labile proteins? What is the state of HSP20-bound polypeptides: unfolded, misfolded, or small aggregates?

Why is the expression of HSP20s so tightly inhibited at low temperatures? Are HSP20s toxic to unstressed plants?

Higher plants are the only eukaryotes that have recently evolved an ER-located HSP20. What is the particular role of this HSP20 in plant thermotolerance ?

Whereas HSP20s can act as passive 'holding chaperones' that prevent protein aggregation during the average lifetime of the lens in a human eye, to what extent do plant HSP20s contribute, mass-wise, to the effective prevention of heat-induced proteins aggregation in plant cells?

What is the mechanism that uses the energy of ATP hydrolysis to enable HSP90 to modify the structure of polypeptides?

Which are the most thermolabile proteins that misfold and aggregate during a heat shock, and are limiting plant growth and survival, and that are repaired by HSP chaperones?

Acknowledgments

We apologize to those whose work was not cited due to space and reference number limitations. This work was funded by grant 31003A_175453 from the Swiss National Science Foundation and grant 90 3150/2 from the Israel Science Foundation to P.G. and a Swiss National Science Foundation grant n°CRSK-3_196689 to A.G. Figure 2 was created with BioRender (biorender.com).

Declaration of interests

The authors have no conflict of interest.

References

1. Masson-Delmotte, V. *et al.* (2019) *Climate Change and Land: An IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems*. Intergovernmental Panel on Climate Change
2. Rodziewicz, P. *et al.* (2013) Influence of abiotic stresses on plant proteome and metabolome changes. *Acta Physiol. Plant.* 36, 1–19
3. Janni, M. *et al.* (2020) Molecular and genetic bases of heat stress responses in crop plants and breeding for increased resilience and productivity. *J. Exp. Bot.* 71, 3780–3802
4. He, M. *et al.* (2018) Abiotic stresses: general defenses of land plants and chances for engineering multistress tolerance. *Front. Plant Sci.* 9, 1771
5. Purschke, M. *et al.* (2010) Thermal injury causes DNA damage and lethality in unheated surrounding cells: active thermal bystander effect. *J. Invest. Dermatol.* 130, 86–92
6. Jarratt-Barnham, E. *et al.* (2021) The complex story of plant cyclic nucleotide-gated channels. *Int. J. Mol. Sci.* 22, 874
7. Dong, B. *et al.* (2016) Understanding the rapid summer warming and changes in temperature extremes since the mid-1990s over Western Europe. *Clim. Dyn.* 48, 1537–1554
8. Finka, A. *et al.* (2016) Experimental milestones in the discovery of molecular chaperones as polypeptide unfolding enzymes. *Annu. Rev. Biochem.* 85, 715–742
9. Rivero, R.M. *et al.* (2021) Developing climate-resilient crops: improving plant tolerance to stress combination. *Plant J.* 109, 373–389
10. Zandalinas, S.I. *et al.* (2020) Systemic signaling during abiotic stress combination in plants. *Proc. Natl. Acad. Sci. U. S. A.* 117, 13810–13820
11. Teskey, R. *et al.* (2015) Responses of tree species to heat waves and extreme heat events. *Plant Cell Environ.* 38, 1699–1712
12. Balazadeh, S. (2021) A 'hot' cocktail: the multiple layers of thermomemory in plants. *Curr. Opin. Plant Biol.* 65, 102147
13. Fauvet, B. *et al.* (2021) Repair or degrade: the thermodynamic dilemma of cellular protein quality-control. *Front. Mol. Biosci.* 8, 768888
14. Llamas, E. *et al.* (2021) The intrinsic chaperone network of *Arabidopsis* stem cells confers protection against proteotoxic stress. *Aging Cell* 20, e13446
15. Larkindale, J. and Vierling, E. (2008) Core genome responses involved in acclimation to high temperature. *Plant Physiol.* 146, 748–761
16. Guihur, A. *et al.* (2020) Quantitative proteomic analysis to capture the role of heat-accumulated proteins in moss plant acquired thermotolerance. *Plant Cell Environ.* 44, 2117–2133
17. Liu, B. *et al.* (2021) Heat priming improved heat tolerance of photosynthesis, enhanced terpenoid and benzenoid emission and phenolics accumulation in *Achillea millefolium*. *Plant Cell Environ.* 44, 2365–2385
18. Kollist, H. *et al.* (2019) Rapid responses to abiotic stress: priming the landscape for the signal transduction network. *Trends Plant Sci.* 24, 25–37
19. Hahn, A. *et al.* (2011) Crosstalk between Hsp90 and Hsp70 chaperones and heat stress transcription factors in tomato. *Plant Cell* 23, 741–755
20. Oyoshi, K. *et al.* (2020) Memory of 5-min heat stress in *Arabidopsis thaliana*. *Plant Signal. Behav.* 15, 1778919
21. Finka, A. *et al.* (2015) Quantitative proteomics of heat-treated human cells show an across-the-board mild depletion of housekeeping proteins to massively accumulate few HSPs. *Cell Stress Chaperones* 20, 605–620
22. Andrasi, N. *et al.* (2021) Diversity of plant heat shock factors: regulation, interactions, and functions. *J. Exp. Bot.* 72, 1558–1575
23. Praat, M. *et al.* (2021) Protein kinase and phosphatase control of plant temperature responses. *J. Exp. Bot.* 72, 7459–7473
24. Cortijo, S. *et al.* (2017) Transcriptional regulation of the ambient temperature response by H2A.Z nucleosomes and HSF1 transcription factors in *Arabidopsis*. *Mol. Plant* 10, 1258–1273
25. Bonnot, T. and Nagel, D.H. (2021) Time of the day prioritizes the pool of translating mRNAs in response to heat stress. *Plant Cell* 33, 2164–2182
26. Liu, H.-C. *et al.* (2018) Distinct heat shock factors and chromatin modifications mediate the organ-autonomous transcriptional memory of heat stress. *Plant J.* 95, 401–413
27. Liu, J. *et al.* (2019) An H3K27me3 demethylase-HSFA2 regulatory loop orchestrates transgenerational thermomemory in *Arabidopsis*. *Cell Res.* 29, 379–390
28. Yamaguchi, N. *et al.* (2021) H3K27me3 demethylases alter HSP22 and HSP17.6C expression in response to recurring heat in *Arabidopsis*. *Nat. Commun.* 12, 3480
29. Gao, K. *et al.* (2014) *Arabidopsis thaliana* phosphoinositide-specific phospholipase C isoform 3 (AtPLC3) and AtPLC9 have an additive effect on thermotolerance. *Plant Cell Physiol.* 55, 1873–1883
30. Ren, H. *et al.* (2017) The role of AtPLC3 and AtPLC9 in thermotolerance in *Arabidopsis*. *Plant Signal. Behav.* 12, e1162368
31. Bourgine, B. and Guihur, A. (2021) Heat shock signaling in land plants: from plasma membrane sensing to the transcription of small heat shock proteins. *Front. Plant Sci.* 12, 710801
32. Lenzoni, G. and Knight, M.R. (2018) Increases in absolute temperature stimulate free calcium concentration elevations in the chloroplast. *Plant Cell Physiol.* 60, 538–548
33. Weigand, C. *et al.* (2021) A ratiometric calcium reporter CGf reveals calcium dynamics both in the single cell and whole plant levels under heat stress. *Front. Plant Sci.* 12, 777975
34. Finka, A. *et al.* (2012) Plasma membrane cyclic nucleotide gated calcium channels control land plant thermal sensing and acquired thermotolerance. *Plant Cell* 24, 3333–3348
35. Saidi, Y. *et al.* (2009) The heat shock response in moss plants is regulated by specific calcium-permeable channels in the plasma membrane. *Plant Cell* 21, 2829–2843
36. Zhang, F. *et al.* (2018) Heat activation is intrinsic to the pore domain of TRPV1. *Proc. Natl. Acad. Sci. U. S. A.* 115, E317–E324
37. Saidi, Y. *et al.* (2010) Membrane lipid composition affects plant heat sensing and modulates Ca(2+)-dependent heat shock response. *Plant Signal. Behav.* 5, 1530–1533
38. Guihur, A. *et al.* (2022) How do humans and plants feel the heat? *Trends Plant Sci.* Published online March 28, 2022. https://doi.org/10.1016/j.tplants.2022.03.006
39. Niu, W.-T. *et al.* (2019) *Arabidopsis* cyclic nucleotide-gated channel 6 is negatively modulated by multiple calmodulin isoforms during heat shock. *J. Exp. Bot.* 71, 90–104
40. Guo, M. *et al.* (2016) The plant heat stress transcription factors (HSFs): structure, regulation, and function in response to abiotic stresses. *Front. Plant Sci.* 7, 114
41. Ohama, N. *et al.* (2016) The transcriptional cascade in the heat stress response of *Arabidopsis* Is strictly regulated at the level of transcription factor expression. *Plant Cell* 28, 181–201
42. Ohama, N. *et al.* (2017) Transcriptional regulatory network of plant heat stress response. *Trends Plant Sci.* 22, 53–65
43. Yoshida, T. *et al.* (2011) *Arabidopsis* HsfA1 transcription factors function as the main positive regulators in heat shock-responsive gene expression. *Mol. Gen. Genomics.* 286, 321–332
44. Liu, H.C. and Charng, Y.Y. (2013) Common and distinct functions of *Arabidopsis* class A1 and A2 heat shock factors in diverse abiotic stress responses and development. *Plant Physiol.* 163, 276–290
45. Friedrich, T. *et al.* (2021) Heteromeric HSFA2/HSFA3 complexes drive transcriptional memory after heat stress in *Arabidopsis*. *Nat. Commun.* 12, 3426
46. Fragkostefanakis, S. *et al.* (2019) The repressor and co-activator HsfB1 regulates the major heat stress transcription factors in tomato. *Plant Cell Environ.* 42, 874–890

47. Masser, A.E. *et al.* (2020) Hsf1 on a leash - controlling the heat shock response by chaperone titration. *Exp. Cell Res.* 396, 112246
48. Bromberg, Z. *et al.* (2013) The membrane-associated transient receptor potential vanilloid channel is the central heat shock receptor controlling the cellular heat shock response in epithelial cells. *PLoS ONE* 8, e57149
49. Zhao, J. *et al.* (2020) Plant responses to heat stress: physiology, transcription, noncoding RNAs, and epigenetics. *Int. J. Mol. Sci.* 22, 117
50. Kumar, S.V. and Wigge, P.A. (2010) H2A.Z-containing nucleosomes mediate the thermosensory response in *Arabidopsis*. *Cell* 140, 136–147
51. Nuruzzaman, M. *et al.* (2013) Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. *Front. Microbiol.* 4, 248
52. Suzuki, N. (2021) High temperature sensing mechanisms and their downstream pathways in plants. In *Plant Growth and Stress Physiology* (Gupta, D.K. and Palma, J.M., eds), pp. 49–71, Springer
53. Li, B. *et al.* (2019) Transcriptional profiling reveals a time-of-day-specific role of REVEILLE 4/8 in regulating the first wave of heat shock-induced gene expression in *Arabidopsis*. *Plant Cell* 31, 2353–2369
54. Mhamdi, A. and Van Breusegem, F. (2018) Reactive oxygen species in plant development. *Development* 145, dev164376
55. Dietz, K.J. *et al.* (2016) Recent progress in understanding the role of reactive oxygen species in plant cell signaling. *Plant Physiol.* 171, 1535–1539
56. Kadota, Y. *et al.* (2015) Regulation of the NADPH oxidase RBOHD during plant immunity. *Plant Cell Physiol.* 56, 1472–1480
57. Li, B. *et al.* (2018) Molecular mechanisms governing plant responses to high temperatures. *J. Integr. Plant Biol.* 60, 757–779
58. Jogawat, A. (2019) Osmolytes and their role in abiotic stress tolerance in plants. In *Molecular Plant Abiotic Stress: Biology and Biotechnology* (Roychoudhury, A. and Tripathi, D., eds), pp. 91–104, Wiley
59. Niu, Y. and Xiang, Y. (2018) An overview of biomembrane functions in plant responses to high-temperature stress. *Front. Plant Sci.* 9, 915
60. Dill, K.A. (1990) Dominant forces in protein folding. *Biochemistry* 29, 7133–7155
61. Dutta, T. *et al.* (2019) Role and regulation of osmolytes as signaling molecules to abiotic stress tolerance. In *Plant Signaling Molecules* (Khan, M.I.R. *et al.*, eds), pp. 459–477, Woodhead Publishing
62. Kempa, S. *et al.* (2008) A central role of abscisic acid in stress-regulated carbohydrate metabolism. *PLoS ONE* 3, e3935
63. Lv, W.T. *et al.* (2011) Proline accumulation is inhibitory to *Arabidopsis* seedlings during heat stress. *Plant Physiol.* 156, 1921–1933
64. Rizhsky, L. *et al.* (2004) When defense pathways collide. The response of *Arabidopsis* to a combination of drought and heat stress. *Plant Physiol.* 134, 1683–1696
65. Zulfiqar, F. *et al.* (2022) Role of glycine betaine in the thermotolerance of plants. *Agronomy* 12, 276
66. Jin, R. *et al.* (2016) Physiological and metabolic changes of purslane (Portulaca oleracea L.) in response to drought, heat, and combined stresses. *Front. Plant Sci.* 6, 1123
67. Song, C. *et al.* (2016) Overexpression of heat shock factor gene HsfA3 increases galactinol levels and oxidative stress tolerance in *Arabidopsis*. *Mol. Cell* 39, 477–483
68. Suprasanna, P. *et al.* (2016) Osmolyte accumulation and implications in plant abiotic stress tolerance. In *Osmolytes and Plants Acclimation to Changing Environment: Emerging Omics Technologies* (Iqbal, N. *et al.*, eds), pp. 1–12, Springer
69. Serrano, N. *et al.* (2019) Thermopriming reprograms metabolic homeostasis to confer heat tolerance. *Sci. Rep.* 9, 181
70. Ellis, J. (1987) Proteins as molecular chaperones. *Nature* 328, 378–379
71. Tiwari, L.D. *et al.* (2021) Stress and development phenotyping of Hsp101 and diverse other Hsp mutants of *Arabidopsis thaliana*. *J. Plant Biochem. Biotechnol.* 30, 889–905
72. Chantarachot, T. and Bailey-Serres, J. (2018) Polysomes, stress granules, and processing bodies: a dynamic triumvirate controlling cytoplasmic mRNA fate and function. *Plant Physiol.* 176, 254–269
73. Amen, T. *et al.* (2021) Resveratrol and related stilbene derivatives induce stress granules with distinct clearance kinetics. *Mol. Biol. Cell* 32, ar18
74. Chodasiewicz, M. *et al.* (2020) Identification and characterization of the heat-induced plastidial stress granules reveal new insight into *Arabidopsis* stress response. *Front. Plant Sci.* 11, 1674
75. McLoughlin, F. *et al.* (2019) HSP101 interacts with the proteasome and promotes the clearance of ubiquitylated protein aggregates. *Plant Physiol.* 180, 1829–1847
76. McLoughlin, F. *et al.* (2016) Class I and II small heat shock proteins together with HSP101 protect protein translation factors during heat stress. *Plant Physiol.* 172, 1221–1236
77. Waters, E.R. and Vierling, E. (2020) Plant small heat shock proteins - evolutionary and functional diversity. *New Phytol.* 227, 24–37
78. Mogk, A. and Bukau, B. (2017) Role of sHsps in organizing cytosolic protein aggregation and disaggregation. *Cell Stress Chaperones* 22, 493–502
79. Torok, Z. *et al.* (2001) Synechocystis HSP17 is an amphitropic protein that stabilizes heat-stressed membranes and binds denatured proteins for subsequent chaperone-mediated refolding. *Proc. Natl. Acad. Sci. U. S. A.* 98, 3098–3103
80. Bernfur, K. *et al.* (2017) The chloroplast-localized small heat shock protein Hsp21 associates with the thylakoid membranes in heat-stressed plants. *Protein Sci.* 26, 1773–1784
81. Sedaghatmehr, M. *et al.* (2016) The plastid metalloprotease FtsH6 and small heat shock protein HSP21 jointly regulate thermomemory in *Arabidopsis*. *Nat. Commun.* 7, 12439
82. Rebeaud, M.E. *et al.* (2021) On the evolution of chaperones and cochaperones and the expansion of proteomes across the Tree of Life. *Proc. Natl. Acad. Sci.* 118, e2020885118
83. Zhao, Q. and Liu, C. (2017) Chloroplast chaperonin: an intricate protein folding machine for photosynthesis. *Front. Mol. Biosci.* 4, 98
84. Usman, M.G. *et al.* (2017) Molecular analysis of Hsp70 mechanisms in plants and their function in response to stress. *Biotechnol. Genet. Eng. Rev.* 33, 26–39
85. Fauvet, B. *et al.* (2021) Bacterial Hsp90 facilitates the degradation of aggregation-prone Hsp70-Hsp40 substrates. *Front. Mol. Biosci.* 8, 653073
86. Wang, T.Y. *et al.* (2021) HSP70-4 and farnesylated AtJ3 constitute a specific HSP70/HSP40-based chaperone machinery essential for prolonged heat stress tolerance in *Arabidopsis*. *J. Plant Physiol.* 261, 153430
87. Sung, D.Y. and Guy, C.L. (2003) Physiological and molecular assessment of altered expression of Hsc70-1 in *Arabidopsis*. Evidence for pleiotropic consequences. *Plant Physiol.* 132, 979–987
88. Imamoglu, R. *et al.* (2020) Bacterial Hsp70 resolves misfolded states and accelerates productive folding of a multi-domain protein. *Nat. Commun.* 11, 365
89. Goloubinoff, P. *et al.* (2018) Chaperones convert the energy from ATP into the nonequilibrium stabilization of native proteins. *Nat. Chem. Biol.* 14, 388–395
90. Kampinga, H.H. *et al.* (2019) Function, evolution, and structure of J-domain proteins. *Cell Stress Chaperones* 24, 7–15
91. Kityk, R. *et al.* (2018) Molecular mechanism of J-domain-triggered ATP hydrolysis by Hsp70 chaperones. *Mol. Cell* 69, 227–237
92. Verma, A.K. *et al.* (2019) The expanding world of plant J-domain proteins. *CRC Crit Rev Plant Sci.* 38, 382–400
93. Nillegoda, N.B. *et al.* (2015) Crucial HSP70 co-chaperone complex unlocks metazoan protein disaggregation. *Nature* 524, 247–251
94. Craig, E.A. and Marszalek, J. (2017) How do J-proteins get Hsp70 to do so many different things? *Trends Biochem. Sci.* 42, 355–368
95. D'Silva, P.R. *et al.* (2008) Interaction of the J-protein heterodimer Pam18/Pam16 of the mitochondrial import motor with the translocon of the inner membrane. *Mol. Biol. Cell* 19, 424–432
96. Xing, Y. *et al.* (2010) Structure of clathrin coat with bound Hsc70 and auxilin: mechanism of Hsc70-facilitated disassembly. *EMBO J.* 29, 655–665

97. Verma, A.K. *et al.* (2017) Evolutionary conservation and emerging functional diversity of the cytosolic Hsp70:J protein chaperone network of *Arabidopsis thaliana*. *G3 (Bethesda)* 7, 1941–1954
98. Ticha, T. *et al.* (2020) Multifaceted roles of HEAT SHOCK PROTEIN 90 molecular chaperones in plant development. *J. Exp. Bot.* 71, 3966–3985
99. Cox, M.B. and Johnson, J.L. (2018) Evidence for Hsp90 co-chaperones in regulating Hsp90 function and promoting client protein folding. In *Chaperones: Methods and Protocols* (Calderwood, S.K. and Prince, T.L., eds), pp. 397–422, Springer
100. Xu, Z.S. *et al.* (2012) Heat shock protein 90 in plants: molecular mechanisms and roles in stress responses. *Int. J. Mol. Sci.* 13, 15706–15723
101. Yamada, K. *et al.* (2007) Cytosolic HSP90 regulates the heat shock response that is responsible for heat acclimation in *Arabidopsis thaliana*. *J. Biol. Chem.* 282, 37794–37804
102. Moran Luengo, T. *et al.* (2018) Hsp90 breaks the deadlock of the Hsp70 chaperone system. *Mol. Cell* 70, 545–552
103. Diamant, S. *et al.* (2000) Size-dependent disaggregation of stable protein aggregates by the DnaK chaperone machinery. *J. Biol. Chem.* 275, 21107–21113
104. Mishra, R.C. and Grover, A. (2016) ClpB/Hsp100 proteins and heat stress tolerance in plants. *Crit. Rev. Biotechnol.* 36, 862–874
105. Sousa, R. (2014) Structural mechanisms of chaperone mediated protein disaggregation. *Front. Mol. Biosci.* 1, 12
106. Lee, U. *et al.* (2007) The *Arabidopsis* ClpB/Hsp100 family of proteins: chaperones for stress and chloroplast development. *Plant J.* 49, 115–127
107. Hong, S.W. and Vierling, E. (2001) Hsp101 is necessary for heat tolerance but dispensable for development and germination in the absence of stress. *Plant J.* 27, 25–35
108. Queitsch, C. *et al.* (2000) Heat shock protein 101 plays a crucial role in thermotolerance in *Arabidopsis*. *Plant Cell* 12, 479–492
109. Gurley, W.B. (2000) HSP101: a key component for the acquisition of thermotolerance in plants. *Plant Cell* 12, 457–460
110. Sedaghatmehr, M. *et al.* (2019) A regulatory role of autophagy for resetting the memory of heat stress in plants. *Plant Cell Environ.* 42, 1054–1064
111. Surguchov, A. *et al.* (2019) Amyloidosis and longevity: a lesson from plants. *Biology (Basel)* 8, 43
112. Raza, A. *et al.* (2021) Integrated analysis of metabolome and transcriptome reveals insights for cold tolerance in rapeseed (*Brassica napus* L.). Front. *Plant Sci.* 12, 721681
113. Che, P. *et al.* (2010) Signaling from the endoplasmic reticulum activates brassinosteroid signaling and promotes acclimation to stress in *Arabidopsis*. *Sci. Signal.* 3, ra69
114. Yu, H.-D. *et al.* (2012) Downregulation of chloroplast RPS1 negatively modulates nuclear heat-responsive expression of HsfA2 and its target genes in *Arabidopsis*. *PLoS Genet.* 8, e1002669
115. Fiorucci, A.-S. *et al.* (2020) PHYTOCHROME INTERACTING FACTOR 7 is important for early responses to elevated temperature in *Arabidopsis* seedlings. *New Phytol.* 226, 50–58
116. Legris, M. *et al.* (2016) Phytochrome B integrates light and temperature signals in *Arabidopsis*. *Science* 354, 897–900
117. Hayes, S. (2020) Interaction of light and temperature signalling in plants. *eLS* Published online January 16, 2020. https://doi.org/10.1002/9780470015902.a0027978
118. Jung, J.-H. *et al.* (2020) A prion-like domain in ELF3 functions as a thermosensor in *Arabidopsis*. *Nature* 585, 256–260
119. Trapero-Mozos, A. *et al.* (2018) A reversible light- and genotype-dependent acquired thermotolerance response protects the potato plant from damage due to excessive temperature. *Planta* 247, 1377–1392
120. Mittler, R. *et al.* (2012) How do plants feel the heat? *Trends Biochem. Sci.* 37, 118–125
121. Zhang, J. *et al.* (2015) Hsf and Hsp gene families in *Populus*: genome-wide identification, organization and correlated expression during development and in stress responses. *BMC Genomics* 16, 181
122. Sarkar, N.K. *et al.* (2013) Functional analysis of Hsp70 superfamily proteins of rice (*Oryza sativa*). *Cell Stress Chaperones* 18, 427–437
123. Zandalinas, S.I. *et al.* (2021) The impact of multifactorial stress combination on plant growth and survival. *New Phytol.* 230, 1034–1048

CHAPTER 6

How do humans and plants feel the heat?

How do humans and plants feel the heat?

Anthony Guihur, Mathieu E. Rebeaud, Baptiste Bourgine and Pierre Goloubinoff

I contributed to this work conceptually, in writing the manuscript and preparing the figures.

CellPress
OPEN ACCESS

Trends in
Plant Science

Special issue: Climate change and sustainability I

Forum

How do humans and plants feel the heat?

Anthony Guihur [1,*], Mathieu E. Rebeaud [1], Baptiste Bourgine,[1] and Pierre Goloubinoff [1,2,*]

The 2021 Nobel prize was awarded for the discovery of the animal thermosensory channel TRPV1. We highlight notable shared features with the higher plant thermosensory channel CNGC2/4. Both channels respond to temperature-induced changes in plasma membrane fluidity, leading to hyperphosphorylation of the HSF1 transcription factor via a specific heat-signaling cascade.

The 2021 Nobel prize in Medicine and Physiology was shared by David Julius and Ardem Patapoutian for the discovery of the thermosensory ion channel TRPV1 and the mechanosensitive ion channel component 2, PIEZO2, in vertebrates (https://www.cell.com/nobelprize#:~:text=in%2DChief%2C%20Patterns-,Nobel%20Prize%20in%20Physiology%20or%20Medicine%202021,receptors%20for%20temperature%20and%20touch). In the current context of human-aggravated global warming, it is important to gain knowledge beyond the vertebrates, on how plants, fungi, protozoa, bacteria, and archaea perceive sudden changes of ambient temperature and react in a timely manner to establish various defenses to avert and repair damages from upcoming environmental stresses. Interestingly, the findings on the human heat sensor provide valuable clues on how land plants may similarly perceive a rise in the ambient temperature and establish comparable molecular defenses, by accumulating a conserved set of heat-shock proteins (HSPs), conferring thermotolerance to both vertebrates and plants [1].

The animal thermosensory channel TRPV1 forms a tetrameric transmembranal ion channel, which at resting low temperature is in a closed polarized state, poised to readily respond to a temperature increase. Cholesterol is a membrane-rigidifying molecule that was found to control membrane phase transitions at the heat-activating temperatures for TRPV1 [2], suggesting that the heat-shock response depends both on the intrinsic thermo-responsive characteristics of the channels, as evidenced by temperature-sensitizing mutations [3], and on the presence of molecules, such as cholesterol or saturated/unsaturated lipids affecting the fluidity of the plasma membrane in which TRPV1s are embedded (Figure 1, left). Under heat shock, the TRPV1 channel transiently opens and mediates the entry of extracellular Ca^{2+} ions, which bind, recruit, and activate calmodulins associated with the N and C terminal cytosolic domains of TRPV1 [4]. This initiates a specific cellular signal that ends in activation of the heat shock transcription factor 1 (HSF1). The widely accepted model for animal cells is that at low temperature, HSF1 is maintained inactive in the cytosol by bound HSP70 and HSP90 molecules. Under heat shock, the chaperones are observed to dissociate, while HSF1 concomitantly becomes hyperphosphorylated and active; therefore, it is generally thought that some unknown thermolabile proteins, that presumably aggregate, can titer away the repressor HSP70s and HSP90s from the HSF1 complex. Consequently, the activated HSF1 is unleashed to translocate to the nucleus, where it recruits RNA polymerase to synthetize HSP mRNA [5]. The consequent massive accumulation of HSPs, many of which are molecular chaperones, establishes effective protective mechanisms against heat damages in thermolabile proteins and membranes.

The heat-depolarized TRPV1 channel soon closes and does not further allow translocation of external Ca^{2+} ions, despite the ongoing heat stimulus. As demonstrated for TRPV3 [6], hours at non-heat shock temperature are necessary for TRPV-type channels to revert into an initial heat-responsive state, poised to respond again to an upcoming heat stimulus. Binding of the agonist capsaicin to TRPV1 induces a localized nociception of heat. Both activation by heat or the binding of capsaicin at 37°C produces a similar profile of HSP accumulation [7]. Because capsaicin is a potent TRPV1 agonist, it is used, somewhat counterintuitively, in anesthesiology [8]. Conversely, and demonstrating that TRPV1 is a central thermo-sensor of human cells, pretreatments of cells with TRPV1 RNAi, or the TRPV1 antagonist capsazepine, or EGTA that chelates external Ca^{2+}, all prevent the heat-induced accumulation of HSPs [7].

Noticeably, higher plants similarly contain specific thermosensory channels, called cyclic nucleotide gated channels (CNGC2/4s) [9–11] (Figure 1, right), which, like TRPV1, form at low-temperature transmembranal ion channels that are closed. This closed polarized state is poised to readily respond to a heat-induced increase in the fluidity of the surrounding plasma membrane. Like TRPV1, upon heat exposure, the heat-responsive CNGC2/4 channels transiently open and mediate the controlled entry of external Ca^{2+} ions, which bind to calmodulins that are associated with the cytosolic C terminal domain of the channel [12]. This triggers a specific heat-signal cascade involving kinases that ultimately phosphorylate and activate HSFA1 [13].

Moreover, like neurons, which by way of propagating the rapid entry of Ca^{2+} ion along axons can transfer a heat signal from a distant organ to the human brain and elicit an escape reaction, plants too may transfer an initial heat signal from a heated leaf to neighboring leaves, by way

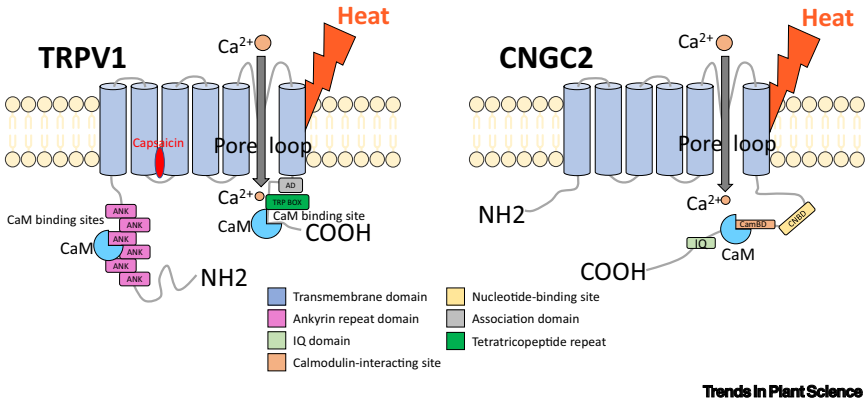


Figure 1. Plant and animal heat sensors are very similar in structure and function. Domain organization of heat-sensory channels. Vertebrate TRPV1 (left) and plant CNGC2 (right) are tetrameric channels in the plasma membrane. Their subunits are composed of six transmembranal helices with a pore loop between helix 5 and 6. Both channels bind calmodulins in their cytosolic C terminal. The N terminal of TRPV1 contains six ankyrin domains that also bind calmodulins. Higher temperature increases the fluidity of the plasma membrane. This, or capsaicin-binding to TRPV1, induces the transient opening of the channels and the binding of extracellular calcium to channel-associated calmodulins on the cytosolic side, which in turn sends a specific signal to activate heat-shock transcription factors and produce protective heat-shock proteins.

of an observed propagation of Ca^{2+} ion entry along the vasculature [14]. Noticeably, the plant vasculature is composed of both dead hollow xylem cells and of live phloem sieves (Figure 2). Interestingly, like animal axons that are highly differentiated live cell extensions fostered by neighboring glial cells, phloem sieves are highly differentiated live cell extensions fostered by companion cells. Whereas the hollow dead xylem cells lack a source of energy to carry out the observed rapid heat-induced propagation of calcium ions (in the order of 20 seconds) against the water flow down the petiole of an arabidopsis (*Arabidopsis thaliana*) leaf (Figure 2), live phloem sieves could, in principle, use energy from ATP hydrolysis to do so by a mechanism yet to be determined.

Yet, at variance with TRPV1-expressing vertebrates with brains and muscles, for which evolution may have found an advantage in developing nociception to escape heat damages, CNGC-expressing plants lack brains and muscles. Therefore, although similar to vertebrates, plant tissues sense heat and react by producing HSPs and by sending a heat signal to neighboring organs, unlike animals, plant have a sessile lifestyle. Plants cannot escape heat damages and, therefore, it is unlikely that evolution has developed nociception as an effective warning system to prompt plants to escape heat stress.

Thus, despite the similarities of sensing and of the signal transduction pathways between the TRPV1 thermo-sensors and the heat CNGC thermo-sensors, it is unlikely that plants feel pain. Yet, being complex organisms and serving as the main food source for the ever-expanding human population of the planet, plants deserve more research to understand their functional

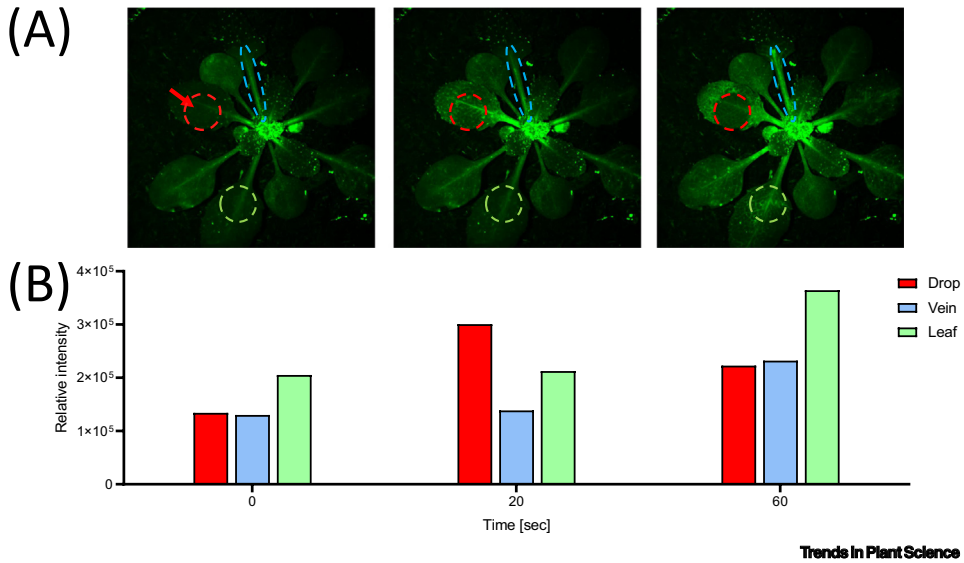


Figure 2. A local heat stress elicits a calcium-dependent heat signal between leaves. (A) Transgenic *Arabidopsis thaliana* plants that constitutively express the calcium-dependent GCaMP3 fluorescent reporter [15]. The time laps (0, 20, and 60 s) following the deposition of a hot drop of water at 45°C on one leaf (red arrow and circle). Grey circles indicate high calcium concentrations in the live tissues of distal vasculature and leaves. (B) The time laps (at 0, 20, and 60 s) fluxes of Ca^{2+} in relative intensity calculated with ImageJ by measuring GFP fluorescence.

characteristics and potential adaptation to global warming, as well as our gratitude and respect [15].

Acknowledgments

We apologize to those whose important work was not cited due to space and reference number limitations. This work was funded by grant 31003A_175453 from the Swiss National Fund and grant 3150/21 from the Israel Science Foundation to P.G., and grant SPARK CRSK-3_196689 from the Swiss National Fund to A.G.

Declaration of interests

The authors have no conflict of interest.

[1]Department of Plant Molecular Biology, Faculty of Biology and Medicine, University of Lausanne, CH-1015 Lausanne, Switzerland
[2]School of Plant Sciences and Food Security, Tel-Aviv University, Tel Aviv, Israel

*Correspondence:
anthony.guihur@gmail.com (A. Guihur) and
Pierre.Goloubinoff@unil.ch (P. Goloubinoff).

https://doi.org/10.1016/j.tplants.2022.03.006

References

1. Mittler, R. *et al.* (2012) How do plants feel the heat? *Trends Biochem. Sci.* 37, 118–125
2. Liu, B. *et al.* (2003) Thermodynamics of heat activation of single capsaicin ion channels VR1. *Biophys. J.* 85, 2988–3006
3. Gracheva, E.O. *et al.* (2011) Ganglion-specific splicing of TRPV1 underlies infrared sensation in vampire bats. *Nature* 476, 88–91
4. Lau, S.Y. *et al.* (2012) Distinct properties of Ca2+-calmodulin binding to N- and C-terminal regulatory regions of the TRPV1 channel. *J. Gen. Physiol.* 140, 541–555
5. Masser, A.E. *et al.* (2020) Hsf1 on a leash - controlling the heat shock response by chaperone titration. *Exp. Cell Res.* 396, 112246
6. Nadezhdin, K.D. *et al.* (2021) Structural mechanism of heat-induced opening of a temperature-sensitive TRP channel. *Nat. Struct. Mol. Biol.* 28, 564–572
7. Bromberg, Z. *et al.* (2013) The membrane-associated transient receptor potential vanilloid channel is the central heat shock receptor controlling the cellular heat shock response in epithelial cells. *PLoS ONE* 8, e57149
8. Koivisto, A.P. *et al.* (2021) Advances in TRP channel drug discovery: from target validation to clinical studies. *Nat. Rev. Drug Discov.* 21, 41–59
9. Finka, A. *et al.* (2012) Plasma membrane cyclic nucleotide gated calcium channels control land plant thermal sensing and acquired thermotolerance. *Plant Cell* 24, 3333–3348
10. Saidi, Y. *et al.* (2009) The heat shock response in moss plants is regulated by specific calcium-permeable channels in the plasma membrane. *Plant Cell* 21, 2829–2843
11. Jarratt-Barnham, E. *et al.* (2021) The complex story of plant cyclic nucleotide-gated channels. *Int. J. Mol. Sci.* 22, 874
12. Fischer, C. *et al.* (2013) An IQ domain mediates the interaction with calmodulin in a plant cyclic nucleotide-gated channel. *Plant Cell Physiol.* 54, 573–584
13. Bourgine, B. and Guihur, A. (2021) Heat shock signaling in land plants: from plasma membrane sensing to the transcription of small heat shock proteins. *Front. Plant Sci.* 12, 710801
14. Nguyen, C.T. *et al.* (2018) Identification of cell populations necessary for leaf-to-leaf electrical signaling in a wounded plant. *Proc. Natl. Acad. Sci. U. S. A.* 115, 10178–10183
15. Knapp, S. (2019) Are humans really blind to plants? *Plants People Planet* 1, 164–168

PART2

Public Health – COVID-19

Public health, treatment controversy and meta-analyses on COVID-19

CHAPTER 7

Moderate Fever Cycles as a Potential Mechanism to Protect the Respiratory System in COVID-19 Patients

Moderate Fever Cycles as a Potential Mechanism to Protect the Respiratory System in COVID-19 Patients

Anthony Guihur, Mathieu E. Rebeaud, Bruno Fauvet, Satyam Tiwari, Yoram G. Weiss and Pierre Goloubinoff

I contributed to this work conceptually and in the manuscript writing.

AUTHOR CONTRIBUTIONS
AG and BF analyzed published transcriptomic and proteomic data. AG, BF, and PG made the figures. All authors conceived the central ideas of the manuscript, interpreted data from literature, contributed to writing, reviewed, edited, and approved its final version of the manuscript.

frontiers in Medicine

MINI REVIEW
published: 11 September 2020
doi: 10.3389/fmed.2020.564170

Moderate Fever Cycles as a Potential Mechanism to Protect the Respiratory System in COVID-19 Patients

Anthony Guihur[1], Mathieu E. Rebeaud[1], Bruno Fauvet[1], Satyam Tiwari[1], Yoram G. Weiss[2] and Pierre Goloubinoff[1]**

[1] Department of Plant Molecular Biology, Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland,
[2] Department of Anesthesiology, Hadassah Hebrew University Medical Center, Jerusalem, Israel

Mortality in COVID-19 patients predominantly results from an acute respiratory distress syndrome (ARDS), in which lungs alveolar cells undergo programmed cell death. Mortality in a sepsis-induced ARDS rat model is reduced by adenovirus over-expression of the HSP70 chaperone. A natural rise of body temperature during mild fever can naturally accumulate high cellular levels of HSP70 that can arrest apoptosis and protect alveolar lung cells from inflammatory damages. However, beyond 1–2 h of fever, no HSP70 is being further produced and a decreased in body temperature required to the restore cell's ability to produce more HSP70 in a subsequent fever cycle. We suggest that antipyretics may be beneficial in COVID-19 patients subsequent to several hours of mild (<38.8°C) advantageous fever, allowing lung cells to accumulate protective HSP70 against damages from the inflammatory response to the virus SARS-CoV-2. With age, the ability to develop fever and accumulate HSP70 decreases. This could be ameliorated, when advisable to do so, by thermotherapies and/or physical training.

Keywords: acute respiratory distress syndrome, COVID-19, SARS-CoV-2, fever, Hsp70, heat- shock response

OPEN ACCESS

Edited by:
Yoram Vodovotz,
University of Pittsburgh, United States

Reviewed by:
Carmen Silvia Valente Barbas,
University of São Paulo, Brazil
Hari H. Subramanian,
Boston Scientific, United States

***Correspondence:**
Yoram G. Weiss
weiss@hadassah.org.il
Pierre Goloubinoff
pierre.goloubinoff@unil.ch

Specialty section:
This article was submitted to Pulmonary Medicine, a section of the journal Frontiers in Medicine

Received: *22 May 2020*
Accepted: *17 August 2020*
Published: *11 September 2020*

Citation:
Guihur A, Rebeaud ME, Fauvet B, Tiwari S, Weiss YG and Goloubinoff P (2020) Moderate Fever Cycles as a Potential Mechanism to Protect the Respiratory System in COVID-19 Patients. Front. Med. 7:564170. doi: 10.3389/fmed.2020.564170

THE EFFECT OF ELEVATED ENVIRONMENTAL TEMPERATURES ON THE COVID-19 PANDEMIC

In February 2020, many health and political officials across the world were still grossly underestimating the severity of the developing COVID-19 pandemic, in part because of the scientifically unproven belief that like seasonal influenza, COVID-19 would disappear by April 2020 with the rise of temperatures in the Northern Hemisphere. Aside from being scientifically improper to extrapolate information from other seasonal viruses to a novel virus propagating in populations lacking prior immune defenses (1), SARS-CoV-2 was since proven to resist warm summer temperatures in countries of the northern hemisphere, experiencing dramatic deadly second waves of infections (https://coronavirus.jhu.edu/map.html). Moreover, the outbreak of another closely related coronavirus, MERS-CoV, occurred in Saudi Arabia despite scorching spring and summer temperatures (2). Unless strong measures are taken against the COVID-19 pandemic through social distancing and by developing effective vaccines and anti-viral drugs, it is likely to become an ongoing plague. Nonetheless, as previously suggested for SARS-CoV-1 (3), warmer temperatures and higher humidity may reduce SARS-CoV-2's viability *ex vivo* on infective surfaces, thereby mitigating the spread (4, 5). *In vivo* evidence is lacking on mitigating or aggravating effects

of high fever on SARS-CoV-2 replication. Yet, cycles of mild harmless (<38.8°C) fever may have strong effects on the disease pathology, as a result of the accumulation of heat-shock proteins (HSPs), in particular of the HSP70 chaperone, which can help the respiratory cells to sustain stress from both the virus and the human inflammatory response to the virus.

ACUTE RESPIRATORY DISTRESS SYNDROME IS THE MAJOR CAUSE OF DEATH FROM SARS-COV-2 INFECTION

By early September 2020, COVID-19 has already caused over 870,000 deaths worldwide. In the most severe cases, the disease progresses into acute respiratory distress syndrome (ARDS), which is among the top three complications after sepsis, causing respiratory failure and death (6). ARDS occurs when protein-rich inflammatory edema fluid builds up in the alveolar space as a result of lung damage, leading to non-cardiogenic pulmonary edema and decreased arterial oxygenation that necessitates mechanical ventilation (7). Early phases of lung pathology in COVID-19 pneumonia show a rather classic edema, with proteinaceous exudates as large protein globules, multinucleated giant cells and hyperplasia of pneumocytes, as with other types of sepsis-induced ARDS. Vascular congestion, combined with inflammatory clusters of fibrinoid material have also been reported, indicating that vascular inflammation and coagulopathy may be more particular hallmarks of the disease (8).

The pulmonary alveoli, which are the main sites of gas exchange with the blood, are composed of a thin alveolar epithelium that covers 99% of the lung surface and includes thin, squamous type I cells (AT1) and cuboid-shaped type 2 cells (AT2). The general hallmark of initial ARDS-induced lung injury is increased capillary leakage and intra-alveolar edema. The AT1 cells that enable gas exchanges undergo irreversible programmed cell death or necrosis, whereas AT2 cells, rather than undergoing limited division and differentiating into new functional AT1 cells, undergo unchecked division and do not differentiate. They accumulate into so-called "ground-glass opacities," filling the lung cavities and leading to lung failure (9). Treatment of severe ARDS from COVID-19 is an ongoing challenge. Protective mechanical ventilation remains the pillar of ARDS management to facilitate oxygenation with the goal of improving oxygenation through the damaged lungs while reducing ventilator-induced lung injury. If mechanical ventilation fails, extracorporeal membrane oxygenation has been used in COVID-19 ARDS patients with promising results (10).

In the absence of prior effective vaccination (11), another important treatment direction is the prevention or reduction of cell infection by the virus through the repurposing of drugs such as remdesivir, chloroquine, lopinavir/ritonavir, which have different mechanisms of action and are still under development and experimental evaluation (12). Additional therapies aiming at enhancing the natural cellular defenses against the onset of ARDS should be considered. Importantly, mortality from SARS-CoV-2 infections is extremely low among young patients and increasing dramatically in patients aged above 65 (13). For example, official numbers from the end of July 2020 showed that the mortality risk in Switzerland is 150-fold higher for COVID-19 patients aged 70–80+ (1,519 deaths), compared to patients aged 30–49 (10 deaths) [data as from July 28 of the Federal Office of Public Health (FOPH), https://www.bag.admin.ch/bag/en/home.html]. This is evidence that strong natural cellular defenses against the virus are at work in youth which, for reasons yet to be clarified, become progressively less effective in late adulthood possibly in association with genetic parameters, such as gender and blood type (14) and are aggravated by health preconditions, such as obesity, smoking, diabetes and heart diseases.

The Mitigating Effect of Mild Fever on ARDS

The heat-shock response (HSR) is an example of the buildup of such natural cellular defenses that are highly effective in youth and become progressively less effective in late adulthood. The HSR is defined by the transient accumulation of so-called heat-shock proteins (HSPs), most of which belonging to the conserved chaperone families HSP70, HSP90, HSP60, and HSP40, in response to a temperature rise. HSPs play a general cytoprotective role, among others, in lung inflammation (15). An effective HSR protects thermolabile proteins and membranes from damage caused by excessive variations in the environment, such as heat stress, oxidative stress, UV light, or infection (16, 17). It typically leads to the onset of acquired thermotolerance, i.e., to the transient resistance to a subsequent otherwise deadly dose of elevated temperature (18). It has been shown that, as with externally applied high temperatures, mild fever also activates the HSR in mammals, thereby accelerating healing and preventing apoptosis of respiratory epithelial cells (19, 20). Fever is a major hallmark of inflammatory diseases. Despite its high metabolic cost, it has been an integral part of vertebrate's immune response to infections for the last 400 million years (21), suggesting that fever provides a strong evolutionary advantage for the survival of the fittest. Yet, for over a century, caregivers generally considered fever dangerous and a source of patient discomfort, leading to the systematic use of antipyretics. There is, however, growing evidence that allowing the onset of mild fever leads to better outcomes (19, 22, 23) and higher survival to infectious diseases, especially in cases of ARDS (24, 25).

High Cellular Concentrations of HSP70s (HSPA1A) Can Repress Inflammation-Induced ARDS

HSP70s belong to a highly conserved family of molecular chaperones constituting up to 1% of the total protein mass of healthy mammalian cells (26). HSP70s can use the energy from ATP hydrolysis to forcefully unfold and dismantle different types of aggregated and functional protein oligomers in the cell. Hence, it can drive conformational changes in various large cytotoxic protein aggregates and convert them into soluble, harmless, functional proteins (27). Interestingly, HSP70s can also drive the

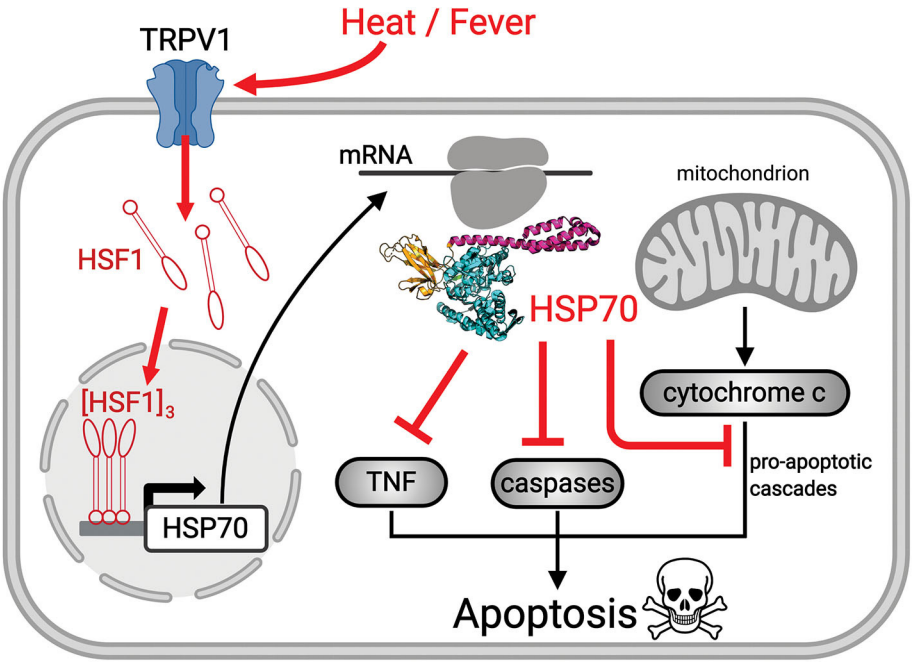


FIGURE 1 | Scheme showing the mechanism of heat-induced HSR and its protective effects. Upon heat-stress such as fever, the fluidized plasma membrane triggers the transient opening of heat sensory calcium channels, such as TRPV1. The ensuing specific calcium-mediated heat-shock signal induces the hyperphosphorylation of cytosolic inactive monomeric HSF1, that trimerizes and translocates to the nucleus where it binds promotor regions of HSP-encoding genes. The mRNA of HSP70 is translated in the cytosol into HSP70 acting to repress sepsis-induced ROS-mediated pro-apoptosis signals (red cross sticks), leading to survival of pre-existing alveolar type 1 cells and to the differentiation of alveolar type 2 cells, into functional, new alveolar type 1 cells.

specific dismantling of various active protein oligomers, such as clathrin cages, active heat-shock transcription factor (HSF1) trimers, and active pro-apoptotic IκB oligomers, which become reversibly inactivated upon HSP70-mediated de-oligomerization (28, 29) (**Figure 1**). Using a rat model for ARDS, it has been shown that an adenoviral vector expressing the stress-inducible form of HSP70, HSPA1A, can effectively protect against sepsis-induced ARDS by limiting neutrophil accumulation in the lungs and causing the inactivation of IκB complexes (30). HSP70 over-expression is also known to efficiently prevent caspase activation, and the heat-induced accumulation of mitochondrial HSP70, HSPA9, also can protect stressed mitochondria (31, 32), thereby conferring cells challenged by pathogens, cytotoxic chemicals, or abiotic stresses, with resistance from ROS-induced mitochondrial- and IκB-associated apoptosis. Hence, cancer cells often resist chemo- and thermo-therapies by over-expressing HSP70 chaperones, HSPA1A in particular (26). Conversely, degenerative neuronal and muscular tissues in aging nematodes and humans that systematically express lower cellular levels of HSP70s than young individuals (33) are particularly fragile and stress-sensitive. Cells with low HSP70 levels tend to spontaneously undergo apoptosis, and consequent tissue losses in aging humans lead to progressive degenerative diseases (34).

The HSR develops once cells have initially sensed a mild rise in temperature by way of converting small fluidity increments in their plasma membrane, into a specific cellular signal that activates HSF1 and ultimately de-represses HSP-encoding genes, leading to the accumulation of HSPs, the foremost of which being HSPA1A (35, 36) (**Figure 1**). Noticeably, beyond 2–3 h of continuous heat-shock, cells become ineffective at further accumulating heat-shock proteins and need to stay several hours back at low temperature to reset their ability to effectively respond again to a temperature rise. This implies that under continuous high fever, in the long term, the protective HSP70 molecules will gradually degrade without being replenished and will possibly reach a critically low level that cannot arrest apoptosis (**Figure 2A**). Interestingly, young COVID-19-infected patients develop ARDS much less frequently than older patients (**Figure 2B**), mirroring the fact that the HSR and the onset of acquired thermotolerance in humans are optimally effective in youth and progressively fail post puberty, in aging adults (**Figure 2C**). This can be attributed, in part, to the gradual stiffening of the plasma membranes in aging individuals as a result of decreased physical activity and the excessive intake of highly caloric food containing saturated lipids and cholesterol (37). In addition, the HSR may become impaired with age in particular in neural, liver and muscle tissues (38–42), likely in response to a hormonal signal that initiates at puberty (33, 43).

The combination of insufficiently elevated fever and a less effective HSR in the lung cells of the elderly may thus

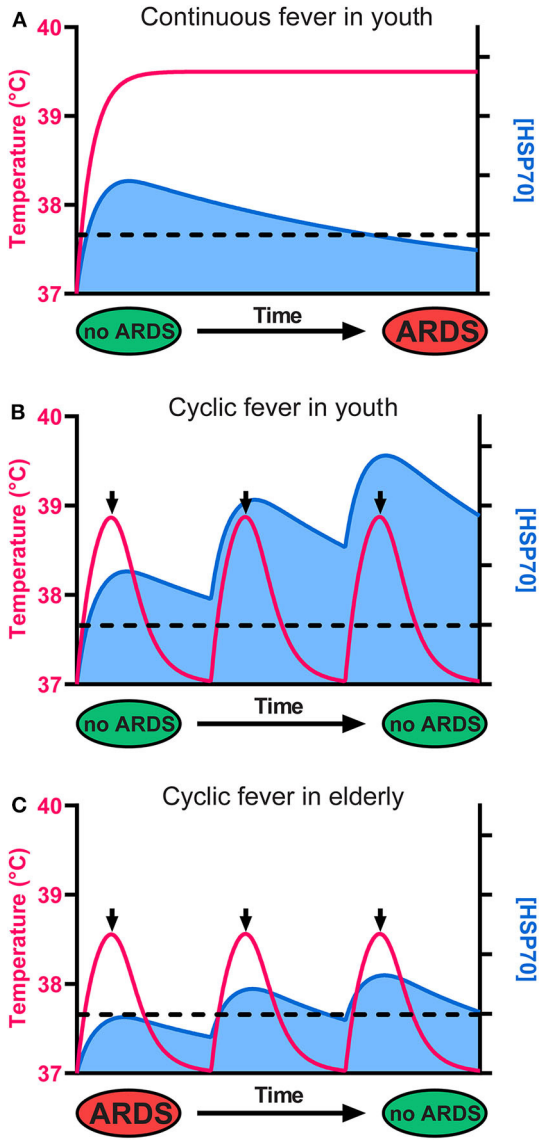


FIGURE 2 | The cellular levels of HSP70s determine the fate of alveolar cells, leading to survival or to ARDS and death. **(A)** Although it initially protects cells from ARDS, a continuous high fever in the long term leads to HSP70 depletion, ARDS, and death. **(B)** In youth, a cyclic mild fever up to 38.8°C (red line) causes in the first hours a strong beneficial accumulation of HSP70s (blue) that can potentially arrest sepsis-induced ARDS in young COVID-19 patients. Because long-term continuous fever leads to HSP70 depletion, the use of antipyretics after 2 h of fever is recommended. Iterative cycles of mild fever can thus maintain and even accumulate high cellular levels of protective HSP70s above a critical threshold (dashed line), preventing ARDS. **(C)** In seniors, the baseline temperature is often lower than 37°C, fever is often less intense, and less protective HSP70s accumulate. However, iterative fever cycles can accumulate protective HSP70s above the critical threshold (dashed line), arresting ARDS even in seniors.

lead to insufficient cellular amounts of protective HSP70s and the consequent failure to repress apoptosis in ARDS (**Figure 2C**) (44).

The HSR Is Transient

Whereas, SARS-CoV-2-infected patients experiencing mild fever may optimally accumulate HSP70s in both AT1 and AT2 cells, it should be noted that the HSR is transient. Following the rapid synthesis of HSP70 mRNA within the first ∼2 h of a temperature rise and the consequent cellular accumulation of HSPs, mRNA levels start decreasing despite the ongoing elevated temperature (**Figure 3A**) (45, 46). Hence, HSPA1A mRNA stops accumulating after about 2 h of high fever, and HSPA1A protein levels peak at around 4 h and thereafter start slowly decreasing, despite the ongoing heat shock (**Figure 3A**) (45). Remarkably, the cells need to return to 37°C for several hours in order for additional HSPA1A to be synthetized in a subsequent fever cycle, to replace the degraded chaperones and thus maintain apoptosis arrest. This behavior results from the fact that the initial step of the heat-shock signaling pathway is the transient opening of heat-sensory calcium channels, called transient receptor potential cation channel subfamily V member type 1 (TRPV1), that become depolarized in response to the heat-induced fluidization of the plasma membrane in which they are embedded (18, 47). Similar to unresponsive pain-depolarized nociceptive channels, and like their heat-sensing plant cognates, the cyclic nucleotide gated channels 2 and 4 (48), the heat-depolarized animal TRPV1 channels need to be returned for several hours at lower temperatures in order to regenerate into fully re-polarized, potent heat-responsive calcium channels (49).

EXTRINSIC HEAT TREATMENTS AND CO-INDUCERS OF THE HSR AS POTENTIAL PROPHYLACTIC AND THERAPEUTIC APPROACHES

A mild fever episode of 2–3 h, not exceeding 38.8°C is considered harmless by most of the medical community (53). Moreover, practitioners of traditional medicine on all continents have customarily provided treatments involving controlled mild rising of body temperature, either environmentally applied, as with warm bath therapy (54) that triggers HSP70 accumulation after 1 h at 40°C. Celastrol, a plant triterpene from the Chinese pharmacopeia, has been shown to have cytoprotective effects in autoimmune and inflammatory diseases (55) and could have protective effects against ARDS (56) through HSF1 activation. Defective heat-induced HSP70 production in seniors could be ameliorated prophylactically by increasing their physical training, during which body temperature naturally increases, or by sauna therapy, which can boost the HSR and is reported to reduce the risk for respiratory diseases (57) and for systemic inflammation in this population (58). The protective effects of iterative thermal exercise and increased heat training in marathon runners, in correlation with the induction of heat-shock proteins such as HSP70s and HSP90s, is well-documented (59). During exercise, the accumulation of HSP70s occurs widely across the organism and is measured up to 8-fold in muscle tissues and in lungs (60). However, because heat treatments may be excessively stressful to severely ill SARS-CoV-2 patients, one might expect them to be considered principally

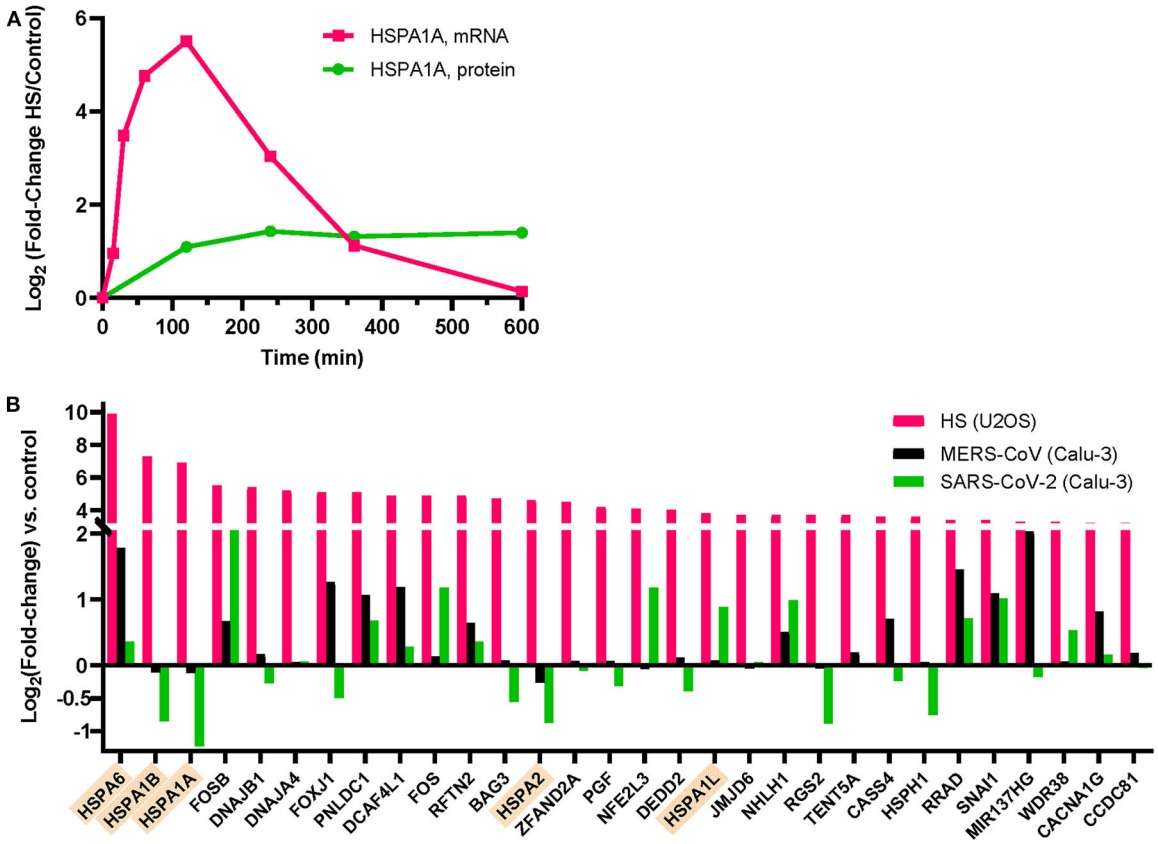


FIGURE 3 | Expression of heat shock proteins (HSPs) in human cells following heat-treatment at 42°C or MERS-CoV infection. **(A)** HSPA1A mRNA (pink) and protein (green) expression in Jurkat cells during a 4 h HS at 41°C, followed by 6 h at 37°C [data from (45)]. HSPA1A mRNA induction is maximal at 2 h HS and thereafter decreases despite the ongoing HS for another 2 h. **(B)** Normalized induction levels of the 30 most heat-induced genes by heat shock (42°C) in U2OS cells (pink) [data from (50)], by MERS-CoV infection (black) of human bronchial epithelial Calu-3 cells (black bars) [data from (51)], or by SARS-CoV-2 infection of Calu-3 cells (green bars) [data from (52)]. Cytosolic HSP70s are indicated in orange.

as prophylactic, in anticipation of a possible infection. In the context of SARS-CoV-2-induced viral pneumonia, the strategy of increasing cellular HSP70 levels by chemical compounds that could induce or co-induce with mild fever a strong HSR, may be of particular interest, given that both MERS-CoV and SARS-CoV-2 infection was found to specifically and significantly reduce HSPA1A, HSPA1B, and HSPA2 mRNAs in bronchial epithelial cells infected with the virus (**Figure 3B**) (51, 52).

In the absence of a vaccine at this stage for the COVID-19 pandemic, drug repurposing of FDA-approved molecules could be a timesaver. Several of these, such as carbenoxolone and arimoclomol, have been shown to have beneficial effects in various diseases, mainly by enhancing the heat-induced expression of HSPs (61–63). Glutamine, a conditionally essential amino acid that also triggers the HSR, can improve survival after sepsis and attenuate ARDS symptoms in a mouse model (64).

Were physicians wrong aiming at reducing moderate fever in sepsis? The current knowledge suggests that the answer is yes and no. On the one hand, fever is discomforting to patients and may have adverse effects, especially above 38.8°C. In addition, fever increases oxygen demand, thereby increasing cardiac and respiratory rates (65). A rise of 1°C can increase metabolic demand by 10% (66) and may therefore be detrimental to patients with heart failure or neurological impairment (67). Moreover, antipyretics that are also anti-inflammatory drugs are expected to reduce lung damage caused by an excessive inflammatory response in the lungs, caused by the viral infection. On the other hand, moderate fever, up to 38.8°C, has been reported to inhibit the replication of viruses, such as influenza and other pathogens, promote immunity and most importantly, cause the beneficial accumulation of anti-apoptotic HSP70s that can repress sepsis-induced ARDS. Yet, owing to the transient nature of the heat sensors in the plasma membrane that become depolarized and unresponsive to the heat beyond 2 h of continuous fever, maintaining a fever beyond that time is vain, as the HSP70 molecules that naturally degrade cannot be replenished (**Figure 2A**). Thus, physicians were not wrong at aiming to reduce fever, as several hours at a low temperature are necessary for the cells to regenerate their heat/fever-depolarized heat sensors and to fully respond again to a new cycle of fever (18, 45). Therefore, based on the above, we hypothesize that the optimal treatment of COVID-19 patients with antipyretic

drugs, such as acetaminophen, would be applied only following a couple of hours of moderate fever (**Figures 2B,C**). Then, several hours at a low temperature would be maintained to allow cells to reset their optimal HSR. Iterative repetitions of such fever cycles, each lasting 8–12 h, may be expected to maintain the highest cellular levels of HSP70s to protect lungs from ARDS damage in COVID-19 patients (**Figure 2B**) and possibly protect the elderly from ARDS and lung failure (**Figure 2C**). Prior to any implementation, our hypothesis must be tested in randomized clinical trials with large samples of patients in confinement of similar age and sex with mild symptoms. Antipyretics and their doses should be standardized and the length of the proposed delay before antipyretic intake, allowing mild fever to develop, should be standardized.

One additional avenue of research would be to take advantage of the tissue samples from the nasopharyngeal epithelial mucosa that are routinely used for PCR-based diagnosis of SARS-Cov-2. Quantitative RNAseq of various HSPs, HSP70s in particular, as well as of hallmark genes for ARDS, such as the pro-inflammatory cytokines (IL-1β, IL-6, KC, and MCP-1) (68) and metalloproteinase 9, which is involved in the degradation of extracellular matrix during ARDS (69), and CBIRC3 which inhibits apoptosis (70, 71), may thus be addressed in correlation with the temperature of the patient at the time of sampling, his/her age, gender, the ongoing evolution and final outcome of the disease.

CONCLUSION AND SUGGESTIONS

A large body of scientific evidence now indicates that the accumulation of cellular HSP70s, especially HSPA1A, in lung alveolar cells is beneficial against ARDS-induced lung damage, as in the case of the most severe COVID-19 pathologies. Because mild fever induces the HSR and the accumulation of cellular HSP70s, one would predict that a therapeutic strategy for fever should not be readily decreased by antipyretics. However, fever would optimally need to be thereafter artificially reduced by antipyretics because several hours at 37°C are needed to restore the cellular ability to produce more protective HSP70s in a subsequent fever cycle. Given that age and viral infection may decrease the basal cellular levels of anti-apoptotic HSP70s and further reduce the ability of lung alveolar cells to accumulate HSP70s under stress, we speculate therapeutic strategies should be sought to restore high HSP70 levels in the lung cells of COVID-19 patients. Prophylactic treatments in anticipation of the disease in the elderly could involve natural repetitive stimulations of the HSR in the whole body through controlled intense physical exercise (72–74), sauna therapies and the regular maintenance of calorie-restricted diets (75) containing minimal amounts of saturated lipids and cholesterol.

Interestingly, a prior period of heat acclimation was found in exercising humans to reduce physiological strain and improve physical performance when exercising in moderate normobaric hypoxia (76). Similar effects were shown in rats for which hypobaric hypoxia invoked a cardioprotective heat shock response, consisting of a significant upregulation of HSP70, HSP90, HSP60, and HSP27 that strongly contributed to their survival under acute sub-lethal hypoxia (77). It is tempting to hypothesize that seniors undergoing prior prophylactic treatments of both mild heat-shock and moderate hypoxia, as in daily intense exercising, might increase their ability, once infected, to withstand the increasing hypoxia associated to the acute phases of the disease.

For lack of yet an effective vaccine, the fundamental role in primary care of the COVID-19 crisis is the diagnosis of the suspected COVID-19 patients. In most developed countries this happens via phone calls to detect warning signs, mainly based on the detection of ARDS components and rarely based also on body temperature fluctuations. Given the emerging key role of fever-induced HSP70 expression in the possible mitigation of ARDS damages in SARS-CoV-2 patients, we more pragmatically advocate a systematic research to set precise criteria for temperature monitoring, as a diagnostic feature for initial telemedicine advises and periodic evaluations during self-isolation.

AUTHOR CONTRIBUTIONS

AG and BF analyzed published transcriptomic and proteomic data. AG, BF, and PG made the figures. All authors conceived the central ideas of the manuscript, interpreted data from literature, contributed to writing, reviewed, edited, and approved its final version of the manuscript.

FUNDING

This work was supported by the Swiss National Fund to PG (31003A_175453).

REFERENCES

1. Jamil T, Alam I, Gojobori T, Duarte CM. No Evidence for Temperature-Dependence of the COVID-19 Epidemic. *Front Public Health*. (2020) 8:436. doi: 10.3389/fpubh.2020.00436
2. Altamimi A, Ahmed AE. Climate factors and incidence of middle east respiratory syndrome coronavirus. *J Infect Public Health*. (2019) 13:704–8. doi: 10.1016/j.jiph.2019.11.011
3. Casanova LM, Jeon S, Rutala WA, Weber DJ, Sobsey MD. Effects of air temperature and relative humidity on coronavirus survival on surfaces. *Appl Environ Microbiol*. (2010) 76:2712–7. doi: 10.1128/AEM.02291-09
4. Bannister-Tyrrell M, Meyer A, Faverjon C, Cameron A. Preliminary evidence that higher temperatures are associated with lower incidence of COVID-19, for cases reported globally up to 29th February 2020. *medRxiv*. (2020) doi: 10.1101/2020.03.18.20036731
5. Wang J, Tang K, Feng K, Lv W. High temperature and high humidity reduce the transmission of COVID-19. *SSRN Electr J*. (2020) doi: 10.2139/ssrn.3551767
6. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. (2020) 395:1054–62. doi: 10.1016/S0140-6736(20)30566-3

7. Force ADT, Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, et al. Acute respiratory distress syndrome: the Berlin definition. *JAMA*. (2012) 307:2526–33. doi: 10.1001/jama.2012.5669
8. Tian S, Hu W, Niu L, Liu H, Xu H, Xiao SY. Pulmonary pathology of early-phase 2019 novel coronavirus (COVID-19) pneumonia in two patients with lung cancer. *J Thorac Oncol*. (2020) 15:700–4. doi: 10.1016/j.jtho.2020.02.010
9. Matthay MA, Zemans RL, Zimmerman GA, Arabi YM, Beitler JR, Mercat A, et al. Acute respiratory distress syndrome. *Nat Rev Dis Prim.* (2018) 5:18. doi: 10.1038/s41572-019-0069-0
10. Ramanathan K, Antognini D, Combes A, Paden M, Zakhary B, Ogino M, et al. Planning and provision of ECMO services for severe ARDS during the COVID-19 pandemic and other outbreaks of emerging infectious diseases. *Lancet Respir Med.* (2020) 8:518–26. doi: 10.1016/S2213-2600(20)30121-1
11. Prompetchara E, Ketloy C, Palaga T. Immune responses in COVID-19 and potential vaccines: lessons learned from SARS and MERS epidemic. *Asian Pac J Allergy Immunol*. (2020) 38:1–9. doi: 10.12932/AP-200220-0772
12. Ahn DG, Shin HJ, Kim MH, Lee S, Kim HS, Myoung J, et al. Current status of epidemiology, diagnosis, therapeutics, and vaccines for novel coronavirus disease 2019 (COVID-19). *J Microbiol Biotechnol*. (2020) 30:313–24. doi: 10.4014/jmb.2003.03011
13. Onder G, Rezza G, Brusaferro S. Case-fatality rate and characteristics of patients dying in relation to COVID-19 in Italy. *JAMA Mar.* (2020) 323:1775–6. doi: 10.1001/jama.2020.4683
14. Latz CA, DeCarlo C, Boitano L, Png CYM, Patell R, Conrad MF, et al. Blood type and outcomes in patients with COVID-19. *Ann Hematol.* (2020) 12:1–6. doi: 10.1007/s00277-020-04169-1
15. Wheeler DS, Wong HR. Heat shock response and acute lung injury. *Free Radic Biol Med.* (2007) 42:1–14. doi: 10.1016/j.freeradbiomed.2006.08.028
16. Barna J, Csermely P, Vellai T. Roles of heat shock factor 1 beyond the heat shock response. *Cell Mol Life Sci.* (2018) 75:2897–916. doi: 10.1007/s00018-018-2836-6
17. Swindell WR, Huebner M, Weber AP. Transcriptional profiling of arabidopsis heat shock proteins and transcription factors reveals extensive overlap between heat and non-heat stress response pathways. *BMC Genomics.* (2007) 8:125. doi: 10.1186/1471-2164-8-125
18. Saidi Y, Finka A, Muriset M, Bromberg Z, Weiss YG, Maathuis FJ, et al. The heat shock response in moss plants is regulated by specific calcium-permeable channels in the plasma membrane. *Plant Cell.* (2009) 21:2829–43. doi: 10.1105/tpc.108.065318
19. Evans EM, Doctor RJ, Gage BF, Hotchkiss RS, Fuller BM, Drewry AM. The association of fever and antipyretic medication with outcomes in mechanically ventilated patients: a cohort study. *Shock*. (2019) 52:152–9. doi: 10.1097/SHK.0000000000001368
20. Nagarsekar A, Tulapurkar ME, Singh IS, Atamas SP, Shah NG, Hasday JD. Hyperthermia promotes and prevents respiratory epithelial apoptosis through distinct mechanisms. *Am J Respir Cell Mol Biol.* (2012) 47:824–33. doi: 10.1165/rcmb.2012-0105OC
21. Hasday JD, Singh IS. Fever and the heat shock response: distinct, partially overlapping processes. *Cell Stress Chaperones.* (2000) 5:471–80. doi: 10.1379/1466-1268(2000)005<0471:fathsr>2.0.co;2
22. Evans SS, Repasky EA, Fisher DT. Fever and the thermal regulation of immunity: the immune system feels the heat. *Nat Rev Immunol.* (2015) 15:335–49. doi: 10.1038/nri3843
23. Schulman CI, Namias N, Doherty J, Manning RJ, Li P, Alhaddad A, et al. The effect of antipyretic therapy upon outcomes in critically ill patients: a randomized, prospective study. *Surg Infect*. (2005) 6:369–75. doi: 10.1089/sur.2005.6.369
24. Petitjeans F, Leroy S, Pichot C, Geloen A, Ghignone M, Quintin L. Hypothesis: fever control, a niche for alpha-2 agonists in the setting of septic shock and severe acute respiratory distress syndrome? *Temperature.* (2018) 5:224–56. doi: 10.1080/23328940.2018.1453771
25. Schell-Chaple HM, Puntillo KA, Matthay MA, Liu KD, Wiedemann HP, Arroliga AC, et al. Body temperature and mortality in patients with acute respiratory distress syndrome. *Am J Crit Care.* (2015) 24:15–23. doi: 10.4037/ajcc2015320
26. Finka A, Goloubinoff P. Proteomic data from human cell cultures refine mechanisms of chaperone-mediated protein homeostasis. *Cell Stress Chaperones.* (2013) 18:591–605. doi: 10.1007/s12192-013-0413-3
27. Diamant S, Ben-Zvi AP, Bukau B, Goloubinoff P. Size-dependent disaggregation of stable protein aggregates by the DnaK chaperone machinery. *J Biol Chem.* (2000) 275:21107–13. doi: 10.1074/jbc.M001293200
28. De Los Rios P, Goloubinoff P. Hsp70 chaperones use ATP to remodel native protein oligomers and stable aggregates by entropic pulling. *Nat Struct Mol Biol.* (2016) 23:766–9. doi: 10.1038/nsmb.3283
29. Weiss YG, Bromberg Z, Raj N, Raphael J, Goloubinoff P, Ben-Neriah Y, et al. Enhanced heat shock protein 70 expression alters proteasomal degradation of IκB kinase in experimental acute respiratory distress syndrome*. *Crit Care Med.* (2007) 35:2128–38. doi: 10.1097/01.CCM.0000278915.78030.74
30. Bromberg Z, Raj N, Goloubinoff P, Deutschman CS, Weiss YG. Enhanced expression of 70-kilodalton heat shock protein limits cell division in a sepsis-induced model of acute respiratory distress syndrome*. *Crit Care Med.* (2008) 36:246–55. doi: 10.1097/01.CCM.0000295473.56522.EF
31. Aschkenasy G, Bromberg Z, Raj N, Deutschman CS, Weiss YG. Enhanced Hsp70 expression protects against acute lung injury by modulating apoptotic pathways. *PLoS ONE.* (2011) 6:e26956. doi: 10.1371/journal.pone.0026956
32. Li X, Yu Y, Gorshkov B, Haigh S, Bordan Z, Weintraub D, et al. Hsp70 suppresses mitochondrial reactive oxygen species and preserves pulmonary microvascular barrier integrity following exposure to bacterial toxins. *Front Immunol.* (2018) 9:1309. doi: 10.3389/fimmu.2018.01309
33. Ben-Zvi A, Miller EA, Morimoto RI. Collapse of proteostasis represents an early molecular event in caenorhabditis elegans aging. *Proc Natl Acad Sci USA*. (2009) 106:14914–9. doi: 10.1073/pnas.0902882106
34. Goloubinoff P. Mechanisms of protein homeostasis in health, aging and disease. *Swiss Med Wkly*. (2016) 146:w14306. doi: 10.4414/smw.2016.14306
35. Saidi Y, Peter M, Fink A, Cicekli C, Vigh L, Goloubinoff P. Membrane lipid composition affects plant heat sensing and modulates Ca2+-dependent heat shock response. *Plant Signal Behav*. (2010) 5:1530–3. doi: 10.4161/psb.5.12.13163
36. Vigh L, Török Z, Crul T, Maresca B, Schütz GJ, Viana F, et al. Plasma membranes as heat stress sensors: from lipid-controlled molecular switches to therapeutic applications. *Biochim Biophys Acta Biomembr.* (2014) 1838:1594–618. doi: 10.1016/j.bbamem.2013.12.015
37. Perona JS. Membrane lipid alterations in the metabolic syndrome and the role of dietary oils. *Biochim Biophys Acta Biomembr.* (2017) 1859:1690–703. doi: 10.1016/j.bbamem.2017.04.015
38. Apostolopoulos A, Nakamura A, Yokoyama S, Aoshima M, Fujimoto R, Nakamura K, et al. Nuclear accumulation of HSP70 in mouse skeletal muscles in response to heat stress, aging, and unloading with or without reloading. *Front Genet.* (2018) 9:617. doi: 10.3389/fgene.2018.00617
39. Hall DM, Xu L, Drake VJ, Oberley LW, Oberley TD, Moseley PL, et al. Aging reduces adaptive capacity and stress protein expression in the liver after heat stress. *J Appl Physiol.* (2000) 89:749–59. doi: 10.1152/jappl.2000.89.2.749
40. Kayani AC, Morton JP, McArdle A. The exercise-induced stress response in skeletal muscle: failure during aging. *Appl Physiol Nutr Metab.* (2008) 33:1033–41. doi: 10.1139/H08-089
41. Sherman MY, Goldberg AL. Cellular defenses against unfolded proteins: a cell biologist thinks about neurodegenerative diseases. *Neuron.* (2001) 29:15–32. doi: 10.1016/S0896-6273(01)00177-5
42. Winklhofer KF, Tatzelt J, Haass C. The two faces of protein misfolding: gain- and loss-of-function in neurodegenerative diseases. *EMBO J.* (2008) 27:336–49. doi: 10.1038/sj.emboj.7601930
43. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. *Cell.* (2013) 153:1194–217. doi: 10.1016/j.cell.2013.05.039
44. Norman DC. Fever in the elderly. *Clin Infect Dis.* (2000) 31:148–51. doi: 10.1086/313896
45. Finka A, Sood V, Quadroni M, De Los Rios PDL, Goloubinoff P. Quantitative proteomics of heat-treated human cells show an across-the-board mild depletion of housekeeping proteins to massively accumulate few HSPs. *Cell Stress Chaperones.* (2015) 20:605–20. doi: 10.1007/s12192-015-0583-2
46. Richter K, Haslbeck M, Buchner J. The heat shock response: life on the verge of death. *Mol Cell.* (2010) 40:253–66. doi: 10.1016/j.molcel.2010.10.006
47. Bromberg Z, Weiss Y. The role of the membrane-initiated heat shock response in cancer. *Front Mol Biosci.* (2016) 3:12. doi: 10.3389/fmolb.2016.00012
48. Finka A, Cuendet AF, Maathuis FJ, Saidi Y, Goloubinoff P. Plasma membrane cyclic nucleotide gated calcium channels control land plant

thermal sensing and acquired thermotolerance. *Plant Cell.* (2012) 24:3333–48. doi: 10.1105/tpc.112.095844
49. Bromberg Z, Goloubinoff P, Saidi Y, Weiss YG. The membrane-associated transient receptor potential vanilloid channel is the central heat shock receptor controlling the cellular heat shock response in epithelial cells. *PLoS ONE.* (2013) 8:e57149. doi: 10.1371/journal.pone.0057149
50. Seifert A, Schofield P, Barton GJ, Hay RT. Proteotoxic stress reprograms the chromatin landscape of SUMO modification. *Sci Signal.* (2015) 8:rs7. doi: 10.1126/scisignal.aaa2213
51. Yuan S, Chu H, Chan JF, Ye ZW, Wen L, Yan B, et al. SREBP-dependent lipidomic reprogramming as a broad-spectrum antiviral target. *Nat Commun.* (2019) 10:120. doi: 10.1038/s41467-018-08015-x
52. Blanco-Melo D, Nilsson-Payant BE, Liu WC, Uhl S, Hoagland D, Moller R, et al. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell.* (2020) 181:1036–45.e1039. doi: 10.1016/j.cell.2020.04.026
53. Kiekkas P, Aretha D, Bakalis N, Karpouhtsi I, Marneras C, Baltopoulos GI. Fever effects and treatment in critical care: literature review. *Aust Crit Care.* (2013) 26:130–5. doi: 10.1016/j.aucc.2012.10.004
54. Faulkner SH, Jackson S, Fatania G, Leicht CA. The effect of passive heating on heat shock protein 70 and interleukin-6: a possible treatment tool for metabolic diseases? *Temperature.* (2017) 4:292–304. doi: 10.1080/23328940.2017.1288688
55. Salminen A, Lehtonen M, Paimela T, Kaarniranta K. Celastrol: molecular targets of thunder god vine. *Biochem Biophys Res Commun.* (2010) 394:439–42. doi: 10.1016/j.bbrc.2010.03.050
56. Wei Y, Wang Y. Celastrol attenuates impairments associated with lipopolysaccharide-induced acute respiratory distress syndrome (ARDS) in rats. *J Immunotoxicol.* (2017) 14:228–34. doi: 10.1080/1547691X.2017.1394933
57. Kunutsor SK, Laukkanen T, Laukkanen JA. Sauna bathing reduces the risk of respiratory diseases: a long-term prospective cohort study. *Eur J Epidemiol.* (2017) 32:1107–11. doi: 10.1007/s10654-017-0311-6
58. Laukkanen JA, Laukkanen T. Sauna bathing and systemic inflammation. *Eur J Epidemiol.* (2018) 33:351–3. doi: 10.1007/s10654-017-0335-y
59. Krüger K, Reichel T, Zeilinger C. Role of heat shock proteins 70/90 in exercise physiology and exercise immunology and their diagnostic potential in sports. *J Appl Physiol.* (2019) 126:916–27. doi: 10.1152/japplphysiol.01052.2018
60. Henstridge DC, Febbraio MA, Hargreaves M. Heat shock proteins and exercise adaptations. Our knowledge thus far and the road still ahead. *J Appl Physiol.* (2016) 120:683–91. doi: 10.1152/japplphysiol.00811.2015
61. Benatar M, Wuu J, Andersen PM, Atassi N, David W, Cudkowicz M, et al. Randomized, double-blind, placebo-controlled trial of arimoclomol in rapidly progressive SOD1 ALS. *Neurology.* (2018) 90:e565–74. doi: 10.1212/WNL.0000000000004960
62. Haldimann P, Muriset M, Vigh L, Goloubinoff P. The novel hydroxylamine derivative NG-094 suppresses polyglutamine protein toxicity in caenorhabditis elegans. *J Biol Chem.* (2011) 286:18784–94. doi: 10.1074/jbc.M111.234773
63. Thakur P, Nehru B. Long-term heat shock proteins (HSPs) induction by carbenoxolone improves hallmark features of parkinson's disease in a rotenone-based model. *Neuropharmacology.* (2014) 79:190–200. doi: 10.1016/j.neuropharm.2013.11.016
64. de Oliveira GP, Kitoko JZ, de Souza Lima-Gomes P, Rochael NC, de Araújo CC, Lugon PN, et al. Glutamine therapy reduces inflammation and extracellular trap release in experimental acute respiratory distress syndrome of pulmonary origin. *Nutrients.* (2019) 11:831. doi: 10.3390/nu11040831
65. Ray JJ, Schulman CI. Fever: suppress or let it ride? *J Thorac Dis.* (2015) 7:E633–36. doi: 10.3978/j.issn.2072-1439.2015.12.28
66. Kluger MJ. Phylogeny of fever. *Fed Proc.* (1979) 38:30–4.
67. Launey Y, Nesseler N, Mallédant Y, Seguin P. Clinical review: fever in septic ICU patients - friend or foe? *Crit Care.* (2011) 15:222. doi: 10.1186/cc10097
68. Muller-Redetzky HC, Felten M, Hellwig K, Wienhold SM, Naujoks J, Opitz B, et al. Increasing the inspiratory time and I:E ratio during mechanical ventilation aggravates ventilator-induced lung injury in mice. *Crit Care.* (2015) 19:23. doi: 10.1186/s13054-015-0759-2
69. Petroni RC, Biselli PJ, de Lima TM, Theobaldo MC, Caldini ET, Pimentel RN, et al. Hypertonic saline (NaCl 7.5%) reduces lps-induced acute lung injury in rats. *Inflammation.* (2015) 38:2026–35. doi: 10.1007/s10753-015-0183-4
70. Qin C, Sai XY, Qian XF, Wu Y, Zou LF, Wang HM, et al. Close relationship between cIAP2 and human ARDS induced by severe H7N9 infection. *Biomed Res Int.* (2019) 2019:2121357. doi: 10.1155/2019/2121357
71. Rodrigue-Gervais IG, Labbe K, Dagenais M, Dupaul-Chicoine J, Champagne C, Morizot A, et al. Cellular inhibitor of apoptosis protein cIAP2 protects against pulmonary tissue necrosis during influenza virus infection to promote host survival. *Cell Host Microbe.* (2014) 15:23–35. doi: 10.1016/j.chom.2013.12.003
72. Archer AE, Von Schulze AT, Geiger PC. Exercise, heat shock proteins and insulin resistance. *Philos Trans R Soc Lond B Biol Sci.* (2018) 373:20160529. doi: 10.1098/rstb.2016.0529
73. Athersuch TJ, Antoine DJ, Boobis AR, Coen M, Daly AK, Possamai L, et al. Paracetamol metabolism, hepatotoxicity, biomarkers and therapeutic interventions: a perspective. *Toxicol Res.* (2018) 7:347–57. doi: 10.1039/C7TX00340D
74. Yuan Y, Pan SS, Shen YJ. Cardioprotection of exercise preconditioning involving heat shock protein 70 and concurrent autophagy: a potential chaperone-assisted selective macroautophagy effect. *J Physiol Sci.* (2018) 68:55–67. doi: 10.1007/s12576-016-0507-7
75. Gat-Yablonski G, Finka A, Pinto G, Quadroni M, Shtaif B, Goloubinoff P. Quantitative proteomics of rat livers shows that unrestricted feeding is stressful for proteostasis with implications on life span. *Aging.* (2016) 8:1735–58. doi: 10.18632/aging.101009
76. Lee BJ, Miller A, James RS, Thake CD. Cross acclimation between heat and hypoxia: heat acclimation improves cellular tolerance and exercise performance in acute normobaric hypoxia. *Front Physiol.* (2016) 7:78. doi: 10.3389/fphys.2016.00078
77. Jain K, Suryakumar G, Ganju L, Singh SB. Differential hypoxic tolerance is mediated by activation of heat shock response and nitric oxide pathway. *Cell Stress Chaperones.* (2014) 19:801–12. doi: 10.1007/s12192-014-0504-9

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

CHAPTER 8

COVID and the Renin-Angiotensin System: Are Hypertension or Its Treatments Deleterious?

COVID and the Renin-Angiotensin System: Are Hypertension or Its Treatments Deleterious?

Florian Zores † and Mathieu E. Rebeaud†

†*Authors contributed equally to this work*

I contributed to this work conceptually and in the manuscript writing.

AUTHOR CONTRIBUTIONS

MR and FZ wrote the manuscript, conceptualized the idea, andmade the figure. All authors reviewed and approved the final version of the manuscript.

frontiers in Cardiovascular Medicine

OPINION
published: 23 April 2020
doi: 10.3389/fcvm.2020.00071

COVID and the Renin-Angiotensin System: Are Hypertension or Its Treatments Deleterious?

Florian Zores[1*†] *and Mathieu E. Rebeaud*[2†]

[1] *Groupe Médical Spécialisé, Strasbourg, France,* [2] *DBMV, Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland*

Keywords: COVID-19, SARS-CoV-2, ACE2, angiotensin, renin, 2019-nCoV

INTRODUCTION

Since its outbreak in December 2019, Severe Acute Respiratory Syndrome CoronaVirus 2 (SARS-CoV-2) has spread worldwide and is considered a pandemic. Coronavirus disease (COVID-19) can lead to acute respiratory distress syndrome (ARDS) or death. Many efforts have been made to identify risk factors predisposing to a severe issue. In the first SARS-CoV epidemic in 2002, hypertension was noted in 9/19 patients who died from SARS-CoV in Toronto (1). In the two largest cohorts of SARS-CoV-2 published, hypertension is the most common comorbidity in patients with severe disease or in those who died or were ventilated (2, 3). Nevertheless, these data are not adjusted for age, although age appears to be a strong predictor of adverse outcome (4) and hypertension is a very common finding in older patients. Finally, cohort studies only show correlation, not causality. In this paper, we hypothesize that the reductions in Angiotensin-Converting Enzyme 2 (ACE-2) observed in hypertension and obesity can explain many abnormalities observed in SARS-CoV-2 and question the role of treatments interfering with ACE2.

ACE2 IN THE CARDIOVASCULAR SYSTEM

Like SARS-CoV, SARS-CoV-2 fuses with human cells after the receptor-binding domain of its S (Spike) protein binds with Angiotensin-Converting Enzyme 2 (ACE-2), an enzyme located on membrane of lung alveolar epithelial cells, renal tubular epithelial cells, enterocytes of the small intestine, and arterial and venous endothelial cells of the kidney (5–10). Cardiomyocytes, fibroblasts, endothelial cells, and pericytes account for the vast majority of cells expressing ACE2 in the heart (10).

ACE-2 is a monocarboxypeptidase homologous to Angiotensin-Converting Enzyme (ACE) whose active site is exposed at the extracellular surface (8, 11). ACE cleaves angiotensin I (ANGI) to generate angiotensin II (ANGII), which binds to and activates Angiotensin Type 1 Receptor (AT1R) to constrict blood vessels and increase salt and fluid retention, thereby elevating blood pressure. ACE2 inactivates ANGII by converting it to angiotensin-(1–7), which has a vasodilator effect when binding to Mas receptor (12) (**Figure 1A**). Moreover, ACE2 cleaves ANGI into angiotensin-(1–9) (albeit with lower affinity than for ANGII), which is further converted into angiotensin-(1–7) by ACE (12). Thus, ACE2 negatively regulates the renin-angiotensin system and modulates the vasoconstriction, fibrosis, and hypertrophy induced by that system (8, 11). In rats, ACE2 deficiency worsens hypertension when ANGII is in excess (8, 13). In human, gene expression and/or ACE2 activity is lower in hypertensive patients than in normotensive ones (13).

OPEN ACCESS

Edited by:
Hendrik Tevaearai Stahel, Bern University Hospital, Switzerland

Reviewed by:
Jan Danser, Erasmus Medical Center, Netherlands

***Correspondence:**
Florian Zores
florian.zores@gmail.com

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Cardiovascular Therapeutics, a section of the journal Frontiers in Cardiovascular Medicine

Received: *02 April 2020*
Accepted: *14 April 2020*
Published: *23 April 2020*

Citation:
Zores F and Rebeaud ME (2020) COVID and the Renin-Angiotensin System: Are Hypertension or Its Treatments Deleterious? Front. Cardiovasc. Med. 7:71. doi: 10.3389/fcvm.2020.00071

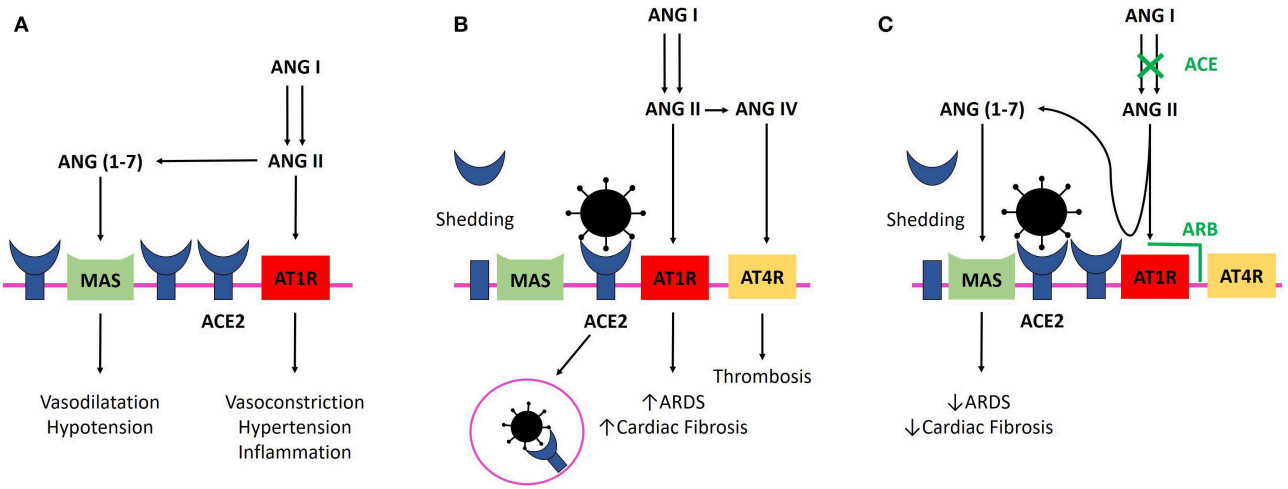


FIGURE 1 | (A) ANGII binding to AT1R elevates blood pressure and promotes inflammation. ACE2 inactivates ANGII by converting it to ang-(1–7) and negatively regulates the renin-angiotensin system, promoting vasodilatation and hypotension. **(B)** SARS-CoV-2 infection. Binding of SARS-CoV-2 with ACE2 leads to their internalization and to ACE2 shedding by ADAM17 (enzyme not shown). Lower availability of ACE2 results in a lower rate of ANGII degradation and excessive stimulation of AT1R, which facilitates ARDS and myocardial injury. Binding of ANGII to AT1R leads to membranous ACE2 internalization, decreasing ACE2 availability even more (not shown). Excessive ANGII is metabolized to ANGIV, which binds to AT4R and promotes thrombosis. Virus replication could also reduce cellular ACE2 expression (not shown). **(C)** SARS-CoV-2 infection and ACEi/ARB treatment. ACEi and ARB upregulate ACE2, and freer ACE2 remains after viral binding. ANGII is still degraded by ACE2 in its beneficial metabolite Ang-(1–7), and AT1R and AT4R are less stimulated. ANGII binding on AT1R prevention with ARB and ANGII synthesis decrease with ACE lead to less AT1R stimulation and persistent interaction with ACE2, avoiding ACE2 internalization. ACE2, Angiotensin Converting Enzyme 2; ACEi, Angiotensin Converting Enzyme Inhibitor; ang-(1–7), Angiotensin-(1–7); ANGII, angiotensin II; ANGIV, angiotensin IV; ARB, Angiotensin Receptor Blocker; ARDS, Acute Respiratory Distress Syndrome; AT1R, Angiotensin II Type 1 Receptor; AT4R, Angiotensin II Type 4 Receptor; SARS-CoV-2, Severe Acute Respiratory Syndrome CoronaVirus 2.

Conversely, ANGII negatively regulates ACE2. AT1R and ACE2 physically interact to form complexes on the cell membrane in the absence of excess Ang II (11). ANGII increase separates AT1R and ACE2 on the cell surface and leads to ACE2 internalization and lysosomal degradation through an AT1R-dependent mechanism (11, 13). Moreover, cellular ACE2 can be cleaved and released (shedding) by the metalloproteinase ADAM17, which is upregulated by ANGII (14). The soluble form of ACE2 circulates in small amounts in the blood, but its physiological role remains elusive, and shedding could be only a mechanism to regulate ACE2 activity on the cell surface (15).

Notably, it has been shown that infection with SARS-CoV can be blocked with soluble ACE2 molecules (6), and some have hypothesized that a soluble recombinant form can be used to overwhelm SARS-CoV-2 to prevent its binding to cellular ACE2 (16). Recombinant human ACE2 has been tested in a phase 2–3 trial in ARDS with interesting results (17), and a pilot trial has recently been launched in COVID-19 (NCT04287686).

ACE inhibitors (ACEi) and AT1R blockers (ARB) are two classes of drugs that are widely used in medicine to treat hypertension or heart failure. ACEi and ARB upregulate ACE2 expression on the cell surface, and ACE2 activity is not prevented by ACEi (8, 11, 18). Accordingly, patients treated with ACEi/ARB could have a higher level of membrane-bound ACE2, providing a more potent binding site to COVID-19 S protein. Nevertheless, in the absence of excess ANGII (either by reduction of ANGII synthesis by ACEi or by AT1R blockade thanks to ARB), AT1R is thought to interact with ACE2 (11). This interaction could reduce the affinity of COVID S protein to ACE2 and then reduce COVID-19 viral entry (11).

In the heart, ACE and ACE2 balance Ang II levels and ACE2 is known to be cardioprotective (8). ACE2 loss leads to a decrease in myocardial function in rodents, likely mediated by ANGII-induced oxidative stress and inflammation through AT1R, but it is unknown whether excess ANGII has a role in an acute setting (8, 19). This decrease is corrected by ARB or ACEi, and these drugs rapidly increase ACE2 activity and mRNA expression in the heart of rats (8, 20). Evidence for such an increase in humans is lacking, but studies checked for variation in the circulating level rather than the tissular level of ACE2 (21). In human failing heart, ACE2 expression is increased, correlating with disease severity, and is thought to be a compensatory mechanism (8, 10).

ROLE OF ACE2 IN SARS-COV-2 INFECTION

SARS-CoV-2 has a 10–20-fold higher affinity for ACE2 than does the 2002 SARS-CoV (22). An increased abundance of cellular ACE2 is associated with a higher susceptibility to SARS-CoV infection in mice (23). However, in both heart and lung, binding of the SARS-CoV to ACE2 leads to the loss of ACE2 by ACE2 internalization with the virus and ACE2 shedding (7, 9, 14). Lower availability of ACE2 results in a

lower rate of ANGII degradation. In rodent lungs, excess ANGII binding to AT1R increases pulmonary vascular permeability and neutrophil accumulation and enhances lung injury (7, 24) (**Figure 1B**). Thus, decreased ACE2 expression promotes increased lung injury and ARB prevents it by limiting ANGII binding to AT1R (7, 8, 24, 25) (**Figure 1C**). This hypothesis is supported *in vivo* by the increased frequency of severe ARDS in patients infected with SARS-CoV with higher levels of ACE determined by genetic predisposition, leading to higher levels of ANGII (26), and by the correlation between viral load, ANGII plasma level, and disease severity in influenza H7N5 (27) and respiratory syncytial virus infection (25). More notably, in a small cohort of patients infected with SARS-CoV-2, viral load was correlated with plasma ANGII level (28). Unfortunately, baseline treatments are unknown in this cohort, and correlation between ARDS severity and plasma ANGII level failed to reach statistical significance, maybe because of the low number of patients.

Moreover, some have suggested that viral replication by itself can reduce cellular ACE2 expression (29). This point is of importance because limitation of ANGII formation by ACEi and binding to AT1R by ARB may yet become the best ways to limit lung injuries if ACE2 is less or not synthetized following viral infection.

SARS-CoV- and SARS-CoV-2-associated cardiac injury contributes significantly to morbidity and mortality and could hit as much as a third of patients with a severe form of the disease (9, 28, 30, 31). SARS-CoV was found in the heart of a third of human autopsy hearts, with a concomitant marked reduction in cellular ACE2 (9). As in lungs, ANGII probably contributes to the deleterious effect of SARS-CoV on the heart and to SARS-associated cardiomyopathy, even if myocardial dysfunction can also be influenced by the strong immune response observed in those patients (9). Inflammatory signals are likely to suppress ACE2 transcription and down-regulate cell-surface expression of ACE2 (8). Thus, inflammatory signals could decrease the cellular susceptibility to SARS-CoV infection but increase the ANGII-mediated tissular injury. Moreover, because pericytes are supposed to play a role in myocardial microcirculation, SARS-CoV-2-induced microcirculation disorder could explain the frequent cardiac marker increase observed in hospitalized patients (2), exacerbated by the reduced oxygen supply caused by lung failure (10).

In summary, a decrease in cellular ACE2 may reduce the susceptibility of cells to SARS CoV-2 but leads to greater activation of AT1R and more severe tissue damage. In contrast, the higher the abundance of ACE2 on the cell membrane, the greater the susceptibility to viral particles but the less the damage, due to less AT1R activation occurring. This latter condition is the one provoked by ACEi/ARB treatment. On the one hand, ACE2 increase under ARB/ACEi treatment could be protective during COVID-19 because some ACE2 remains free to degrade ANGII, but on the other hand, this ACE2 increase could be deleterious by favoring cellular infection by COVID-19, leading to potent myocarditis (**Figure 1C**). The protective or deleterious role of ACEi/ARB in COVID-19 is harder to modelize, as ACE2 is not the only protein required for SARS-COV-2 penetration (5).

ARE ACEI AND ARB DELETERIOUS IN SARS-COV-2 INFECTION?

It has been shown that both ACEi and ARB upregulates ACE2, and a hypothesis was proposed by several authors of a potential deleterious effect of treatment with ARB and ACEi in the course of SARS-CoV-2 infection (32, 33). Since these molecules are widely used to treat hypertension or heart failure, such a fact could be a huge matter of concern.

Obesity seems to be a major determinant of adverse outcome in COVID-19 (34). Besides the altered pulmonary function associated with obesity, it must be noted that obesity is associated with a decrease in membranous ACE2 (35, 36). Moreover, empirical observations are suggestive of an abnormally high prevalence of pulmonary embolism in patients with COVID-19 (37), and prophylactic curative anticoagulation is recommended in severe patients (38). Severe infections are a known precipitant factor for acute venous thrombo-embolism because of epithelial damage and platelet and endothelial cell dysfunction, but does it by itself explain the observed high prevalence of pulmonary embolism in these patients? When ANGII is increased, it can be metabolized to angiotensin IV (ANGIV) by aminopeptidase A and binds to Angiotensin Type 4 Receptor (AT4R) (39). Multiple datasets underline the enhancement of thrombosis development by ANGII and ANGIV (40, 41), and it can be hypothesized that a reduction in ACE2 can increase thrombotic risk.

Despite the many potential cofounders, reduction in membranous ACE2 expression could be an explanation for numerous abnormalities observed in SARS-CoV-2 infection. Thus, even if both ARB and ACEi increase the level of ACE2, more ACE2 could be better rather than worse: more ACE2 remains on the cell surface after virus binding, maintaining ANGII degradation and less stimulation of AT1R. Furthermore, treatment with ARB inhibits AT1R and limits the damage induced by its overstimulation. It is not clear whether continuation or discontinuation of ARB or ACEi is a good option in COVID-19 infection, as there is a lack of clinical data to support an increased risk of contracting a severe form of COVID-19. In addition, we do not even know whether renin angiotensin system inhibitor therapy is beneficial or harmful for virally mediated lesions, and switching to other drugs may worsen the patient's condition, especially for heart failure patients with reduced ejection fraction (42). Clinical trials are ongoing to analyze the beneficial effect of LOSARTAN in COVID-19 (NCT04311177 and NCT04312009), and a trial will start soon to analyze the consequences of discontinuation or continuation of ACEi/ARB (NCT04338009).

ACEi and ARB are not the only treatments for hypertension or heart failure, but other classes only have a limited impact on ACE2. Beta blockers suppress plasma angiotensin II levels by inhibiting prorenin processing to renin and probably do not

interfere with ACE or ACE2 (43). Calcium channel blockers seem to reduce ANGII-induced downregulation of ACE2, but data are limited to those presented in one paper on the effect of nifedipine on fractionated cell extracts (44). In hypertensive rats, neither thiazides nor mineralocorticoid-receptor antagonists (MRAs) improve the spontaneous low ACE2 activity (18, 45), but MRA could decrease ACE expression (18). Conversely, MRAs increase membranous ACE2 activity in patients (46) with heart failure. If the reduction of membranous ACE2 observed in hypertension and obesity plays an important role in the pathophysiology of severe COVID-19, can it be hypothesized that non-ACEi/BRA drugs (beta-blockers, calcium channel blockers, diuretics) are more likely to increase the risk of deleterious outcomes than ACEi/BRA drugs that increase ACE2 and provide theoretical protection? Data on baseline treatments are urgently needed but are lacking to date in published cohorts.

CONCLUSION

The downregulation of ACE2 induced by viral binding, resulting in increased stimulation of AT1R, may be an important element in explaining severe COVID-19. Overall, the ACEi/ARB-mediated increase in ACE2 is not obviously deleterious and may even be protective. Only a well-conducted trial will provide a valid answer to this question. To date, stopping this treatment solely on the basis of presumed considerations does not seem to be a good option.

AUTHOR CONTRIBUTIONS

MR and FZ wrote the manuscript, conceptualized the idea, and made the figure. All authors reviewed and approved the final version of the manuscript.

REFERENCES

1. Farcas GA, Poutanen SM, Mazzulli T, Willey BM, Butany J, Asa SL, et al. Fatal severe acute respiratory syndrome is associated with multiorgan involvement by coronavirus. *J Infect Dis*. (2005) 191:193–7. doi: 10.1086/426870
2. Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med*. (2020). doi: 10.1056/NEJMoa2002032. [Epub ahead of print].
3. Grasselli G, Zangrillo A, Zanella A, Antonelli M, Cabrini L, Castelli A, et al. Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy Region, Italy. *JAMA*. (2020). doi: 10.1001/jama.2020.5394. [Epub ahead of print].
4. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*. (2020) 395:1054–62. doi: 10.1016/S0140-6736(20)30566-3
5. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*. (2020). doi: 10.1016/j.cell.2020.02.052. [Epub ahead of print].
6. Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature*. (2003) 426:450–4. doi: 10.1038/nature02145
7. Kuba K, Imai Y, Rao S, Gao H, Guo F, Guan B, et al. A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus–induced lung injury. *Nat Med*. (2005) 11:875–9. doi: 10.1038/nm1267
8. Kuba K, Imai Y, Ohto-Nakanishi T, Penninger JM. Trilogy of ACE2: a peptidase in the renin–angiotensin system, a SARS receptor, and a partner for amino acid transporters. *Pharmacol Ther*. (2010) 128:119–28. doi: 10.1016/j.pharmthera.2010.06.003
9. Oudit GY, Kassiri Z, Jiang C, Liu PP, Poutanen SM, Penninger JM, et al. SARS-coronavirus modulation of myocardial ACE2 expression and inflammation in patients with SARS. *Eur J Clin Invest*. (2009) 39:618–25. doi: 10.1111/j.1365-2362.2009.02153.x
10. Chen L, Li X, Chen M, Feng Y, Xiong C. The ACE2 expression in human heart indicates new potential mechanism of heart injury among patients infected with SARS-CoV-2. *Cardiovasc Res*. (2020) cvaa078. doi: 10.1093/cvr/cvaa078
11. Deshotels MR, Xia H, Sriramula S, Lazartigues E, Filipeanu CM. Angiotensin II mediates angiotensin converting enzyme type 2 internalization and degradation through an angiotensin II type I receptor–dependent mechanism. *Hypertension*. (2014) 64:1368–75. doi: 10.1161/HYPERTENSIONAHA.114.03743
12. Hamming I, Cooper M, Haagmans B, Hooper N, Korstanje R, Osterhaus A, et al. The emerging role of ACE2 in physiology and disease. *J Pathol*. (2007) 212:1–11. doi: 10.1002/path.2162
13. Soler MJ, Wysocki J, Batlle D. ACE2 alterations in kidney disease. *Nephrol Dial Transplant*. (2013) 28:2687–97. doi: 10.1093/ndt/gft320
14. Heurich A, Hofmann-Winkler H, Gierer S, Liepold T, Jahn O, Pöhlmann S. TMPRSS2 and ADAM17 cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome coronavirus spike protein. *J Virol*. (2014) 88:1293–307. doi: 10.1128/JVI.02202-13
15. Lambert DW, Hooper NM, Turner AJ. Angiotensin-converting enzyme 2 and new insights into the renin–angiotensin system. *Biochem Pharmacol*. (2008) 75:781–6. doi: 10.1016/j.bcp.2007.08.012
16. Batlle D, Wysocki J, Satchell K. Soluble angiotensin-converting enzyme 2: a potential approach for coronavirus infection therapy? *Clin Sci*. (2020) 134:543–5. doi: 10.1042/CS20200163
17. Khan A, Benthin C, Zeno B, Albertson TE, Boyd J, Christie JD, et al. A pilot clinical trial of recombinant human angiotensin-converting enzyme 2 in acute respiratory distress syndrome. *Crit Care*. (2017) 21:234. doi: 10.1186/s13054-017-1823-x
18. Takeda Y, Zhu A, Yoneda T, Usukura M, Takata H, Yamagishi M. Effects of aldosterone and angiotensin II receptor blockade on cardiac angiotensinogen and angiotensin-converting enzyme 2 expression in Dahl salt-sensitive hypertensive rats. *Am J Hypertens*. (2007) 20:1119–24. doi: 10.1016/j.amjhyper.2007.05.008
19. Crackower MA, Sarao R, Oudit GY, Yagil C, Kozieradzki I, Scanga SE, et al. Angiotensin-converting enzyme 2 is an essential regulator of heart function. *Nature*. (2002) 417:822–8. doi: 10.1038/nature00786
20. Burrell LM, Risvanis J, Kubota E, Dean RG, MacDonald PS, Lu S, et al. Myocardial infarction increases ACE2 expression in rat and humans. *Eur Heart J*. (2005) 26:369–75. doi: 10.1093/eurheartj/ehi114
21. Ferrario CM, Jessup J, Chappell MC, Averill DB, Brosnihan KB, Tallant EA, et al. Effect of angiotensin-converting enzyme inhibition and angiotensin II receptor blockers on cardiac angiotensin-converting enzyme 2. *Circulation*. (2005) 111:2605–10. doi: 10.1161/CIRCULATIONAHA.104.510461
22. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh C-L, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*. (2020) 367:1260–3. doi: 10.1126/science.abb2507
23. Yang X, Deng W, Tong Z, Liu Y, Zhang L, Zhu H, et al. *Mice Transgenic for Human Angiotensin-Converting Enzyme 2 Provide a Model for SARS Coronavirus Infection*. American Association for Laboratory Animal Science (2007).
24. Imai Y, Kuba K, Rao S, Huan Y, Guo F, Guan B, et al. Angiotensin-converting enzyme 2 protects from severe acute lung failure. *Nature*. (2005) 436:112–6. doi: 10.1038/nature03712
25. Gu H, Xie Z, Li T, Zhang S, Lai C, Zhu P, et al. Angiotensin-converting enzyme 2 inhibits lung injury induced by respiratory syncytial virus. *Sci Rep*. (2016) 6:19840. doi: 10.1038/srep19840

26. Itoyama S, Keicho N, Quy T, Phi NC, Long HT, Ha LD, et al. ACE1 polymorphism and progression of SARS. *Biochem Biophys Res Commun*. (2004) 323:1124–9. doi: 10.1016/j.bbrc.2004.08.208
27. Huang F, Guo J, Zou Z, Liu J, Cao B, Zhang S, et al. Angiotensin II plasma levels are linked to disease severity and predict fatal outcomes in H7N9-infected patients. *Nat Commun*. (2014) 5:1–7. doi: 10.1038/ncomms4595
28. Liu Y, Yang Y, Zhang C, Huang F, Wang F, Yuan J, et al. Clinical and biochemical indexes from 2019-nCoV infected patients linked to viral loads and lung injury. *Sci China Life Sci*. (2020) 63:364–74. doi: 10.1007/s11427-020-1643-8
29. Dijkman R, Jebbink MF, Deijs M, Milewska A, Pyrc K, Buelow E, et al. Replication-dependent downregulation of cellular angiotensin-converting enzyme 2 protein expression by human coronavirus NL63. *J Gen Virol*. (2012) 93:1924–9. doi: 10.1099/vir.0.043919-0
30. Arentz M, Yim E, Klaff L, Lokhandwala S, Riedo FX, Chong M, et al. Characteristics and outcomes of 21 critically ill patients with COVID-19 in Washington State. *JAMA*. (2020). doi: 10.1001/jama.2020.4326. [Epub ahead of print].
31. Shi S, Qin M, Shen B, Cai Y, Liu T, Yang F, et al. Association of cardiac injury with mortality in hospitalized patients with COVID-19 in Wuhan, China. *JAMA Cardiol*. (2020). doi: 10.1001/jamacardio.2020.0950. [Epub ahead of print].
32. Sommerstein, R, Gräni, C. Re: preventing a covid-19 pandemic: ACE inhibitors as a potential risk factor for fatal Covid-19. *BMJ*. (2020) 368:m810. doi: 10.1136/bmj.m810
33. Fang L, Karakiulakis G, Roth M. Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? *Lancet Respir Med*. (2020). doi: 10.1016/S2213-2600(20)30116-8. [Epub ahead of print].
34. Simonnet A, Chetboun M, Poissy J, Raverdy V, Noulette J, Duhamel A, et al. High prevalence of obesity in severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) requiring invasive mechanical ventilation. *Obesity*. (2020) doi: 10.1002/oby.22831. [Epub ahead of print].
35. Patel VB, Mori J, McLean BA, Basu R, Das SK, Ramprasath T, et al. ACE2 deficiency worsens epicardial adipose tissue inflammation and cardiac dysfunction in response to diet-induced obesity. *Diabetes*. (2016) 65:85–95. doi: 10.2337/dbi15-0037
36. Shoemaker R, Tannock LR, Su W, Gong M, Gurley SB, Thatcher SE, et al. Adipocyte deficiency of ACE2 increases systolic blood pressures of obese female C57BL/6 mice. *Biol Sex Differ*. (2019) 10:45. doi: 10.1186/s13293-019-0260-8
37. Chen J, Wang X, Zhang S, Liu B, Wu X, Wang Y, et al. *Findings of Acute Pulmonary Embolism in COVID-19 Patients*. Rochester, NY: Social Science Research Network. Report No.: ID 3548771 (2020).
38. Société Française d'Anesthésie et de Réanimation. *Traitement Anticoagulant pour la Prévention du Risque Thrombotique chez un Patient Hospitalisé avec Covid-19 et Surveillance de l'Hémostase*. Société Fr D'Anesthésie Réanimation.
39. Wolf G, Wenzel U, Assmann KJM, Stahl RAK. Renal expression of aminopeptidase A in rats with two-kidney, one-clip hypertension. *Nephrol Dial Transplant*. (2000) 15:1935–42. doi: 10.1093/ndt/15.12.1935
40. Mogielnicki A, Chabielska E, Pawlak R, Szemraj J, Buczko W. Angiotensin II enhances thrombosis development in renovascular hypertensive rats. *Thromb Haemost*. (2005) 93:1069–76. doi: 10.1160/TH04-10-0701
41. Senchenkova EY, Russell J, Esmon CT, Granger DN. Roles of coagulation and fibrinolysis in angiotensin II-enhanced microvascular thrombosis. *Microcirculation*. (2014) 21:401–7. doi: 10.1111/micc.12120
42. Halliday BP, Wassall R, Lota AS, Khalique Z, Gregson J, Newsome S, et al. Withdrawal of pharmacological treatment for heart failure in patients with recovered dilated cardiomyopathy (TRED-HF): an open-label, pilot, randomised trial. *Lancet*. (2019) 393:61–73. doi: 10.1016/S0140-6736(18)32484-X
43. Vilas-Boas WW, Ribeiro-Oliveira A Jr., da Cunha Ribeiro R, Vieira RLP, Almeida J, Nadu AP, et al. Effect of propranolol on the splanchnic and peripheral renin angiotensin system in cirrhotic patients. *World J Gastroenterol*. (2008) 14:6824–30. doi: 10.3748/wjg.14.6824
44. Iizuka K, Kusunoki A, Machida T, Hirafuji M. Angiotensin II reduces membranous angiotensin-converting enzyme 2 in pressurized human aortic endothelial cells. *J Renin Angiotensin Aldosterone Syst*. (2009) 10:210–5. doi: 10.1177/1470320309343710
45. Jessup JA, Brosnihan KB, Gallagher PE, Chappell MC, Ferrario CM. Differential effect of low dose thiazides on the renin angiotensin system in genetically hypertensive and normotensive rats. *J Am Soc Hypertens*. (2008) 2:106–15. doi: 10.1016/j.jash.2007.10.005
46. Keidar S, Gamliel-Lazarovich A, Kaplan M, Pavlotzky E, Hamoud S, Hayek T, et al. Mineralocorticoid receptor blocker increases angiotensin-converting enzyme 2 activity in congestive heart failure patients. *Circ Res*. (2005) 97:946–53. doi: 10.1161/01.RES.0000187500.24964.7A

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

CHAPTER 9

SARS-CoV-2 and the Use of Chloroquine as an Antiviral Treatment

SARS-CoV-2 and the Use of Chloroquine as an Antiviral Treatment

Mathieu E. Rebeaud† and Florian Zores†

†*Authors contributed equally to this work*

I contributed to this work conceptually and in the manuscript writing.

AUTHOR CONTRIBUTIONS
MR and FZ wrote the manuscript and collected the data of the existing clinical trials. All authors reviewed and approved the final version of the manuscript.

frontiers in Medicine

OPINION
published: 24 April 2020
doi: 10.3389/fmed.2020.00184

SARS-CoV-2 and the Use of Chloroquine as an Antiviral Treatment

Mathieu E. Rebeaud[1*†] *and Florian Zores*[2†]

[1] *DBMV, Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland,* [2] *Groupe médical spécialisé, Strasbourg, France*

Keywords: COVID-19, SARS-CoV-2, chloroquine, antiviral, 2019-nCoV

INTRODUCTION

In December 2019, a newly discovered SARS-CoV-2 virus emerged from China and propagated worldwide as a pandemic, becoming a major global public health issue. Different publications have discussed the possible efficacy of the antimalarial drug chloroquine (CQ) and its derivatives as a possible treatment against the disease, and, as the drug has often been recommended, we would like to shed a light on the previous experiments and trials conducted with CQ and its derivatives on several viruses, the outcomes being based on *in vitro* and *in vivo* results, and call for a well-designed clinical evaluation.

OPEN ACCESS

Edited by:
Zisis Kozlakidis, International Agency For Research On Cancer (IARC), France

Reviewed by:
Kevin Tan, National Neuroscience Institute (NNI), Singapore

****Correspondence:***
Mathieu E. Rebeaud
mathieu.rebeaud@unil.ch

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Infectious Diseases Ioroquine as a COVID-19 Treat- Surveillance, Prevention and Treatment, a section of the journal Frontiers in Medicine

Received: *31 March 2020*
Accepted: *17 April 2020*
Published: *24 April 2020*

Citation:
Rebeaud ME and Zores F (2020) SARS-CoV-2 and the Use of Chloroquine as an Antiviral Treatment. Front. Med. 7:184. doi: 10.3389/fmed.2020.00184

CHLOROQUINE, HYDROXYCHLOROQUINE, AND OTHER QUININE-DERIVATIVE DRUGS

As a semisynthetic derivative of quinine, CQ has for decades been the drug of choice to treat malaria because of its relative safety, good efficacy, and for being relatively inexpensive. CQ is a lysosome-penetrating antimalarial drug that neutralizes lysosomal acidification and prevents autophagosomal degradation. Hydroxychloroquine (HCQ) is a 4-aminoquinoline that differs by the addition of a hydroxyl group, decreasing its toxicity while conserving its efficacy. Nevertheless, CQ has a narrow therapeutic window and can cause life-threating cardiovascular issues, documented since the early 80s, especially for patients with underlying cardiac diseases (1). Cardiomyopathies, fatal arrhythmia, or even complete heart block have been described for 40 years, for chronic as well as acute treatment, even in patients with normal underlying cardiac function (2, 3). Another issue is represented by the possibility of vision-threatening toxic retinopathy (4). Thus, major contraindications are related to ocular (pre-existing maculopathy and retinopathy) and cardiac abnormalities [recent myocardial infarction and heart failure, corrected QT interval (QTc) >500 ms] but also include hypersensitivity to the active ingredient, porphyria, or glucose-6-phosphate dehydrogenase (G6PD) deficiency. It is also not recommended to combine these drugs with macrolides such as Azithromycin, which are known to have a synergistic effect on QTc prolongation, as QTc prolongation is associated with an increased risk of life-threatening arrhythmia (5). For the same reason, CQ and HCQ should not be used concomitantly with lopinavir/ritonavir and remdesivir. However, these drugs are not contraindicated during pregnancy (6).

SARS-COV-2

In December 2019, COVID19, a novel pneumonia caused by a previously unknown pathogen, emerged in Wuhan, China. The pathogen was soon identified as a novel coronavirus (2019-nCoV, later called SARS-CoV-2), closely related to the one responsible for severe acute respiratory syndrome SARS (SARS-CoV). SARS-CoV-2 infection is triggered by the binding of the spike

protein of the virus to angiotensin converting enzyme 2 (ACE2), which is highly expressed in the heart, gut, oral cavity, and lungs (7–9). SARS-CoV-2 mainly invades alveolar epithelial cells, resulting in respiratory symptoms. Briefly, in the cases where it is required, the median duration of hospitalization is 12 days (mean, 12.8) (10). Whereas, many people infected by SARS-CoV-2 develop mild, inconsequential respiratory symptoms, some individuals may develop more severe forms. During hospital stay, pneumonia is the most frequent diagnosis (91.1%), followed by acute respiratory distress syndrome (ARDS) (3.4%), but other organ dysfunctions can occur, leading to shock, multiple organ failure, and eventually death. Despite a lower case fatality rate than either SARS-CoV or Middle East respiratory syndrome-related coronavirus (MERS-CoV) (11, 12), the high number of infected patients can lead to a critical healthcare crisis, as depicted recently in China, Italy, France, and other countries. Currently, there is no specific treatment against the new virus other than supportive care. Therefore, identifying effective agents is urgently needed, either to combat the acute and severe forms of the disease, or to reduce infectiousness in less severe forms in order to reduce the burden for healthcare systems.

CHLOROQUINE AS A COVID-19 TREATMENT: *IN VITRO* AND *IN VIVO* DATA

CQ efficacy has been tested *in vitro* since the late 60s in different animal cells and viruses (13, 14). Thirty years ago, when comparing *in vitro* and *in vivo* trials and experiments, Hellgren et al. (15) already raised doubts concerning extrapolation drawn between the two systems and bench to bedside reproducibility. The sensitivity and therapeutic range of CQ, even in antimalarial treatment, cannot be easily derived from *in vitro* to *in vivo*. Hellgren et al. studied the *in vivo* response to a standard (25 mg/kg) dosage of chloroquine in a group of semi-immunized children from Tanzania. The EC99 (99% inhibition of schizont maturation) *in vitro* was 2.7 μg/L, and *in vivo* minimum inhibitory concentrations (MIC) median values were 44.29 (13–202; $n = 22$) μg/l, for a clearance of parasites, but recrudescence 1–4 weeks later and 237 (range 133–261; $n = 7$) μg/L for a response when parasitemia failed to clear after 1 week of treatment.

CQ, by inhibiting pH-dependent steps of the replication of several viruses, has already been quite extensively tested *in vitro* and *in vivo* on different virus strains: African swine fever virus (16), HIV (17), SARS-CoV (18, 19), Influenza A (20), Chikungunya (21), Ebola (22, 23), Zika (24), and, recently, on SARS-CoV-2 (25–27). Treatment with CQ has showed interesting results but also strong differences of application between live animals and cell lines. The major conclusion of these studies was that, if CQ exhibited promising results on virus and cells, the *in vivo* application is not that straightforward. In the case of Influenza A, the effectiveness of CQ *in vitro* on limiting the replication of viruses does not extend to *in vivo* models of influenza. For Ebola virus, the replication was inhibited by chloroquine *in vitro* but failed to protect Guinea pigs, mice, and hamsters. The most important warning on the difficulties to translate *in vitro* success into clinical reality is provided by the paradoxical results against Chikungunya. Despite inhibiting Chikungunya *in vitro*, CQ decreases cytokines levels and thus delays adaptive immune responses (28). De Lamballerie et al. (21) subsequently showed in a double blind randomized control trial that CQ has no more effect than a placebo in the acute phase but, in spite of this, increases late onset symptoms.

DISCUSSION

Despite these discrepancies between *in vitro* and *in vivo* data on all other tested viruses, CQ has been called a potential effective treatment for COVID-19. Many commentators have urged the use of CQ to lower the COVID-19 mortality rate after the publication of a Chinese expert consensus on CQ use in COVID-19 (29) and the result of a first trial (30). Nevertheless, this consensus did not provide any clinical data and is only based on *in vitro* assumptions. The trial by Gautret et al. suffers from several strong methodological problems, which preclude any conclusion (31). To date, only one small randomized unblinded prospective trial of 30 patients comparing CQ + standard of care vs. standard of care alone has been published and failed to show a difference between both arms for the primary endpoint [negative conversion rate of COVID-19 nucleic acid in respiratory pharyngeal swab on day 7 after randomization (32)].

Like in other major previous viral outbreak, treatment of COVID-19 is largely based on off-label, and compassionate therapies based on physiopathological or *in vitro* considerations. Likewise, is CQ, as suggested, a good treatment option given that it is presented as a well-known drug that has been used for decades? Thus, it is assumed it cannot be worse than the disease itself. For ethical reasons, this statement can equally be used to refute the need of a trial or the need of a control arm. Nevertheless, at the end of March, we counted 30 (**Table 1**) ongoing trials listed in Chinese, European, and US clinical trial registries, with a large variety found in the design or endpoint (EP).

Several drugs have failed in the past to confirm, in a randomized control trial, a putative efficacy seen in observational or phase 2 studies. Some have even been found to increase mortality despite promising results on physiological endpoints and safe use in other diseases. Since CQ has well-known potentially life-threatening cardiac side effects due to its quinidinic-like properties and the cardiac involvement of COVID-19 is now well-documented (33), the CAST study example (34) is of particular interest. It underlines the deleterious effect of class 1 antiarrhythmics in case of cardiac ischemia or left ventricular dysfunction despite its apparent safety in other medical conditions.

Some argue the mortality rate is too high to ethically run a controlled trial. Firstly, this assumes placebo is always worse than active treatment (that is untrue). Secondly, even if the global mortality rate is perceived as high because of the large number of infected patients, it is far lower than the terrible outcome associated with out-of-hospital cardiac arrest. Nevertheless, a randomized double-blind trial was performed to establish epinephrine effectiveness in out-of-hospital arrest (35). As reminded by Kalil in a recent paper (36), randomized control trials are the only way to precisely determine the harms of the

TABLE 1 | List of clinical trials listed in Chinese, European, and US clinical trial registries.

	Nb. of studies	Nb. of patients (median [IQR25-75])	Control arm*				Endpoint				
			SOC	Placebo	Active treatment	None	Viral surrogate EP	Length of stay	Clinical status	Time to improvement	Mortality
Cohort study	7 (23,3%)	100.0 [22.5; 129.0]	3	0	2 (antiviral)	2	4	2	1	0	0
Open-label randomized trial	17 (56,7%)	150.0 [100.0; 647.5]	10	0	10 (antiviral, immunomodulator, hydroxyCQ vs. CQ)	0	5	0	4	6	3
Single blind randomized trial	2 (6,7%)	125.0 [112.5; 137;5]	1	1	0	0	1	1	0	0	0
Double blind randomized trial	4 (13,3%)	278.0 [224;5; 372;5]	0	4	0	0	2	0	0	0	2
Total	30	150.0 [100.0; 440.0]	14	5	12	2	12	3	5	6	5

**some trials can have multiple control arms.*

drug and its safety in all medical situation and in the precise context of COVID-19. Only a quarter of ongoing trials are cohort studies, and a vast majority are controlled ones and will probably provide a good enough level of evidence for the effectiveness and the safety of CQ in COVID-19, if EPs are well-chosen.

A valuable EP is of particular importance to establish the efficiency. The first two published trials (30, 32) used a surrogate endpoint (the viral clearance). The sensitivity of SARS-COV2 PCR is quite low (37) and it can preclude any translation of the effectiveness on viral clearance to mortality or morbidity benefits. It is easier and less expensive to show that a treatment improves a surrogate endpoint than a clinical one (like clinical status or, at best, mortality). Nevertheless, a CAST trial showed us an improvement in a surrogate endpoint does not necessary translate into a decrease in clinical events or in mortality. As demonstrated many years ago by Prasad and Cifu (38), such surrogate endpoints, especially for unblinded trials, are the way to medical reversal and can lead to patient harm. The weaker the endpoint, the stronger the trial design to avoid inconclusive results. More importantly, falsely reassuring results based on surrogate endpoints can slow down the research of an effective treatment. Concerns have been raised about enrolment in the major European randomized trial DISCOVERY because of the mediatized claimed CQ effectiveness. Nine (69.2%) of the actual ongoing studies using a viral surrogate EP are of poor methodological quality (cohort studies or open-label trials) and will hardly give a valuable answer for the therapeutic value of CQ. All-cause mortality is the ideal endpoint but can be hard to reach due to economic, temporal, and demographic considerations. EPs, such as vital status evolution or length of stay, are more pragmatic to have a rapid and quite robust answer in a randomized trial and are used by near half of ongoing studies. Nevertheless, these EPs are potentially more subjective and more subject to bias than an objective one like death (39). Thus, particular attention should be paid to the design of these trials and the definitions of theses EP when interpreting the future results.

CONCLUSION

Since the late 60s, the option to use CQ and quinine derivative drugs as antivirals has been considered in a wide range of diseases (40). Based on the recent announcements of Gao et al. (25), Wang et al. (26), and Colson et al. (27), Chloroquine may be the first successful attempt to use this drug as an *in vivo* (human) antiviral.

However, despite the increased knowledge accumulated in recent decades, CQ has never been selected as a definitive or effective treatment in humans, as it failed to translate *in vitro* efficacity to *in vivo* efficiency. Moreover, the narrow therapeutic windows, along with possible side effects, have often interceded against its use. The ongoing SARS-CoV2 pandemic is a huge challenge for the whole world. Its relatively moderate mortality rate is aggravated by its high infectivity and the burden it causes on healthcare system in many countries. The will to give patients a treatment option even if proof is lacking is a human natural behavior in this time of need. Though scientific

precision may seem insensitive, it is the best way to avoid harming patients. Medical history is made of unmet hopes, and potential beneficial drugs have shown at best no effectiveness and have even been associated with increased adverse events. Failure to translate *in vivo* the *in vitro* success of CQ on Chikungunya is another reminder of the need of a careful clinical evaluation. To date, no published data support the use of CQ in COVID19. Well-designed clinical trials (randomized and controlled) with valuable and less as possible subjective EPs are urgently needed to clearly establish safety and effectiveness of quinine derivatives like Chloroquine as antiviral treatments.

AUTHOR CONTRIBUTIONS

MR and FZ wrote the manuscript and collected the data of the existing clinical trials. All authors reviewed and approved the final version of the manuscript.

ACKNOWLEDGMENTS

The authors would like to thank Christian Lehmann, MD; Julien Gere, MD; Filipe De Vadder, PhD; and Gwenaëlle Lemoine for constructive criticism of the manuscript.

REFERENCES

1. Frisk-Holmberg M, Bergqvist Y, Englund U. Chloroquine intoxication. *Br J Clin Pharmacol*. (1983) 15:502–3. doi: 10.1111/j.1365-2125.1983.tb01540.x
2. Piette JC, Guillevin L, Chapelon C, Wechsler B, Bletry O, Godeau P. Chloroquine cardiotoxicity. *N Engl J Med*. (1987) 317:710–1. doi: 10.1056/NEJM198709103171114
3. Don Michael TA, Aiwazzadeh S. The effects of acute chloroquine poisoning with special reference to the heart. *Am Heart J*. (1970) 79:831–42. doi: 10.1016/0002-8703(70)90371-6
4. Jorge A, Ung C, Young LH, Melles RB, Choi HK. Hydroxychloroquine retinopathy — implications of research advances for rheumatology care. *Nat Rev Rheumatol*. (2018) 14:693–703. doi: 10.1038/s41584-018-0111-8
5. Algra A, Tijssen JG, Roelandt JR, Pool J, Lubsen J. QTc prolongation measured by standard 12-lead electrocardiography is an independent risk factor for sudden death due to cardiac arrest. *Circulation*. (1991) 83:1888–94. doi: 10.1161/01.CIR.83.6.1888
6. Singh AK, Singh A, Shaikh A, Singh R, Misra A. Chloroquine and hydroxychloroquine in the treatment of COVID-19 with or without diabetes: a systematic search and a narrative review with a special reference to India and other developing countries. *Diabetes Metab Syndr*. (2020) 14:241–6. doi: 10.1016/j.dsx.2020.03.011
7. Turner AJ, Hiscox JA, Hooper NM. ACE2: from vasopeptidase to SARS virus receptor. *Trends Pharmacol Sci*. (2004) 25:291–4. doi: 10.1016/j.tips.2004.04.001
8. Hamming I, Timens W, Bulthuis ML, Lely AT, Navis G, van Goor H. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. a first step in understanding SARS pathogenesis. *J Pathol*. (2004) 203:631–7. doi: 10.1002/path.1570
9. Xu H, Zhong L, Deng J, Peng J, Dan H, Zeng X, et al. High expression of ACE2 receptor of 2019-nCoV on the epithelial cells of oral mucosa. *Int J Oral Sci*. (2020) 12:8. doi: 10.1038/s41368-020-0074-x
10. Guan WJ, Ni Z-Y, Hu Y, Liang W-H, Ou C-Q, He J-X, et al. Clinical Characteristics of coronavirus disease 2019 in China. *N Engl J Med*. (2020). doi: 10.1056/NEJMoa2002032
11. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the chinese center for disease control and prevention. *JAMA*. (2020) 13:1239–1242. doi: 10.1001/jama.2020.2648
12. Zheng YY, Ma Y-T, Zhang J-Y, Xie X. COVID-19 and the cardiovascular system. *Nat Rev Cardiol*. (2020). doi: 10.1038/s41569-020-0360-5
13. Inglot AD. Comparison of the antiviral activity *in vitro* of some non-steroidal anti-inflammatory drugs. *J Gen Virol*. (1969) 4:203–14. doi: 10.1099/0022-1317-4-2-203
14. Shimizu Y, Yamamoto S, Homma M, Ishida N. Effect of chloroquine on the growth of animal viruses. *Arch Gesamte Virusforsch*. (1972) 36:93–104. doi: 10.1007/BF01250299
15. Hellgren U, Kihamia CM, Mahikwano LF, Björkman A, Eriksson O, Rombo L. Response of plasmodium falciparum to chloroquine treatment: relation to whole blood concentrations of chloroquine and desethylchloroquine. *Bull World Health Organ*. (1989) 67:197–202.
16. Geraldes A, Valdeira ML. Effect of chloroquine on African swine fever virus infection. *J Gen Virol*. (1985) 66:1145–8. doi: 10.1099/0022-1317-66-5-1145
17. Tsai WP, Nara PL, Kung HF, Oroszlan S. Inhibition of human immunodeficiency virus infectivity by chloroquine. *AIDS Res Hum Retroviruses*. (1990) 6:481–9. doi: 10.1089/aid.1990.6.481
18. Keyaerts E, Vijgen L, Maes P, Neyts J, Van Ranst M. *In vitro* inhibition of severe acute respiratory syndrome coronavirus by chloroquine. *Biochem Biophys Res Commun*. (2004) 323:264–8. doi: 10.1016/j.bbrc.2004.08.085
19. Vincent MJ, Bergeron E, Benjannet S, Erickson BR, Rollin PE, Ksiazek TG, et al. Chloroquine is a potent inhibitor of SARS coronavirus infection and spread. *Virol J*. (2005) 2:69. doi: 10.1186/1743-422X-2-69
20. Vigerust DJ, McCullers JA. Chloroquine is effective against influenza a virus *in vitro* but not *in vivo*. *Influenza Other Respir Viruses*. (2007) 1:189–92. doi: 10.1111/j.1750-2659.2007.00027.x
21. De Lamballerie X, Boisson V, Reynier JC, Enault S, Charrel RN, Flahault A, et al. On chikungunya acute infection and chloroquine treatment. *Vector Borne Zoonotic Dis*. (2008) 8:837–9. doi: 10.1089/vbz.2008.0049
22. Dowall SD, Bosworth A, Watson R, Bewley K, Taylor I, Rayner E, et al. Chloroquine inhibited Ebola virus replication *in vitro* but failed to protect against infection and disease in the *in vivo* guinea pig model. *J Gen Virol*. (2015) 96:3484–92. doi: 10.1099/jgv.0.000309
23. Falzarano D, Safronetz D, Prescott J, Marzi A, Feldmann F, Feldmann H. Lack of protection against ebola virus from chloroquine in mice and hamsters. *Emerg Infect Dis*. (2015) 21:1065–7. doi: 10.3201/eid2106.150176
24. Delvecchio R, Higa LM, Pezzuto P, Valadão AL, Garcez PP, Monteiro FL, et al. Chloroquine, an endocytosis blocking agent, inhibits Zika virus infection in different cell models. *Viruses*. (2016) 8:322. doi: 10.3390/v8120322
25. Gao J, Tian Z, Yang X. Breakthrough: chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. *Biosci Trends*. (2020) 14:72–3 doi: 10.5582/bst.2020.01047
26. Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, et al. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) *in vitro*. *Cell Res*. (2020) 30:269–71. doi: 10.1038/s41422-020-0282-0
27. Colson P, Rolain JM, Raoult D. Chloroquine for the 2019 novel coronavirus SARS-CoV-2. *Int J Antimicrob Agents*. (2020) 55:105923. doi: 10.1016/j.ijantimicag.2020.105923
28. Roques P, Thiberville SD, Dupuis-Maguiraga L, Lum FM, Labadie K, Martinon F, et al. Paradoxical effect of chloroquine treatment in enhancing chikungunya virus infection. *Viruses*. (2018) 10:268. doi: 10.3390/v10050268
29. multicenter collaboration group of Department of S, P. Technology of Guangdong, and p. Health Commission of Guangdong Province for chloroquine in the treatment of novel coronavirus. [Expert consensus on chloroquine phosphate for the treatment of novel coronavirus pneumonia]. *Zhonghua Jie He He Hu Xi Za Zhi*. (2020) 43:185–8. doi: 10.3760/cma.j.issn.1001-0939.2020.03.009
30. Gautret P, Lagier JC, Parola P, Hoang VT, Meddeb L, Mailhe M, et al. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. *Int J Antimicrob Agents*. (2020) 105949. doi: 10.1016/j.ijantimicag.2020.105949. [Epub ahead of print].
31. Kim AHJ, Sparks JA, Liew JW, Putman MS, Berenbaum F, Duarte-García A, et al. A rush to judgment? Rapid reporting and dissemination of results and its consequences regarding the use of hydroxychloroquine for COVID-19.

Annal Internal Med. (2020) M20-1223. doi: 10.7326/M20-1223. [Epub ahead of print].
32. Chen Jun LD, Li L, Ping L, Qingnian X, Lu X, Yun L, et al. A pilot study of hydroxychloroquine in treatment of patients with common coronavirus disease-19 (COVID-19). *J Zhejiang Univ*. (2020) 49. doi: 10.3785/j.issn.1008-9292.2020.03.03
33. Madjid M, Safavi-Naeini P, Solomon SD, Vardeny O. Potential effects of coronaviruses on the cardiovascular system: a review. *JAMA Cardiol*. (2020). doi: 10.1001/jamacardio.2020.1286. [Epub ahead of print].
34. Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. the cardiac arrhythmia suppression trial. *N Engl J Med*. (1991) 324:781–8. doi: 10.1056/NEJM199103213241201
35. Perkins GD, Ji C, Deakin CD, Quinn T, Nolan JP, Scomparin C, et al. A randomized trial of epinephrine in out-of-hospital cardiac arrest. *N Eng J Med*. (2018) 379:711–21. doi: 10.1056/NEJMoa1806842
36. Kalil AC. Treating COVID-19—off-label drug use, compassionate use, and randomized clinical trials during pandemics. *JAMA*. (2020). doi: 10.1001/jama.2020.4742. [Epub ahead of print].
37. Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 Cases. *Radiology*. (2020) 200642. doi: 10.1148/radiol.2020200642
38. Prasad V, Cifu A. Medical reversal: why we must raise the bar before adopting new technologies. *Yale J Biol Med*. (2011) 84:471–8.
39. Evans SR. Fundamentals of clinical trial design. *J Exp Stroke Translat Med*. (2010) 3:19–27. doi: 10.6030/1939-067X-3.1.19
40. Touret F, de Lamballerie X. Of chloroquine and COVID-19. *Antiviral Res*. (2020) 177:104762. doi: 10.1016/j.antiviral.2020.104762

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

CHAPTER 10

Diabetes, hypertension, body mass index, smoking and COVID-19- related mortality: a systematic review and meta-analysis of observational studies

Diabetes, hypertension, body mass index, smoking and COVID-19- related mortality: a systematic review and meta-analysis of observational studies

Yahya Mahamat-Saleh, Thibault Fiolet, Mathieu Edouard Rebeaud, Matthieu Mulot, Anthony Guihur, Douae El Fatouhi, Nasser Laouali, Nathan Peiffer-Smadja, Dagfinn Aune, Gianluca Severi

I contributed to this work conceptually and in the manuscript writing.

Contributors

YM-S conceived and designed the research. YM-S performed statistical analysis and wrote the first draft of the paper. YM-S, TF and MER performed the literature search and literature screening. YM-S and TF assessed the risk of bias of the studies and assessed the certainty of evidence of the associations. MM assisted with the statistical analysis. YM-S, TF, MER, MM, AG, DEF, NL, NP-S, DA and GS contributed to data interpretation and revision of the manuscript for intellectual content. YM-Stakes responsibility for the integrity of the data and the accuracy of the data analysis. YM-S is responsible for the overall content as guarantor.

BMJ Open

Diabetes, hypertension, body mass index, smoking and COVID-19-related mortality: a systematic review and meta-analysis of observational studies

Yahya Mahamat-Saleh ,[1] Thibault Fiolet,[1] Mathieu Edouard Rebeaud,[2] Matthieu Mulot,[3] Anthony Guihur ,[2] Douae El Fatouhi,[1] Nasser Laouali,[1] Nathan Peiffer-Smadja,[4,5,6] Dagfinn Aune ,[7,8,9,10] Gianluca Severi[1,11]

To cite: Mahamat-Saleh Y, Fiolet T, Rebeaud ME, *et al*. Diabetes, hypertension, body mass index, smoking and COVID-19-related mortality: a systematic review and meta-analysis of observational studies. *BMJ Open* 2021;**11**:e052777. doi:10.1136/bmjopen-2021-052777

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-052777).

DA and GS are joint senior authors.

Received 26 April 2021
Accepted 07 October 2021

For numbered affiliations see end of article.

Correspondence to
Dr Yahya Mahamat-Saleh; Yahya.MAHAMAT-SALEH@gustaveroussy.fr

ABSTRACT

Objectives We conducted a systematic literature review and meta-analysis of observational studies to investigate the association between diabetes, hypertension, body mass index (BMI) or smoking with the risk of death in patients with COVID-19 and to estimate the proportion of deaths attributable to these conditions.

Methods Relevant observational studies were identified by searches in the PubMed, Cochrane library and Embase databases through 14 November 2020. Random-effects models were used to estimate summary relative risks (SRRs) and 95% CIs. Certainty of evidence was assessed using the Cochrane methods and the Grading of Recommendations, Assessment, Development and Evaluations framework.

Results A total of 186 studies representing 210 447 deaths among 1 304 587 patients with COVID-19 were included in this analysis. The SRR for death in patients with COVID-19 was 1.54 (95% CI 1.44 to 1.64, I^2=92%, n=145, low certainty) for diabetes and 1.42 (95% CI 1.30 to 1.54, I^2=90%, n=127, low certainty) for hypertension compared with patients without each of these comorbidities. Regarding obesity, the SSR was 1.45 (95% CI 1.31 to 1.61, I^2=91%, n=54, high certainty) for patients with BMI ≥30 kg/m^2 compared with those with BMI <30 kg/m^2 and 1.12 (95% CI 1.07 to 1.17, I^2=68%, n=25) per 5 kg/m^2 increase in BMI. There was evidence of a J-shaped non-linear dose–response relationship between BMI and mortality from COVID-19, with the nadir of the curve at a BMI of around 22–24, and a 1.5–2-fold increase in COVID-19 mortality with extreme obesity (BMI of 40–45). The SRR was 1.28 (95% CI 1.17 to 1.40, I^2=74%, n=28, low certainty) for ever, 1.29 (95% CI 1.03 to 1.62, I^2=84%, n=19) for current and 1.25 (95% CI 1.11 to 1.42, I^2=75%, n=14) for former smokers compared with never smokers. The absolute risk of COVID-19 death was increased by 14%, 11%, 12% and 7% for diabetes, hypertension, obesity and smoking, respectively. The proportion of deaths attributable to diabetes, hypertension, obesity and smoking was 8%, 7%, 11% and 2%, respectively.

Conclusion Our findings suggest that diabetes, hypertension, obesity and smoking were associated with higher COVID-19 mortality, contributing to nearly 30% of COVID-19 deaths.

Trial registration number CRD42020218115.

Strengths and limitations of this study

- ► We did a systematic review and meta-analysis of 186 observational studies (210 447 deaths among 1 304 587 patients with COVID-19) to investigate the association between diabetes, hypertension, obesity and smoking and risk of death in patients with COVID-19 and to estimate the proportion of deaths attributable to these four conditions.
- ► We performed several subgroup analyses, and assessed certainty of evidence using the Cochrane methods and the Grading of Recommendations, Assessment, Development and Evaluations framework. A non-linear dose–response relation between body mass index and COVID-19 mortality was also explored.
- ► However, we did not investigate the association between presence of two or more coexisting comorbidities and risk of death in patients with COVID-19.

INTRODUCTION

COVID-19 is a viral infectious disease caused by SARS-CoV-2, which was first reported in Wuhan City, China, in December 2019.[1] SARS-CoV-2 has since spread to all countries worldwide and COVID-19 has been declared a pandemic by the WHO.[2] As of 24 August 2021, over 212.3 million cases and 4.4 million deaths have been reported globally since the start of the pandemic.[3] Age is the main risk factor for poor outcome in people with COVID-19 infection,[4 5] as it is correlated with more comorbidities. About 70%–87% of COVID-19 deaths are among people aged 70 years or older.[6 7] Patients with comorbidities, including diabetes, cardiovascular disease, respiratory disease, chronic kidney disease and others chronic diseases are at increased risk of developing severe or critical COVID-19,[8 9] which may partly explain a greater mortality in hospital.[10–14] Studies suggest that

about 20%–51% of patients hospitalised with COVID-19 have at least one comorbidity.[10 15]

Previous meta-analyses reported a higher mortality rate from COVID-19 in patients with comorbidities.[16–22] Ssentongo *et al*, based on 25 studies published from December 2019 to 9 July 2020, suggested that diabetes and hypertension were respectively associated with a 1.48-fold and 1.82-fold greater risk of COVID-19 death compared with those without these comorbidities.[16] Based on studies published during the same period, Luo *et al* reported similar results.[17] Du *et al* found that patients with obesity had a 2.68-fold risk of dying from COVID-19 compared with non-obese patients.[18] Most of the published meta-analyses did not investigate the shape of the dose–response relationship between body mass index (BMI) and risk of death in order to clarify whether the association is dose-dependent or if there are threshold effects.[20 23–25] In addition, evidence suggests that smoking may increase risk of severe disease and death from COVID-19.[19] However, it is not clear whether such an increase in COVID-19 mortality is different in current and past smokers since previous meta-analyses have not performed separate analyses.[19 26 27]

However, since the publication of these meta-analyses, several observational studies have been published on diabetes, hypertension, obesity or smoking and risk of death in patients with COVID-19[4 28–122] and the strength of the associations differed greatly between studies. Moreover, the proportion of deaths attributable to diabetes, hypertension, obesity or smoking habits has not been estimated. This last aspect may help adapting public health measures and vaccination strategies to populations at risk of severe COVID-19.

Given the rapidly increasing death from COVID-19 globally, and since diabetes, hypertension, obesity and tobacco smoking represent the most important public health problems worldwide, which contributed to higher risk of death globally; we thus conducted a systematic review and meta-analysis of published observational studies to investigate the association between diabetes, hypertension, smoking and obesity and risk of death in patient with COVID-19 and to clarify the strength of these associations. We further estimated the proportion of deaths attributable to these conditions.

MATERIALS AND METHODS

Search strategy and selection criteria

The meta-analysis was reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement[123] and Meta-Analysis of Observational Studies in Epidemiology.[124] This study was registered and accepted in the International Prospective Register of PROSPERO in October 2020. PubMed (MEDLINE), Cochrane library and Embase databases were searched to identify relevant articles published in English from December 2019 to 14 November 2020. The search terms that we used are provided in online supplemental file. In addition, we searched the reference lists of the relevant publications, reviews and meta-analyses to identify additional potentially relevant studies. We only included observational studies (cohort studies and cross-sectional) that reported relative risk estimates (such as hazard ratios (HRs), relative risk (RR) or odds ratios (ORs)) with the 95% CIs with or without adjustment for potential confounders. The search was independently screened by two researchers (YM-S and TF) and discrepancies were resolved by discussion with a third researcher (MER).

Data collection

From each included publication, we extracted results and study characteristics which included first author's last name, publication year, country where the research was conducted in, study design, study description or name, study period, sample size with number of deaths, exposure, categories, risk estimate and 95% CIs, and adjustment factors. Data were extracted by YM-S and extractions were checked for accuracy by TF. Discrepancies were resolved through discussion with a third researcher (MER).

Quality assessment and risk of bias

The quality of individual studies was assessed independently by two researchers (YM-S and TF) using the Cochrane risk of bias tool ROBINS-I, which grades studies on a scale from critical risk of bias to low risk of bias considering bias due to confounding, selection of study participants, exposure measurement, misclassification of exposure during follow-up, missing data, measurement of outcomes and bias due to selection of reported results.[125] Following the assessment of risk of bias, the body of evidence for each comorbidity and risk of death was rated independently using the Grading of Recommendations, Assessment, Development and Evaluations approach.[126] Discrepancies were resolved through discussion with a third researcher (DA).

Data analysis

We used random effects models that consider both within study and between-study variation to calculate summary RRs (SRRs) (95% CIs) of COVID-19 mortality for patients with diabetes compared with those without diabetes, for patients with hypertension versus those without hypertension, for obese versus non-obese and for current, former and ever smoker compared with never smokers. The natural logarithm of the RRs was weighted using random effects weights.[127] Statistical heterogeneity between studies was assessed by the Cochrane Q test and the I^2 statistic.[128] We calculated the absolute risk difference (RD) from the baseline risk of mortality (BR) from Docherty *et al*,[129] large cohort and relative risk (RR) using the formula RD=BR×(RR−1).[130]

We further performed a dose–response analysis for the associations between BMI and COVID-19 mortality using the method described by Greenland and Longnecker to compute the linear trend from the natural logs of the RRs and CIs across categories of BMI.[131] We calculated SRRs

and 95% CIs for a 5-unit increment in BMI using random effects models. This method required mean or median of BMI, RRs and 95% CIs for at least three categories. The mean or median BMI level per category was used if provided in the publication, and if not, the midpoint of the upper and lower boundaries was estimated as a range in each category. When the highest and lowest categories were open-ended, we used the width of the adjacent interval to estimate the upper and lower boundaries for the category. For studies that reported results separately for young and adults, for current and former smoker, but not overall, we pooled the results using a fixed-effects model as reported by the Hamling procedure to obtain an overall estimate to be used in the meta-analysis.[132]

To explore the potential non-linear dose–response relation between BMI and mortality among patients with COVID-19, we used fractional polynomial models.[133] We determined the best fitting second order fractional polynomial regression model, defined as the one with the lowest deviance. Only studies which presented more than two categories of BMI were included in the non-linear analysis. Subgroup and meta-regression analyses were conducted to investigate potential sources of heterogeneity. Small-study effects, such as publication bias, were visually assessed by examining funnel plots for asymmetry, and with Egger's test,[134] and the results were considered to indicate potential small-study bias when p values were <0.10. We conducted sensitivity analyses excluding one study at a time to clarify whether the results were driven by one large study or a study with an extreme result.

We finally calculated the population attributable fraction (PAF) of mortality among patients with COVID-19 due to diabetes, hypertension, obesity and smoking, worldwide using the following formula[135]:

$$PAF = \frac{p \times (RR-1)}{[p \times (RR-1)+1]}$$

Where RR was the relative risk, p was the prevalence of the exposure in patient with COVID-19. The prevalence of diabetes (11.5%), hypertension (22.9%), obesity (29%) and smoking (9%) were obtained from previous meta-analyses.[136–138]

Patient and public involvement

Patients or the public were not involved in any aspect of the study design, conduct or in the development of the research question or outcome measures.

RESULTS

A total of 6007 records were identified in MEDLINE, Cochrane library and in EMBASE (figure 1). A total of 4665 publications were excluded after reading title and abstract or because of duplicates. Among 1342 full-text articles retrieved, 994 were excluded as not meeting the inclusion criteria, leaving a total of 348 publications. Of these, 162 articles were not eligible because they lacked sufficient data,[139 140] reported no risk estimate or irrelevant data,[139 140] because they had identical populations[141–151] or were retracted.[152] Finally, a total of 186 observational studies were included in this meta-analysis. Of the included studies, 58 were from Europe, 58 from North America, 60 from Asia, 6 from South America and 4 from Africa (online supplemental table 1). From the 186 publications assessed using the ROBINS-I tool, 92 were evaluated as being at low risk of bias, 49 at moderate risk of bias, 21 at serious risk of bias and 24 at critical risk of bias because of insufficient adjustment of relevant confounders (online supplemental table 2).

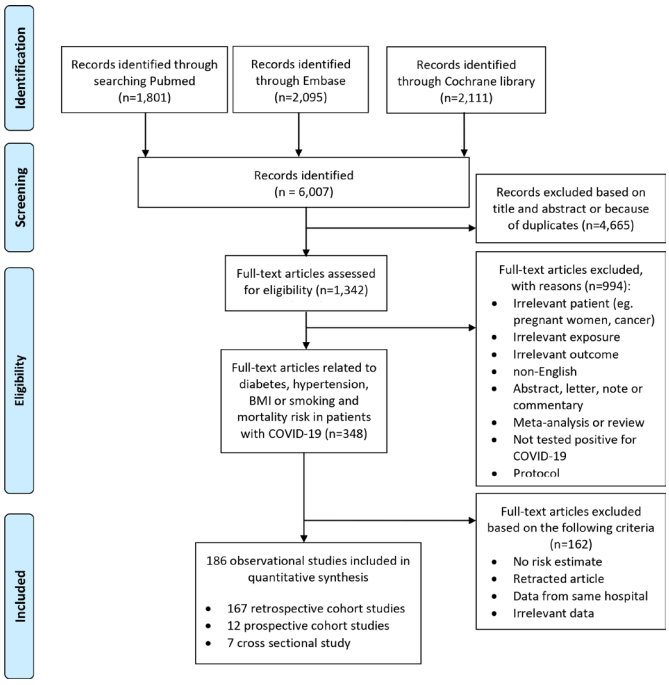


Figure 1 Flow-chart of study selection.

Diabetes and mortality in patient with COVID-19

A total of 145 studies[4 28–119 121 122 129 153–201] were included in the analysis of the association between diabetes and mortality, including a total of 198 491 deaths among 1 165 897 patients with COVID-19. The SRR for diabetes patients compared with those without diabetes was 1.54 (95% CI 1.44 to 1.64) and there was high heterogeneity among studies (I^2=92%). We found that diabetes increased the absolute risk of death by 14% (table 1). The funnel plot did not provide evidence of publication bias (online supplemental figure 1) and we found no evidence of small study effects (Egger's test, p value=0.54).

The positive association persisted across all subgroups analyses stratified by study design, number of patients, geographic location and adjustment for some confounding factors (table 2). However, meta-regression analyses suggested some indication of heterogeneity between studies that adjusted for multiple factors such as age, sex and other comorbidities versus studies without such adjustment or adjusted for age only ($P_{heterogeneity}$=0.003), with a stronger association for the studies with no adjustment factors or those adjusted for age only. The

Table 1 Summary relative risk, risk difference and fraction of deaths attributable to diabetes, hypertension, obesity and smoking in patients with COVID-19

	Number of studies	SRR (95% CI)	I² (%)	Risk difference (95% CI) (%)	PAF (95% CI) (all included studies) (%)	PAF (95% CI) (for fully adjusted studies) (%)	Number of patients	Number of deaths
Prognostic factors								
Diabetes	145	1.54 (1.44 to 1.64)	92	+14 (+11 to +17)	10.6 (8.8 to 12.4)	8.1 (6.2 to 9.7)	1 165 897	198 491
Hypertension	127	1.42 (1.30 to 1.54)	90	+11 (+8 to +14)	11.8 (8.7 to 14.7)	6.9 (4.0 to 9.6)	934 958	113 243
Obesity	54	1.45 (1.31 to 1.61)	91	+12 (+8 to +16)	11.5 (8.2 to 15.0)	11.1 (7.1 to 15.0)	858 374	145 605
Ever smoking	28	1.28 (1.17 to 1.40)	74	+7 (+4 to +10)	2.45 (1.5 to 3.5)	2.0 (0.8 to 3.3)	47 096	11 333

I^2 (%) is a measure of the proportion of the heterogeneity attributed to between study variation rather than due to chance. I^2 values of 25%, 50% and 75% indicates low, moderate and high between study heterogeneity, respectively.
CI, confidence interval; PAF, population attributable fraction of mortality; SRR, summary relative risk.

positive association was also stronger among studies with critical or serious risk of bias compared with those with moderate or low risk of bias, with heterogeneity detected across studies ($P_{heterogeneity}$=0.001).

The influence analysis showed no substantial influence of any of the included studies on the global estimate of diabetes and mortality. The SRR ranged from 1.51 (95% CI 1.42 to 1.61) when excluding the Barbu *et al* study[35] to 1.55 (95% CI 1.45 to 1.64) when excluding the Meng *et al* study[114] (online supplemental figure 2).

Hypertension and mortality in patient with COVID-19

We included 127 observational studies[4 29 30 32 33 35–40 42–45 47–50 52–56 58 59 61–63 65 66 69 71–76 78 81–84 86 88–90 92–95 97 103 105 106 108–112 114–118 120–122 153 154 156–159 161–167 169 173–180 182–188 193–200 202–219] in the analysis of the association between hypertension and mortality, including a total of 113 243 deaths among 934 958 patients with COVID-19. The SRR for hypertension patients versus those without hypertension was 1.42 (95% CI 1.30 to 1.54) with high evidence of heterogeneity (I^2=90%) (table 1). The absolute risk of death for patients with COVID-19 with hypertension compared with without hypertension was increased by 11%. Small study effects, such as publication bias, were not indicated with the funnel plot (online supplemental figure 3) or Egger's test (p value=0.26).

Here again, the positive association persisted in most subgroup analyses (table 2). Our meta-regression analysis showed that the study design and the geographic location did not significantly influence the magnitude of the overall association. However, heterogeneity between subgroup analyses was observed in analyses stratified by risk of bias with higher association for studies with critical compared with those with low, moderate or serious risk of bias ($P_{heterogeneity}$≤0.0001) and by adjustment for confounding factors with higher association for studies without adjustment compared with those that adjusted for age only or for multiple factors ($P_{heterogeneity}$≤0.0001).

When excluding the most influential studies, we found no substantial influence of any of the included studies (online supplemental figure 4).

BMI and mortality in patient with COVID-19

Fifty-four studies[29 32 38 44 45 48 49 52 61 62 64 65 67 69 73 74 79 84–86 88–90 93 99 103 107 118 120–122 129 154 164 165 178 181 182 184 192 196 197 202 206 220–228] investigated BMI and mortality risk including 145 605 deaths among 858 374 patients with COVID-19 were identified. The analysis yielded an SRR of 1.45 (95% CI 1.31 to 1.61) for obese (BMI ≥30 kg/m^2) versus non-obese (BMI <30 kg/m^2) patients, with high heterogeneity detected between studies (I^2=91%) (figure 2) (table 1). We found that obesity increased the absolute risk of death by 12%. There was no statistically significant evidence of publication bias (p value Egger's test=0.92) (online supplemental figure 5). The meta-regression analysis showed that study characteristics did not significantly alter the overall estimate (online supplemental table 3). When excluding the most influential studies one by one,

Table 2 Subgroup analyses of association between diabetes and hypertension and mortality risk in patients with COVID-19

	Diabetes					Hypertension				
	n	SRR (95% CI)	I^2 (%)	P_{within}*	$P_{between}$†	n	SRR (95% CI)	I^2 (%)	P_{within}*	$P_{between}$†
All studies	145	1.54 (1.44 to 1.64)	91.7	<0.0001		127	1.42 (1.30 to 1.54)	90.1	<0.0001	
Study design										
Retrospective	129	1.54 (1.44 to 1.64)	91.7	<0.0001		112	1.42 (1.30 to 1.56)	90.9	<0.0001	
Prospective	11	1.28 (1.12 to 1.48)	70.0	<0.0001		9	1.21 (0.98 to 1.50)	73.4	<0.0001	
Cross-sectional	5	2.47 (1.56 to 3.94)	81.4	<0.0001	0.05	6	1.70 (1.18 to 2.44)	61.8	0.02	0.50
Geographical location										
North America	45	1.33 (1.19 to 1.48)	90.3	<0.0001		40	1.30 (1.12 to 1.50)	93.4	<0.0001	
South America	6	1.54 (1.27 to 1.87)	93.7	<0.0001		2	2.03 (0.99 to 4.18)	69.1	0.072	
Europe	46	1.53 (1.36 to 1.72)	93.8	<0.0001		43	1.36 (1.20 to 1.54)	82.8	<0.0001	
Asia	43	1.94 (1.65 to 2.29)	75.1	<0.0001		38	1.62 (1.28 to 2.04)	86.1	<0.0001	
Africa	5	1.56 (0.92 to 2.62)	87.3	<0.0001	0.10	4	1.08 (0.86 to 1.35)	14.9	0.318	0.33
Number of patients										
<1000	79	1.73 (1.52 to 1.97)	74.8	<0.0001	0.001	78	1.56 (1.36 to 1.79)	74.8	<0.0001	
≥1000	66	1.43 (1.32 to 1.55)	95.4	<0.0001		49	1.28 (1.14 to 1.44)	94.8	<0.0001	<0.0001
Patients admission unit										
Non-ICU admitted	142	1.55 (1.45 to 1.66)	91.6	<0.0001	0.24	125	1.43 (1.31 to 1.56)	90.2	<0.0001	0.28
ICU admitted	3	1.22 (1.14 to 1.30)	0.0	0.81		2	0.98 (0.82 to 1.17)	0.0	0.81	
Risk of bias										
Low	72	1.46 (1.34 to 1.59)	92.4	<0.0001		63	1.17 (1.08 to 1.29)	80.2	<0.0001	
Moderate	37	1.28 (1.19 to 1.39)	64.9	<0.0001		34	1.34 (1.13 to 1.59)	88.9	<0.0001	
Serious	17	1.90 (1.44 to 2.51)	86.4	<0.0001		15	1.83 (1.42 to 2.37)	84.4	<0.0001	
Critical	19	2.11 (1.61 to 2.77)	93.4	<0.0001	0.001	15	2.64 (1.61 to 4.34)	96.3	<0.0001	<0.0001
Adjustment for confounders										
No	21	2.09 (1.62 to 2.69)	89.4	<0.0001		17	2.54 (1.62 to 3.99)	95.9	<0.0001	
Age only	18	1.90 (1.56 to 2.31)	85.4	<0.0001	0.003	14	1.78 (1.36 to 2.32)	85.1	<0.0001	
Multiple	106	1.40 (1.30 to 1.49)	87.5	<0.0001		96	1.23 (1.13 to 1.33)	85.4	<0.0001	<0.0001

I^2 (%) is a measure of the proportion of the heterogeneity attributed to between study variation rather than due to chance. I^2 values of 25%, 50% and 75% indicates low, moderate and high between study heterogeneity, respectively.
*P value for heterogeneity within each subgroup.
†P value for heterogeneity between subgroups with meta-regression analysis.
CI, confidence interval; ICU, intensive care unit; SRR, summary relative risk.

the global estimate did not substantially change (online supplemental figure 6).

Twenty-five studies (32 072 deaths among 95 852 patients with COVID-19)[38 56 61 74 86 89 91–93 97 108 111 112 161 162 175 177 221 222 229–235] were included in the dose–response meta-analysis of BMI and mortality risk. The summary RR for a 5 kg/m^2 increment in BMI was 1.12 (95% CI 1.07 to 1.17, I^2=68%) (figure 3) and no statistically significant evidence of publication bias (p value=0.11) or by inspection of the funnel plot was observed (online supplemental figure 7). However, evidence of a J-shaped non-linear relation between BMI and mortality risk was observed ($P_{non\text{-}linearity}$≤0.0001), suggesting a flat dose–response curve at a BMI around 22–24 kg/m^2 with a slight increase in risk of death below that range and a 1.5–2-fold increase in risk of death with a BMI of 40–45 versus 22–24 kg/m^2 (figure 4).

In sensitivity analyses excluding one study at a time from the analysis, the summary for a 5 kg/m^2 increment in BMI ranged from 1.11 (95% CI 1.06 to 1.15) when excluding the Czernichow *et al.* Study[108] to 1.13 (95% CI 1.08 to 1.19) when excluding the Ferrando-Vivas *et al* study[230] (online supplemental figure 8).

Smoking and mortality in patient with COVID-19

Twenty-eight studies[38 40 58 61 63 66 67 76 93 97 103 108 110 111 118 121 129 161 164–166 175 177 183 232 236–238] were included in the analysis of ever smoking versus never smoking and mortality with a total of 11 333 deaths among 47 096 patients with COVID-19. The SRR for hospital death in patient with COVID-19 was 1.28 (95% CI 1.17 to 1.40, I^2=74%) for ever smokers versus never smokers (figure 5). The absolute risk of death for smoking was increased by 7% (table 1). There was no evidence of publication or small study bias

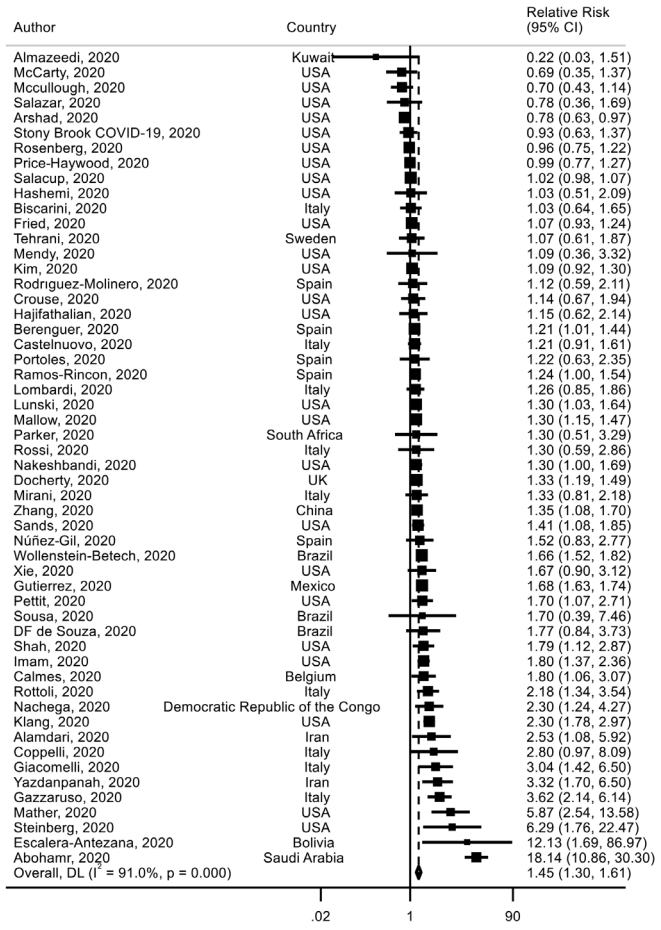


Figure 2 Association between obesity and mortality risk in patients with COVID-19.

(p value=0.91) (online supplemental figure 9). In sensitivity analyses, the results persisted when excluding one study at a time (online supplemental figure 10).

For current smoking versus never smoking, nineteen studies,[4 58 66 76 97 103 110 111 117 118 121 129 161 162 165 177 203 236 239] which included 9845 deaths among 33 147 patients with COVID-19, were identified. The SRR of current smoking was 1.29 (95% CI 1.03 to 1.62, I^2=84%) (figure 6) and no statistically significant evidence of publication bias was observed (p value=0.86) (online supplemental figure 11).

A total of 14 studies[58 66 76 97 103 110 111 118 121 129 161 165 177 236] were included in the analysis of former smoking versus never smoking and mortality risk, including 8121 deaths among 25 340 patients with COVID-19. The SRR was 1.25 (95% CI 1.11 to 1.42) with moderate to high heterogeneity (I^2=75%) (figure 7). There was no evidence of publication bias with Egger's test (p value=0.70). In sensitivity analyses excluding one study at a time from the analyses of current and former smoking, the results were not materially altered (online supplemental figures 12-13). Globally, results did not change in nearly all subgroup analyses (online supplemental table 4).

Finally, a total of 15 studies[29 44 52 58 107 174 186 192 208 218 240–243] did not provide a definition of the smoking variable (>76 400 deaths, 682 310 patients) and the SRR was 1.31 (95% CI 1.07 to 1.62, I^2=88%) (online supplemental figure 14) which seemed to be similar to current and ever smoking.

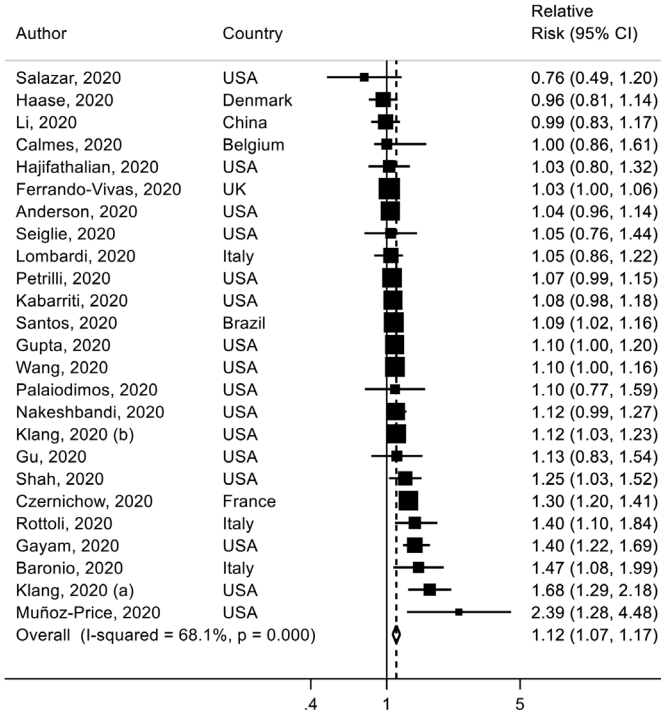


Figure 3 Association between body mass index (BMI) and mortality risk in patients with COVID-19, linear dose–response analysis, per 5 kg/m^2 increment of BMI.

PAF of deaths and assessment of certainty of the body of evidence

The estimated PAF was 10.6%, 11.8%, 11.5% and 2.5% for diabetes, hypertension, obesity and ever smoking, respectively when considering all studies included in this meta-analysis (table 1). Based on studies that adjusted for multiple risk factors, attributable death was 8% for

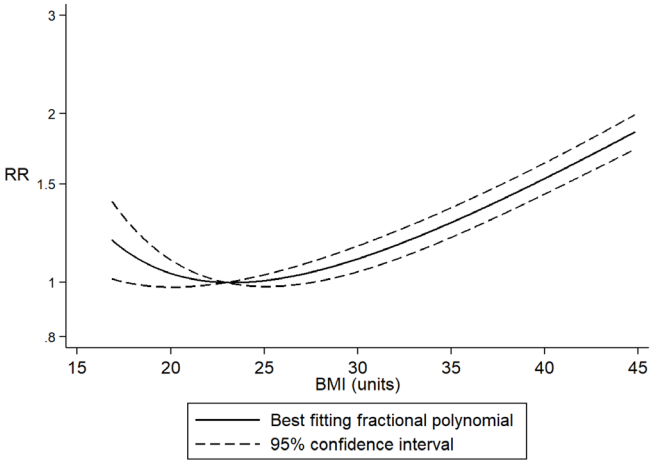


Figure 4 Association between body mass index (BMI) and mortality risk in patients with COVID-19: non-linear dose–response analysis.

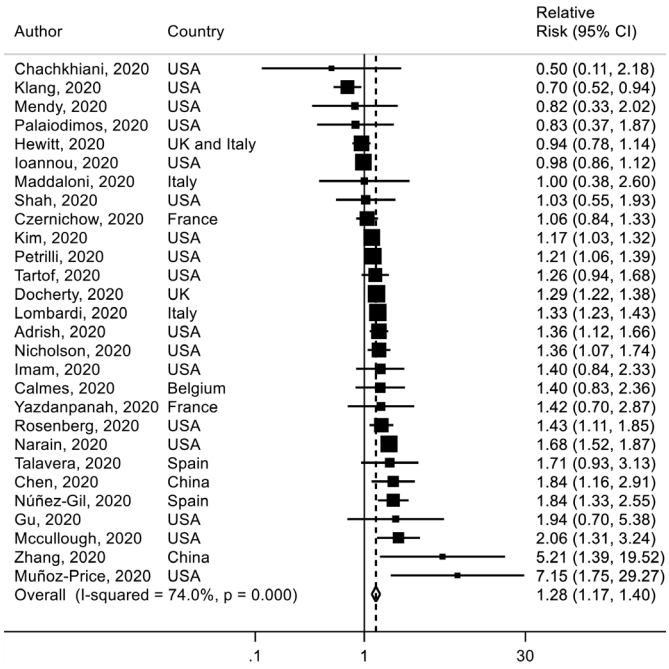


Figure 5 Association between ever smoking (vs never smoking) and mortality risk in patients with COVID-19.

diabetes, 7% for hypertension, 11% for obesity and 2% for ever smoking.

Regarding the body of evidence, certainty of evidence was rated 'low' for diabetes, hypertension and smoking and 'high' for obesity (online supplemental table 5).

DISCUSSION

Principal findings

To our knowledge, this study is one of the largest meta-analyses of association between diabetes, hypertension, obesity, smoking and COVID-19 mortality. We found that

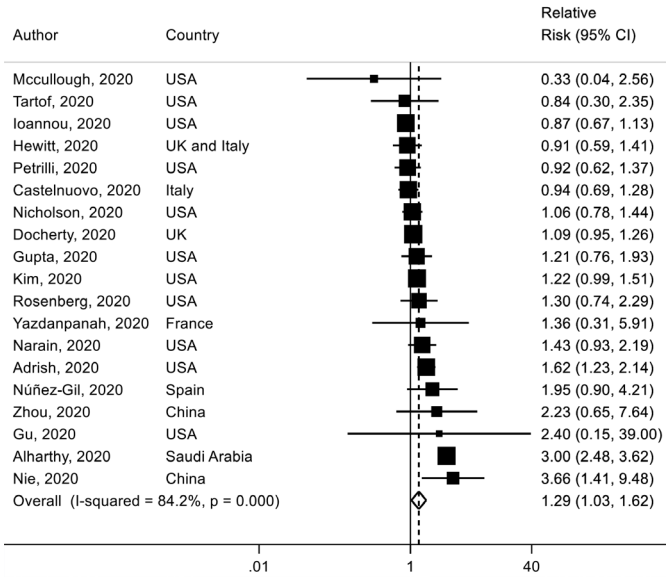


Figure 6 Association between current smoking (vs never smoking) and mortality risk in patients with COVID-19.

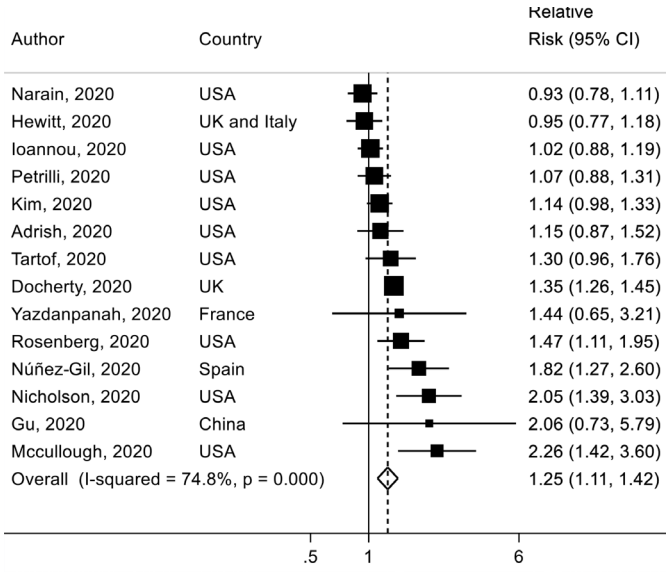


Figure 7 Association between former smoking (vs never smoking) and mortality risk in patients with COVID-19.

patients with diabetes had a 54% higher risk of death from COVID-19 compared with patients without diabetes; those with hypertension had a 42% increase in the relative risk of death from COVID-19 compared with patients without hypertension and those with obesity have a 45% greater in the relative risk of COVID-19 death compared with non-obese patients. In addition, we found that ever, current and former smoking was associated with 28%, 29% and 25% increases in the relative risk of death in patients with COVID-19. Our linear dose–response meta-analysis suggested that each 5 kg/km^2 increment in BMI was associated with a 12% greater risk of COVID-19 death. However, evidence of non-linearity was observed in the analysis of BMI and risk of COVID-19 death, with a J-shaped dose–response relation with flattening of the dose–response curve between 22 and 24 of BMI level and a slight increase below that range and a 1.5–2-fold increase in risk with a BMI of 40–45. While there was no publication bias, study heterogeneity was high for all exposure and this persisted in most of the subgroup analyses. However, the heterogeneity appeared to be driven to a larger extent by differences in the strength of the associations, than differences in the direction of the effect, as the vast majority of studies reported significant or non-significant positive associations between these exposures and increased mortality, and relatively few studies reported risk estimates in the direction of an inverse association. Given this meta-analysis included more studies than a typical meta-analysis, I^2 and heterogeneity were high as the likelihood of divergent findings increases with increasing number of studies.

Comparisons to findings from previous epidemiological studies and biological mechanisms

Since the first reports of COVID-19, several studies have shown that patients with COVID-19 with comorbidities have a higher risk of death.[4 17 244 245] However, these studies

have differed greatly in term of sample size, and results are conflicting and heterogeneous. Previous meta-analyses have shown that patients with diabetes, hypertension and obesity had an increased risk of mortality.[16–18 21 246 247] Unfortunately, these previous meta-analyses were limited by the lack of subgroup analyses, which is crucial to evaluate heterogeneity and no previous meta-analysis has estimated the number of deaths attributed to these comorbidities or conditions.

This meta-analysis summarises the results of 186 observational studies published up to November 2020, including 210 447 deaths among 1 304 587 million patients with COVID-19. Our findings are similar to results from previously published systematic reviews, suggesting a higher mortality rate of COVID-19 in patients with cardiovascular or chronic condition.[16 17 21 22] We found a 1.54-fold greater mortality from COVID-19 among patient with diabetes compared with those without (n=145 studies), which is similar to those yielded in previous meta-analyses,[16] whereas the estimate magnitude is weaker than those in de Almeida-Pititto *et al*,[21] Kumar *et al*,[246] Shang *et al*[248] and Luo *et al*.[17] In our subgroup analyses, we found that the positive association was stronger in studies without any adjustment or adjusted for age only compared with studies that adjusted for multiple factors. Thus, the higher magnitude observed in the previous meta-analyses may be in part due to the important number of studies without adjustment for confounding factors such as age and comorbidities, which are mostly with critical or serious risk of bias. In addition, we estimated that 8% of deaths by COVID-19 were attributed to diabetes; this aspect has to our knowledge not been investigated previously. The absolute risk of death associated with diabetes was increased by 14%. The mechanism underlying the increased mortality from COVID-19 in patients with diabetes may be explained by chronic inflammatory conditions. Patients with COVID-19 with diabetes have a significantly higher inflammatory markers such as C reactive protein (CRP), interleukin 6 (IL-6) compared with patients without diabetes.[249] Inflammatory markers such as IL-6, CRP, IL-10, lactate dehydrogenase and tumor necrosis factor-α, which are indicative of different aspects of COVID-19 severity, requirements of intensive care support including dialysis and ventilation, are associated with higher risk of death.[250 251]

Regarding hypertension, our findings yielded a 1.42-fold higher risk of death from COVID-19 in patient with hypertension (n=127 studies), which also is slightly weaker than the results from previous meta-analyses.[16 22 247] Nearly 7% of death in patients with COVID-19 could be attributed to hypertension and we found that hypertension increased the absolute risk of death by 11%. Previous clinical studies showed that hypertension is a major risk factor for worse outcome in patients infected with SARS and Middle East respiratory syndrome.[252 253] Although, the exact mechanism by which hypertension increase mortality rate remains unclear, chronic inflammation may play an active role in increasing risk of death.

Du *et al* suggested that patients with obesity had a 2.68-fold risk for COVID-19 mortality compared with non-obese patients (n=7 studies).[18] In dose–response analysis, they showed that for each 1 kg/m^2 increase in BMI, the risk of death increased by 6%. Partly consistent with this study, we found that a 5 kg/m^2 increase in BMI level was associated with a 12% increase in the risk of death in patient with COVID-19 (n=25 studies). In addition, we found evidence of a non-linear J-shaped association between BMI and mortality from COVID-19, with a flattening of the dose–response curve for BMI values between 22 and 24 kg/m^2 and with a slightly higher risk below that range and a moderate to strong increase in mortality with severe obesity (BMI 40–45). Obesity is associated with a low-grade systemic inflammation, which plays a major role in the pathogenesis of respiratory conditions. Patients with COVID-19 and with preexisting obesity may have an overactivated inflammation response, which may induce excessive inflammatory response.[254] Obesity is also strongly associated with increased risk of diabetes, hypertension and several other chronic diseases[255] that increases risk of COVID-19 mortality.[256]

In addition, our finding showed that ever, current and former smoking was associated with 28%, 29% and 25% increases in the relative risk of death in patients with COVID-19 compared with never smokers. Tobacco smoking is known to alter the function of the immune system; therefore smokers are more likely to get a severe infection of COVID-19 due to their poor mucociliary clearance, which could lead to the release of pro-inflammatory markers and oxidative stress and thereby contribute to higher mortality rates. The risk of death in former smokers was higher than in never smokers and only slightly lower than risk in current smoker. However, more studies are required to clarify the impact of longer durations of smoking cessation in former smokers on risk of death in patients with COVID-19.

Although we found that diabetes, hypertension, BMI and smoking were associated with greater COVID-19 mortality, a recent meta-analysis suggested that mortality was more frequently observed in patients with COVID-19 with cardiovascular disease, cerebrovascular accident and chronic kidney disease.[257] The authors observed that COVID-19 mortality among all comorbidities was high in European and Latin American patients compared with the US patients. It is possible that geographical differences in therapeutic practice of COVID-19 such as the use of antibiotics, antivirals and others drugs may partly explain the greater COVID-19 death in some regions,[258] while there was no evidence of heterogeneity in findings across geographic location in our study. The review also suggested that COVID-19 mortality among those with underlying medical diseases was high in mostly elderly patients.[258] However, we did not perform subgroup analysis by age because this information was lacking in most of the included studies.

Strengths, limitations and public health implications

This present meta-analysis of observational studies on diabetes, hypertension, obesity and smoking and risk of death in patients with COVID-19 has several strengths, including the large sample size and number of COVID-19 deaths, the detailed subgroup and sensitivity analyses, as well as the linear and non-linear dose–response analyses, which clarified the strength and shape of the dose–response relationship. Original aspects of our study included the estimation of the number of deaths attributable to these conditions. This is, to our knowledge, the first meta-analysis that perform a separate analysis of ever, current and former smoking versus never smoking in relation to COVID-19 mortality. As any previous published meta-analyses, the current analysis has some limitations that should be considered in the interpretation of our findings. First, we did not investigate the association between presence of two or more coexisting comorbidities and risk of death in patients with COVID-19. Finally, subgroup analyses stratified by clinical or lifestyle factors such as medications, diabetes type and duration, adherence to specific diet, or physical activity were not possible because of the lack of such data from the studies included.

Despite these limitations, our findings may have important public health implications in the context of increasing numbers of severe COVID-19 cases, overburdened hospitals and leading to higher hospital death due to COVID-19 and suggest that people with cardiovascular risk factors, in particular those with diabetes, hypertension and obesity, should be considered as a high priority to get vaccinated. In addition, since smoking is a risk factor for several chronic diseases, including cancer and cardiovascular disease, our finding lend support to the importance of smoking prevention and smoking cessation and support policies and public health efforts to reduce the prevalence smoking in the general population.

CONCLUSION

Our finding suggests that presence of diabetes, hypertension, obesity and smoking in patients with COVID-19 is associated with a 1.54-fold, 1.42-fold, 1.45-fold and 1.28-fold greater risk of mortality, respectively. We have found that the proportion of death attributable to diabetes, hypertension, obesity and smoking was 8%, 7%, 11% and 2%, respectively. These findings support that people with diabetes, hypertension, obesity should be prioritised for vaccination in order to limit the higher death rates in hospital. Public policies should promote a healthier lifestyle including healthier diets and regular physical activity to reduce patient risk factors and comorbidities.

Author affiliations
[1]Paris-Saclay University, UVSQ, Inserm, Gustave Roussy, "Exposome and Heredity" team, CESP, F-94805, Villejuif, France
[2]Department of Plant Molecular Biology, Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland
[3]Laboratory of Soil Biodiversity, Faculty of Science, University of Neuchatel, Neuchâtel, Switzerland
[4]Universite de Paris, IAME, INSERM, Paris, France
[5]National Institute for Health Research, Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance, Imperial College London, London, UK
[6]Infectious and Tropical Diseases Department, Bichat-Claude Bernard Hospital, AP-HP, Paris, France
[7]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK
[8]Department of Nutrition, Bjørknes University College, Oslo, Norway
[9]Department of Endocrinology, Morbid Obesity and Preventive Medicine, Oslo University Hospital, Oslo, Norway
[10]Unit of Cardiovascular and Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
[11]Department of Statistics, Computer Science and Applications "G. Parenti", University of Florence, Florence, Italy

Twitter Yahya Mahamat-Saleh @MS_Yahya_, Thibault Fiolet @T_Fiolet, Mathieu Edouard Rebeaud @Damkyan_Omega and Anthony Guihur @AnthonyGuihur

Contributors YM-S conceived and designed the research. YM-S performed statistical analysis and wrote the first draft of the paper. YM-S, TF and MER performed the literature search and literature screening. YM-S and TF assessed the risk of bias of the studies and assessed the certainty of evidence of the associations. MM assisted with the statistical analysis. YM-S, TF, MER, MM, AG, DEF, NL, NP-S, DA and GS contributed to data interpretation and revision of the manuscript for intellectual content. YM-S takes responsibility for the integrity of the data and the accuracy of the data analysis. YM-S is responsible for the overall content as guarantor.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iDs
Yahya Mahamat-Saleh http://orcid.org/0000-0002-5892-8886
Anthony Guihur http://orcid.org/0000-0001-5353-1428
Dagfinn Aune http://orcid.org/0000-0002-4533-1722

REFERENCES

1 Zhu N, Zhang D, Wang W, *et al*. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med Overseas Ed* 2020;382:727–33.
2 WHO. Coronavirus disease (COVID-19) pandemic, 2020. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019?gclid=CjwKCAiA57D_BRAZEiwAZcfCxWRma5FQmvw6WgYEsv692kL7rMdTuLqwaaSN7-Hfq0rjcu-rrfGYpBoCT1IQAvD_BwE
3 World Health Organization (WHO). Who coronavirus disease (COVID-19) Dashboard. Available: https://covid19.who.int/?gclid=CjwKCAiA57D_BRAZEiwAZcfCxWQWth4sNqEZ1J34KInDRNOjiLuwzkFwpDNCU7V2CPqCPGDU2jLYfhoCb6UQAvD_BwE

4 Zhou F, Yu T, Du R, *et al*. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 2020;395:1054–62.
5 Onder G, Rezza G, Brusaferro S. Case-Fatality rate and characteristics of patients dying in relation to COVID-19 in Italy. *JAMA* 2020;323:1775-1776.
6 Goldstein JR, Lee RD. Demographic perspectives on the mortality of COVID-19 and other epidemics. *Proc Natl Acad Sci U S A* 2020;117:22035–41.
7 Demombynes G. COVID-19 age-mortality curves are flatter in developing countries 2020.
8 Richardson S, Hirsch JS, Narasimhan M, *et al*. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the new York City area. *JAMA* 2020;323:2052–9.
9 Wang D, Hu B, Hu C, *et al*. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* 2020;323:1061.
10 Huang C, Wang Y, Li X, *et al*. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* 2020;395:497–506.
11 Yang J, Zheng Y, Gou X, *et al*. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. *Int J Infect Dis* 2020;94:91–5.
12 Siordia JA. Epidemiology and clinical features of COVID-19: a review of current literature. *J Clin Virol* 2020;127:104357.
13 Guan W-J, Liang W-H, Zhao Y, *et al*. Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. *Eur Respir J* 2020;55:2000547.
14 Khan MMA, Khan MN, Mustagir MG, *et al*. Effects of underlying morbidities on the occurrence of deaths in COVID-19 patients: a systematic review and meta-analysis. *J Glob Health* 2020;10:020503.
15 Liu K, Fang Y-Y, Deng Y, *et al*. Clinical characteristics of novel coronavirus cases in tertiary hospitals in Hubei Province. *Chin Med J* 2020;133:1025–31.
16 Ssentongo P, Ssentongo AE, Heilbrunn ES, *et al*. Association of cardiovascular disease and 10 other pre-existing comorbidities with COVID-19 mortality: a systematic review and meta-analysis. *PLoS One* 2020;15:e0238215.
17 Luo L, Fu M, Li Y, *et al*. The potential association between common comorbidities and severity and mortality of coronavirus disease 2019: a pooled analysis. *Clin Cardiol* 2020;43:1478–93.
18 Du Y, Lv Y, Zha W, *et al*. Association of body mass index (BMI) with critical COVID-19 and in-hospital mortality: a dose-response meta-analysis. *Metabolism* 2021;117:154373.
19 Karanasos A, Aznaouridis K, Latsios G, *et al*. Impact of smoking status on disease severity and mortality of hospitalized patients with COVID-19 infection: a systematic review and meta-analysis. *Nicotine Tob Res* 2020;22:1657–9.
20 Hussain A, Mahawar K, Xia Z, *et al*. Obesity and mortality of COVID-19. meta-analysis. *Obes Res Clin Pract* 2020;14:295–300.
21 de Almeida-Pititto B, Dualib PM, Zajdenverg L, *et al*. Severity and mortality of COVID 19 in patients with diabetes, hypertension and cardiovascular disease: a meta-analysis. *Diabetol Metab Syndr* 2020;12:75.
22 Noor FM, Islam MM. Prevalence and associated risk factors of mortality among COVID-19 patients: a meta-analysis. *J Community Health* 2020;45:1270–82.
23 Chu Y, Yang J, Shi J, *et al*. Obesity is associated with increased severity of disease in COVID-19 pneumonia: a systematic review and meta-analysis. *Eur J Med Res* 2020;25:64.
24 Popkin BM, Du S, Green WD, *et al*. Individuals with obesity and COVID-19: a global perspective on the epidemiology and biological relationships. *Obes Rev* 2020;21:e13128.
25 Seidu S, Gillies C, Zaccardi F, *et al*. The impact of obesity on severe disease and mortality in people with SARS-CoV-2: a systematic review and meta-analysis. *Endocrinol Diabetes Metab* 2020:e00176.
26 Mesas AE, Cavero-Redondo I, Álvarez-Bueno C, *et al*. Predictors of in-hospital COVID-19 mortality: a comprehensive systematic review and meta-analysis exploring differences by age, sex and health conditions. *PLoS One* 2020;15:e0241742.
27 Alqahtani JS, Oyelade T, Aldhahir AM, *et al*. Prevalence, severity and mortality associated with COPD and smoking in patients with COVID-19: a rapid systematic review and meta-analysis. *PLoS One* 2020;15:e0233147.
28 Neumann-Podczaska A, Chojnicki M, Karbowski LM, *et al*. Clinical characteristics and survival analysis in a small sample of older COVID-19 patients with defined 60-Day outcome. *Int J Environ Res Public Health* 2020;17:8362.
29 Gutierrez JP, Bertozzi SM. Non-Communicable diseases and inequalities increase risk of death among COVID-19 patients in Mexico. *PLoS One* 2020;15:e0240394.
30 Albitar O, Ballouze R, Ooi JP, *et al*. Risk factors for mortality among COVID-19 patients. *Diabetes Res Clin Pract* 2020;166:108293.
31 Al-Salameh A, Lanoix J-P, Bennis Y, *et al*. Characteristics and outcomes of COVID-19 in hospitalized patients with and without diabetes. *Diabetes Metab Res Rev* 2021;37:e3388.
32 Hashemi N, Viveiros K, Redd WD, *et al*. Impact of chronic liver disease on outcomes of hospitalized patients with COVID-19: a multicentre United States experience. *Liver Int* 2020;40:2515–21.
33 Atkins JL, Masoli JAH, Delgado J, *et al*. Preexisting comorbidities predicting COVID-19 and mortality in the UK Biobank community cohort. *J Gerontol A Biol Sci Med Sci* 2020;75:2224–30.
34 Barron E, Bakhai C, Kar P, *et al*. Associations of type 1 and type 2 diabetes with COVID-19-related mortality in England: a whole-population study. *Lancet Diabetes Endocrinol* 2020;8:813–22.
35 Barbu MG, Thompson RJ, Thompson DC, *et al*. The impact of SARS-CoV-2 on the most common Comorbidities-A retrospective study on 814 COVID-19 deaths in Romania. *Front Med* 2020;7:567199.
36 Boulle A, Davies M-A, Hussey H, *et al*. Risk factors for coronavirus disease 2019 (COVID-19) death in a population cohort study from the Western Cape Province, South Africa. *Clin Infect Dis* 2021;73:e2005–15.
37 Bousquet G, Falgarone G, Deutsch D, *et al*. ADL-dependency, D-dimers, LDH and absence of anticoagulation are independently associated with one-month mortality in older inpatients with Covid-19. *Aging* 2020;12:11306–13.
38 Calmes D, Graff S, Maes N, *et al*. Asthma and COPD are not risk factors for ICU stay and death in case of SARS-CoV2 infection. *J Allergy Clin Immunol Pract* 2021;9:160-169.
39 Carrasco-Sánchez FJ, López-Carmona Mª Dolores, Martínez-Marcos FJ, *et al*. Admission hyperglycaemia as a predictor of mortality in patients hospitalized with COVID-19 regardless of diabetes status: data from the Spanish SEMI-COVID-19 registry. *Ann Med* 2021;53:103–16.
40 Chachkhiani D, Soliman MY, Barua D, *et al*. Neurological complications in a predominantly African American sample of COVID-19 predict worse outcomes during hospitalization. *Clin Neurol Neurosurg* 2020;197:106173.
41 Chang MC, Hwang J-M, Jeon J-H, *et al*. Fasting plasma glucose level independently predicts the mortality of patients with coronavirus disease 2019 infection: a multicenter, retrospective cohort study. *Endocrinol Metab* 2020;35:595–601.
42 Chilimuri S, Sun H, Alemam A, *et al*. Predictors of mortality in adults admitted with COVID-19: retrospective cohort study from New York City. *West J Emerg Med* 2020;21:779–84.
43 Ciardullo S, Zerbini F, Perra S, *et al*. Impact of diabetes on COVID-19-related in-hospital mortality: a retrospective study from northern Italy. *J Endocrinol Invest* 2021;44:843–50.
44 pp Coppelli A, Giannarelli R, Aragona M, *et al*. Hyperglycemia at hospital admission is associated with severity of the prognosis in patients hospitalized for COVID-19: the Pisa COVID-19 study. *Diabetes Care* 2020;43:2345–8.
45 Crouse A, Grimes T, Li P, *et al*. Metformin use is associated with reduced mortality in a diverse population with COVID-19 and diabetes. *medRxiv* 2020. doi:10.1101/2020.07.29.20164020. [Epub ahead of print: 31 Jul 2020].
46 Dennis JM, Mateen BA, Sonabend R, *et al*. Type 2 diabetes and COVID-19-Related mortality in the critical care setting: a national cohort study in England, March-July 2020. *Diabetes Care* 2021;44:50–7.
47 Desai A, Voza G, Paiardi S, *et al*. The role of anti-hypertensive treatment, comorbidities and early introduction of LMWH in the setting of COVID-19: a retrospective, observational study in northern Italy. *Int J Cardiol* 2021;324:249–54.
48 de Souza Silva GA, da Silva SP, da Costa MAS, *et al*. Sars-Cov, MERS-CoV and SARS-CoV-2 infections in pregnancy and fetal development. *J Gynecol Obstet Hum Reprod* 2020;101846:101846.
49 Escalera-Antezana JP, Lizon-Ferrufino NF, Maldonado-Alanoca A, *et al*. Risk factors for mortality in patients with coronavirus disease 2019 (COVID-19) in Bolivia: an analysis of the first 107 confirmed cases. *Infez Med* 2020;28:238–42.
50 Esme M, Koca M, Dikmeer A, *et al*. Older adults with coronavirus disease 2019: a nationwide study in turkey. *J Gerontol A Biol Sci Med Sci* 2021;76:e68–75.
51 Fernández-Cruz A, Ruiz-Antorán B, Muñoz-Gómez A, *et al*. A retrospective controlled cohort study of the impact of glucocorticoid treatment in SARS-CoV-2 infection mortality. *Antimicrob Agents Chemother* 2020;64:e01168–20.

52 Fried MW, Crawford JM, Mospan AR, *et al*. Patient characteristics and outcomes of 11 721 patients with coronavirus disease 2019 (COVID-19) hospitalized across the United States. *Clin Infect Dis* 2021;72:e558–65.
53 Galloway JB, Norton S, Barker RD, *et al*. A clinical risk score to identify patients with COVID-19 at high risk of critical care admission or death: an observational cohort study. *J Infect* 2020;81:282–8.
54 Ghany R, Palacio A, Chen G. Prior cardiovascular risk and screening echocardiograms predict hospitalization and severity of coronavirus infection among elderly Medicare patients. *AmJPrevCardiol* 2020;3:100090.
55 Guisado-Vasco P, Valderas-Ortega S, Carralón-González MM, *et al*. Clinical characteristics and outcomes among hospitalized adults with severe COVID-19 admitted to a tertiary medical center and receiving antiviral, antimalarials, glucocorticoids, or immunomodulation with tocilizumab or cyclosporine: a retrospective observational study (COQUIMA cohort). *EClinicalMedicine* 2020;28:100591.
56 Haase N, Plovsing R, Christensen S, *et al*. Characteristics, interventions, and longer term outcomes of COVID-19 ICU patients in Denmark-A nationwide, observational study. *Acta Anaesthesiol Scand* 2021;65:68–75.
57 Harrison SL, Fazio-Eynullayeva E, Lane DA, *et al*. Comorbidities associated with mortality in 31,461 adults with COVID-19 in the United States: a federated electronic medical record analysis. *PLoS Med* 2020;17:e1003321.
58 Narain S, Stefanov DG, Chau AS, *et al*. Comparative survival analysis of immunomodulatory therapy for coronavirus disease 2019 cytokine storm. *Chest* 2021;159:933–48.
59 Iaccarino G, Grassi G, Borghi C, *et al*. Age and multimorbidity predict death among COVID-19 patients: results of the SARS-RAS study of the Italian Society of hypertension. *Hypertension* 2020;76:366–72.
60 Lampasona V, Secchi M, Scavini M, *et al*. Antibody response to multiple antigens of SARS-CoV-2 in patients with diabetes: an observational cohort study. *Diabetologia* 2020;63:2548–58.
61 pp Lombardi CM, Carubelli V, Iorio A, *et al*. Association of troponin levels with mortality in Italian patients hospitalized with coronavirus disease 2019: results of a multicenter study. *JAMA Cardiol* 2020;5:1274–7.
62 Lunski MJ, Burton J, Tawagi K, *et al*. Multivariate mortality analyses in COVID-19: comparing patients with cancer and patients without cancer in Louisiana. *Cancer* 2021;127:266–74.
63 Maddaloni E, D'Onofrio L, Alessandri F. Cardiometabolic multimorbidity is associated with a worse Covid-19 prognosis than individual cardiometabolic risk factors: a multicentre retrospective study (CoViDiab II). *CardiovascDiabetol* 2020;19:164.
64 Mallow PJ, Belk KW, Topmiller M, *et al*. Outcomes of hospitalized COVID-19 patients by risk factors: results from a United States Hospital claims database. *J Health Econ Outcomes Res* 2020;7:165–75.
65 McCarty TR, Hathorn KE, Redd WD. How do presenting symptoms and outcomes differ by Race/Ethnicity among hospitalized patients with coronavirus disease 2019 infection? experience in Massachusetts. *Clinical Infectious Diseases* 2020;69:458–64.
66 McCullough SA, Goyal P, Krishnan U, *et al*. Electrocardiographic findings in coronavirus Disease-19: insights on mortality and underlying myocardial processes. *J Card Fail* 2020;26:626–32.
67 Mendy A, Apewokin S, Wells AA. Factors associated with hospitalization and disease severity in a racially and ethnically diverse population of COVID-19 patients. *medRxiv* 2020.
68 Soares RdeCM, Mattos LR, Raposo LM. Risk factors for hospitalization and mortality due to COVID-19 in Espírito Santo state, Brazil. *Am J Trop Med Hyg* 2020;103:1184–90.
69 Mirani M, Favacchio G, Carrone F. Impact of Comorbidities and Glycemia at Admission and Dipeptidyl Peptidase 4 Inhibitors in Patients With Type 2 Diabetes With COVID-19: A Case Series From an Academic Hospital in Lombardy, Italy. Diabetes care(no pagination 2020;06.
70 Moon SJ, Rhee E-J, Jung J-H, *et al*. Independent impact of diabetes on the severity of coronavirus disease 2019 in 5,307 patients in South Korea: a nationwide cohort study. *Diabetes Metab J* 2020;44:737–46.
71 Moon SS, Lee K, Park J, *et al*. Clinical characteristics and mortality predictors of COVID-19 patients hospitalized at Nationally-Designated treatment hospitals. *J Korean Med Sci* 2020;35:e328.
72 Munblit D, Nekliudov NA, Bugaeva P, *et al*. Stop COVID cohort: an observational study of 3480 patients admitted to the Sechenov university hospital network in Moscow City for suspected coronavirus disease 2019 (COVID-19) infection. *Clin Infect Dis* 2021;73:1–11.
73 Nachega JB, Ishoso DK, Otokoye JO, *et al*. Clinical characteristics and outcomes of patients hospitalized for COVID-19 in Africa: early insights from the Democratic Republic of the Congo. *Am J Trop Med Hyg* 2020;103:2419–28.
74 Nakeshbandi M, Maini R, Daniel P, *et al*. The impact of obesity on COVID-19 complications: a retrospective cohort study. *Int J Obes* 2020;44:1832–7.
75 Anzola GP, Bartolaminelli C, Gregorini GA, *et al*. Neither ACEIs nor Arbs are associated with respiratory distress or mortality in COVID-19 results of a prospective study on a hospital-based cohort. *Intern Emerg Med* 2020;15:1477–84.
76 Nicholson CJ, Wooster L, Sigurslid HH. Estimating risk of mechanical ventilation and mortality among adult COVID-19 patients admitted to mass General Brigham: the vice and dice scores. *medRxiv*2020.
77 Nogueira PJ, de Araújo Nobre M, Costa A, *et al*. The role of health preconditions on COVID-19 deaths in Portugal: evidence from surveillance data of the first 20293 infection cases. *J Clin Med* 2020;9:2368.
78 Pan W, Zhang J, Wang M, *et al*. Clinical features of COVID-19 in patients with essential hypertension and the impacts of renin-angiotensin-aldosterone system inhibitors on the prognosis of COVID-19 patients. *Hypertension* 2020;76:732–41.
79 Portolés J, Marques M, López-Sánchez P, *et al*. Chronic kidney disease and acute kidney injury in the COVID-19 Spanish outbreak. *Nephrol Dial Transplant* 2020;35:1353–61.
80 Rastad H, Karim H, Ejtahed H-S, *et al*. Risk and predictors of in-hospital mortality from COVID-19 in patients with diabetes and cardiovascular disease. *Diabetol Metab Syndr* 2020;12:57.
81 Recinella G, Marasco G, Serafini G, *et al*. Prognostic role of nutritional status in elderly patients hospitalized for COVID-19: a monocentric study. *Aging Clin Exp Res* 2020;32:2695–701.
82 Reilev M, Kristensen KB, Pottegård A, *et al*. Characteristics and predictors of hospitalization and death in the first 11 122 cases with a positive RT-PCR test for SARS-CoV-2 in Denmark: a nationwide cohort. *Int J Epidemiol* 2020;49:1468–81.
83 Rivera-Izquierdo M, Del Carmen Valero-Ubierna M, R-delAmo JL, *et al*. Sociodemographic, clinical and laboratory factors on admission associated with COVID-19 mortality in hospitalized patients: a retrospective observational study. *PLoS One* 2020;15:e0235107.
84 Rodríguez-Molinero A, Gálvez-Barrón C, Miñarro A, *et al*. Association between COVID-19 prognosis and disease presentation, comorbidities and chronic treatment of hospitalized patients. *PLoS One* 2020;15:e0239571.
85 Giorgi Rossi P, Marino M, Formisano D, *et al*. Characteristics and outcomes of a cohort of COVID-19 patients in the province of Reggio Emilia, Italy. *PLoS One* 2020;15:e0238281.
86 pp Rottoli M, Bernante P, Belvedere A, *et al*. How important is obesity as a risk factor for respiratory failure, intensive care admission and death in hospitalised COVID-19 patients? results from a single Italian centre. *Eur J Endocrinol* 2020;183:389–97.
87 Rozaliyani A, Savitri AI, Setianingrum F, *et al*. Factors associated with death in COVID-19 patients in Jakarta, Indonesia: an epidemiological study. *Acta Med Indones* 2020;52:246–54.
88 Salacup G, Lo KB, Gul F, *et al*. Characteristics and clinical outcomes of COVID-19 patients in an underserved-inner City population: a single tertiary center cohort. *J Med Virol* 2021;93:416–23.
89 Salazar E, Christensen PA, Graviss EA, *et al*. Treatment of coronavirus disease 2019 patients with convalescent plasma reveals a signal of significantly decreased mortality. *Am J Pathol* 2020;190:2290–303.
90 Sands KE, Wenzel RP, McLean LE, *et al*. Patient characteristics and admitting vital signs associated with coronavirus disease 2019 (COVID-19)-related mortality among patients admitted with noncritical illness. *Infect Control Hosp Epidemiol* 2021;42:399–405.
91 Santos MM, Lucena EES, Lima KC, *et al*. Survival and predictors of deaths of patients hospitalised due to COVID-19 from a retrospective and multicentre cohort study in Brazil. *Epidemiol Infect* 2020;148:e198.
92 Seiglie J, Platt J, Cromer SJ, *et al*. Diabetes as a risk factor for poor early outcomes in patients hospitalized with COVID-19. *Diabetes Care* 2020;43:2938–44.
93 Shah P, Owens J, Franklin J, *et al*. Demographics, comorbidities and outcomes in hospitalized Covid-19 patients in rural Southwest Georgia. *Ann Med* 2020;52:354–60.
94 Li P, Chen L, Liu Z, *et al*. Clinical features and short-term outcomes of elderly patients with COVID-19. *Int J Infect Dis* 2020;97:245–50.

95 Shang J, Wang Q, Zhang H, *et al*. The relationship between diabetes mellitus and COVID-19 prognosis: a retrospective cohort study in Wuhan, China. *Am J Med* 2021;134:e6–14.
96 Smith AA, Fridling J, Ibrahim D, *et al*. Identifying patients at greatest risk of mortality due to COVID-19: a new England perspective. *West J Emerg Med* 2020;21:785–9.
97 Tartof SY, Qian L, Hong V, *et al*. Obesity and mortality among patients diagnosed with COVID-19: results from an integrated health care organization. *Ann Intern Med* 2020;173:773–81.
98 Turcotte JJ, Meisenberg BR, MacDonald JH, *et al*. Risk factors for severe illness in hospitalized Covid-19 patients at a regional hospital. *PLoS One* 2020;15:e0237558.
99 Wollenstein-Betech S, Silva AAB, Fleck JL, *et al*. Physiological and socioeconomic characteristics predict COVID-19 mortality and resource utilization in Brazil. *PLoS One* 2020;15:e0240346.
100 Yahyavi A, Hemmati N, Derakhshan P, *et al*. Angiotensin enzyme inhibitors and angiotensin receptor blockers as protective factors in COVID-19 mortality: a retrospective cohort study. *Intern Emerg Med* 2021;16:883–93.
101 Yan Y, Yang Y, Wang F, *et al*. Clinical characteristics and outcomes of patients with severe covid-19 with diabetes. *BMJ Open Diabetes Res Care* 2020;8:e001343.
102 Yan Q, Zuo P, Cheng L, *et al*. Acute kidney injury is associated with in-hospital mortality in older patients with COVID-19. *J Gerontol A Biol Sci Med Sci* 2021;76:456–62.
103 Yazdanpanah Y, French COVID cohort investigators and study group. Impact on disease mortality of clinical, biological, and virological characteristics at hospital admission and overtime in COVID-19 patients. *J Med Virol* 2021;93:2149–59.
104 Zandkarimi E, Moradi G, Mohsenpour B. The prognostic factors affecting the survival of Kurdistan Province COVID-19 patients: a cross-sectional study from February to may 2020. *Int J Health Policy Manag* 2020. doi:10.34172/ijhpm.2020.155. [Epub ahead of print: 22 Aug 2020].
105 An C, Lim H, Kim D-W, *et al*. Machine learning prediction for mortality of patients diagnosed with COVID-19: a nationwide Korean cohort study. *Sci Rep* 2020;10:18716.
106 pp Bepouka BI, Mandina M, Makulo JR, *et al*. Predictors of mortality in COVID-19 patients at Kinshasa university Hospital, Democratic Republic of the Congo (from March to June 2020). *Pan Afr Med J* 2020;37:1105–17.
107 Abohamr SI, Abazid RM, Aldossari MA, *et al*. Clinical characteristics and in-hospital mortality of COVID-19 adult patients in Saudi Arabia. *Saudi Med J* 2020;41:1217–26.
108 Czernichow S, Beeker N, Rives-Lange C, *et al*. Obesity doubles mortality in patients hospitalized for severe acute respiratory syndrome coronavirus 2 in Paris hospitals, France: a cohort study on 5,795 patients. *Obesity* 2020;28:2282–9.
109 Harmouch F, Shah K, Hippen JT, *et al*. Is it all in the heart? myocardial injury as major predictor of mortality among hospitalized COVID-19 patients. *J Med Virol* 2021;93:973–82.
110 Hewitt J, Carter B, Vilches-Moraga A, *et al*. The effect of frailty on survival in patients with COVID-19 (cope): a multicentre, European, observational cohort study. *Lancet Public Health* 2020;5:e444–51.
111 Ioannou GN, Locke E, Green P, *et al*. Risk Factors for Hospitalization, Mechanical Ventilation, or Death Among 10 131 US Veterans With SARS-CoV-2 Infection. *JAMA Netw Open* 2020;3:e2022310.
112 Kabarriti R, Brodin NP, Maron MI, *et al*. Association of race and ethnicity with comorbidities and survival among patients with COVID-19 at an urban medical center in New York. *JAMA Netw Open* 2020;3:e2019795.
113 Kolhe NV, Fluck RJ, Selby NM, *et al*. Acute kidney injury associated with COVID-19: a retrospective cohort study. *PLoS Med* 2020;17:e1003406.
114 Meng Y, Lu W, Guo E, *et al*. Cancer history is an independent risk factor for mortality in hospitalized COVID-19 patients: a propensity score-matched analysis. *J Hematol Oncol* 2020;13:75.
115 Ñamendys-Silva SA, Alvarado-Ávila PE, Domínguez-Cherit G, *et al*. Outcomes of patients with COVID-19 in the intensive care unit in Mexico: a multicenter observational study. *Heart Lung* 2021;50:28–32.
116 Li G, Deng Q, Feng J, *et al*. Clinical characteristics of diabetic patients with COVID-19. *J Diabetes Res* 2020;2020:1652403.
117 Nie S-F, Yu M, Xie T, *et al*. Cardiac troponin I is an independent predictor for mortality in hospitalized patients with COVID-19. *Circulation* 2020;142:608–10.
118 Núñez-Gil IJ, Fernández-Pérez C, Estrada V, *et al*. Mortality risk assessment in Spain and Italy, insights of the hope COVID-19 registry. *Intern Emerg Med* 2021;16:957–66.
119 Park JG, Kang MK, Lee YR, *et al*. Fibrosis-4 index as a predictor for mortality in hospitalised patients with COVID-19: a retrospective multicentre cohort study. *BMJ Open* 2020;10:e041989.
120 Parker WH. Bilateral oophorectomy versus ovarian conservation: effects on long-term women's health. *J Minim Invasive Gynecol* 2010;17:161–6.
121 Rosenberg ES, Dufort EM, Udo T, *et al*. Association of treatment with hydroxychloroquine or azithromycin with in-hospital mortality in patients with COVID-19 in New York state. *JAMA* 2020;323:2493–502.
122 Tehrani S, Killander A, Åstrand P, *et al*. Risk factors for death in adult COVID-19 patients: frailty predicts fatal outcome in older patients. *Int J Infect Dis* 2021;102:415–21.
123 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.
124 Stroup DF, Berlin JA, Morton SC, *et al*. Meta-Analysis of observational studies in epidemiology: a proposal for reporting. meta-analysis of observational studies in epidemiology (moose) group. *JAMA* 2000;283:2008–12.
125 Sterne JA, Hernán MA, Reeves BC, *et al*. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;355:i4919.
126 Schünemann HJ, Cuello C, Akl EA, *et al*. Grade guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. *J Clin Epidemiol* 2019;111:105–14.
127 DerSimonian R, Laird N. Meta-Analysis in clinical trials. *Control Clin Trials* 1986;7:177–88.
128 Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002;21:1539–58.
129 Docherty AB, Harrison EM, Green CA, *et al*. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ* 2020;369:m1985.
130 Newcombe RG, Bender R. Implementing grade: calculating the risk difference from the baseline risk and the relative risk. *Evid Based Med* 2014;19:6–8.
131 Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol* 1992;135:1301–9.
132 Hamling J, Lee P, Weitkunat R, *et al*. Facilitating meta-analyses by deriving relative effect and precision estimates for alternative comparisons from a set of estimates presented by exposure level or disease category. *Stat Med* 2008;27:954–70.
133 Bagnardi V, Zambon A, Quatto P, *et al*. Flexible meta-regression functions for modeling aggregate dose-response data, with an application to alcohol and mortality. *Am J Epidemiol* 2004;159:1077–86.
134 Egger M, Davey Smith G, Schneider M, *et al*. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629–34.
135 Spiegelman D, Hertzmark E, Wand HC. Point and interval estimates of partial population attributable risks in cohort studies: examples and software. *Cancer Causes Control* 2007;18:571–9.
136 Moazzami B, Chaichian S, Kasaeian A, *et al*. Metabolic risk factors and risk of Covid-19: a systematic review and meta-analysis. *PLoS One* 2020;15:e0243600.
137 Farsalinos K, Barbouni A, Poulas K, *et al*. Current smoking, former smoking, and adverse outcome among hospitalized COVID-19 patients: a systematic review and meta-analysis. *Ther Adv Chronic Dis* 2020;11:2040622320935765.
138 Shastri MD, Shukla SD, Chong WC, *et al*. Smoking and COVID-19: what we know so far. *Respir Med* 2021;176:106237.
139 Acharya D, Lee K, Lee DS, *et al*. Mortality rate and predictors of mortality in hospitalized COVID-19 patients with diabetes. *Health Care* 2020;8:338.
140 Jackson BR, Gold JAW, Natarajan P, *et al*. Predictors at admission of mechanical ventilation and death in an observational cohort of adults hospitalized with COVID-19. *Clin Infect Dis* 2020:ciaa1459.
141 Braude P, Carter B, Short R, *et al*. The influence of ACE inhibitors and Arbs on hospital length of stay and survival in people with COVID-19. *Int J Cardiol Heart Vasc* 2020;31:100660.
142 Cai Y, Shi S, Yang F, *et al*. Fasting blood glucose level is a predictor of mortality in patients with COVID-19 independent of diabetes history. *Diabetes Res Clin Pract* 2020;169:108437.
143 Hwang J-M, Kim J-H, Park J-S, *et al*. Neurological diseases as mortality predictive factors for patients with COVID-19: a retrospective cohort study. *Neurol Sci* 2020;41:2317–24.
144 Williamson EJ, Walker AJ, Bhaskaran K, *et al*. Factors associated with COVID-19-related death using OpenSAFELY. *Nature* 2020;584:430–6.

145 Bruce E, Barlow-Pay F, Short R, *et al*. Prior routine use of non-steroidal anti-inflammatory drugs (NSAIDs) and important outcomes in hospitalised patients with covid-19. *J Clin Med* 2020;9:2586.
146 Carrillo-Vega MF, Salinas-Escudero G, García-Peña C, *et al*. Early estimation of the risk factors for hospitalization and mortality by COVID-19 in Mexico. *PLoS One* 2020;15:e0238905.
147 Hernández-Galdamez DR, González-Block Miguel Ángel, Romo-Dueñas DK, *et al*. Increased Risk of Hospitalization and Death in Patients with COVID-19 and Pre-existing Noncommunicable Diseases and Modifiable Risk Factors in Mexico. *Arch Med Res* 2020;51:683–9.
148 Parra-Bracamonte GM, Lopez-Villalobos N, Parra-Bracamonte FE. Clinical characteristics and risk factors for mortality of patients with COVID-19 in a large data set from Mexico. *Ann Epidemiol* 2020;52:93–8.
149 Wollenstein-Betech S, Cassandras CG, Paschalidis IC. Personalized predictive models for symptomatic COVID-19 patients using basic preconditions: hospitalizations, mortality, and the need for an ICU or ventilator. *Int J Med Inform* 2020;142:104258.
150 Hernández-Vásquez A, Azañedo D, Vargas-Fernández R, *et al*. Association of comorbidities with pneumonia and death among COVID-19 patients in Mexico: a nationwide cross-sectional study. *J Prev Med Public Health* 2020;53:211–9.
151 Shi S, Qin M, Shen B, *et al*. Association of cardiac injury with mortality in hospitalized patients with COVID-19 in Wuhan, China. *JAMA Cardiol* 2020;5:802–10.
152 Mehra MR, Desai SS, Kuy S, *et al*. Cardiovascular disease, drug therapy, and mortality in COVID-19. *N Engl J Med Overseas Ed* 2020;382:e102.
153 Cao Y, Han X, Gu J, *et al*. Prognostic value of baseline clinical and HRCT findings in 101 patients with severe COVID-19 in Wuhan, China. *Sci Rep* 2020;10:16.
154 Di Castelnuovo A, Bonaccio M, Costanzo S, *et al*. Common cardiovascular risk factors and in-hospital mortality in 3,894 patients with COVID-19: survival analysis and machine learning-based findings from the multicentre Italian CORIST study. *Nutr Metab Cardiovasc Dis* 2020;30:1899–913.
155 Chen Y, Yang D, Cheng B, *et al*. Clinical characteristics and outcomes of patients with diabetes and COVID-19 in association with glucose-lowering medication. *Diabetes Care* 2020;43:1399–407.
156 Chen F, Sun W, Sun S, *et al*. Clinical characteristics and risk factors for mortality among inpatients with COVID-19 in Wuhan, China. *Clin Transl Med* 2020;10:e40.
157 Cummings MJ, Baldwin MR, Abrams D, *et al*. Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study. *The Lancet* 2020;395:1763–70.
158 Fang H, Liu Q, Xi M, *et al*. Impact of comorbidities on clinical prognosis in 1280 patients with different types of COVID-19. *J Investig Med* 2021;69:75–85.
159 Grasselli G, Greco M, Zanella A, *et al*. Risk factors associated with mortality among patients with COVID-19 in intensive care units in Lombardy, Italy. *JAMA Intern Med* 2020;180:1345–55.
160 Gu T, Chu Q, Yu Z, *et al*. History of coronary heart disease increased the mortality rate of patients with COVID-19: a nested case-control study. *BMJ Open* 2020;10:e038976.
161 Gu T, Mack JA, Salvatore M, *et al*. COVID-19 outcomes, risk factors and associations by race: a comprehensive analysis using electronic health records data in Michigan medicine. *medRxiv* 2020. doi:10.1101/2020.06.16.20133140. [Epub ahead of print: 18 Jun 2020].
162 Gupta S, Hayek SS, Wang W, *et al*. Factors associated with death in critically ill patients with coronavirus disease 2019 in the US. *JAMA Intern Med* 2020;180:1436–47.
163 Hu W, Dong M, Xiong M, *et al*. Clinical courses and outcomes of patients with chronic obstructive pulmonary disease during the COVID-19 epidemic in Hubei, China. *Int J Chron Obstruct Pulmon Dis* 2020;15:2237–48.
164 Imam Z, Odish F, Gill I, *et al*. Older age and comorbidity are independent mortality predictors in a large cohort of 1305 COVID-19 patients in Michigan, United States. *J Intern Med* 2020;288:469–76.
165 Kim DW, Byeon KH, Kim J, *et al*. The correlation of comorbidities on the mortality in patients with COVID-19: an observational study based on the Korean National health insurance big data. *J Korean Med Sci* 2020;35:e243.
166 Klang E, Soffer S, Nadkarni G, *et al*. Sex differences in age and comorbidities for COVID-19 mortality in urban New York City. *SN Compr Clin Med* 2020:1319–22.
167 Kim L, Garg S, O'Halloran A, *et al*. Risk Factors for Intensive Care Unit Admission and In-hospital Mortality Among Hospitalized Adults Identified through the US Coronavirus Disease 2019 (COVID-19)-Associated Hospitalization Surveillance Network (COVID-NET). *Clin Infect Dis* 2021;72:e206–14.
168 Li H, Tian S, Chen T, *et al*. Newly diagnosed diabetes is associated with a higher risk of mortality than known diabetes in hospitalized patients with COVID-19. *Diabetes Obes Metab* 2020;22:1897–906.
169 Lee JY, Kim HA, Huh K, *et al*. Risk factors for mortality and respiratory support in elderly patients hospitalized with COVID-19 in Korea. *J Korean Med Sci* 2020;35:e223.
170 Li L, Li F, Fortunati F, *et al*. Association of a prior psychiatric diagnosis with mortality among hospitalized patients with coronavirus disease 2019 (COVID-19) infection. *JAMA Netw Open* 2020;3:e2023282.
171 Lui GC-Y, Yip TC-F, Wong VW-S, *et al*. Significantly lower case-fatality ratio of coronavirus disease 2019 (COVID-19) than severe acute respiratory syndrome (SARS) in Hong Kong—A Territory-Wide cohort study. *Clin Infect Dis* 2021;72:e466–75.
172 Liu J, Zhang S, Wu Z, *et al*. Clinical outcomes of COVID-19 in Wuhan, China: a large cohort study. *Ann Intensive Care* 2020;10:99.
173 Liu Z, Li J, Huang J, *et al*. Association between diabetes and COVID-19: a retrospective observational study with a large sample of 1,880 cases in Leishenshan Hospital, Wuhan. *Front Endocrinol* 2020;11:478.
174 Mikami T, Miyashita H, Yamada T, *et al*. Risk factors for mortality in patients with COVID-19 in New York City. *J Gen Intern Med* 2021;36:17–26.
175 Palaiodimos L, Chamorro-Pareja N, Karamanis D, *et al*. Diabetes is associated with increased risk for in-hospital mortality in patients with COVID-19: a systematic review and meta-analysis comprising 18,506 patients. *Hormones* 2021;20:305–14.
176 Perkin MR, Heap S, Crerar-Gilbert A, *et al*. Deaths in people from black, Asian and minority ethnic communities from both COVID-19 and non-COVID causes in the first weeks of the pandemic in London: a hospital case note review. *BMJ Open* 2020;10:e040638.
177 Petrilli CM, Jones SA, Yang J, *et al*. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. *BMJ* 2020;369:m1966.
178 Pettit NN, MacKenzie EL, Ridgway JP, *et al*. Obesity is associated with increased risk for mortality among hospitalized patients with COVID-19. *Obesity* 2020;28:1806–10.
179 Poblador-Plou B, Carmona-Pírez J, Ioakeim-Skoufa I, *et al*. Baseline chronic comorbidity and mortality in Laboratory-Confirmed COVID-19 cases: results from the PRECOVID study in Spain. *Int J Environ Res Public Health* 2020;17:5171.
180 Shi Q, Zhang X, Jiang F, *et al*. Clinical characteristics and risk factors for mortality of COVID-19 patients with diabetes in Wuhan, China: a two-center, retrospective study. *Diabetes Care* 2020;43:1382–91.
181 Sousa GJB, Garces TS, Cestari VRF, *et al*. Mortality and survival of COVID-19. *Epidemiol Infect* 2020;148:e123.
182 Stony Brook COVID-19 Research Consortium. Geospatial distribution and predictors of mortality in hospitalized patients with COVID-19: a cohort study. *Open Forum Infect Dis* 2020;7:ofaa436.
183 Talavera B, García-Azorín D, Martínez-Pías E, *et al*. Anosmia is associated with lower in-hospital mortality in COVID-19. *J Neurol Sci* 2020;419:117163.
184 Xie J, Zu Y, Alkhatib A, *et al*. Metabolic syndrome and COVID-19 mortality among adult black patients in New Orleans. *Diabetes Care* 2020. doi:10.2337/dc20-1714. [Epub ahead of print: 25 Aug 2020].
185 Wu C, Chen X, Cai Y, *et al*. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Intern Med* 2020;180:934–43.
186 Wu J, Huang J, Zhu G, *et al*. Elevation of blood glucose level predicts worse outcomes in hospitalized patients with COVID-19: a retrospective cohort study. *BMJ Open Diabetes Res Care* 2020;8:e001476.
187 Yang P, Wang N, Wang J, *et al*. Admission fasting plasma glucose is an independent risk factor for 28-day mortality in patients with COVID-19. *J Med Virol* 2021;93:2168–76.
188 Yu C, Lei Q, Li W, *et al*. Clinical characteristics, associated factors, and predicting COVID-19 mortality risk: a retrospective study in Wuhan, China. *Am J Prev Med* 2020;59:168–75.
189 Zhang J, Kong W, Xia P, *et al*. Impaired fasting glucose and diabetes are related to higher risks of complications and mortality among patients with coronavirus disease 2019. *Front Endocrinol* 2020;11:525.

190 Zhu L, She Z-G, Cheng X, *et al*. Association of blood glucose control and outcomes in patients with COVID-19 and pre-existing type 2 diabetes. *Cell Metab* 2020;31:1068–77.
191 Zimering MB, Razzaki T, Tsang T, *et al*. Inverse association between serotonin 2A receptor antagonist medication use and mortality in severe COVID-19 infection. *Endocrinol Diabetes Metab J* 2020;4:1–5.
192 Almazeedi S, Al-Youha S, Jamal MH, *et al*. Characteristics, risk factors and outcomes among the first consecutive 1096 patients diagnosed with COVID-19 in Kuwait. *EClinicalMedicine* 2020;24:100448.
193 Foy BH, Carlson JCT, Reinertsen E, *et al*. Association of red blood cell distribution width with mortality risk in hospitalized adults with SARS-CoV-2 infection. *JAMA Netw Open* 2020;3:e2022058.
194 Shi S, Qin M, Cai Y, *et al*. Characteristics and clinical significance of myocardial injury in patients with severe coronavirus disease 2019. *Eur Heart J* 2020;41:2070–9.
195 Wang F, Cao J, Yu Y, *et al*. Epidemiological characteristics of patients with severe COVID-19 infection in Wuhan, China: evidence from a retrospective observational study. *Int J Epidemiol* 2021;49:1940–50.
196 Arshad S, Kilgore P, Chaudhry ZS, *et al*. Treatment with hydroxychloroquine, azithromycin, and combination in patients hospitalized with COVID-19. *Int J Infect Dis* 2020;97:396–403.
197 Berenguer J, Ryan P, Rodríguez-Baño J, *et al*. Characteristics and predictors of death among 4035 consecutively hospitalized patients with COVID-19 in Spain. *Clin Microbiol Infect* 2020;26:1525–36.
198 Ciceri F, Castagna A, Rovere-Querini P, *et al*. Early predictors of clinical outcomes of COVID-19 outbreak in Milan, Italy. *Clin Immunol* 2020;217:108509.
199 Deng G, Yin M, Chen X, *et al*. Clinical determinants for fatality of 44,672 patients with COVID-19. *Crit Care* 2020;24:179.
200 Wang K, Zuo P, Liu Y, *et al*. Clinical and laboratory predictors of in-hospital mortality in 305 patients with COVID-19: a cohort study in Wuhan, China. *SSRN Journal* 2020.
201 Biscarini S, Colaneri M, Ludovisi S, *et al*. The obesity paradox: analysis from the SMAtteo COvid-19 registry (SMACORE) cohort. *Nutr Metab Cardiovasc Dis* 2020;30:1920–5.
202 Biscarini S, Colaneri M, Ludovisi S, *et al*. The obesity paradox: analysis from the SMAtteo COvid-19 registry (SMACORE) cohort. *Nutr Metab Cardiovasc Dis* 2020;30:30.
203 CA D, Costanzo S, Antinori A, *et al*. Raas inhibitors are not associated with mortality in COVID-19 patients: findings from an observational multicenter study in Italy and a meta-analysis of 19 studies. *Vascul Pharmacol* 2020;135:106805.
204 De Vito A, Geremia N, Fiore V, *et al*. Clinical features, laboratory findings and predictors of death in hospitalized patients with COVID-19 in Sardinia, Italy. *Eur Rev Med Pharmacol Sci* 2020;24:7861–8.
205 Gao C, Cai Y, Zhang K, *et al*. Association of hypertension and antihypertensive treatment with COVID-19 mortality: a retrospective observational study. *Eur Heart J* 2020;41:2058–66.
206 Giacomelli A, Ridolfo AL, Milazzo L, *et al*. 30-Day mortality in patients hospitalized with COVID-19 during the first wave of the Italian epidemic: a prospective cohort study. *Pharmacol Res* 2020;158:104931.
207 Liu J, Liu Z, Jiang W, *et al*. Clinical predictors of COVID-19 disease progression and death: analysis of 214 hospitalised patients from Wuhan, China. *Clin Respir J* 2021;15:293–309.
208 Liu D, Wang Y, Wang J, *et al*. Characteristics and outcomes of a sample of patients with COVID-19 identified through social media in Wuhan, China: observational study. *J Med Internet Res* 2020;22:e20108.
209 Magleby R, Westblade LF, Trzebucki A, *et al*. Impact of SARS-CoV-2 viral load on risk of intubation and mortality among hospitalized patients with coronavirus disease 2019. *Clin Infect Dis* 2020;30:ciaa851.
210 Matsuzawa Y, Ogawa H, Kimura K, *et al*. Renin-Angiotensin system inhibitors and the severity of coronavirus disease 2019 in Kanagawa, Japan: a retrospective cohort study. *Hypertens Res* 2020;43:1257–66.
211 Muñoz P, Galar A, Catalán P, *et al*. The first 100 cases of COVID-19 in a hospital in Madrid with a 2-month follow-up. *Rev Esp Quimioter* 2020;33:369–78.
212 Russo E, Esposito P, Taramasso L, *et al*. Kidney disease and all-cause mortality in patients with COVID-19 hospitalized in Genoa, Northern Italy. *J Nephrol* 2021;34:173–83.
213 Zangrillo A, Beretta L, Scandroglio AM. Characteristics, treatment, outcomes and cause of death of invasively ventilated patients with COVID-19 ARDS in Milan, Italy. *Crit Care Resusc* 2020. [Epub ahead of print: 23 Apr 2020].
214 Almazeedi S, Al-Youha S, Jamal MH, *et al*. Characteristics, risk factors and outcomes among the first consecutive 1096 patients diagnosed with COVID-19 in Kuwait. *EClinicalMedicine* 2020;24:100448.
215 Brouns SH, Brüggemann R, Linkens AEMJH, *et al*. Mortality and the use of antithrombotic therapies among nursing home residents with COVID-19. *J Am Geriatr Soc* 2020;68:1647–52.
216 Lopez-Mendez I, Aquino-Matus J, Gall SM-B, *et al*. Association of liver steatosis and fibrosis with clinical outcomes in patients with SARS-CoV-2 infection (COVID-19). *Ann Hepatol* 2021;20:100271.
217 Du R-H, Liang L-R, Yang C-Q, *et al*. Predictors of mortality for patients with COVID-19 pneumonia caused by SARS-CoV-2: a prospective cohort study. *Eur Respir J* 2020;55:2000524.
218 Mowla SGM, Azad KAK, Kabir A, *et al*. Clinical profile of 100 confirmed COVID-19 patients admitted in Dhaka medical college Hospital, Dhaka, Bangladesh. *Journal of Bangladesh College of Physicians and Surgeons* 2020:29–36.
219 Sun H, Ning R, Tao Y, *et al*. Risk factors for mortality in 244 older adults with COVID-19 in Wuhan, China: a retrospective study. *J Am Geriatr Soc* 2020;68:E19–23.
220 Alamdari NM, Afaghi S, Rahimi FS, *et al*. Mortality risk factors among hospitalized COVID-19 patients in a major referral center in Iran. *Tohoku J Exp Med* 2020;252:73–84.
221 Hajifathalian K, Kumar S, Newberry C, *et al*. Obesity is associated with worse outcomes in COVID-19: analysis of early data from New York City. *Obesity* 2020;28:1606–12.
222 Klang E, Kassim G, Soffer S, *et al*. Severe obesity as an independent risk factor for COVID-19 mortality in hospitalized patients younger than 50. *Obesity* 2020;28:1595–9.
223 Price-Haywood EG, Burton J, Fort D, *et al*. Hospitalization and mortality among black patients and white patients with Covid-19. *N Engl J Med* 2020;382:2534–43.
224 Ramos-Rincon J-M, Buonaiuto V, Ricci M, *et al*. Clinical characteristics and risk factors for mortality in very old patients hospitalized with COVID-19 in Spain. *J Gerontol A Biol Sci Med Sci* 2021;76:e28–37.
225 Steinberg E, Wright E, Kushner B. In young adults with COVID-19, obesity is associated with adverse outcomes. *West J Emerg Med* 2020;21:752–5.
226 Zhang F, Xiong Y, Wei Y, *et al*. Obesity predisposes to the risk of higher mortality in young COVID-19 patients. *J Med Virol* 2020;92:2536–42.
227 Gazzaruso C, Mariani G, Ravetto C, *et al*. Lupus anticoagulant and mortality in patients hospitalized for COVID-19. *J Thromb Thrombolysis* 2021;52:85–91.
228 Mather JF, Seip RL, McKay RG. Impact of famotidine use on clinical outcomes of hospitalized patients with COVID-19. *Am J Gastroenterol* 2020;115:1617–23.
229 Anderson MR, Geleris J, Anderson DR, *et al*. Body Mass Index and Risk for Intubation or Death in SARS-CoV-2 Infection : A Retrospective Cohort Study. *Ann Intern Med* 2020;173:782–90.
230 Ferrando-Vivas P, Doidge J, Thomas K, *et al*. Prognostic factors for 30-day mortality in critically ill patients with coronavirus disease 2019: an observational cohort study. *Crit Care Med* 2021;49:102–11.
231 Gayam V, Chobufo MD, Merghani MA, *et al*. Clinical characteristics and predictors of mortality in African-Americans with COVID-19 from an inner-city community teaching hospital in New York. *J Med Virol* 2021;93:812–9.
232 Muñoz-Price LS, Nattinger AB, Rivera F, *et al*. Racial disparities in incidence and outcomes among patients with COVID-19. *JAMA Netw Open* 2020;3:e2021892.
233 Li G, Zhou C-L, Ba Y-M, *et al*. Nutritional risk and therapy for severe and critical COVID-19 patients: a multicenter retrospective observational study. *Clin Nutr* 2021;40:2154–61.
234 Wang Z, Zheutlin A, Kao Y-H, *et al*. Hospitalised COVID-19 patients of the Mount Sinai health system: a retrospective observational study using the electronic medical records. *BMJ Open* 2020;10:e040441.
235 Baronio M, Freni-Sterrantino A, Pinelli M, *et al*. Italian SARS-CoV-2 patients in intensive care: towards an identikit for subjects at risk? *Eur Rev Med Pharmacol Sci* 2020;24:9698–704.
236 Adrish M, Chilimuri S, Mantri N, *et al*. Association of smoking status with outcomes in hospitalised patients with COVID-19. *BMJ Open Respir Res* 2020;7:e000716.
237 Chen L, Yu J, He W, *et al*. Risk factors for death in 1859 subjects with COVID-19. *Leukemia* 2020;34:2173–83.
238 Zhang J-J, Cao Y-Y, Tan G, *et al*. Clinical, radiological, and laboratory characteristics and risk factors for severity and mortality of 289 hospitalized COVID-19 patients. *Allergy* 2021;76:533–50.

239 Alharthy A, Aletreby W, Faqihi F, *et al*. Clinical characteristics and predictors of 28-day mortality in 352 critically ill patients with COVID-19: a retrospective study. *J Epidemiol Glob Health* 2021;11:98.
240 Qin L, Li X, Shi J, *et al*. Gendered effects on inflammation reaction and outcome of COVID-19 patients in Wuhan. *J Med Virol* 2020;92:2684–92.
241 Wang Z-H, Shu C, Ran X, *et al*. Critically ill patients with coronavirus disease 2019 in a designated ICU: clinical features and predictors for mortality. *Risk Manag Healthc Policy* 2020;13:833–45.
242 Farrell RJ, O'Regan R, O'Neill E, *et al*. Sociodemographic variables as predictors of adverse outcome in SARS-CoV-2 infection: an Irish Hospital experience. *Ir J Med Sci* 2021;190:893–903.
243 Wang K, Zuo P, Liu Y, *et al*. Clinical and laboratory predictors of in-hospital mortality in patients with coronavirus Disease-2019: a cohort study in Wuhan, China. *Clin Infect Dis* 2020;71:2079–88.
244 Chen T, Dai Z, Mo P, *et al*. Clinical characteristics and outcomes of older patients with coronavirus disease 2019 (COVID-19) in Wuhan, China: a Single-Centered, retrospective study. *J Gerontol A Biol Sci Med Sci* 2020;75:1788–95.
245 Yang X, Yu Y, Xu J, *et al*. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir Med* 2020;8:475–81.
246 Kumar A, Arora A, Sharma P, *et al*. Is diabetes mellitus associated with mortality and severity of COVID-19? A meta-analysis. *Diabetes Metab Syndr* 2020;14:535–45.
247 Pranata R, Huang I, Lim MA, *et al*. Impact of cerebrovascular and cardiovascular diseases on mortality and severity of COVID-19–systematic review, meta-analysis, and meta-regression. *J Stroke Cerebrovasc Dis* 2020;29:104949.
248 Shang L, Shao M, Guo Q, *et al*. Diabetes mellitus is associated with severe infection and mortality in patients with COVID-19: a systematic review and meta-analysis. *Arch Med Res* 2020;51:700–9.
249 Varikasuvu SR, Varshney S, Dutt N. Markers of coagulation dysfunction and inflammation in diabetic and non-diabetic COVID-19. *J Thromb Thrombolysis* 2021;51:941–6.
250 Maddaloni E, Buzzetti R. Covid-19 and diabetes mellitus: unveiling the interaction of two pandemics. *Diabetes Metab Res Rev* 2020:e33213321.
251 Kermali M, Khalsa RK, Pillai K, *et al*. The role of biomarkers in diagnosis of COVID-19 - A systematic review. *Life Sci* 2020;254:117788.
252 Morra ME, Van Thanh L, Kamel MG, *et al*. Clinical outcomes of current medical approaches for middle East respiratory syndrome: a systematic review and meta-analysis. *Rev Med Virol* 2018;28:e1977.
253 Matsuyama R, Nishiura H, Kutsuna S, *et al*. Clinical determinants of the severity of middle East respiratory syndrome (MERS): a systematic review and meta-analysis. *BMC Public Health* 2016;16:1203.
254 Yang J, Hu J, Zhu C. Obesity aggravates COVID-19: a systematic review and meta-analysis. *J Med Virol* 2021;93:257–61.
255 Guh DP, Zhang W, Bansback N, *et al*. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. *BMC Public Health* 2009;9:88.
256 Semenzato L, Botton J, Drouin J, *et al*. Chronic diseases, health conditions and risk of COVID-19-related hospitalization and in-hospital mortality during the first wave of the epidemic in France: a cohort study of 66 million people. *Lancet Reg Health Eur* 2021;8:100158.
257 Thakur B, Dubey P, Benitez J, *et al*. A systematic review and meta-analysis of geographic differences in comorbidities and associated severity and mortality among individuals with COVID-19. *Sci Rep* 2021;11:8562.
258 Dubey P, Thakur B, Reddy S, *et al*. Current trends and geographical differences in therapeutic profile and outcomes of COVID-19 among pregnant women - a systematic review and meta-analysis. *BMC Pregnancy Childbirth* 2021;21:247.

CHAPTER 11

Effect of hydroxychloroquine with or without azithromycin on the mortality of coronavirus disease 2019 (COVID-19) patients: a systematic review and meta-analysis

Effect of hydroxychloroquine with or without azithromycin on the mortality of coronavirus disease 2019 (COVID-19) patients: a systematic review and meta-analysis

Thibault Fiolet , Anthony Guihur, Mathieu Edouard Rebeaud, Matthieu Mulot, Nathan Peiffer-Smadja, Yahya Mahamat-Saleh

I contributed to this work conceptually and in the manuscript writing.

Authors' contributions
TF designed the research. TF, MR, AG, MM, NPS and YMS conducted the research. TF performed the statistical analysis and wrote the first draft of the paper. MR, AG, MM, NPS and YMS contributed to the writing of the paper. All authors contributed to the data interpretation, revised each draft for important intellectual content, and read and approved the final manuscript.

Clinical Microbiology and Infection 27 (2021) 19–27

Contents lists available at ScienceDirect

Clinical Microbiology and Infection

journal homepage: www.clinicalmicrobiologyandinfection.com

Systematic review

Effect of hydroxychloroquine with or without azithromycin on the mortality of coronavirus disease 2019 (COVID-19) patients: a systematic review and meta-analysis

Thibault Fiolet [1, 2, *], Anthony Guihur [3], Mathieu Edouard Rebeaud [3], Matthieu Mulot [4], Nathan Peiffer-Smadja [5, 6, 7], Yahya Mahamat-Saleh [1, 2]

[1] CESP (Centre for Research in Epidemiology and Population Health), Faculté de Médecine—Université Paris-Sud, Faculté de Médecine—UVSQ, INSERM, Université Paris Saclay, Villejuif, France
[2] Gustave Roussy, Villejuif, France
[3] Department of Plant Molecular Biology, Faculty of Biology and Medicine, University of Lausanne, Switzerland
[4] Laboratory of Soil Biodiversity, Faculty of Science, University of Neuchâtel, Switzerland
[5] Université de Paris, IAME, INSERM, Paris, France
[6] National Institute for Health Research Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance, Imperial College, London, UK
[7] Infectious and Tropical Diseases Department, Bichat-Claude Bernard Hospital, AP-HP, Paris, France

ARTICLE INFO

Article history:
Received 6 July 2020
Received in revised form
10 August 2020
Accepted 15 August 2020
Available online 26 August 2020

Editor: L. Leibovici

Keywords:
Azithromycin
Chloroquine
Coronavirus
Coronavirus disease 2019
Hydroxychloroquine
Meta-analysis
Mortality
Severe acute respiratory syndrome coronavirus 2

ABSTRACT

Background: Hydroxychloroquine or chloroquine with or without azithromycin have been widely promoted to treat coronavirus disease 2019 (COVID-19) following early *in vitro* antiviral effects against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).
Objective: The aim of this systematic review and meta-analysis was to assess whether chloroquine or hydroxychloroquine with or without azithromycin decreased COVID-19 mortality compared with the standard of care.
Data sources: PubMed, Web of Science, Embase Cochrane Library, Google Scholar and MedRxiv were searched up to 25 July 2020.
Study eligibility criteria: We included published and unpublished studies comparing the mortality rate between patients treated with chloroquine or hydroxychloroquine with or without azithromycin and patients managed with standard of care.
Participants: Patients ≥18 years old with confirmed COVID-19.
Interventions: Chloroquine or hydroxychloroquine with or without azithromycin.
Methods: Effect sizes were pooled using a random-effects model. Multiple subgroup analyses were conducted to assess drug safety.
Results: The initial search yielded 839 articles, of which 29 met our inclusion criteria. All studies except one were conducted on hospitalized patients and evaluated the effects of hydroxychloroquine with or without azithromycin. Among the 29 articles, three were randomized controlled trials, one was a non-randomized trial and 25 were observational studies, including 11 with a critical risk of bias and 14 with a serious or moderate risk of bias. After excluding studies with critical risk of bias, the meta-analysis included 11 932 participants for the hydroxychloroquine group, 8081 for the hydroxychloroquine with azithromycin group and 12 930 for the control group. Hydroxychloroquine was not significantly associated with mortality: pooled relative risk (RR) 0.83 (95% CI 0.65–1.06, $n = 17$ studies) for all studies and RR = 1.09 (95% CI 0.97–1.24, $n = 3$ studies) for randomized controlled trials. Hydroxychloroquine with azithromycin was associated with an increased mortality (RR = 1.27; 95% CI 1.04–1.54, $n = 7$ studies). We found similar results with a Bayesian meta-analysis.

* Corresponding author: T. Fiolet, Centre for Research in Epidemiology and Population Health, Inserm U1018 'Health across Generations' Team and Paris-Sud 11 University/Paris-Saclay University, 114 rue Edouard Vaillant, 94805, Villejuif Cedex, France.
E-mail address: Thibault.fiolet@gustaveroussy.fr (T. Fiolet).
(T. Fiolet)

https://doi.org/10.1016/j.cmi.2020.08.022

Conclusion: Hydroxychloroquine alone was not associated with reduced mortality in hospitalized COVID-19 patients but the combination of hydroxychloroquine and azithromycin significantly increased mortality. **Thibault Fiolet, Clin Microbiol Infect 2021;27:19**

Introduction

On 31 December 2019, the WHO identified an unknown pneumonia caused by a new coronavirus, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), in Wuhan, China. By 30 July 2020, WHO confirmed more than 17 million cases and 667 935 deaths [1]. Chloroquine (CQ) and its derivative hydroxychloroquine were rapidly identified as potential drug candidates because chloroquine had an antiviral activity against Middle East respiratory syndrome and severe acute respiratory syndrome *in vitro* [2]. *In vitro* antiviral activity of the aminoquinolines hydroxychloroquine and chloroquine was confirmed against SARS-CoV-2 and a study reported a synergistic effect of hydroxychloroquine with azithromycin against SARS-CoV-2 [3]. These drugs appeared as potential low-cost treatments for individuals with coronavirus disease 2019 (COVID-19) [4–7] and received wide and speculative coverage by the international press and the US President [8].

Subsequently, hydroxychloroquine and azithromycin were tested in a study where macaques were infected by SARS-CoV-2 and received either a high dose of hydroxychloroquine (90 mg/kg on day 1 then 45 mg/kg) or a low hydroxychloroquine dose (30 mg/kg on day 1 then 15 mg/kg) [9]. Hydroxychloroquine with or without azithromycin did not improve the time to viral clearance regardless of the stage of disease: prophylaxis, early treatment or late treatment.

Among the ongoing trials, chloroquine or hydroxychloroquine are among the most studied drugs [10,11]. Until today, most of the published studies on hydroxychloroquine with a comparative group (standard care) were observational and non-randomized with inconsistent results [12–18]. Given the magnitude of the COVID-19 pandemic and the need for effective therapeutics, timely meta-analyses can play an important role in assessing the impacts of chloroquine and hydroxychloroquine compared with standard of care on reliable clinical outcomes such as mortality. Previous meta-analyses on COVID-19 included a limited number of studies and used unadjusted risk ratios [19–21].

The aim of this systematic review and meta-analysis was to assess whether chloroquine or hydroxychloroquine with or without azithromycin decreased the mortality of COVID-19 compared with standard of care.

Methods

The research question was: in individuals with confirmed COVID-19, is the addition of hydroxychloroquine or chloroquine with or without azithromycin to the standard of care effective in improving survival?

PICO question

Population patients with confirmed COVID-19.
Intervention hydroxychloroquine or chloroquine, with or without azithromycin.
Comparison a standard of care.
Outcomes the survival rate of COVID-19 patients.

Data sources, search strategy

A search was performed using PubMed, Web of Science, Embase and Cochrane Review up to 25 July 2020 with the following string search: (COVID-19 OR SARS-CoV-2) AND (MORTALITY OR DEATH) AND (HYDROXYCHLOROQUINE OR hydroxychloroquine) (see Supplementary material, Text S1). Given that the number of articles about hydroxychloroquine and COVID-19 is rapidly growing, we also manually searched for additional references on the MedRxiv preprint server and on Google Scholar with the same terms. An additional search on PubMed, Web of Science and Cochrane Review was conducted for CQ with the search terms described in the Supplementary materials (Text S1): (COVID-19 OR SARS-CoV-2) AND (MORTALITY OR DEATH) AND (CHLOROQUINE OR chloroquine). This meta-analysis was conducted following the PRISMA statements in the Supplementary material (Text S2). This study has been recorded on the international database of prospectively registered systematic reviews, PROSPERO (Registration number: CRD42020190801).

Study selection

Study selection was conducted by two investigators (TF and YM) who screened the titles and the abstracts. Discrepancies were resolved by a third investigator (AG). Inclusion criteria were (a) reports containing original data with available risk estimates (hazard ratios (HR), odds ratios (OR), relative risk (RR) and/or with data on the number of deaths in hydroxychloroquine/chloroquine and control groups; (b) any publication dates; (c) comparative studies with a control group with no hydroxychloroquine nor chloroquine; and (d) PCR-confirmed cases of COVID-19. Studies reporting no deaths, reviews and meta-analyses, commentaries, editorials and *in vitro* and *in vivo* animal studies were excluded.

Data extraction

Two investigators (TF and YM) extracted the following data for each study: study design, publication date, journal, location, number of participants and deaths (in treatment and control groups), hydroxychloroquine or chloroquine doses when available, effect size (HR, OR or RR) and 95% CI for reported risk estimates. The estimates from the model, adjusted for the maximum number of covariates, were used to control potential confounders, according to Cochrane Methodology [22]. For each study, risk factors associated with higher mortality were taken into account through the reported adjusted effect sizes.

When studies did not report an effect size for mortality risk [17,23,24], we used the number of deaths per group to calculate an unadjusted relative risk using *metabin* function in *meta* package in R Software [25].

For all the other studies, reported adjusted OR, RR or HR were used.

Individual risk of bias

The quality of each study was assessed with the ROBIN-I tool following Cochrane guidelines for non-randomized studies and with Rob2 for randomized studies [26,27].

Outcome

The outcome was the mortality of COVID-19 patients.

Statistical analysis

Effect of chloroquine/hydroxychloroquine alone and hydroxychloroquine + azithromycin

A primary meta-analysis was performed to compare the survival rate (or mortality) between patients treated with chloroquine or hydroxychloroquine and standard of care. Then, the relationship between hydroxychloroquine associated with azithromycin and mortality was assessed. HR, OR and RR were treated as equivalent measures of mortality risk. Pooled RR were determined by using a random effect model with inverse variance weighting (DerSimonian–Laird method) [28]. Significance was checked using a *Z*-test, where $p < 0.05$ is considered as significant. The absolute risk difference (RD) was calculated from the UK baseline hospital mortality risk (BR) of 26% (according to ISARIC WHO CCP-UK cohort based on 20 133 patients) using the formula $RD = BR \times (RR - 1)$ [29].

Heterogeneity was assessed by the Cochrane Q test and I^2 test [30]. $30\% < I^2 < 60\%$ was interpreted as moderate heterogeneity and $I^2 > 60$ as substantial heterogeneity. A funnel plot was constructed to assess the publication bias. Begg's and Egger's tests were conducted to assess the publication bias [31,32]. RR or HR were used to assess mortality risk within a 95% CI. In the main analysis, studies with critical bias were excluded. A sensitivity analysis including these studies was conducted. A Bayesian meta-analysis was performed to test the robustness of our results, allowing incorporation of full uncertainty in all parameters [33]. The traditional random-effect model has fixed parameters for the distribution of the true treatment effect RR with an unknown mean θ, within-study variance σ^2 and between-study variance τ^2. The Bayesian random-effect model assumes that these parameters are random with a probability distribution. Two prior distributions were tested μ~Normal (1,100) with a large variance and τ~Half-Cauchy (0,0.5) and a second scenario with μ~Normal (1,1) and τ~Half-Cauchy (0,0.5). The Bayesian analysis was conducted with the R package *brms* [34].

Subgroup analysis

Subgroup analyses were conducted according to the quality assessment to explore the source of heterogeneity among observational studies. We performed stratified analyses by type of article (peer-reviewed versus unpublished), use of an adjustment on confounding factors (studies with $RR_{unadjusted}$ versus $RR_{adjusted}$), mean daily dose of hydroxychloroquine or chloroquine (continuous), median population age across the studies, level of bias risk identified with ROBIN-I (moderate/serious/critical) [26] and when we excluded studies with cancer and dialysis patients. Mean daily dose of hydroxychloroquine or chloroquine was the daily average between the loading dose and the maintenance doses. Additionally, influence analysis was conducted by omitting each study to find potential outliers [34]. Influence analysis is used to detect studies that influence the overall estimate of a meta-analysis the most, omitting one study at a time (leave-one-out method).

A two-sided p-value <0.05 was considered statistically significant. All analyses were conducted using R version 3.6.1 with *meta* package and *robvis* package [35].

Results

Literature search

A flow chart is presented in Fig. 1. After searching PubMed, Cochrane Review and Web of Science, 839 articles were identified. After screening the title and the abstract, only 21 articles about hydroxychloroquine and COVID-19 were included for further consideration. We excluded 564 articles that did not meet the inclusion criteria. We did not find any non-English articles meeting our inclusion criteria. Two duplicate studies on the same cohort were excluded [12,36]. Two Chinese randomized controlled trials (RCT) on hydroxychloroquine reported zero deaths in both treatment and control groups [37,38] and so their results were not included in our meta-analysis. Ten articles from Medrxiv/Google Scholar were added, so 29 articles were included, of which 25 were observational studies, one was an interventional non-randomized study and three were RCT. These studies included 27 articles for hydroxychloroquine [14–19,23,24,36,39–56] and 12 articles for hydroxychloroquine + azithromycin [18,36,41,42,47,48,50,51,57–60]. For chloroquine, after searching PubMed, Cochrane Review, Embase and Web of Science, 449 articles were identified. After screening the title and the abstract, only one Brazilian RCT and three observational studies described chloroquine and COVID-19. However, among these studies, those by Borba et al. and Saleh et al. did not have a standard of care comparative group [61,62]. Khamis et al. did not report death data related to CQ and Huang et al. did not report any death [63,64]. Consequently, no study on chloroquine met our inclusion criteria.

Study characteristics

This meta-analysis included 15 190 patients in the hydroxychloroquine group, 8081 patients in the hydroxychloroquine with azithromycin group and 14 060 patients in the standard of care group with 3152 deaths, 1063 deaths and 2857 deaths, respectively. Individual studies are described in the Supplementary material (Tables S1 and S2). All included studies were carried out on hospitalized patients except for one [39]. Mean (±SD) age of participants was 62.1 ± 8.5 years. Ten studies were conducted in the USA [15,18,23,41,42,49,50,53,56,58], four in Spain [16,17,44,57], seven in France [13,24,46,48,54,59,60], one in the UK [40], two in Italy [43,65], one in China [14], one in Brazil [51] and three in several other countries (USA, Canada, Italy and Spain) [39,47,52]. Twenty-two articles were published [13–15,17,18,24,39,41,43,44,46,47,49–54,56,57,59,60,65] and six articles were preprints [16,23,40,42,48,58]. Mean daily dose of hydroxychloroquine ranged from 333 mg/day to 945 mg/day. Few studies precisely described concomitant use of corticosteroids (see Supplementary material, Table S3) [15–17,44,48,50–52,65]. Only the RECOVERY trial precisely reported the use of dexamethasone (8% versus 9% in both arms) [40].

Study quality

Risk of bias was assessed with ROBIN-I for non-randomized studies ($n = 26$) and Rob2 for RCT ($n = 3$) (see Supplementary material, Figs S1 and S2). Three RCT had some concerns [39,40,51] and one interventional non-randomized study had critical risk of bias [24]. Among the observational studies, fourteen articles had a moderate or serious risk of bias [13–18,41,42,44,46–48,56,58] and eleven studies had a critical risk of bias

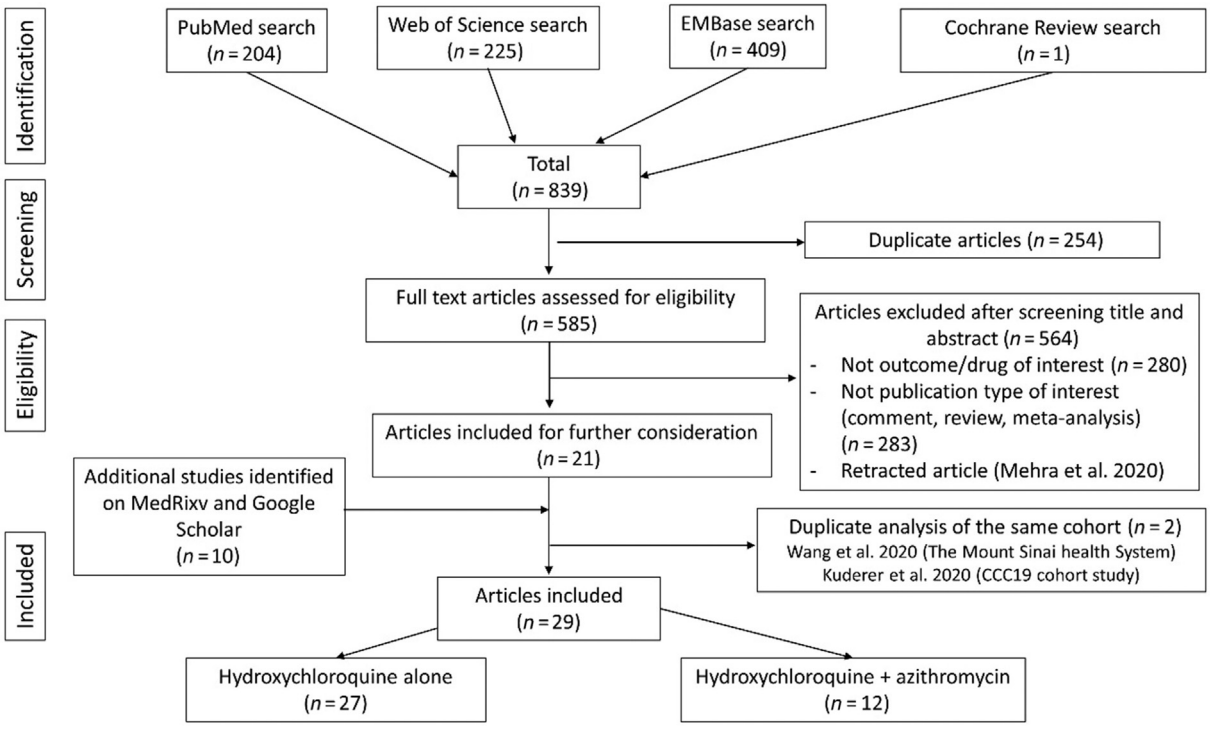


Fig. 1. Flow diagram of study selection process.

[23,43,49,50,52–54,57,59,60,65]. Eleven observational studies did not report adjusted effect sizes to control confusion and selection bias [23,24,43,44,49,53,54,57,59,60,65]. Quality of studies was lowered by the lack of information about the assignment of treatment, the time between start of follow up and start of intervention, some unbalanced co-intervention with other antiviral and antibiotic drugs and imbalance between groups for confounders such as co-morbidities and age.

Hydroxychloroquine and mortality

After excluding studies with critical bias, the pooled RR for COVID-19 mortality was 0.83 (95% CI 0.65–1.06, $n = 17$ studies) indicating no significant association between hydroxychloroquine and COVID-19 mortality (Fig. 2). Under the hypothesis of having a baseline mortality risk of 26% (based on ISARIC WHO CCP-UK cohort [29]), these pooled relative risk values would correspond to a non-significant risk difference of −4.4% [29] (Table 1). There was a significant subgroup difference between RCT and non-randomized studies ($P_{\text{heterogeneity between}} = 0.03$) with respectively $RR_{RCT} = 1.09$ (95% CI 0.97–1.24) and $RR_{\text{non-randomized}} = 0.79$ (95% CI 0.60–1.04) (Fig. 2). Among observational studies with a moderate risk of bias, we found no association between hydroxychloroquine and mortality $RR_{\text{moderate bias}} = 1.03$ (95% CI 0.91–1.17, $I^2 = 0\%$, $n = 7$ studies) with no subgroup heterogeneity (see Supplementary material, Table S4, Fig. S3). Results remained non-significant with influence analysis (see Supplementary material, Fig. S4). The Bayesian meta-analysis led to similar results with a pooled RR for mortality of 0.93 (95% CI 0.72–1.14, $n = 17$ studies) (see Supplementary material, Table S5, Fig. S5). In sensitivity analysis, after inclusion of studies with critical risk of bias, the global RR was marginally not significant 0.80 (95% CI 0.65–1.00) (see Supplementary material, Table S6).

There was a significant higher heterogeneity among non-randomized studies compared with RCT ($I^2 = 84\%$, $P_{\text{heterogeneity within}} < 0.01$). In fact, heterogeneity was null for RCT. Egger's test (p 0.68) and Begg's test (p 0.13) were not significant for asymmetry of the funnel plot, indicating that there was no major publication bias for non-randomized studies (see Supplementary material, Fig. S6).

Hydroxychloroquine with azithromycin and mortality

After exclusion of studies with critical bias, the pooled RR for COVID-19 mortality was 1.27 (95% CI 1.04–1.54, $n = 7$), indicating an increased mortality linked to the use of hydroxychloroquine with azithromycin. With a baseline hospital mortality of 26%, we identified a significant absolute risk difference of +7%. We found an increased risk of mortality in patients treated with hydroxychloroquine and azithromycin compared with standard of care (RR 1.29, 95% CI 1.06–1.58, $n = 6$) among non-randomized studies, but this relationship was not found in the single Brazilian RCT, with no heterogeneity observed across the study design ($P_{\text{heterogeneity between}} = 0.28$) (Fig. 3). There was a low heterogeneity across the included studies ($I^2 = 38\%$, p 0.14). Egger's test (p 0.70) and Begg's test (p 0.65) were not significant but the asymmetry in the funnel plot indicates that a publication bias could be present (see Supplementary material, Fig. S7). However, the number of included studies was small. Subgroup analyses are described in the Supplementary material (Table S4, Fig. S8). The Bayesian meta-analysis led to similar results with a pooled RR for mortality of 1.32 (95% CI 0.97–1.68, $n = 7$ studies) (see Supplementary material, Table S5, Fig. S9). The increase in mortality was also significant with influence analysis (see Supplementary material, Fig. S10).

Discussion

This meta-analysis summarized the results of 25 observational studies, three RCT and one interventional non-randomized study on the effect of hydroxychloroquine with or without azithromycin on the mortality of COVID-19 patients (Table 1). Despite our

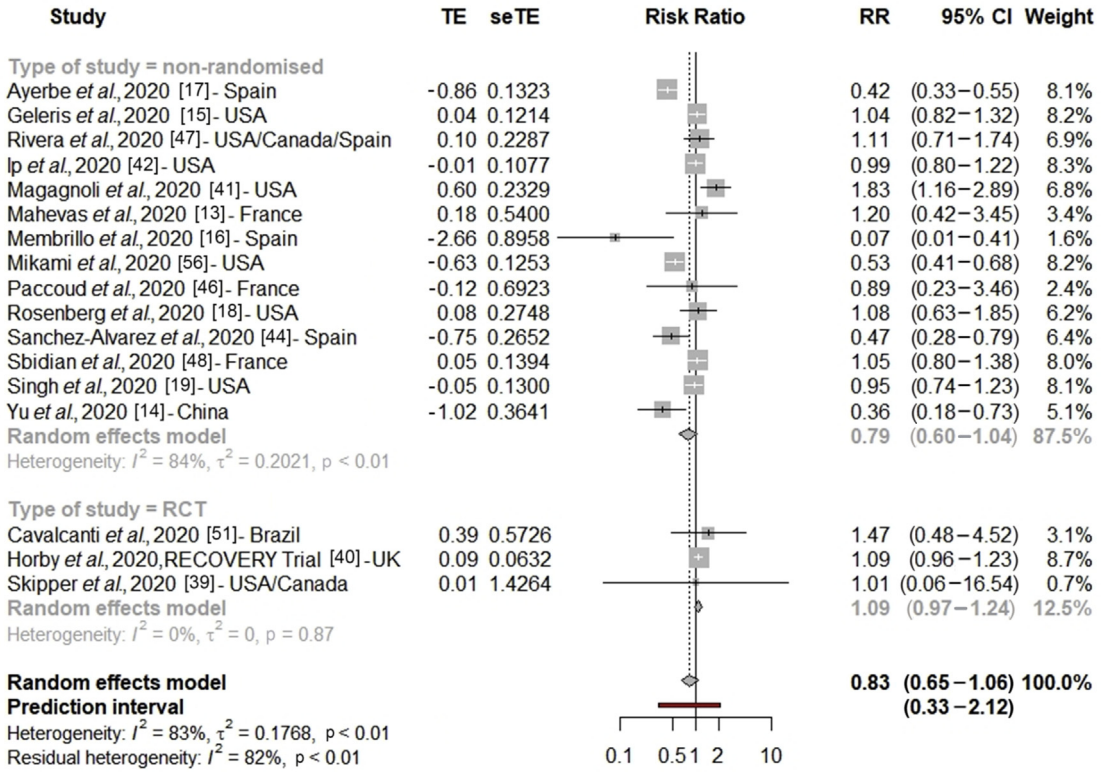


Fig. 2. Forest plot of the association between hydroxychloroquine alone and COVID-19 mortality (excluding studies with critical risk of bias). RR, risk ratio.

inclusion criteria that did not specify the stage of the disease, all the studies were conducted with hospitalized patients except the RCT by Skipper et al. [39]. Our results show that hydroxychloroquine alone was not associated with reduced mortality in COVID-19 patients, but the combination of hydroxychloroquine and azithromycin significantly increased mortality. We found similar results with a Bayesian analysis.

Our meta-analysis reported a high heterogeneity for hydroxychloroquine alone, but this heterogeneity was lowered among RCT, studies with moderate risk of bias and for the association of hydroxychloroquine + azithromycin. The variable quality of the studies (not reporting hydroxychloroquine dose, the lack of adjustment in reported estimates) may explain one part of the heterogeneity observed according to our subgroup analysis (see Supplementary material, Table S4).

A previous systematic review only included eight studies on all-cause mortality in COVID-19 patients [13–16,23,38,41,66] and concluded that the level of evidence for a hydroxychloroquine effect was very weak [67]. A preprint meta-analysis, using routinely collected records from clinical practice in Germany, Spain, the UK, Japan and the USA compared the use of hydroxychloroquine with sulfasalazine [68]. This study observed an increased risk of 30-day cardiovascular mortality (HR = 2.19, 95% CI 1.22–3.94), although the study lacked a standard of care comparative group. Some previous meta-analyses were also conducted on hydroxychloroquine and various health end points including mortality. However, these studies did not report all the published and unpublished literature, including a very limited number of studies: from three articles [19,20] to six articles [21]. These previous meta-analyses did not perform subgroup and sensitivity analyses to test the effect of pooling RCT and observational studies, nor did they study the source of heterogeneity. They used unadjusted risk ratios (calculated with the number of events in each group) whereas in our meta-analysis, we used adjusted relative risk [69] and we ran sensitivity analyses on the adjustment of effect size. Statistical adjustments for key prognostic variables limit confusion bias, especially in observational studies, which are not randomized. This meta-analysis confirmed the partial preliminary results of these other meta-analyses about the absence of effect for hydroxychloroquine on survival and found an increased mortality with the

Table 1
Relative risk and risk difference for mortality associated with hydroxychloroquine with or without azithromycin, assuming a UK mortality rate in hospital of 26% according to the ISARIC WHO CCP-UK cohort

Outcome: All-cause mortality	Number of studies	Pooled relative risk (95% CI)	Risk difference (95% CI)
Hydroxychloroquine alone			
All studies	17	0.83 (0.65–1.06)	-4.4% (−9% to +1.5%)
Non-randomized studies	14	0.79 (0.60–1.04)	-5.5% (−10% to +1%)
Randomized studies	3	1.09 (0.97–1.24)	+2.3% (−0.8% to +6.2%)
Hydroxychloroquine with azithromycin			
All studies	7	1.27 (1.04–1.54)	+7% (+1% to +14%)
Non-randomized studies	6	1.29 (1.06–1.58)	+7.5% (+1.6% to +15%)
Randomized studies	1	0.64 (0.18–2.24)	-9% (−21% to +32%)

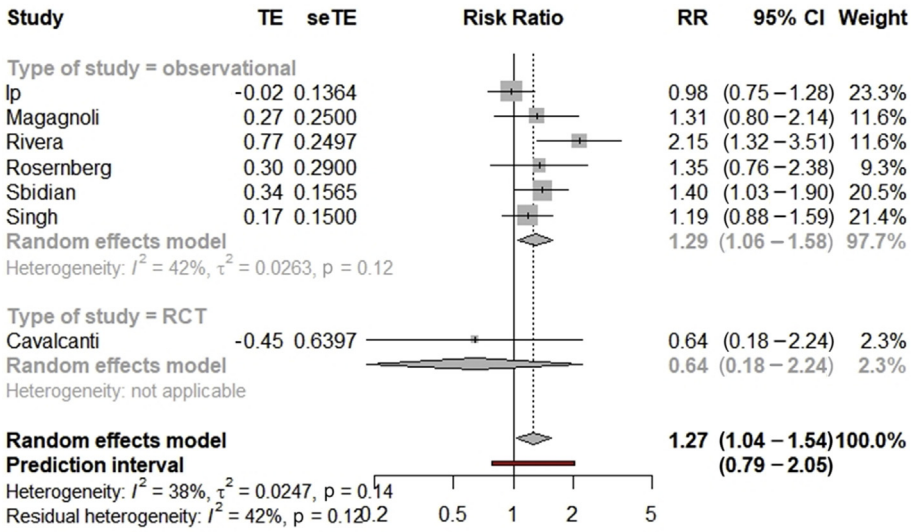


Fig. 3. Forest plot of the association between hydroxychloroquine with azithromycin and COVID-19 mortality (excluding studies with critical risk of bias). RR, risk ratio.

use of the combination of hydroxychloroquine with azithromycin in COVID-19 patients. These results confirm the preliminary findings of several observational studies, which have shown that the combination of hydroxychloroquine and azithromycin might increase the risk of acute, life-threatening cardiovascular events [70]. A first study found that, among individuals treated with this combination, 6 out of 18 (33%) developed a significant increase in the QTc interval [70]. Another work found that in 84 patients treated with hydroxychloroquine + azithromycin, nine had a severe prolongation of QTc [71]. The combination of hydroxychloroquine + azithromycin was associated with a greater variation in the QTc interval compared with hydroxychloroquine alone in a study with 90 patients [72]. In a study conducted in New York on 1438 patients, cardiac arrest was significantly more likely in patients receiving hydroxychloroquine with azithromycin compared to patients receiving neither of the two drugs (adjusted OR 2.13, 95% CI 1.12–4.05) [18]. Finally, a study conducted on the WHO database bringing together more than 167 000 patients found an increased risk of potentially fatal acute cardiac events in patients treated with azithromycin alone or with hydroxychloroquine alone [73]. The combination of the two drugs posed an even greater risk of life-threatening acute cardiac effects [18,72,73].

Several national health organizations (US Food and Drug Administration [74], French Agency for the Safety of Health Products [75], European Medicine Agency [76]) raised concerns about using unapproved drugs for COVID-19. The French Agency for the Safety of Health Products and the US Food and Drug Administration removed the authorization for the use of hydroxychloroquine outside clinical trials. The Indian Council of Medical Research took the opposite position and recommended chemoprophylaxis with hydroxychloroquine for asymptomatic individuals [77]. Finally, in the comparative peer-reviewed studies, a clear conclusion on hydroxychloroquine is not possible because of the small sample size, the lack of well-performed RCT (mainly non-randomized and retrospective studies) and inconsistent results. Many preprints without a comparative group and without randomization added to confusion surrounding this highly politicized topic [78]. There is a gap between the speed of clinical research and the expectation of a clear solution to treat people with COVID-19. Indeed, producing robust clinical trials is necessarily time-consuming. In a press communication, on 20 June 2020, the US National Institutes of Health stopped the clinical trial of hydroxychloroquine because this drug was very unlikely to be efficient for treatment of individuals with COVID-19 [79]. Based on SOLIDARITY trial results, the WHO previously took the same decision [80].

A Bayesian meta-analysis confirmed our findings from classical random-effect meta-analysis. We included several unpublished papers to minimize the publication bias. Our subgroup analysis by published studies (versus unpublished studies) found that the inclusion of preprints did not change the results. Exclusion of grey literature (unpublished studies, with limited distribution) could lead to an exaggeration of the intervention effect by 15% [81]. There is limited evidence to identify whether grey studies have a poorer methodological quality than published studies [82].

A major limitation is the inclusion of individuals with different levels of COVID-19 severity. However, we could not conduct subgroup analysis for severity because most study reports do not use the same definition of severity and do not report the same biological and clinical outcomes. We also noted a high level of heterogeneity in the administration of hydroxychloroquine (dosing, timing between hospital administration and intervention, duration). In some studies, these data were not reported at all. Another limitation comes from the studies that did not report adjusted effect size when mortality was not the primary end point, leading to a high risk of confounding bias. As is usually done, this meta-analysis was based on aggregated data, without access to original patient data. Most of the included studies were observational studies, which are not adapted to identify a causal association. Indeed, some of the included studies had very low quality of evidence (missing data, small sample size, confusion bias, bias in classification of intervention and selection bias), although our supplementary analyses and the exclusion of these articles did not change the results. Finally, this meta-analysis did not include results from the European DisCoVeRy trial and the WHO Solidarity trial, which are not yet published or communicated [80].

In conclusion, this meta-analysis clearly shows that hydroxychloroquine alone is not effective for the treatment of people with COVID-19 and that the combination of hydroxychloroquine and azithromycin increases the risk of mortality. These data support current clinical recommendations such as those of the National Institutes of Health [83], which do not recommend the use of hydroxychloroquine alone or in combination with azithromycin for COVID-19. There is already a great number of studies that have evaluated hydroxychloroquine alone or in combination [10] and it

seems unlikely at this stage that any efficacy will emerge. Our results suggest that there is no need for further studies evaluating these molecules, and the European DisCoveRy clinical trial or the WHO international Solidarity clinical trial have already discontinued treatment arms using hydroxychloroquine [80,84].

Transparency declaration

All authors declare no support from any organization for the submitted work other than that described above; no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years; and no other relationships or activities that could appear to have influenced the submitted work.

Funding

There was no specific funding for this meta-analysis.

Authors' contributions

TF designed the research. TF, MR, AG, MM, NPS and YMS conducted the research. TF performed the statistical analysis and wrote the first draft of the paper. MR, AG, MM, NPS and YMS contributed to the writing of the paper. All authors contributed to the data interpretation, revised each draft for important intellectual content, and read and approved the final manuscript.

Acknowledgements

The authors would like to thank Dominique Meroux for proofreading the manuscript. We also thank Drifa Belhadi for her helpful comments on the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cmi.2020.08.022.

References

[1] World Health Organization (WHO). Coronavirus disease (COVID-19). Situation Report - 153. 21 June 2020. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports. [Accessed 8 September 2020].
[2] Dyall J, Gross R, Kindrachuk J, Johnson RF, Olinger GG, Hensley LE, et al. Middle East respiratory syndrome and severe acute respiratory syndrome: current therapeutic options and potential targets for novel therapies. Drugs 2017;77:1935–66.
[3] Andreani J, Le Bideau M, Duflot I, Jardot P, Rolland C, Boxberger M, et al. *In vitro* testing of combined hydroxychloroquine and azithromycin on SARS-CoV-2 shows synergistic effect. Microb Pathog 2020;145:104228.
[4] Hill A, Wang J, Levi J, Heath K, Fortunak J. Minimum costs to manufacture new treatments for COVID-19. J Virus Erad 2020;6:61–9.
[5] Yao X, Ye F, Zhang M, Cui C, Huang B, Niu P, et al. In Vitro antiviral activity and projection of optimized dosing design of hydroxychloroquine for the treatment of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Clin Infect Dis n.d. https://doi.org/10.1093/cid/ciaa237.
[6] Liu J, Cao R, Xu M, Wang X, Zhang H, Hu H, et al. Hydroxychloroquine, a less toxic derivative of chloroquine, is effective in inhibiting SARS-CoV-2 infection in vitro. Cell Discov 2020;6:1–4.
[7] Maisonnasse P, Guedj J, Contreras V, Behillil S, Solas C, Marlin R, et al. Hydroxychloroquine in the treatment and prophylaxis of SARS-CoV-2 infection in non-human primates. Nature 2020. https://doi.org/10.1038/s41586-020-2558-4.
[8] DeJong C, Wachter RM. The risks of prescribing hydroxychloroquine for treatment of COVID-19—first, do no harm. JAMA Intern Med 2020. https://doi.org/10.1001/jamainternmed.2020.1853.
[9] Maisonnasse P, Guedj J, Contreras V, Behillil S, Solas C, Marlin R, et al. Hydroxychloroquine use against SARS-CoV-2 infection in non-human primates. Nature 2020:1–8.
[10] Fragkou PC, Belhadi D, Peiffer-Smadja N, Moschopoulos CD, Lescure F-X, Janocha H, et al. Review of trials currently testing treatment and prevention of COVID-19. Clin Microbiol Infect 2020;26:988–98.
[11] Peiffer-Smadja N, Lescure F-X, Sallard E, Ravaud P, Vegreville B, Zeitoun J-D. Anticovid, a comprehensive open-access real-time platform of registered clinical studies for COVID-19. J Antimicrob Chemother n.d. https://doi.org/10.1093/jac/dkaa223.
[12] Kuderer NM, Choueiri TK, Shah DP, Shyr Y, Rubinstein SM, Rivera DR, et al. Clinical impact of COVID-19 on patients with cancer (CCC19): a cohort study. Lancet 2020. https://doi.org/10.1016/S0140-6736(20)31187-9.
[13] Mahévas M, Tran V-T, Roumier M, Chabrol A, Paule R, Guillaud C, et al. Clinical efficacy of hydroxychloroquine in patients with COVID-19 pneumonia who require oxygen: observational comparative study using routine care data. BMJ 2020;369. https://doi.org/10.1136/bmj.m1844.
[14] Yu B, Li C, Chen P, Zhou N, Wang L, Li J, et al. Low dose of hydroxychloroquine reduces fatality of critically ill patients with. Sci China Life Sci 2020:1–7.
[15] Geleris J, Sun Y, Platt J, Zucker J, Baldwin M, Hripcsak G, et al. Observational study of hydroxychloroquine in hospitalized patients with COVID-19. N Engl J Med 2020. https://doi.org/10.1056/NEJMoa2012410.
[16] Membrillo FJ, Ramírez-Olivencia G, Estébanez M, Dios B de, Herrero MD, Mata T, et al. Early hydroxychloroquine is associated with an increase of survival in COVID-19 patients: an observational study. 2020. https://doi.org/10.20944/preprints202005.0057.v2. Preprints.
[17] Ayerbe L, Risco C, Ayis S. The association between treatment with heparin and survival in patients with Covid-19. J Thromb Thrombolysis 2020:1–4. https://doi.org/10.1007/s11239-020-02162-z.
[18] Rosenberg ES, Dufort EM, Udo T, Wilberschied LA, Kumar J, Tesoriero J, et al. Association of treatment with hydroxychloroquine or azithromycin with in-hospital mortality in patients with COVID-19 in New York State. JAMA 2020. https://doi.org/10.1001/jama.2020.8630.
[19] Singh AK, Singh A, Singh R, Misra A. Hydroxychloroquine in patients with COVID-19: a systematic review and meta-analysis. Diabetes Metab Syndr 2020;14:589–96.
[20] Sarma P, Kaur H, Kumar H, Mahendru D, Avti P, Bhattacharyya A, et al. Virological and clinical cure in COVID-19 patients treated with hydroxychloroquine: a systematic review and meta-analysis. J Med Virol 2020;92:776–85.
[21] Patel TK, Barvaliya M, Kevadiya BD, Patel PB, Bhalla HL. Does adding of hydroxychloroquine to the standard care provide any benefit in reducing the mortality among COVID-19 patients?: a systematic review. J Neuroimmune Pharmacol 2020. https://doi.org/10.1007/s11481-020-09930-x.
[22] Cochrane Training. Cochrane handbook for systematic reviews of interventions. Part 3: special topics. 13.6.2.2. Combining studies (version 6). 2019. Available from: https://handbook-5-1.cochrane.org/chapter_13/13_6_2_2_combining_studies.htm. [Accessed 28 July 2020].
[23] Barbosa J, Kaitis D, Freedman R, Le K, Lin X. Clinical outcomes of hydroxychloroquine in hospitalized patients with COVID-19: a quasi-randomized comparative study. Dropbox 2020. Available from: https://www.dropbox.com/s/urzapkyij542qx5/NEJM_Clinical%20Outcomes%20of%20Hydroxychlorquine%20in%20Patients%20with%20COVID19.pdf.pdf.pdf.pdf.pdf.pdf.pdf?dl=0. [Accessed 8 June 2020].
[24] Gautret P, Lagier J-C, Parola P, Hoang VT, Meddeb L, Mailhe M, et al. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. Int J Antimicrob Agents 2020:105949. https://doi.org/10.1016/j.ijantimicag.2020.105949.
[25] Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health 2019;22:153–60.
[26] Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016;355.
[27] Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366.
[28] DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trial 1986;7:177–88.
[29] Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, et al. Features of 20 133 UK patients in hospital with COVID-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. BMJ 2020:369.
[30] Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539–58.
[31] Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994;50:1088–101.
[32] Al-Bari MdAA. Targeting endosomal acidification by chloroquine analogs as a promising strategy for the treatment of emerging viral diseases. Pharmacol Res Perspect 2017;5:e00293.
[33] Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. J R Stat Soc Ser A Stat Soc 2009;172:137–59.
[34] Viechtbauer W, Cheung MW-L. Outlier and influence diagnostics for meta-analysis. Res Synth Method 2010;1:112–25.
[35] McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): an R package and Shiny web app for visualizing risk-of-bias assessments. Res Synth Methods 2020. https://doi.org/10.1002/jrsm.1411.
[36] Wang A-L, Zhong X, Hurd Y. Comorbidity and sociodemographic determinants in COVID-19 mortality in a US urban healthcare system. Infectious Diseases (except HIV/AIDS) 2020. https://doi.org/10.1101/2020.06.11.20128926.

[37] Tang W, Cao Z, Han M, Wang Z, Chen J, Sun W, et al. Hydroxychloroquine in patients with mainly mild to moderate coronavirus disease 2019: open label, randomised controlled trial. BMJ 2020;369. https://doi.org/10.1136/bmj.m1849.

[38] Chen Jun LD, Chen Jun LD. A pilot study of hydroxychloroquine in treatment of patients with moderate COVID-19. J Zhejiang Univ (Med Sci 2020;49:215–9.

[39] Skipper CP, Pastick KA, Engen NW, Bangdiwala AS, Abassi M, Lofgren SM, et al. Hydroxychloroquine in nonhospitalized adults with early COVID-19. Ann Intern Med 2020. https://doi.org/10.7326/M20-4207.

[40] Horby P, Mafham M, Linsell L, Bell JL, Staplin N, Emberson JR, et al. Effect of hydroxychloroquine in hospitalized patients with COVID-19: preliminary results from a multi-centre, randomized, controlled trial. MedRxiv 2020. https://doi.org/10.1101/2020.07.15.20151852.

[41] Magagnoli J, Narendran S, Pereira F, Cummings TH, Hardin JW, Sutton SS, et al. Outcomes of hydroxychloroquine usage in United States veterans hospitalized with COVID-19. Med 2020. https://doi.org/10.1016/j.medj.2020.06.001.

[42] Ip A, Berry DA, Hansen E, Goy AH, Pecora AL, Sinclaire BA, et al. Hydroxychloroquine and tocilizumab therapy in COVID-19 patients—an observational study. Infectious Diseases (except HIV/AIDS) 2020. https://doi.org/10.1101/2020.05.21.20109207.

[43] Alberici F, Delbarba E, Manenti C, Econimo L, Valerio F, Pola A, et al. A report from the Brescia Renal COVID Task Force on the clinical characteristics and short-term outcome of hemodialysis patients with SARS-CoV-2 infection. Kidney Int 2020. https://doi.org/10.1016/j.kint.2020.04.030.

[44] Sánchez-Álvarez JE, Fontán MP, Martín CJ, Pelícano MB, Reina CJC, Prieto ÁMS, et al. Status of SARS-CoV-2 infection in patients on renal replacement therapy. In: Report of the COVID-19 registry of the Spanish society of nephrology (SEN). Nefrología. English Edition; 2020. https://doi.org/10.1016/j.nefroe.2020.04.002.

[45] Wilkinson E. RECOVERY trial: the UK COVID-19 study resetting expectations for clinical trials. BMJ 2020;369. https://doi.org/10.1136/bmj.m1626.

[46] Paccoud O, Tubach F, Baptiste A, Bleibtreu A, Hajage D, Monsel G, et al. Compassionate use of hydroxychloroquine in clinical practice for patients with mild to severe COVID-19 in a French university hospital. Clin Infect Dis n.d. https://doi.org/10.1093/cid/ciaa791.

[47] Rivera DR, Peters S, Panagiotou OA, Shah DP, Kuderer NM, Hsu C-Y, et al. Utilization of COVID-19 treatments and clinical outcomes among patients with cancer: a COVID-19 and Cancer Consortium (CCC19) cohort study. Cancer Discov 2020. https://doi.org/10.1158/2159-8290.CD-20-0941.

[48] Sbidian E, Josse J, Lemaitre G, Mayer I, Bernaux M, Gramfort A, et al. Hydroxychloroquine with or without azithromycin and in-hospital mortality or discharge in patients hospitalized for COVID-19 infection: a cohort study of 4,642 in-patients in France. MedRxiv 2020. https://doi.org/10.1101/2020.06.16.20132597.

[49] Luo J, Rizvi H, Preeshagul IR, Egger JV, Hoyos D, Bandlamudi C, et al. COVID-19 in patients with lung cancer. Ann Oncol 2020. https://doi.org/10.1016/j.annonc.2020.06.007.

[50] Arshad S, Kilgore P, Chaudhry ZS, Jacobsen G, Wang DD, Huitsing K, et al. Treatment with hydroxychloroquine, azithromycin, and combination in patients hospitalized with COVID-19. Int J Infect Dis 2020;97:396–403.

[51] Cavalcanti AB, Zampieri FG, Rosa RG, Azevedo LCP, Veiga VC, Avezum A, et al. Hydroxychloroquine with or without azithromycin in mild-to-moderate COVID-19. N Engl J Med 2020. https://doi.org/10.1056/NEJMoa2019014.

[52] Cravedi P, Suraj SM, Azzi Y, Haverly M, Farouk S, Pérez-Sáez MJ, et al. COVID-19 and kidney transplantation: results from the TANGO International Transplant Consortium. Am J Transplant n.D;n/a. https://doi.org/10.1111/ajt.16185.

[53] Gupta S, Hayek SS, Wang W, Chan L, Mathews KS, Melamed ML, et al. Factors associated with death in critically ill patients with coronavirus disease 2019 in the USA. JAMA Intern Med 2020. https://doi.org/10.1001/jamainternmed.2020.3596.

[54] Lecronier M, Beurton A, Burrel S, Haudebourg L, Deleris R, Le Marec J, et al. Comparison of hydroxychloroquine, lopinavir/ritonavir, and standard of care in critically ill patients with SARS-CoV-2 pneumonia: an opportunistic retrospective analysis. Crit Care 2020;24:418.

[55] Fontana F, Giaroni F, Frisina M, Alfano G, Mori G, Lucchi L, et al. SARS-CoV-2 infection in dialysis patients in northern Italy: a single-centre experience. Clin Kidney J 2020;13:334–9.

[56] Mikami T, Miyashita H, Yamada T, Harrington M, Steinberg D, Dunn A, et al. Risk factors for mortality in patients with COVID-19 in New York City. J Gen Intern Med 2020:1–10. https://doi.org/10.1007/s11606-020-05983-z.

[57] Rogado J, Obispo B, Pangua C, Serrano-Montero G, Martín Marino A, Pérez-Pérez M, et al. Covid-19 transmission, outcome and associated risk factors in cancer patients at the first month of the pandemic in a Spanish hospital in Madrid. Clin Transl Oncol 2020:1–5. https://doi.org/10.1007/s12094-020-02381-z.

[58] Singh S, Khan A, Chowdhry M, Chatterjee A. Outcomes of hydroxychloroquine treatment among hospitalized COVID-19 patients in the United States—real-world evidence from a federated electronic medical record network. Infect Dis (except HIV/AIDS) 2020. https://doi.org/10.1101/2020.05.12.20099028.

[59] Lagier J-C, Million M, Gautret P, Colson P, Cortaredona S, Giraud-Gatineau A, et al. Outcomes of 3,737 COVID-19 patients treated with hydroxychloroquine/azithromycin and other regimens in Marseille, France: a retrospective analysis. Travel Med Infect Dis 2020:101791. https://doi.org/10.1016/j.tmaid.2020.101791.

[60] Bousquet G, Falgarone G, Deutsch D, Derolez S, Lopez-Sublet M, Goudot F-X, et al. ADL-dependency, D-Dimers, LDH and absence of anticoagulation are independently associated with one-month mortality in older inpatients with COVID-19. Aging (Albany NY) 2020;12:11306–13.

[61] Borba MGS, Val FFA, Sampaio VS, Alexandre MAA, Melo GC, Brito M, et al. Effect of high vs low doses of chloroquine diphosphate as adjunctive therapy for patients hospitalized with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection: a randomized clinical trial. JAMA Netw Open 2020;3:e208857.

[62] Saleh M, Gabriels J, Chang D, Kim BS, Mansoor A, Mahmood E, et al. The effect of chloroquine, hydroxychloroquine and azithromycin on the corrected QT interval in patients with SARS-CoV-2 infection. Circ Arrhythm Electrophysiol 2020. https://doi.org/10.1161/CIRCEP.120.008662.

[63] Huang M, Li M, Xiao F, Pang P, Liang J, Tang T, et al. Preliminary evidence from a multicenter prospective observational study of the safety and efficacy of chloroquine for the treatment of COVID-19. Natl Sci Rev n.d. https://doi.org/10.1093/nsr/nwaa113.

[64] Khamis F, Al-Zakwani I, Al Naamani H, Al Lawati S, Pandak N, Omar MB, et al. Clinical characteristics and outcomes of the first 63 adult patients hospitalized with COVID-19: an experience from Oman. J Infect Public Health 2020. https://doi.org/10.1016/j.jiph.2020.06.002.

[65] Fontana F, Alfano G, Mori G, Amurri A, Tei L, Ballestri M, et al. COVID-19 pneumonia in a kidney transplant recipient successfully treated with tocilizumab and hydroxychloroquine. Am J Transplant 2020;20:1902–6.

[66] Mallat J, Hamed F, Balkis M, Mohamed MA, Mooty M, Malik A, et al. Hydroxychloroquine is associated with slower viral clearance in clinical COVID-19 patients with mild to moderate disease: a retrospective study. MedRxiv 2020:2020. https://doi.org/10.1101/2020.04.27.20082180.

[67] Hernandez AV, Roman YM, Pasupuleti V, Barboza JJ, White CM. Hydroxychloroquine or chloroquine for treatment or prophylaxis of COVID-19: a living systematic review. Ann Intern Med 2020. https://doi.org/10.7326/M20-2496.

[68] Lane JCE, Weaver J, Kostka K, Duarte-Salles T, Abrahao MTF, Alghoul H, et al. Safety of hydroxychloroquine, alone and in combination with azithromycin, in light of rapid wide-spread use for COVID-19: a multinational, network cohort and self-controlled case series study. Rheumatology 2020. https://doi.org/10.1101/2020.04.08.20054551.

[69] Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). J Clin Epidemiol 2011;64:407–15.

[70] Bessière F, Roccia H, Delinière A, Charrière R, Chevalier P, Argaud L, et al. Assessment of QT intervals in a case series of patients with coronavirus disease 2019 (COVID-19) infection treated with hydroxychloroquine alone or in combination with azithromycin in an intensive care unit. JAMA Cardiol 2020. https://doi.org/10.1001/jamacardio.2020.1787.

[71] Chorin E, Wadhwani L, Magnani S, Dai M, Shulman E, Nadeau-Routhier C, et al. QT interval prolongation and Torsade De Pointes in patients with COVID-19 treated with hydroxychloroquine/azithromycin. Heart Rhythm 2020. https://doi.org/10.1016/j.hrthm.2020.05.014.

[72] Mercuro NJ, Yen CF, Shim DJ, Maher TR, McCoy CM, Zimetbaum PJ, et al. Risk of QT Interval prolongation associated with use of hydroxychloroquine with or without concomitant azithromycin among hospitalized patients testing positive for coronavirus disease 2019 (COVID-19). JAMA Cardiol 2020. https://doi.org/10.1001/jamacardio.2020.1834.

[73] Nguyen LS, Dolladille C, Drici M-D, Fenioux C, Alexandre J, Mira J-P, et al. Cardiovascular toxicities associated with hydroxychloroquine and azithromycin: an analysis of the world health organization pharmacovigilance database. Circulation 2020;142:303–5.

[74] Commissioner O of the. Coronavirus (COVID-19) Update: FDA revokes emergency use authorization for chloroquine and hydroxychloroquine. FDA. 2020. Available from: https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-revokes-emergency-use-authorization-chloroquine-and. [Accessed 15 June 2020].

[75] Agence Nationale de Sécurité du médicament et des produits de santé (Ansm). COVID-19 : l'ANSM souhaite suspendre par précaution les essais cliniques évaluant l'hydroxychloroquine dans la prise en charge des patients—point d'Information—ANSM: agence nationale de sécurité du médicament et des produits de santé n.d. Available from: https://ansm.sante.fr/S-informer/Points-d-information-Points-d-information/COVID-19-l-ANSM-souhaite-suspendre-par-precaution-les-essais-cliniques-evaluant-l-hydroxychloroquine-dans-la-prise-en-charge-des-patients-Point-d-Information (accessed 15 June 2020).

[76] Dimitrova EK. COVID-19: reminder of risk serious side effects with chloroquine and hydroxychloroquine. European Medicines Agency; 2020. Available from: https://www.ema.europa.eu/en/news/covid-19-reminder-risk-serious-side-effects-chloroquine-hydroxychloroquine. [Accessed 15 June 2020].

[77] Rathi S, Ish P, Kalantri A, Kalantri S. Hydroxychloroquine prophylaxis for COVID-19 contacts in India. Lancet Infect Dis 2020;0. https://doi.org/10.1016/S1473-3099(20)30313-3.

[78] Alexander PE, Debono VB, Mammen MJ, Iorio A, Aryal K, Deng D, et al. COVID-19 coronavirus research has overall low methodological quality thus far: case in point for chloroquine/hydroxychloroquine. J Clin Epidemiol 2020. https://doi.org/10.1016/j.jclinepi.2020.04.016.

[79] NIH halts clinical trial of hydroxychloroquine. National Institutes of Health (NIH); 2020. https://www.nih.gov/news-events/news-releases/nih-halts-clinical-trial-hydroxychloroquine Available from. [Accessed 22 June 2020].

[80] WHO (World Health Organization). "Solidarity" clinical trial for COVID-19 treatments n.d. Available from: https://www.who.int/emergencies/diseases/

novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov/solidarity-clinical-trial-for-covid-19-treatments (accessed 8 June 2020).
[81] McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? Lancet 2000;356:1228–31.
[82] Hopewell S, McDonald S, Clarke MJ, Egger M. Grey literature in meta-analyses of randomized trials of health care interventions. Cochrane Database Syst Rev 2007. https://doi.org/10.1002/14651858.MR000010.pub3.
[83] Us NIH. Hydroxychloroquine plus azithromycin | coronavirus disease COVID-19. COVID-19 treatment guidelines. 2020. Available from: https://www.covid19treatmentguidelines.nih.gov/antiviral-therapy/hydroxychloroquine-plus-azithromycin/. [Accessed 10 August 2020].
[84] INSERM. Discovery stopping inclusions in two treatment groups. 2020. Available from: https://presse.inserm.fr/en/discovery-stopping-inclusions-in-two-treatment-groups/40087/. [Accessed 10 August 2020].

CLM-20-18519.R1

Supplementary tables and figures

Figure S1: Assessment of quality of studies using ROBIN-I for randomized studies

D1: Bias arising from the randomization process. D2: Bias due to deviations from intended intervention. D3: Bias due to missing outcome data. D4: Bias in measurement of the outcome. D5: Bias in selection of the reported result.

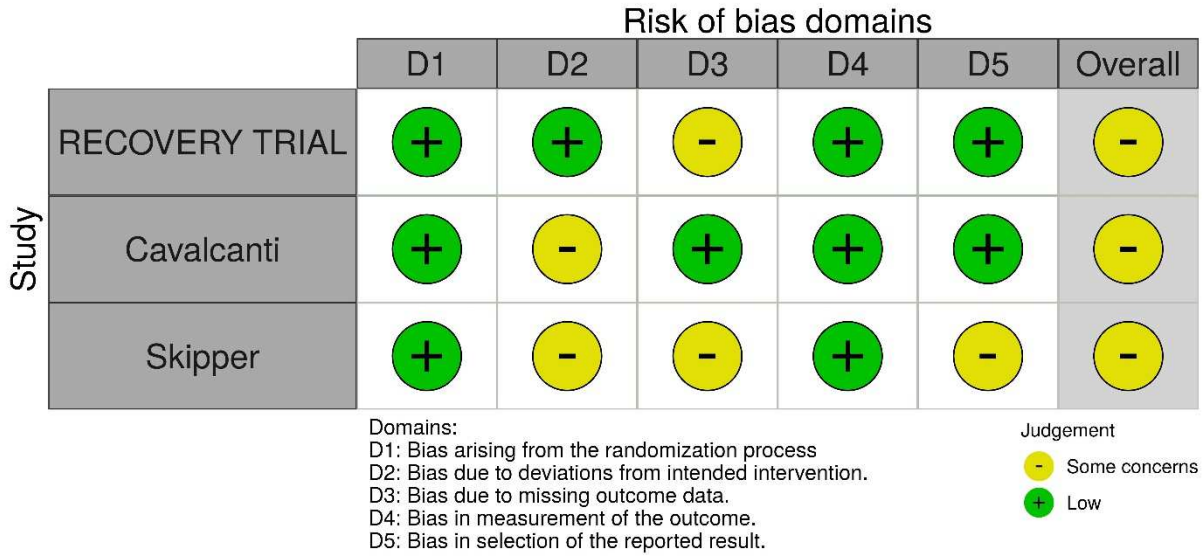


Figure S2: Assessment of quality of studies using ROBIN-I for non-randomized studies

D1: Bias due to confounding. D2: Bias due to selection of participants. D3: Bias in classification of interventions. D4: Bias due to deviations from intended interventions. D5: Bias due to missing data. D6: Bias in measurement of outcomes. D7: Bias in selection of the reported result.

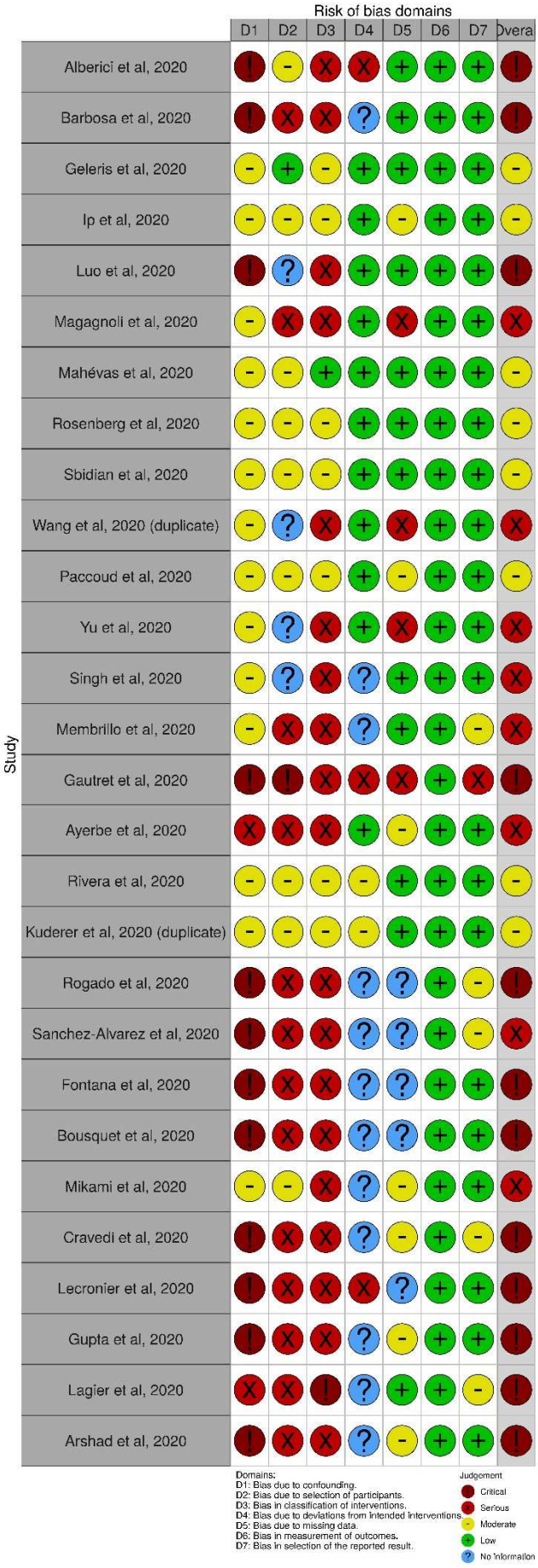
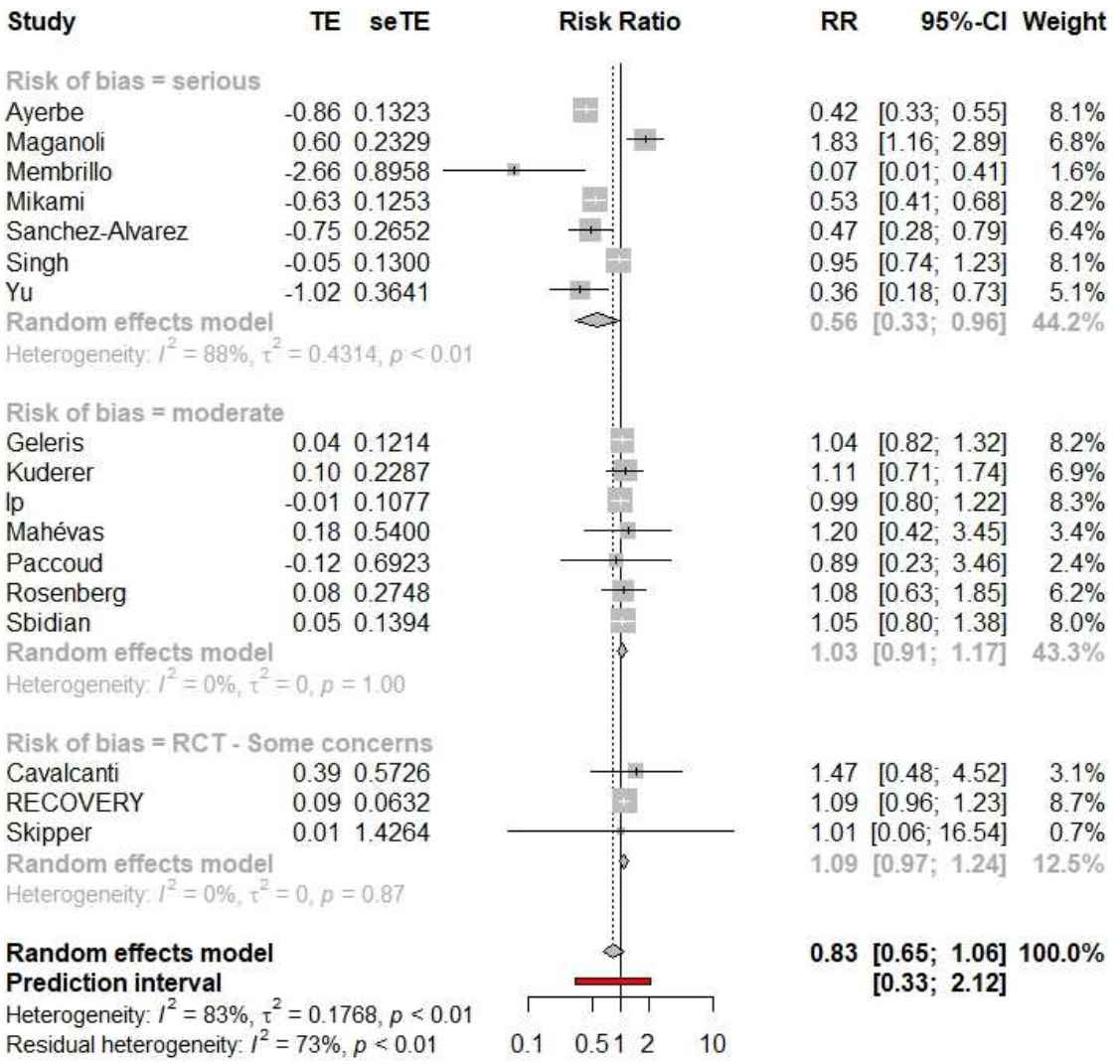
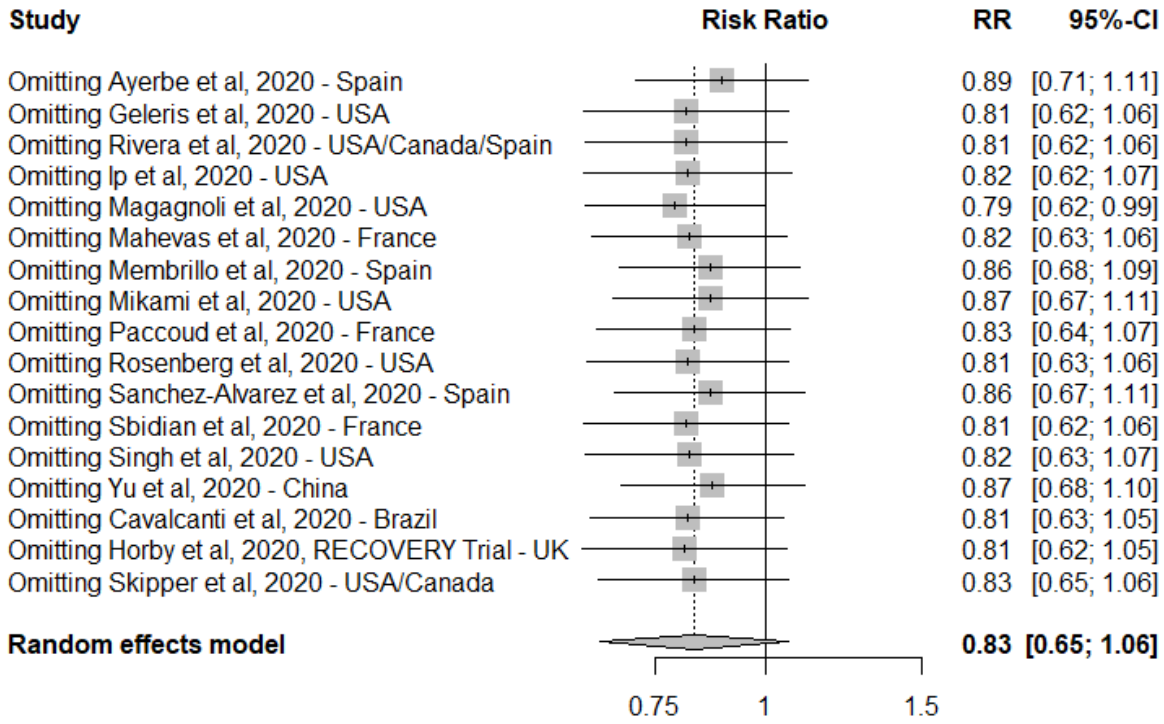

Risk of bias domains
D1
D2
D3
D4
D5
D6
D7
Overa
Study
Alberici et al, 2020
Barbosa et al, 2020
Geleris et al, 2020
Ip et al, 2020
Luo et al, 2020
Magagnoli et al, 2020
Mahévas et al, 2020
Rosenberg et al, 2020
Sbidian et al, 2020
Wang et al, 2020 (duplicate)
Paccoud et al, 2020
Yu et al, 2020
Singh et al, 2020
Membrillo et al, 2020
Gautret et al, 2020
Ayerbe et al, 2020
Rivera et al, 2020
Kuderer et al, 2020 (duplicate)
Rogado et al, 2020
Sanchez-Alvarez et al, 2020
Fontana et al, 2020
Bousquet et al, 2020
Mikami et al, 2020
Cravedi et al, 2020
Lecronier et al, 2020
Gupta et al, 2020
Lagier et al, 2020
Arshad et al, 2020
Domains:
D1: Bias due to confounding.
D2: Bias due to selection of participants.
D3: Bias in classification of interventions.
D4: Bias due to deviations from intended interventions.
D5: Bias due to missing data.
D6: Bias in measurement of outcomes.
D7: Bias in selection of the reported result.
Judgement
Critical
Serious
Moderate
Low
No Information

Figure S3. Forest plot for hydroxychloroquine alone and COVID-19 mortality risk, subgroup analysis per risk of bias*



*Excluding studies with critical risk of bias

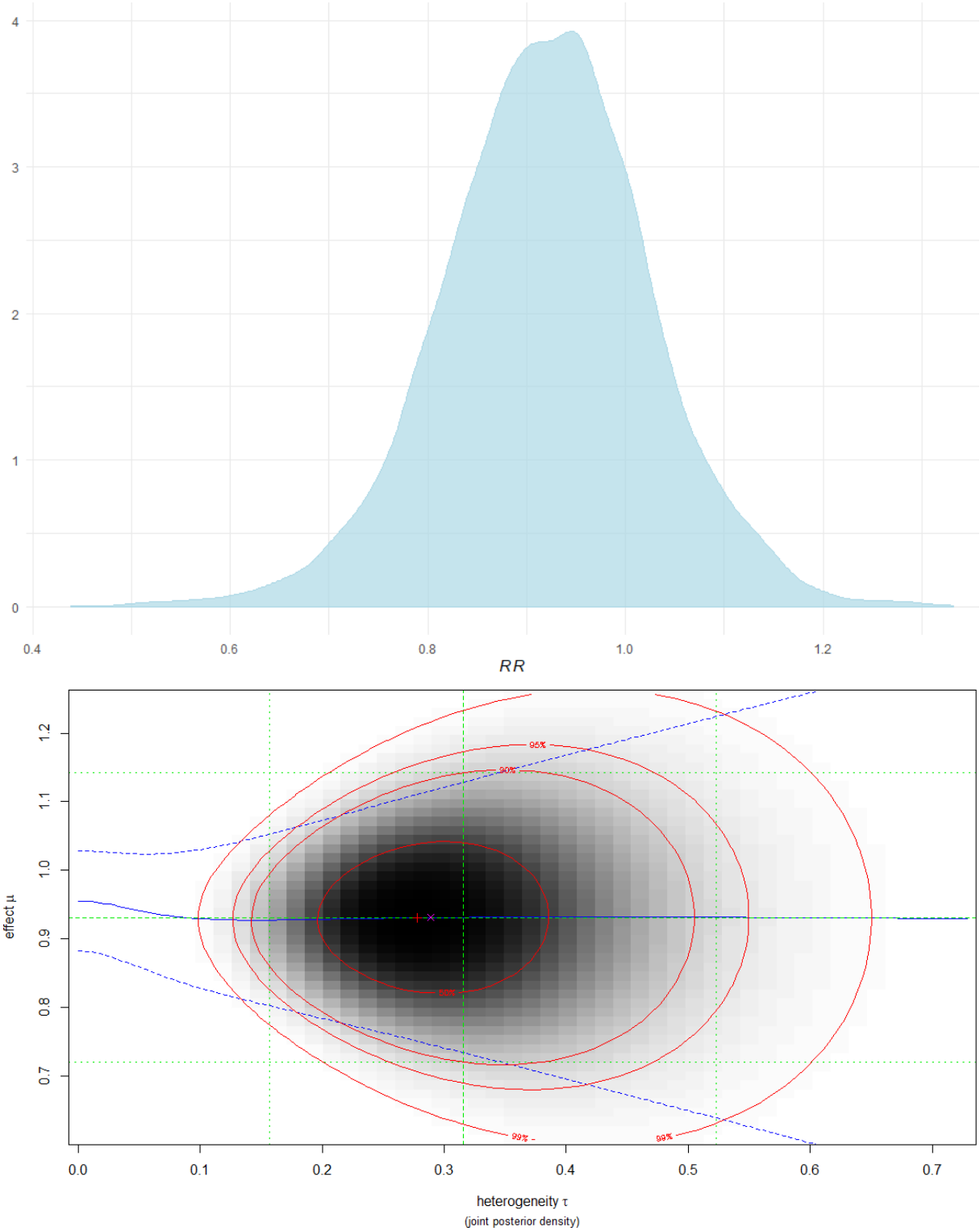
Figure S4. Influence analysis for hydroxychloroquine and COVID-19 mortality*



*Excluding studies with critical risk of bias

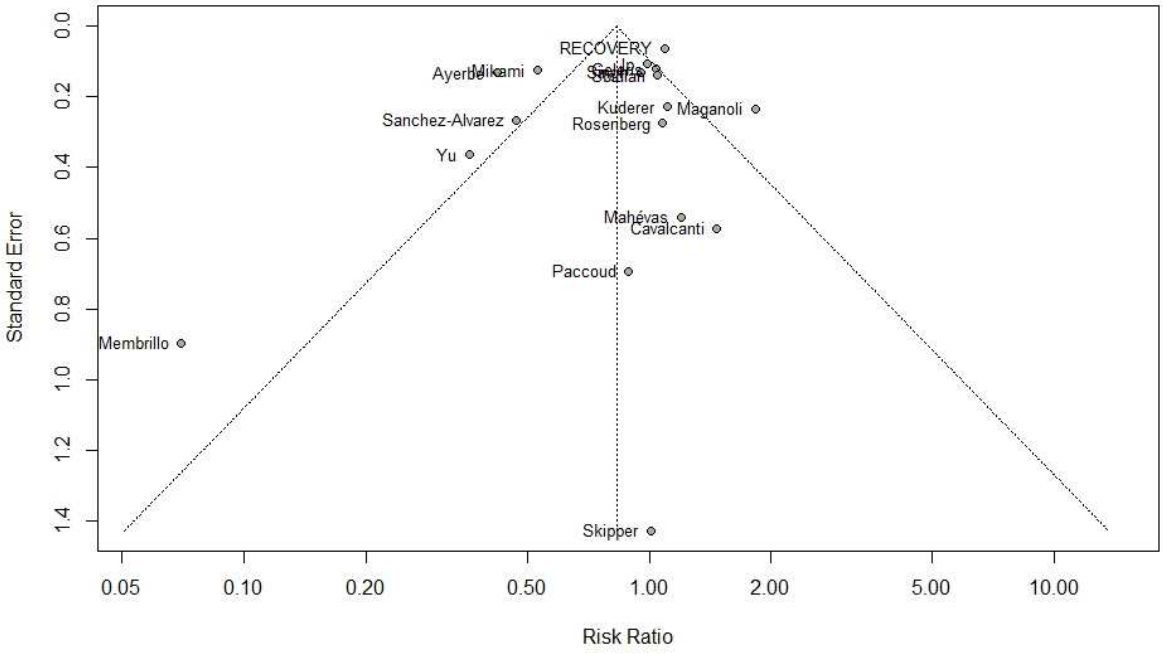
Figure S5. Posterior distribution obtained with 4,000 iterations by the Bayesian multilevel meta-analysis to assess the pooled effect (RR, Relative Risk) for mortality and hydroxychloroquine alone and the joint posterior density of heterogeneity (τ) and effect (μ)*

Darker shading corresponds to higher probability density. The red lines indicate (approximate) 2-dimensional credible regions, and the green lines show marginal posterior medians and 95% credible intervals. The blue lines show the conditional posterior mean effect $\hat{\mu}(\tau)$ as a function of the heterogeneity τ along with a 95% interval based on its conditional standard error $\hat{\sigma}(\tau)$. The red cross (+) indicates the posterior mode, while the pink cross (×) shows the ML estimate



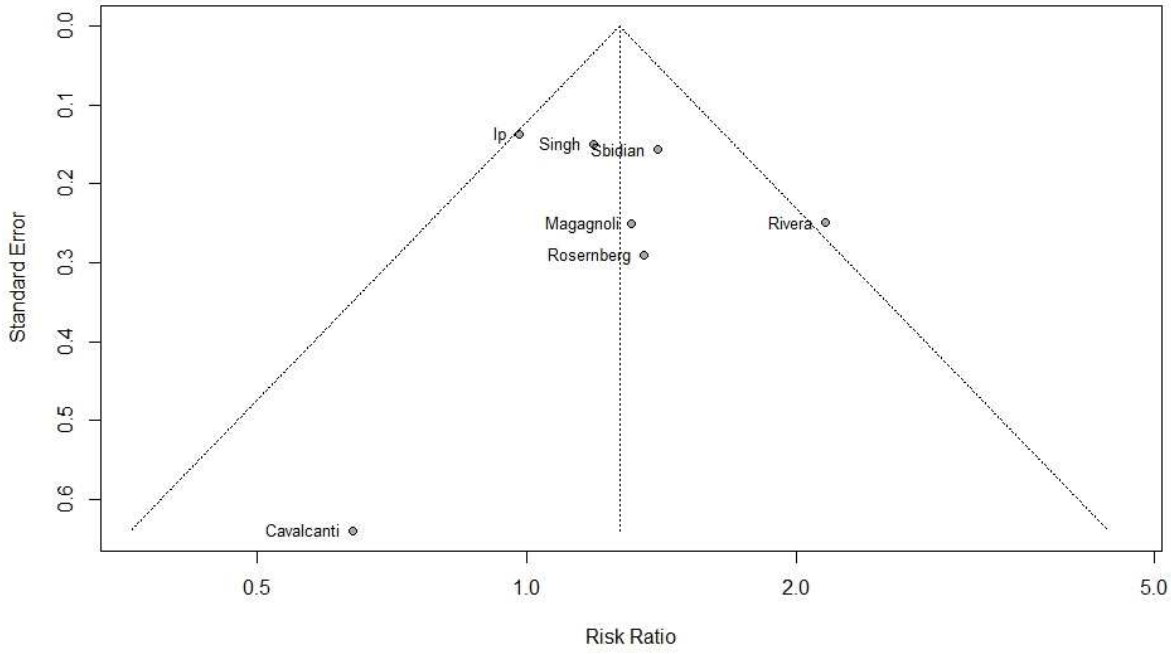
*Excluding studies with critical risk of bias

Figure S6. Funnel plot for hydroxychloroquine alone and COVID-19 mortality risk*



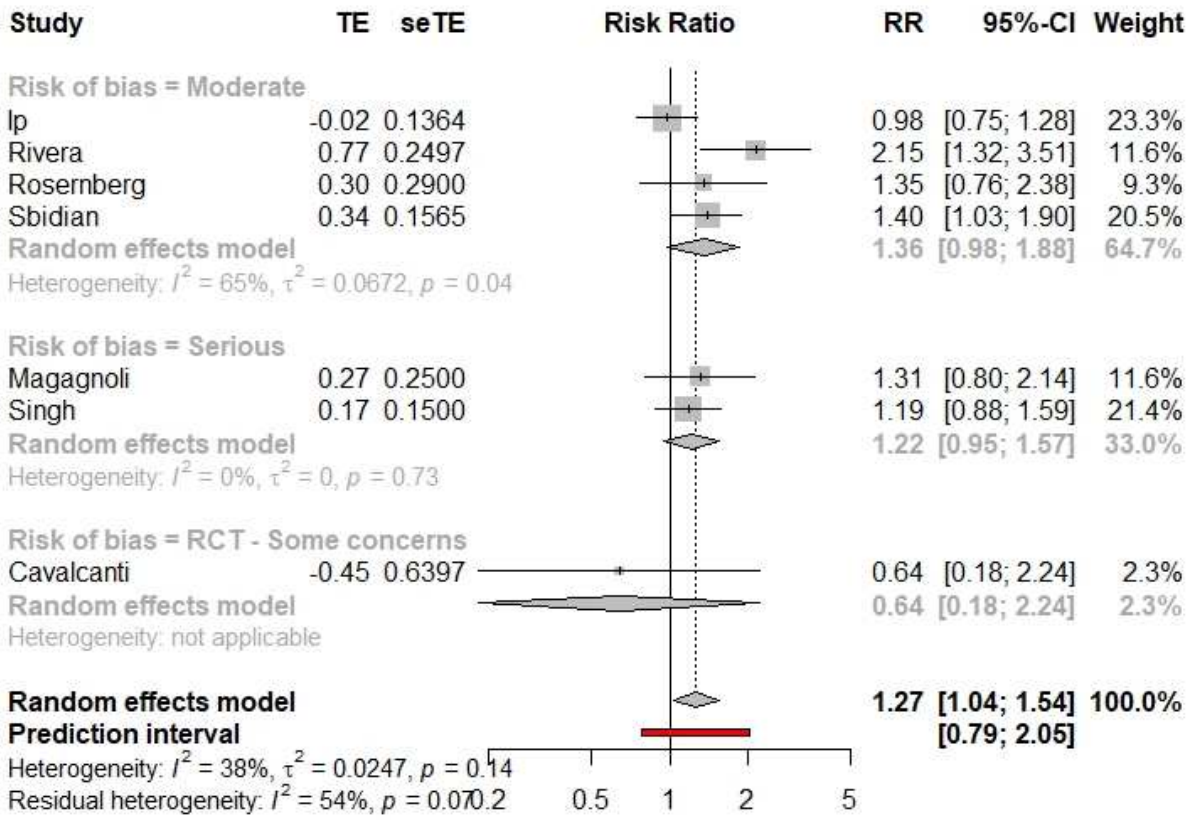
*Excluding studies with critical risk of bias

Figure S7: Funnel plot for hydroxychloroquine with azithromycin and COVID-19 mortality risk (without studies with critical risk of bias)*



*Excluding studies with critical risk of bias

Figure S8. Forest plot for hydroxychloroquine with azithromycin and COVID-19 mortality risk, subgroup analysis per risk of bias*

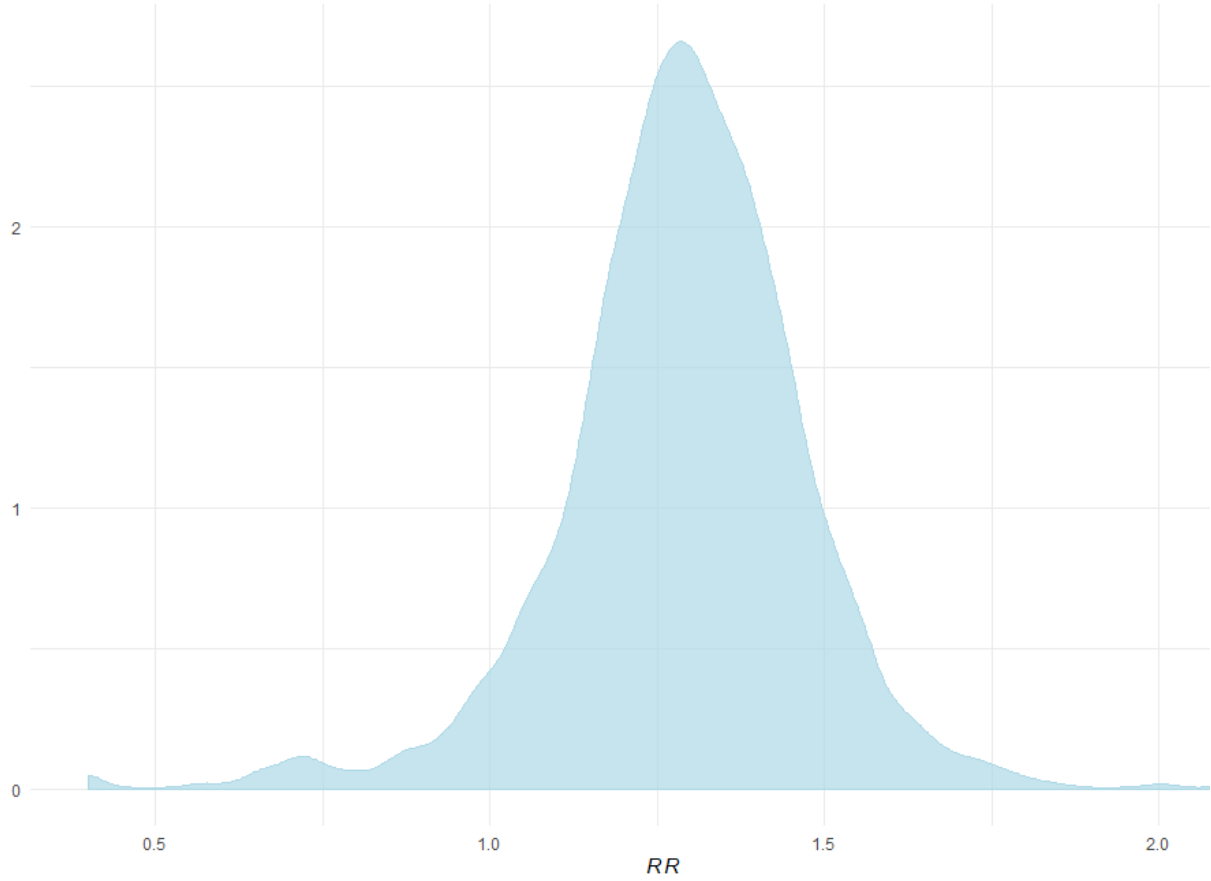


*Excluding studies with critical risk of bias

Figure S9. Posterior distribution obtained with 4,000 iterations by the Bayesian multilevel meta-analysis to assess the pooled effect for mortality and hydroxychloroquine with azithromycin*

Darker shading corresponds to higher probability density. The red lines indicate (approximate) 2-dimensional credible regions, and the green lines show marginal posterior medians and 95% credible intervals. The blue lines show the conditional posterior mean effect $\hat{\mu}(\tau)$ as a function of the heterogeneity τ along with a 95% interval based on its conditional standard error $\hat{\sigma}(\tau)$. The red cross (+) indicates the posterior mode, while the pink cross (×) shows the ML estimate

*Excluding studies with critical risk of bias



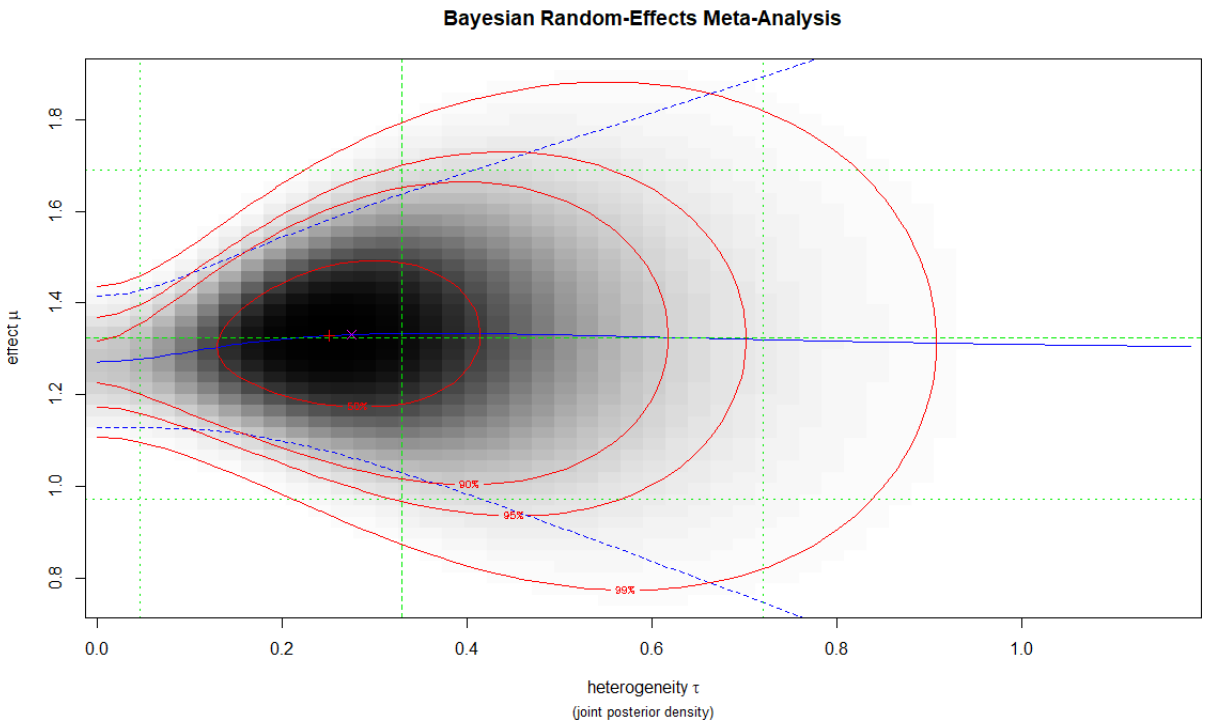
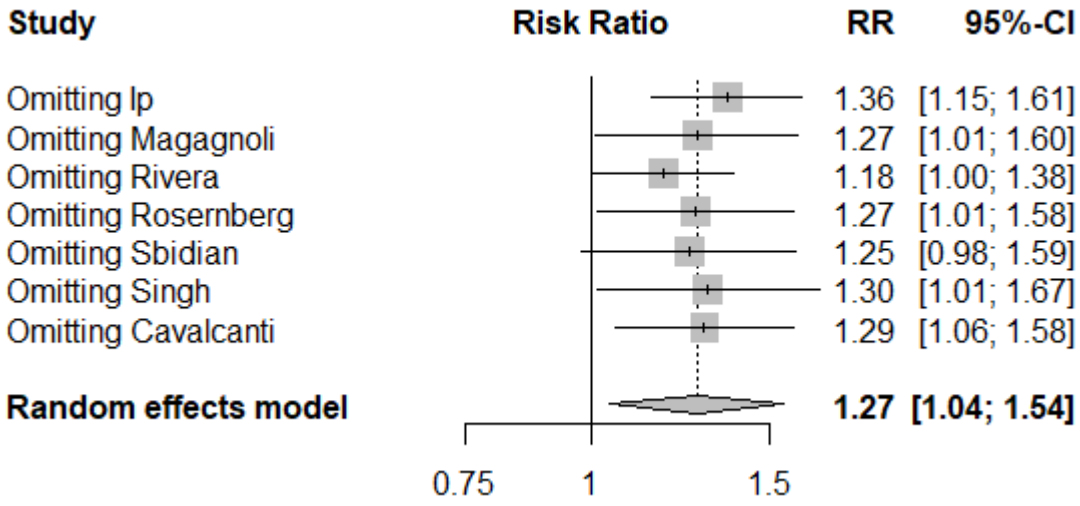

Bayesian Random-Effects Meta-Analysis
effect μ
0.8
1.0
1.2
1.4
1.6
1.8
50%
90%
95%
99%
0.0
0.2
0.4
0.6
0.8
1.0
heterogeneity τ
(joint posterior density)

Figure S10. Influence analysis for hydroxychloroquine with azithromycin and COVID-19 mortality*



*Excluding studies with critical risk of bias

Supplementary Excel file 1: Risk of bias assessment using ROBIN-I for non-randomized studies and RoB2 for RCT

S1. Full electronic search strategy
For hydroxychloroquine
Cochrane Library
Website: https://www.cochranelibrary.com/advanced-search
Cochrane Review matching (Hydroxychloroquine or HCQ) in Title Abstract Keyword AND (mortality or death) in Title Abstract Keyword AND (COVID-19 or SARS-CoV-2) in Title Abstract Keyword - (Word variations have been searched)

PubMed
Website:
https://pmlegacy.ncbi.nlm.nih.gov/pubmed/?term=(hydroxychloroquine+or+HCQ)+AND+(COVID-19+OR+SARS-CoV-2+OR+coronavirus)+AND+(Mortality+OR+death)
((hydroxychloroquine or HCQ) AND (COVID-19 OR SARS-CoV-2 OR coronavirus) AND (Mortality OR death)

Additional search on PubMed with controlled vocabulary	Number of results
((((covid19[MeSH Terms])) OR (sars-CoV-2[MeSH Terms])) AND (hydroxychloroquine[MeSH Terms])) AND ((mortality[MeSH Terms]) OR (death[MeSH Terms]))	0
(sars)[MeSH Terms])	0
(covid19[MeSH Terms]) AND (hydroxychloroquine[MeSH Terms])	1
(COVID-19[MeSH Major Topic]) AND (hydroxychloroquine[MeSH Major Topic])	1
(COVID-19[MeSH Major Topic])	0
(coronavirus[MeSH Major Topic]) AND (hydroxychloroquine[MeSH Major Topic])	71
(coronavirus[MeSH Terms]) AND (hydroxychloroquine[MeSH Terms])	263

Web of Science
Website:
http://apps.webofknowledge.com.proxy.insermbiblio.inist.fr/Search.do?product=UA&SID=F6KgcWI7K6kjXJwhAoH&search_mode=GeneralSearch&prID=9a27b347-ecf8-4832-9206-db1bbd2cc9a8
You searched for: TOPIC: (covid-19 OR SARS-CoV-2) AND TOPIC: (hydroxychloroquine or HCQ) AND TOPIC: (mortality or death)

EMBase
('hydroxychloroquine'/exp OR hydroxychloroquine OR hcq) AND ('covid 19'/exp OR 'covid 19' OR 'sars cov 2'/exp OR 'sars cov 2' OR 'coronavirus'/exp OR coronavirus) AND ('mortality'/exp OR mortality OR 'death'/exp OR death)"

Manual additional searches:
MedRxiv
https://www.medrxiv.org/
Search: Hydroxychloroquine COVID-19 mortality
Google scholar:
https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=hydroxychloroquine+COVID-19&btnG=
Search: Hydroxychloroquine COVID-19 mortality

For chloroquine
Cochrane Library
Cochrane Review matching (chloroquine or CQ) in Title Abstract Keyword AND (mortality or death) in Title Abstract Keyword AND (COVID-19 or SARS-CoV-2) in Title Abstract Keyword - (Word variations have been searched)

PubMed
((chloroquine or CQ) AND (COVID-19 OR SARS-CoV-2 OR coronavirus) AND (Mortality OR death)

Web of Science
You searched for: TOPIC: (covid-19 OR SARS-CoV-2) AND TOPIC: (chloroquine or CQ) AND TOPIC: (mortality or death)

EMBase
('chloroquine'/exp OR chloroquine OR hcq) AND ('covid 19'/exp OR 'covid 19' OR 'sars cov 2'/exp OR 'sars cov 2' OR 'coronavirus'/exp OR coronavirus) AND ('mortality'/exp OR mortality OR 'death'/exp OR death)"

S2. PRISMA Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	3-4
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	6
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	6 PROSPERO (Registration number: CRD42020190801
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6-7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	6 S1. Full electronic search strategy
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6-7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome	8

		level), and how this information is to be used in any data synthesis.	
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7-8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	8-9
Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	8-9 for subgroup analysis
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	9
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	10 Tables S1,S2
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	11-12 Figures S1 S2
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	11-12 Figures 2;-3
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	11-12 Figures 2-3
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Figures S6 S
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Figures S4, S5, S7, S8,S9,S10
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	12-14
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	15
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	15-16

FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	16

Table S1 (continued): Characteristics of studies included in the meta-analysis for COVID-19 mortality

First Author	Type of study	Country	Treatment	Number of participants		Number of deaths		HCQ dosing	Age[a] (years)	Patients	Study quality
				Control	HCQ alone	Control	HCQ alone				
Alberici et al[43], 2020	Observational, cohort	Italy	HCQ alone	22	72	Not reported	Not reported	Not specified	72 (median) IQ=62-79	Hospitalized patients with haemodialysis	Critical
Arshad et al[50], 2020	Observational, cohort	USA	HCQ alone HCQ+AZI	409	1202	108	162	800mg/day on day 1, 400mg for 4 additional days	64 (median) IQ=53-76	Hospitalized patients	Critical
Ayerbe et al[17], 2020	Observational, cohort	Spain	HCQ alone	162	1857	49	237	Not specified	67,57 (mean)	Hospitalized patients	Serious
Barbosa Joshua et al[23], 2020	Observational, cohort	USA	HCQ alone	21	17	1	2	800 mg for 2 days then 200-400mg for 3-4 days	62.7 (mean) SD=15.1	Hospitalized patients (mild/moderate symptoms)	Critical
Bousquet al[60], 2020	Observational, cohort	France	HCQ+AZI	27	81	Not reported	Not reported	Dose not specified for 1-9 days	78.4 (mean) SD=7.8	Hospitalized patients (23% mild, 38% moderate, 29% severe)	Critical
Cavalcanti et al[51], 2020	RCT	Brazil	HCQ alone HCQ+AZI	173	159	6	7	800mg/day for 7 days	50.3 (mean) SD=14.6	Hospitalized patients with mild-to-moderate symptoms	Some concerns
Cravedi et al[52],	Observational, cohort	USA, Italy and Spain	HCQ alone	43	101	10	36	Not specified	60 (mean) SD=12	Hospitalized kidney transplant patients with immunosuppression	Critical
Gupta et al[53],	Observational, cohort	USA	HCQ alone HCQ+AZI	454	1761	153	631	Not specified	60.5 (mean) SD=14.5	Hospitalized patients in ICU	Critical
Fontana et al[65], 2020	Observational, cohort	Italy	HCQ alone	3	12	2	4	125-400mg for a median time of 6.5 days	75.96 (mean) SD=11.09	Dialysis patients	Critical

Gautret et al[24], 2020	Non-randomised controlled trial	France	HCQ alone	16	26	0	1	A maintenance dose of 600 mg/day	45.1 (mean) SD=22	Hospitalized patients (mild symptoms)	Critical
Geleris et al[15], 2020	Observational, cohort	USA	HCQ alone	565	811	75	157	1200mg at day 1 then 400mg for 4 days	From <40 to >80	Hospitalized patients (moderate/severe symptoms)	Moderate
Ip et al[42], 2020	Observational, cohort	USA	HCQ alone HCQ+AZI	598	1914	115	432	800mg at day 1 then 400mg on day 2-5 (80%)	64 (median) IQ=52-76	Hospitalized patients (44% moderate/severe symptoms)	Moderate
Lagier et al[59], 2020	Observational, cohort	France	HCQ alone HCQ+AZI	162	101	4	2	600mg/day for 10 days	45.3 (mean) SD=16.8	Hospitalized patients (mild symptoms)	Critical
Lecronier et al[54], 2020	Observational, cohort	France	HCQ alone	22	38	9	9	400mg/day Duration is not specified	57 (median) IQ=53-68	Hospitalized patients in ICU (severe cases)	Critical
Luo et al[49], 2020	Observational, cohort	USA	HCQ alone	102		25		NA	68 (median) Min-max 31-91 years	Hospitalized patients with a lung cancer	Critical
Kuderer et al[12], 2020	Observational, cohort (CCC19 cohort study)	USA Canada Spain	HCQ alone HCQ+AZ	486	89	41	11	Not specified	66 (median) IQ=57-76	Hospitalized patients with who have a current or past diagnosis of cancer	Moderate
Rivera et al[47], 2020	(CCC19 cohort study, updated)	USA, Canada, Spain	HCQ alone HCQ+any other drug	1321	179	115	44	Not specified	67 (median) IQ=57-77	Hospitalized patients with who have a current or past diagnosis of cancer	Moderate
Magagnoli et al[41], 2020	Observational, cohort	USA	HCQ alone HCQ+AZI	395	198	37	38	Median HCQ dose: 400mg/day Median HCQ+AZ dose: 422.2 mg/day	70 (median) IQ=60-75	Hospitalized patients	Serious

Mahevas et al[13], 2020	Observational, cohort	France	HCQ alone	89	84	8	9	600mg/day	60 (median) IQ=52-68	Hospitalized patients with covid-19 pneumonia who require oxygen:	Moderate
Membrillo et al[16], 2020	Observational, cohort	Spain	HCQ alone	43	123	21	27	A loading dose of 800 mg + 400 mg in following days (ten days for moderate cases)	HCQ: 61.5 No HCQ: 68.7 (mean)	Hospitalized patients	Serious
Mikami et al[56], 2020	Observational, cohort	USA	HCQ alone	895	2813	231	575	Not specified	66 (median) IQR=55-78	Hospitalized patients	Serious
Paccoud et al[46], 2020	Observational, cohort	France	HCQ alone	46	38	6	3	Compassionate basis: 600mg/day for 10 days	65.5 SD=16	Hospitalized patients	Moderate
Horby, et al[40], 2020 RECOVERY TRIAL	Randomized controlled trial	UK	HCQ alone	3132	1542	736	396	A loading dose of 2400mg at day 1, then 800mg/day for 10 days	Not specified	Hospitalized patients	Some concerns
Rogado et al[57], 2020	Observational, cohort	Spain	HCQ+AZI	8	18	Not reported	Not reported	Not specified	71 (median) Range:34-90	Hospitalized patients (64% severe cases)	Critical
Rosenberg et al[18], 2020	Observational, cohort	USA	HCQ alone HCQ+AZI	221	271	28	54	400mg then 200-400mg at 2nd prescription then 200-400mg at 3rd	63 (median)	Hospitalized patients	Moderate
Sanchez-Alvarez et al[44], 2020	Observational, cohort	Spain	HCQ alone	53	322	32	166	Not specified	71 SD=15	(85%) required hospital admission, 8% in intensive care units, with haemodialysis	Serious
Sbidian et al[48], 2020	Observational, cohort	France	HCQ alone HCQ+AZI	623	3792	830	111	A loading dose of 600mg on day 1 then 400mg/day for 9 days	66.1 (mean) SD=18	Hospitalized patients (moderate/severe cases)	Moderate
Singh et al[58], 2020	Observational, cohort	USA	HCQ alone HCQ+AZI	910	910	104	109	Not specified	62 SD=17	Hospitalized patients	Serious

Skipper et al[39], 2020	RCT	USA and Canada	HCQ alone	247	244	1	1	A loading dose of 800mg then 600 mg in 6-8h, then 600mg/day for 4 days	40 (median) IQ=32-50	Non-hospitalized adults	Some concerns
Wang et al[36], 2020	Observational, cohort	USA	HCQ alone HCQ+AZI	4165	591	168	101	Not specified	NA	Hospitalized patients	Serious
Yu et al[14], 2020	Observational, cohort	China	HCQ alone	502	48	238	9	400mg during 7-10 days	68 (median) IQ: 59-77	Critically ill patients	Serious

IQ=Interquartile range, SD=Standard Deviation, HCQ=Hydroxychloroquine, AZI=Azithromycin, NA=Not available

[a]Some studies did not report mean or median age

*Excluding studies with critical risk of bias

Table S2: Reported and calculated effect sizes of studies included in the meta-analysis for COVID-19 mortality*

First Author	Effect size reported in each study[b]	Adjustments
Alberici et al[43], 2020	OR=0,44 [0,16-1,24]	Not adjusted
Arshad et al[50], 2020	HR=0.34 [0.254-0.455] HR=0.294 [0.218-0.396]	Adjusted on demographics (e.g., age, gender), preexisting medical conditions (e.g. CVD, lung disease) and clinical disease severity (mSOFA, O2 saturation
Ayerbe et al[17], 2020	OR=0,422 [0,325-0,546]	Adjusted for age, gender, temperature > 37C and Oxygen Saturation < 90% (complementary information from ResearchGate)
Barbosa Joshua et al[23], 2020	$RR_{calculated}$=2,47 [0,24-24,98]	Not adjusted
Bousquet et al[60], 2020	HR=0.49 [0.19-1.29]	Not adjusted
Cravedi et al[52], 2020	$RR_{calculated}$ =1.53 [0.84-2.80]	Not adjusted. Not included in the reported logistic regression
Cavalcanti et al[51], 2020	HR=1.47 [0.48-4.53] HR=0.64 [0.18-2.21]	Cox model adjusted for age and the use of supplemental oxygen at admission.
Fontana et al[65], 2020	$RR_{calculated}$=0.50 [0.16-1.55]	Not adjusted

Gautret et al[24], 2020	$RR_{calculated}$ =3,41 [0,1505,77,45]	Not adjusted
Geleris et al[15], 2020	HR=1,04 [0,82-1,32]	Cox model adjusted on demographic factors, clinical factors, laboratory tests, and medications. Inverse probability weighting from a propensity-score based on these same covariates.
Gupta et al[53], 2020	$RR_{calculated}$ = 1.06 [0.92-1.23] $RR_{calculated}$ = 0.96 [0.86-1.08]	Not adjusted. Not included in the reported Cox models
Horby et al[40], 2020 RECOVERY TRIAL	HR=1.09 [0.96-1.23]	Log-rank 'observed minus expected' statistic. There was nearly no imbalance of baseline characteristics between HCQ and usual care groups
Ip et al[42], 2020	HR=0.99 [0.8-1.22] HR=0.98 [0.75-1.28]	Cox model adjusted on the propensity-score variable: gender, coronary disease, stroke, heart failure, arrhythmia, African American, COPD, , renal failure, rheumatologic disorder, inflammatory bowel disease, advanced liver disease, age, diabetes mellitus, insulin use prior to hospitalization, asthma, HIV/hepatitis, any cancer, and log ferritin
Kuderer et al[12], 2020[c]	OR=1,06 [0,51,2,2] OR=2.93 [1.79-4.79]	Adjusted for age, sex, smoking status, and obesity
Rivera et al[47], 2020	OR=1.11 [0.71-1.74] OR=2.15 [1.15-3.06]	Multivariable logistic regression, adjusted on age, sex, self-reported race and ethnicity, region of patient residence, smoking status, obesity, hypertension, diabetes, cardiovascular, pulmonary and renal comorbidities, ECOG PS, cancer status, baseline severity of COVID-19. A Propensity score matching regression model was conducted to assess the relationship between the treatment e xposure and 30-day all-cause mortality
Lagier et al[59], 2020	$RR_{calculated}$ = 0.8 [0.15-4.3] $RR_{calculated}$ = 0.29 [0.1-0.83]	Not adjusted. The article reported an effect size a comparison between HCQ+AZ>3 days vs "other drugs". Other drugs included HQ alone, HQ+AZ<3 days, neither drug.
Lecronier et al[54], 2020	$RR_{calculated}$= 0.58 [0.27-1.24]	Not adjusted
Luo et al, [49] 2020	OR=1.03 [0.26-3.55]	Not adjusted (univariate logistic regression)
Magagnoli et al[41], 2020	HR=1.83 [1.16-2.89] HR=1.31 [0.80-2.15]	Propensity score adjustment. All baseline covariates were included in the propensity score models (age, race, BMI, SpO2, breaths per minute, heart rate, T°, systolic blood pressure, ALT, AST, serum albumin, Total bilirubin, Creatinine, Erythrocytes, Haematocrit, Leukocytes, Lymphocytes, Platelets, Blood urea nitrogen, C-reactive protein
Mahevas et al[13], 2020	HR=1,2 [0,4,3,3]	Inverse probability of treatment weighting in Cox model. age, sex, comorbidities (presence of chronic respiratory insufficiency during oxygen treatment, or asthma, cystic fibrosis, or any chronic respiratory disease likely to result in decompensation during a viral infection; heart failure

		(New York Heart Association class III or IV); chronic kidney disease; liver cirrhosis with Child-Pugh class B or more; personal history of cardiovascular disease (hypertension, stroke, coronary artery disease, or cardiac surgery); insulin dependent diabetes mellitus, or diabetic microangiopathy or macroangiopathy; treatment with immunosuppressive drugs, including anticancer chemotherapy; uncontrolled HIV infection or HIV infection with CD4 cell counts <200/µL; or a haematological malignancy); body mass index (≥30 or not); third trimester of pregnancy; treatment by angiotensin converting enzyme inhibitors or angiotensin receptor blockers13; time since symptom onset; and severity of condition at admission (percentage of lung affected: ≥50% or not; presence of confusion; respiratory frequency; oxygen saturation without oxygen; oxygen flow; systolic blood pressure; and C reactive protein level).
Membrillo et al[16], 2020	OR=0,07 [0,012,0,402]	Adjusted on variables with p<0,25in univariate analysis
Mikami et al[56], 2020	HR=0.53 [0.41-0.67]	Cox model adjusted for age, sex, race, cigarette use history, past medical history of asthma, hypertension, diabetes, or cancer, systolic BP, RR, SpO2, BMI, initial laboratory values (lymphocyte proportion, D-dimer, IL-6), and hydroxychloroquine use. IPTW-adjusted Cox with doubly robust methods (inverse probability weighting (IPTW) based on propensity scoring to control for observed differences in baseline characteristics between treatment group and control group)
Paccoud et al[46], 2020	HR=0.89 [0.23-3.47]	Propensity score based on time between symptom onset and admission ≤ 7, Charlson comorbidity index, NEWS2 score at admission, pneumonia severity and medical history of arterial hypertension or obesity.
Rogado et al[57], 2020	OR=0,02 [0,01,0,73]	Adjusted by median age, histology, staging, cancer treatment received and hypertension
Rosenberg et al[18], 2020	HR=1,08 [0,63,1,85] HR=1.35 [0.76-2.4]	Multiple adjustments on potential confounders (age>65, sex, hospital, comorbidities, respiratory capacities
Sanchez-Alvarez et al[44], 2020	OR=0,471 [0,28,0,792]	No information about adjustments in logistic regression
Sbidian et al[48], 2020	HR=1.05 [0.77-1.33] HR=1.40 [0.98-1.81]	Baseline covariables considered for adjustment were sex, age, current smoker, diabetes, obesity, hypertension, dyslipidemia, Ischaemic heart disease, rhythmic heart diseases, Chronic renal failure & Chronic end-stage kidney failure, any chronic lung disease, hepatic failure, cancer, hemopathies, chemotherapy, current steroid use, oxygen saturation, partial pressure of oxygen,

		paCO2, lymphocytes, neutrophils,D-Dimer, Creatine, C Reactive protein, Dehydrogenase lactate, prothrombin time
Singh et al[58], 2020	HR=0,95 [0,74,1,23] HR=1.19 [0.89-1.60]	Creation of groups based on propensity score matching for age, gender, race, confounding comorbidities
Skipper et al[39], 2020	$RR_{calculated}$ = 1.01 [0.06-16.09]	Not adjusted. HR for mortality were not reported
Wang et al[36], 2020	OR=0.96 [0.69-1.34] OR=0.94 [0.73-1.21]	Adjustment on covariables with p<0.05 in the bivariate analysis with mortality, then application of a backward selection algorithm
Yu et al[14], 2020	HR=0,36 [0,18,0,75]	Adjustment: respiratory rate, shortness of breath, alanine aminotransferase (when p<0,01 in univariate Cox model)

HR and OR are the most adjusted effect size reported in each study. Some studies did not report effect size. $RR_{calculated}$ were calculated using the number of deaths in the treatment and the control groups

*Excluding studies with critical risk of bias

Table S3. Reported concomitant use of corticosteroids in studies included in our meta-analysis

First author	Use of dexamethasone?
Rosenberg et al, 2020	Not used
Geleris et al, 2020	yes, some glucocorticoids (26.6% in HQ group vs 10.1% in control)
Magagnoli et al, 2020	Not specified
Yu et al, 2020	Not specified
lp et al, 2020	Not specified
Luo et al, 2020	Not specified
Sbidian et al, 2020	Steroids have been administrated in 17% of HCQ group, 18.9 of HCQ-AZI group and 10.5% in the control group
Wang et al, 2020	Not used
Paccoud et al, 2020	No patient received glucorticosteroid therapy
Mahevas et al, 2020	Not used
Alberici et al, 2020	Not used
Barbosa et al, 2020	Not specified
Gautret et al, 2020	Not specified
Singh et al, 2020	Not specified
Membrillo et al, 2020	Steroids have been used
Ayerbe et al, 2020	Steroids have been used (46% ; 960/2075)
Kuderer et al, 2020	Not used
Sanchez-Alvarez et al, 2020	Steroids were used
RECOVERY Trial	Use of dexamethasone was similar in both arms (8% vs 9%)
Lagier et al, 2020	Not specified
Arshad et al, 2020	Steroids have been administrated in 79% of HQ group, 74% of HCQ-AZI group and 36% in the control group
Fontana et al. 2020	Two patients (13%) received steroids infusion.
Bousquet t al, 2020	Not used
Mikami et al, 2020	Not used
Cravedi et al, 2020	Steroids were used (66%)
Lecronier et al, 2020	Not used
Gupta et al, 2020	Not used

Skipper et al, 2020	Not specified
Cavalcanti et al, 2020	Use of corticosteroids in 19.8% in HCQ-AZI group, 19.5% in HCQ group and 20.2% in control group

Table S4. Subgroup analysis for the associations between HCQ alone or HCQ associated with AZI and mortality risk of patients with COVID-19*

	N	Pooled Relative Risk	Heterogeneity		
			I^2 (%)	P_{within}	$P_{between}$ (test for subgroup differences)
HCQ alone					
All Studies	17	0.83 [0.65-1.06]	83%	<0.01	
Study Design					
Non-randomised	14	0.79 [0.60-1.04]	84%	<0.01	0.03
RCT	3	1.09 [0.97-1.24]	0%	0.487	
Among non-randomised studies	14				
Type of article					
Peer-reviewed	10	0.77 [0.5-1.13]	85%	<0.01	0.22
Unpublished	4	0.98 [0.82-1.13]	67%	0.03	
Adjusted estimate					
Yes	13	0.82 [0.62-1.09]	84%	<0.01	0.06
No	1	0.47 [0.28-0.79]	NA	NA	
Risk of bias					
Moderate	7	1.03 [0.91-1.17]	0%	0.99	0.03
Serious	7	0.56 [0.33-0.96]]	88%	<0.01	
Mean daily dose					
<500 mg/d	5	1.08 [0.79-1.46]	74%	<0.01	0.02
>500 mg/d	3	1.04 [0.83-1.31]	0%	0.9	
Not specified	6	0.60 [0.42-0.86]	84%	<0.01	

Age					
63 years or less	5	0.99 [0.84-1.16]	57%	0.25	0.17
64 years or more	9	0.75 [0.52-1.08]	87%	<0.01	
Cancer or haemodialysis patient based-population					
No	12	0.80 [0.58-1.09]	85%	<0.01	0.85
Yes	2	0.73 [0.32-1.69]	83%	0.01	
HCQ+AZI					
All Studies	7	1.27 [1.04-1.54]	38%	0.14	
Study Design					
non-randomised	6	1.29 [1.06-1.58]	42%	<0.01	0.28
RCT	1	0.64 [0.18-2.24]	NA	NA	
Among non-randomised studies					
Type of article	6				
Peer-reviewed	3	1.58 [1.14-2.20]	69%	0.04	0.12
Unpublished	3	1.16 [0.95-1.43]	33%%	0.22	
Adjusted estimate					
Yes	6	1.29 [1.06-1.58]	42%	<0.01	NA
No	0				
Risk estimated					
Reported in the paper	6	1.29 [1.06-1.58]	42%	<0.01	
Calculated	0				
Risk of bias					
Moderate	4	1.36 [0.98-1.88]	80%	<0.01	0.59
Serious	2	1.22 [0.95-1.57]	0%	0.73	

Mean daily dose					
<500 mg/d	4	1.20 [0.96-1.50]	13%	0.33	0.4
>500 mg/d	2	1.54 [0.86-2.76]	76%	<0.01	
Age					
63 years or less	2	1.22 [0.94-1.58]	0%	0.69	0.62
64 years or more	4	1.35 [0.99-.1.84]	65%	<0.01	
Cancer or haemodialysis patient based-population					
No	6	1.29 [1.06-1.58]	42%	<0.01	
Yes	0	NA	NA	NA	

N: number of studies. NA: Not applicable for a single study
*Excluding studies with critical risk of bias

Table S5. Bayesian meta-analysis and traditional random effect meta-analysis for the association between HCQ alone with or without azithromycin and mortality in patients with COVID-19*

	Bayesian meta-analysis				**Traditional random-effect model**	
	Prior distribution					
	μ: Normal (1,100) τ: Cauchy (0,0.5)		μ: Normal (1,1) τ: Cauchy (0,0.5)			
	RR	τ	RR	τ	RR	τ
HCQ alone n=17	0.93 [0.72-1.14]	0.33	0.93 [0.72-1.14]	0.33	0.83 (0.65-1.06)	0.18
HCQ n=14 (exclusion of RCT)	0.90 [0.65-1.13]	0.12	0.90 [0.66-1.14]	0.35	0.79 [0.60-1.04]	0.20
HCQ n=3 RCT	1.14 [0.64-1.69]	0.24	1.13 [0.56-1.70]	0.31	1.09 [0.97-1.24]	0
HCQ +AZI n=7	1.32 [0.97-1.68]	0.17	1.32 (0.96-1.65]	0.35	1.27 [1.04-1.54]	0.03
HCQ +AZI n=6 (exclusion of RCT)	1.37 [0.97-1.79]	0.21	1.36 [1.01-1.74]	0.36	1.29 [1.06-1.58]	0.03

n=number of studies τ =between-study heterogeneity RR=Relative Risk

*Excluding studies with critical risk of bias

Table S6. Subgroup analysis for the associations between hydroxychloroquine with or without azithromycin and mortality in patients with COVID-19

	N	Pooled Relative Risk	Heterogeneity		
			I^2 (%)	P_{within}	$P_{between}$ (test for subgroup differences)
HCQ alone					
All Studies	27	0.80 [0.65-1.00]	83%	<0.01	
Study Design					
non-randomised	24	0.77 [0.61-0.98]	83%	<0.01	0.01
RCT	3	1.09 [0.97-1.24]	0%	0.87	
Among non-randomised studies	24				
Type of article					
Peer-reviewed	19	0.75 [0.57-0.98]	85%	<0.01	0.08
Unpublished	5	0.98 [0.85-1.13]	58%	0.05	
Adjusted estimate					
Yes	14	0.76 [0.56-1.03]	88%	<0.01	0.8
No	10	0.81 [0.56-1.17]	49%	0.04	
Risk of bias					
Moderate	7	1.03 [0.91-1.17]	0%	0.9	0.04
Serious	7	0.56 [0.33-0.96]	88%	<0.01	
Critical	10	0.75 [0.48-1.16]	84%	<0.01	
Mean daily dose					
<500 mg/d	8	0.72 [0.43-1.19]	89%	<0.01	0.09
>500 mg/d	6	1.05 [0.84-1.32]	0%	0.9	
Not specified	10	0.71 [0.52-0.97]	87%	<0.01	
Age					
63 years or less	11	1.03 [0.93-1.14]	32%	0.14	0.01
64 years or more	13	0.67 [0.49-0.92]	87%	<0.01	

Cancer or haemodialysis patient based-population					
No	18	0.77 [0.58-1.01]	86%	<0.01	0.9
Yes	6	0.79 [0.49-1.27]	60%	0.03	
HCQ+AZI					
All Studies	12	0.87 [0.59-1.27]	88%	<0.01	
Study Design					
Non-randomised	11	0.88 [0.59-1.32]	89%	<0.01	0.6
RCT	1	0.64 [0.18-2.24]	NA	NA	
Among non-randomised studies	11				
Type of article					
Peer-reviewed	8	0.73 [0.40-1.34]	91%	<0.01	0.15
Unpublished	3	1.16 [0.95-1.43]	33%	0.22	
Adjusted estimate					
Yes	7	1.27 [1.04-1.56]	56%	<0.01	0.004
No	4	0.48 [0.25-0.91]	95%	<0.01	
Risk estimated					
Reported in the paper	9	0.94 [0.60-1.49]	91%	<0.01	0.5
Calculated	2	0.60 [0.19-1.88]	79%	0.03	
Risk of bias					
Moderate	4	1.36 [0.98-1.88]	65%	<0.01	0.008
Serious	2	1.22 [0.95-1.57]	0%	0.73	
Critical	5	0.43 [0.22-0.84]	94%	<0.01	
Mean daily dose					
Not specified	5	1.04 [0.62-1.74]	78%	<0.01	0.09
<500 mg/d	5	0.92 [0.51-1.67]	94%	<0.01	

>500 mg/d	1	0.29 [0.10-0.84]	NA	NA	
Age					
63 years or less	4	1.01 [0.84-1.22]	63%	0.05	0.6
64 years or more	7	0.84 [0.45-1.56]	93%	<0.01	
Cancer or haemodialysis patient based-population					
No	11	0.88 [0.59-1.32]	89%	<0.01	NA
Yes	0	NA	NA	NA	

N=Number of studies NA=Not Available

CHAPTER 12

Hydroxychloroquine and COVID-19: a tale of populism and obscurantism

Hydroxychloroquine and COVID-19: a tale of populism and obscurantism

Nathan Peiffer-Smadja, Mathieu E. Rebeaud, Anthony Guihur, Yahya Mahamat-Saleh, Thibault Fiolet

I contributed to this work conceptually and in the manuscript writing.

Published Online
November 13, 2020
https://doi.org/10.1016/
S1473-3099(20)30866-5

Hydroxychloroquine and COVID-19: a tale of populism and obscurantism

We read with interest the article by Estella Ektorp, which describes the death threats received by Marcus Lacerda following a trial on chloroquine for COVID-19 in Brazil.[1] We give Lacerda our full support and herein report our experience in France and Switzerland following publication of a meta-analysis[2] on hydroxychloroquine, with or without azithromycin, for COVID-19.

The meta-analysis included 11 932 participants treated with hydroxychloroquine, 8081 with hydroxychloroquine and azithromycin, and 12 930 patients in a control group. Hydroxychloroquine was not significantly associated with mortality: pooled relative risk (RR) was 0·83 (95% CI 0·65–1·06) across all 17 studies and 1·09 (0·97–1·24) across three randomised controlled trials. Hydroxychloroquine with azithromycin was associated with increased mortality (RR 1·27, 95% CI 1·04–1·54; seven studies).

Several authors of this work have suffered a violent campaign of cyber-harassment on social networks, receiving hundreds of insults, xenophobic messages, anonymous phone calls, and intimidation, including death threats. These actions were accompanied by the public sharing of contact details, including the postal address of authors, on Facebook groups with hundreds of thousands of members. In the same way Ektorp describes the response to Lacerda's trial, aggressive communication and an online campaign of misinformation against the meta-analysis were shared by certain medical and scientific professors, as well as French politicians, going beyond the framework of scientific debate and involving the political sphere.

This behaviour has a goal: to scare researchers and doctors and to silence them. However, silence would be the worst response to this type of behaviour, making societies vulnerable to populism and obscurantism. In a context of uncertainty and anxiety about the pandemic, and when expectations of clear and accessible medical information were immense, silence left medical communication to the champions of unfounded certainties and outrageous simplifications who were perfectly aware of new forms of communication via social networks and YouTube. Against these communicators, most doctors and researchers were unable to explain either the complexity of the medical process or that doubt, differences, and dialogues between peers are the guarantees of quality medicine. The credibility of medical speech emerges deeply shaken out of this, and the false hopes and disillusions generated by unfounded announcements have undermined confidence in medical research.

Citizens are entitled to expect transparent and honest medical information, and we believe this is crucial to address the contemporary challenge of medical communication aimed at the general public, in order to succeed in restoring confidence in medicine and science.

We declare no competing interests.

*Nathan Peiffer-Smadja, Mathieu E Rebeaud, Anthony Guihur, Yahya Mahamat-Saleh, Thibault Fiolet
nathan.peiffer-smadja@inserm.fr

Université de Paris, INSERM, IAME, F-75006 Paris, France (NP-S); National Institute for Health Research Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance, Imperial College London, London, UK (NP-S); Infectious and Tropical Diseases Department, Bichat-Claude Bernard Hospital, AP-HP, Paris, France (NP-S); Department of Plant Molecular Biology, Faculty of Biology and Medicine, University of Lausanne, Switzerland (MER, AG); Centre for Research in Epidemiology and Population Health, INSERM "Health across Generations" Team, Paris-Sud 11 University/Paris Saclay, Villejuif, France (YM-S, TF); and Gustave Roussy, Villejuif, France (YM-S, TF)

1 Ektorp E. Death threats after a trial on chloroquine for COVID-19. *Lancet Infect Dis* 2020; **20:** 661.

2 Fiolet T, Guihur A, Rebeaud ME, Mulot M, Peiffer-Smadja N, Mahamat-Saleh Y. Effect of hydroxychloroquine with or without azithromycin on the mortality of COVID-19 patients: a systematic review and meta-analysis. *Clin Microbiol Infect* 2020; published online Aug 26. https://doi.org/10.1016/j.cmi.2020.08.022.

PART3

Scientific Integrity

Exposing predatory publishers and strengthening science whistleblowers

CHAPTER 13

Raising public awareness about the misuse of predatory journals: One year after the "hydroxychloroquine and push-scooters accidents" hoax

Raising public awareness about the misuse of predatory journals: One year after the "hydroxychloroquine and push-scooters accidents" hoax

Mathieu E. Rebeaud, Florian Cova, Valentin Ruggeri, Michaël Rochoy

I contributed to this work conceptually and in the manuscript writing.

Authors' contributions
MER wrote the first draft of the paper. MER, FC, VR and MR contributed to the writing of the paper. All authors contributed to the writing of the paper and read and approved the final manuscript.

References

[1] Medicines Healthcare products Regulatory Agency (MHRA). Oxford University/AstraZeneca COVID-19 vaccine approved. GOVUK; 2020, https://www.gov.uk/government/news/oxford-universityastrazeneca-covid-19-vaccine-approved [Accessed 22 July 2021].

[2] Europan Medicines Agency (EMA). EMA recommends COVID-19 Vvaccine AstraZeneca for authorisation in the EU. Eur Med Agency 2021, https://www.ema.europa.eu/en/news/ema-recommends-covid-19-vaccine-astrazeneca-authorisation-eu [Accessed 22 July 2021].

[3] Voysey M, Clemens SAC, Madhi SA, Weckx LY, Folegatti PM, Aley PK, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. Lancet 2021;397(10269):99—111.

[4] Deplanque D, Launay O. Efficacy of COVID-19 vaccines: From clinical trials to real life. Therapies 2021;76(4):227—83, http://dx.doi.org/10.1016/j.therap.2021.05.004.

[5] Folegatti PM, Ewer KJ, Aley PK, Angus B, Becker S, Belij-Rammerstorfer S, et al. Safety and immunogenicity of the ChAdOx1 nCoV-19 vaccine against SARS-CoV-2: a preliminary report of a phase 1/2, single-blind, randomised controlled trial. Lancet Lond Engl 2020;396:467—78.

[6] Lacroix C, Salvo F, Gras-Champel V, Gautier S, Massy N, Valnet-Rabier MB, et al. French organization for the pharmacovigilance of COVID-19 vaccines: a major challenge. Therapies 2021;76(4):297—303, http://dx.doi.org/10.1016/j.therap.2021.05.003.

[7] Agence Nationale de Sécurité des Médicaments (ANSM). Point de situation sur la surveillance des vaccins contre la COVID-19. ANSM: Agence nationale de sécurité du médicament et des produits de santé; 2021, https://www.ansm.sante.fr/S-informer/Points-d-information-Points-d-information/Point-de-situation-sur-la-surveillance-des-vaccins-contre-la-COVID-197 [Accessed 22 July 2021].

[8] Chapin-Bardales J, Gee J, Myers T. Reactogenicity following receipt of mRNA-based COVID-19 vaccines. JAMA 2021;325(21):2201—2.

Laurent Chouchana[a,*], Etienne Canouï[b], Rui Batista[c], Adrien Contejean[b,d], Alain Cariou[d,e], Jean-Marc Treluyer[a,d], Caroline Charlier[b,d,f]

[a] *Regional Center of Pharmacovigilance, Cochin Hospital, AP—HP Centre, Université de Paris, 75014 Paris, France*

[b] *Antimicrobial Stewardship Team, Cochin Hospital, AP—HP Centre, Université de Paris, 75014 Paris, France*

[c] *Pharmacy Department, Cochin Hospital, AP-HP Centre, Université de Paris, 75014 Paris, France*

[d] *Université de Paris, 75006 Paris, France*

[e] *Medical ICU, Cochin Hospital, AP—HP Centre, Université de Paris, 75014 Paris, France*

[f] *Institut Pasteur, Biology of Infection Unit, French National Reference Center and WHO Collaborating Center Listeria, Inserm U1117, 75014 Paris, France*

* Corresponding author. Cochin Hospital, AP—HP Centre, 27, rue du Faubourg Saint-Jacques, 75014 Paris, France.

E-mail address: laurent.chouchana@aphp.fr (L. Chouchana)

Received 23 May 2021;
accepted 20 July 2021
Available online 27 July 2021

https://doi.org/10.1016/j.therap.2021.07.004

Raising public awareness about the misuse of predatory journals: One year after the ''hydroxychloroquine and push-scooters accidents'' hoax

KEYWORDS
COVID-19;
Hydroxychloroquine;
Peer review;
Predatory journals;
Push-scooters

Abbreviations

COVID-19	coronavirus disease 2019
HCQ	hydroxychloroquine
WHO	World Health Organization

Predatory publishers are a threat to the good functioning of scientific research. Leading scholars and several publishers agreed on a standard definition of what are predatory journals and publishers: ''Predatory journals and publishers are entities that prioritize self-interest at the expense of scholarship and are characterized by false or misleading information, deviation from best editorial and publication practices, a lack of transparency, and/or the use of aggressive and indiscriminate solicitation practices'' [1]. The problems raised by predatory journals have been known for a number of years and several people have been working on trying to alert researchers about it or finding ways of addressing it. For example, Jeaffrey Beall published a list in his blog in 2010 with the aim of identifying the different predatory publishers [2]. His work and his list had to be abandoned in 2017 following complaints from various publishers and documentalists. Since then, the work has been taken over by Cabell's International, which offers publishing support services to universities and also a list of potentially predatory publishers based on different criteria [3]. Predatory journals can be considered a part of a bigger problem that also includes the surge of preprints and their misuse, as well as the multiple issues that can plague publication in non-predatory journals, such as expedite reviewing, conflicts of interests, and methodological and statistical issues [4].

In the past, predatory journals have mainly been a problem for researchers and have rarely affected the general public in a direct way. Indeed, in normal circumstances, the

general public and journalists' access to scientific data is filtered by representatives of the scientific community, who are usually able to make the distinction between trustworthy scientific journals and predatory ones, even if a grey zone still exists, making the task sometimes complex even for academics and scientists. However, in the context of the coronavirus disease 2019 (COVID-19) pandemic, which was accompanied by an infodemic and the propagation of fake news [5], this filter proved to be no longer sufficient. In competition with the voices of competent scientific experts, we witnessed the rise of a new type of clientelism that thrived on people's ignorance and misunderstanding of the process of scientific publication and tried to capitalize on the legitimacy offered by the notions of peer review and publication of one's work in a scientific journal to push their agenda [6]. Indeed, for those who want to buy themselves a veneer of scientific credibility, predatory journals might constitute a quick and easy solution. In exchange for a little money, anyone can publish anything they want in very little time [7]. As such, it has not been uncommon for articles published in predatory journals to be put forward as indisputable proof that a molecule X or Y brought spectacular results in the treatment of COVID-19. The use of these journals can therefore be used to precipitate an unreasonable hype for treatments, leading researchers to increase the number of low-quality studies that do not provide clear results due to lack of statistical power and questionable methodology, rather than more informative large-scale randomized clinical trials [8].

One example of such misuses of predatory journals came from heated debates about the use of hydroxychloroquine (HCQ) in treating COVID-19. For several months before randomized clinical trials such as World Health Organization (WHO) SOLIDARITY, RECOVERY, and others allowed to conclude to the lack of convincing results for this molecule, the use of HCQ was the topic of an international controversy, as summarized by the editorial on the chaotic and political uses of HCQ that accompanies the Cochrane meta-analysis that compile all these results [9]. In this context, a French collective of physicians called ''*Laissons les médecins prescrire*'' argued for physicians' right to prescribe HCQ as a treatment against COVID-19 on the basis of a study they published in July 2020 (''Azithromycin and hydroxychloroquine accelerate recovery of outpatients with mild/moderate COVID-19''). However, this study was published in the *Asian Journal of Medicine and Health*, which showed all signs of being a predatory journal. This did not stop the study's authors to mention their study in mainstream media, and to present it as serious. When interrogated by journalists about the journal they published in, they argued that it was a genuine scientific journal, in which they had complete trust. Thus, the main public was presented directly with the results of a study published in a predatory journal — a study that was used to push a political agenda.

In reaction to this publication, we decided to publish our own article in the *Asian Journal of Medicine and Health* [10]. Our aim was to publish a ''scientific'' paper that was absurd enough for non-specialists to identify it as a hoax. Our goal for doing this was twofold. Our first and more specific goal was to expose the collective ''*Laissons les médecins prescrire*'' for their scientific misconduct and abuse of the general public's lack of awareness of the existence of predatory journals. The second and more general goal was to raise public awareness about the existence of predatory journals. Indeed, the use of predatory journals constitutes a greater threat than the abuse of preprints, which still constitute an efficient way of making results available to the scientific community before publication and fostering interest for the research. Making research accessible ahead of acceptance and publication can lead to very interesting discussions between scientists from different backgrounds on some possible weaknesses of the study and offer a community peer review before or during the submission of the article to the journal of the authors' choice. Publication in predatory journals shows none of these benefits.

Finally, on July 24th 2020, we submitted our paper to the *Asian Journal of Medicine and Health*. The paper was written to be obviously bogus. Entitled ''SARS-CoV-2 was unexpectedly deadlier than push-scooters: could hydroxychloroquine be the unique solution?'', its main claim was that the prophylactic use of HCQ could prevent push-scooters' accidents and had actually been instrumental in diminishing the number of push-scooters' accidents during the first French lockdown. In addition to its implausible conclusion, the paper also contained several outrageous claims, such as HCQ being a solution to the ''Israelo-Palestinian problem''. It also contained egregious methodological errors (such as putting people who died early in the trial in the control group), ethical issues (such as literally killing people or trying to dig up participants' corpses), and statistical mistakes (such as correlation > 1 or unadapted statistical tests). It also contained references to irrelevant, inexistant, or straight-up unscientific materials, an absurd list of authors' contributions (including 2 authors who did nothing, and a dog who barked to support the others), and also included quotes from various movies (such as the *Dark Knight* or *Apocalypse Now*). All in all, it was blatant to any careful reader that the paper was not serious.

Despite these issues, our paper was ultimately accepted after one turn of revision in which 3 reviewers only pointed to small issues in our papers (such as the fact that we cited Wikipedia). One of the 3 editors was worried by the fact that we initially claimed that we received ethical approval from a committee composed of ourselves and our friends, but his worries were assuaged when we answered that we stayed out of the room during the deliberation. Finally, our paper was published online on August 15th 2020. All in all, we were able to publish our paper in three weeks for the sum of $85: $55 for the publication with an extra $30 to speed up peer review ''without losing quality''.

The article was retracted by the journal one day later, on August 16th 2020, after requests for explanations by journalists. However, this did not prevent our hoax from receiving wide media coverage, probably with a *Streisand* effect. But what lesson did journalists (and the general public) draw from our hoax? One worry was that our hoax might actually contribute to damage the credibility of scientific publishing and scientific papers in general because journalists might have failed to convey the idea that AJMAH was not representative of genuine scientific journals. To determine what was the focus of media coverage, we surveyed media coverage of our hoax in French-speaking venues and established a

list of 29 items. The reason for this choice was that, in other languages, media coverage of our hoax was mainly destined to a scientific audience rather than tailored to general audiences (except in India, Brazil and Finland). From these, 22 (around 75%) explained the distinction between genuine scientific journals and predatory journals. Only 12 (around 40%) explained that the hoax was also a reply to the collective's original paper. Thus, most journalists took our hoax as an opportunity to inform the general public of the existence of predatory scientific journals (Table S1).

Thus, it seems that the hoax contributed to increasing public awareness of the existence of ''fake'' scientific journals. However, we do not suggest that scientists should regularly engage in such hoaxes, as their repetition might lead them to lose their ''shock value'' while undermining public trust in scientific publication. Of course, our hoax was not the first one to denounce the problems raised by predatory journals: for example, in collaboration with the journal Science, Bohannon submitted fake scientific articles with serious and obvious scientific flaws to more than 300 journals belonging to as many fee-charging open access publishers, with an impressive 60% acceptance rate [11]. But we think our hoax was the first to reach so many non-scientists, due to its topic and the debates about hydroxychloroquine. This shows that hoaxes can be used as pedagogical material to illustrate and explain the notion of predatory scientific journals. Indeed, non-scientists are rarely familiar with the functioning of scientific journals, leaving them clueless when having to discriminate between legitimate and dubious scientific sources. One way to fight the infodemic might thus be to increase knowledge about the ways scientific knowledge is produced and disseminated.

Finally, we should note that there was an additional positive outcome to our hoax, as we received a lot of positive feedback from other scientists thanking us for giving them the occasion to have a good laugh. This might not be much but, in these trying times, it is still something.

Acknowledgments

In memory of science journalist Jean-Daniel Flaysakier, who suddenly passed away on 7 October 2021. The authors send their condolences to his family and friends.

Disclosure of interest

The authors declare that they have no competing interest.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.therap.2021.10.009.

References

[1] Grudniewicz A, Moher D, Cobey KD, Bryson GL, Cukier S, Allen K, et al. Predatory journals: no definition, no defence. Nature 2019;576:210—2.
[2] Beall J. Predatory publishers are corrupting open access. Nature 2012;489:179.
[3] Mimouni M, Braun E, Mimouni FB, Mimouni D, Blumenthal EZ. Beall's list removed: what stands between us and open access predators? Am J Med 2017;130:e371—2.
[4] Besançon L, Peiffer-Smadja N, Segalas C, Jiang H, Masuzzo P, Smout C, et al. Open science saves lives: lessons from the COVID-19 pandemic. BMC Med Res Methodol 2021;21:117.
[5] Orso D, Federici N, Copetti R, Vetrugno L, Bove T. Infodemic and the spread of fake news in the COVID-19-era. Eur J Emerg Med 2020;27:327—8.
[6] Maisonneuve H. Les revues prédatrices : une menace pour la science. À propos de la COVID-19. Médecine 2020;16:389—91.
[7] Vervoort D, Ma X, Shrime MG. Money down the drain: predatory publishing in the COVID-19 era. Can J Public Health 2020;111:665—6, http://dx.doi.org/10.17269/s41997-020-00411-5.
[8] West JD, Bergstrom CT. Misinformation in and about science. Proc Natl Acad Sci USA 2021;118 [e1912444117].
[9] Gould S, Norris SL. Contested effects and chaotic policies: the 2020 story of (hydroxy) chloroquine for treating COVID-19. Cochrane Database Syst Rev 2021;3:ED000151.
[10] Oodendijk W, Rochoy M, Ruggeri V, Cova F, Lembrouille D, Trottinetta S, et al. SARS-CoV-2 was unexpectedly deadlier than push-scooters: could hydroxychloroquine be the unique solution? Asian J Med Health 2020;18(9):14—21.
[11] Bohannon J. Who's afraid of peer review? Science 2013;342:60—5.

Mathieu E. Rebeaud[a], Florian Cova[b], Valentin Ruggeri[c], Michaël Rochoy[d,e,*]
[a] *Department of Plant Molecular Biology, Faculty of Biology and Medicine, University of Lausanne, CH-1015 Lausanne, Switzerland*
[b] *Philosophy Department, University of Geneva, 1211 Geneva, Switzerland*
[c] *Nuclear Medicine Department, Grenoble-Alpes University Hospital, 38000 Grenoble, France*
[d] *Collectif Du Côté de la Science, 62230 Outreau France*
[e] *ULR 2694 — METRICS, CERIM, Département de Médecine Générale, University Lille, CHU Lille, 59000 Lille, France*

* Corresponding author. 20, rue André-Pantigny, 62230 Outreau, France.
E-mail address: michael.rochoy@gmail.com (M. Rochoy)

Received 24 August 2021;
accepted 25 October 2021
Available online 29 October 2021

https://doi.org/10.1016/j.therap.2021.10.009

Prise en charge de l'hypersensibilité immédiate aux β-lactamines : étude rétrospective au centre de pharmacovigilance de Sfax

Management of immediate hypersensitivity to β-lactams: Retrospective study in Sfax pharmacovigilance centre

Mots clés β-lactamines ; Hypersensibilité immédiate ; Test cutané ; Test de provocation orale ; Réactions croisées

CHAPTER 14

SARS-CoV-2 was Unexpectedly Deadlier than Push-scooters: Could Hydroxychloroquine be the Unique Solution?

SARS-CoV-2 was Unexpectedly Deadlier than Push-scooters: Could Hydroxychloroquine be the Unique Solution?

Mathieu E. Rebeaud (Willard Oodendijk), Michaël Rochoy, Valentin Ruggeri, Florian Cova, Didier Lembrouille, Sylvano Trottinetta, Otter F. Hantome, Nemo Macron and Manis Javanica

I contributed to this work conceptually and in the manuscript writing.

Authors' contributions

This work was carried out in collaboration among all authors. Author WO launched idea on Twitter, added some sentences, submitted the paper, corresponded with the kind publisher. Author MR launched MP group on Twitter and Google Docs, performed study 1, added some sentences here and there, responded to reviewers. Author VR needed SIGAPS points, did the minimum. Found a beautiful picture for figure 3. Author FC wrote a lot of sentences but didn't need too many SIGAPS points, so fourth place was not as bad. Also, performed Study 3 in his head (philosophers are good at thought experiments). Author DL was on holidays and added his name at the last time. Author ST wrote nothing but provided the push-scooters and did the outside job. Author ÖFH did nothing but is a very good friend of us; he helped us get some administrative paperwork. Author NM said "waouf" when the authors started to doubt (doubts are common in science, don't let them win you over, believe in yourself and what you do, don't let anyone distract you from the truth you know). Author MJ wrote sentences and said that the last place will be "enough for him". He does it every time, and it works pretty well on interns, we have to admit. All authors read and approved the final manuscript.

Asian Journal of Medicine and Health

18(9): 14-21, 2020; Article no.AJMAH.60013
ISSN: 2456-8414

SARS-CoV-2 was Unexpectedly Deadlier than Push-scooters: Could Hydroxychloroquine be the Unique Solution?

Willard Oodendijk[1*], Michaël Rochoy[2], Valentin Ruggeri[3], Florian Cova[4], Didier Lembrouille[5], Sylvano Trottinetta[6], Otter F. Hantome[7], Nemo Macron[8] and Manis Javanica[9]

[1]Belgian Institute of Technology and Education (BITE), Couillet, Belgium.
[2]General Practitioner and Independent Seeker of Science, Ankh, Morpork, France.
[3]Observatoire of Zététique, Grenoble, France.
[4]Institute for Quick and Dirty Science, Neuneuchâtel, Switzerland.
[5]Département de Médecine Nucléaire Compliante de la SFR, île de Guyane, France.
[6]Collectif Laissons les Vendeurs de Trottinette Prescrire, France.
[7]University of Melon, Melon, France.
[8]Palais de l'Élysée, Paris, France.
[9]Institute of Chiropteran Studies, East Timor.

Authors' contributions

This work was carried out in collaboration among all authors. Author WO launched idea on Twitter, added some sentences, submitted the paper, corresponded with the kind publisher. Author MR launched MP group on Twitter and Google Docs, performed study 1, added some sentences here and there, responded to reviewers. Author VR needed SIGAPS points, did the minimum. Found a beautiful picture for figure 3. Author FC wrote a lot of sentences but didn't need too many SIGAPS points, so fourth place was not as bad. Also, performed Study 3 in his head (philosophers are good at thought-experiments). Author DL was on holidays and added his name at the last time. Author ST wrote nothing but provided the push-scooters and did the outside job. Author ÖFH did nothing but is a very good friend of us; he helped us get some administrative paperwork. Author NM said "waouf" when the authors started to doubt (doubts are common in science, don't let them win you over, believe in yourself and what you do, don't let anyone distract you from the truth you know). Author MJ wrote sentences and said that the last place will be "enough for him". He does it every time, and it works pretty well on interns, we have to admit. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJMAH/2020/v18i930232
Editor(s):
(1) Dr. Nicolas Padilla-Raygoza, University of Celaya, Mexico.
(2) Dr. Janvier Gasana, International University, USA.
(3) Dr. Ashish Anand, GV Montgomery Veteran Affairs Medical Center, USA.
Reviewers:
(1) S. Danish Kadir, University of Texas Rio Grande Velley, USA.
(2) Maria del Carmen Marquett, Institute of Tropical Medicine Pedro Kouri, Cuba.
(3) Ahmed Abdullah Khalleefah, Tripoli University, Libya.
Complete Peer review History: http://www.sdiarticle4.com/review-history/60013

Original Research Article

Received 24 July 2020
Accepted 11 August 2020
Published 15 August 2020

**Corresponding author: E-mail: willardoodendijk@protonmail.com;*

ABSTRACT

Background: YouTube™ and Dropbox© studies have warned against the deadly potential of push-scooters.
Aims: Through three studies, we evaluate the potential of a combination of hydroxychloroquine and azithromycin for preventing push-scooter accidents.
Study Design: Studies 1 and 2 are retrospective observational studies in which we rely on archival data to explore the relationship between push-scooters accidents (PSA) and usage of hydroxychloroquine plus azithromycin (HCQ + AZT) in France in 2020 and 2019 respectively (7 participants). Study 3 is a partly randomized clinical trial (pRCT), retrospective, in which the use of HCQ + AZT for preventing PSA was assessed in a radically "direct" way (6 participants).
Place and Duration of Study: Studies 1 and 2 were conducted in the authors' office chair (Ikea) in France (multicentric), on July 20th, 2020. Study 3 was conducted in the parking lot of an abandoned factory (Montcuq, Occitan region, France).
Methodology: For Studies 1 and 2, we used data from OpenMEDIC to determine usage of hydroxychloroquine in France in 2020 and Google Actuality to determine the rate of PSA in France in 2020. For Study 3, we adopted an experimental approach and had participants exposed to HCQ + AZ (treatment group) or homeopathy (control group) before having them perform a standard push-scooter exercise. Advanced statistical models were used to assess the prophylactic effect of the HCQ + AZT combination on PSA.
Results: Wide use of hydroxychloroquine is strongly associated with a very low level of PSA, both in time (2020 VS 2019) and in space (Marseille, Bouches-du-Rhônes versus the rest of France). Moreover, the results of our retrospective pRCT prove without any doubt that prophylactic use of a HCQ + AZT combination helped to prevent PSA.
Conclusion: The HCQ + AZT combo should urgently be used in prevention of PSA all around the world.

Keywords: Hydroxychloroquine; azithromycin; zinc; soup; COVID-19; motion sickness; push-scooters.

ACRONYMS

HCQ : Hydroxychloroquine
AZT : Azithromycin
PSA : Push-scooters Accidents

1. INTRODUCTION

As the number of push-scooters has been rising in France, so has the number of push-scooters accidents. Some of these accidents have proven to be deadly and previous YouTube™ and Dropbox© studies have warned against the deadly potential of push-scooters [1]. For a comparison, only three Chinese people had died from the novel coronavirus SARS-CoV-2 at the end of 2019 [2]. It is therefore important to reflect on the use of push-scooters through an accurate and ethical cost-benefit analysis.

Use and promotion of push-scooters have been advocated on the basis that they would contribute to the reduction and slowing of global warming. In fact, the French scientific elite has been working on the subject and has recently argued that there was no proof of global warming, as he could not see the ice cap melt on his computer [3]. So, even if global warming was real, there are serious reasons to think that France is not affected, as global warming clearly stopped at the closed border [4]. Unfortunately, the debate is being polluted by bots, trolls and so-called experts funded by Big Trottinette to spread misinformation. Indeed, an independent study (in press on the third author's Google Drive®) found a positive correlation between experts' positive advocacy of push-scooters and the amount of money they received from Decathlon® (r = 3.14). The fact that push-scooters is now a 'generic' means of locomotion that can be produced by anyone for a cheap price might lead people to the conclusion that no private interest is involved, but we're not fooled, we know the truth [5]. So, it is important to diminish the increasing number of push-scooter drivers who are sacrificed on a daily basis.

In the present paper, we investigate an unexpected way of mitigating the death toll of the push-scooter craze: a combination of hydroxychloroquine and azithromycin (zinc can be added to improve the flavor, but is only necessary if the study fails to provide any significant effect). We combine two observational

studies (Study 1) and one randomized clinical trial (Study 3) to justify our hypothesis that HCQ + AZT is the key to all world's problems (see second author's personal diary for an application to the Israel-Palestine problem). Indeed, association of hydroxychloroquine and azithromycin was proposed in a brilliant paper to treat the novel coronavirus disease (later named COVID-19), with a spectacular effect [6]. It is obvious that the use of an antibiotic in viral infection could not cause antimicrobial resistance (because of the absence of bacteria), so that's not a problem in our study; plus, these treatments are old (first mentioned two millenaries ago by Galen in *De Remediis Parabilibus*, as an extract of *Cinchona officinalis*), absolutely safe (some adverse effect are possible - the Summaries of Product Characteristics are as much a lie as the cake) and inexpensive ($400 per ton [7]), so we need to test them [8]. In 2019, hydroxychloroquine was given for rheumatoid arthritis, lupus erythematosus (discoid and subacute), systemic lupus and lucites; since 2020, it was given since March 2020, it's given because we have to let doctors prescribe what they want, damn it! Given the obvious similarities between COVID-19 and push-scooters accidents (i.e. both can have deadly outcomes in which the patient might even die), it seemed natural to expand the use of HCQ + AZT to push-scooters accidents (henceforth: PSA), even if no *in-vitro* study ever found an effect of HCQ on PSA.

2. MATERIALS AND METHODS

2.1 Study 1 – Relationship between the Use of HCQ + AZT and Frequency of PSA in France in 2020

Our objective was to evaluate the relationship between PSA and usage of hydroxychloroquine (HCQ) in France in 2020.

The computers and the Internet services were provided by the authors or by private institutions exclusively. The state of public research in France is not really compatible with the purchase of equipment quickly (except after justifying needs with the help of twelve forms, a photocopy of the rental lease of your student room from 18 to 25 years old and your car registration, in 3 copies). We used data from OpenMEDIC (https://www.data.gouv.fr/fr/datasets/open-medic-base-complete-sur-les-depenses-de-medicaments-interregimes/) to determine usage of hydroxychloroquine in France in 2020. We used Google Actuality to determine the rate of PSA in French press in 2020 (query equation: "accidents de trottinette"). We have not classified the accidents by type, date or anything else, essentially by laziness.

2.2 Study 2 – Use of HCQ + AZT and PSA before 2020

Study 2 was excluded from analysis and from this paper, as it did not provide informative results (i.e. the results we wanted).

2.3 Study 3 – Use of HCQ + AZT in Prophylaxis against PSA

In Study 3, we tested the efficiency of our protocol in the prevention of PSA: HCQ + AZT (or spiramycin or nothing) +/− Zinc (or Magnesium or a teaspoon of Benco (C or R in a circle, or maybe TM) +/− Vitamin D (or Selenium). We have sometimes added apples, as their therapeutic effectiveness is popularly recognized [9]. This study was supported by the "Laissons les Vendeurs de Trottinette Prescrire" Collectif, the National Assembly and the Independent Push-Scooter Salesmen's Pension Fund.

2.3.1 Participants

Two groups of volunteers (friends and relatives of the authors) were constituted. Push-scooters were invented in 1960; some researchers says 1930, but it wasn't born at that time, so it doesn't know. So, all participants with age > 60 (born before 1960) were placed in the control group (M_{age} = 75.09, SD_{age} = 5.21). Apart from that, the assignment to either condition was random -- we swear it on the "*La vérité sur la maladie de Lyme*" book.

In a pre-test phase, we asked each participant in the treatment group to roll 500m in a straight line on a push-scooter. Participants who fell or died during the pre-test were reallocated to our control group (two falls, one death). Thus, we were left with six volunteers in our treatment group (M_{age} = 13.13, SD_{age} = 1.11).

2.3.2 Justification for sample size and analysis plan

Following the methodological rule according to which the smaller the sample, the higher the statistical significance [9], we decided to stop recruitment as soon as a significant effect at 84% was detected.

2.3.3 Method

Participants in both groups were asked to go down a 45° slope with a steep brick wall at the end on a push-scooter. They were instructed to go as fast as possible and brake at the last moment before hitting the brick wall. To reinforce ecological validity, sounds of cars and insults from other push-scooter drivers were broadcasted from the experimenters' phones (sounds were recorded in Paris prior to lockdown) [10]. Due to limited resources and fundings, only two push-scooters (one very old, one brand new) were available. The old, rusty push-scooter was randomly attributed to participants in the control group. It should be noted that the brand new push-scooter was in zinc, which might have contributed to potentialize the HCQ + AZ combination.

This study was retrospective, which is why we did not need an opinion from the ethics committee.

3. RESULTS

3.1 Study 1 – Relationship between the Use of HCQ + AZT and Frequency of PSA in France in 2020

On Google Actuality (page 1), we noted 1 PSA on 20th July (Val d'Oise), 1 on 20th October 2019 (Bordeaux), 2 on 2th and 22th september 2019 (Reims and Levallois-Perret), 1 on 26th april 2020 (Nord-sur-Edre), 1 on 2th december 2019 (Nancy) and 1 on 20th january 2020 (Villefranche-sur-Saône).

As our results didn't find PSA in March 2020, we concluded that hydroxychloroquine was an effective preventive therapy for PSA with a RR = 0 ($p<0.0001$).

We only consulted page 1 of Google Actuality: following the methodological rule according to which the smaller the sample, the higher the significance [11], we decided to stop recruitment as soon as we found a significant effect.

In order to increase sensibility and specificity, we decided to run additional, exploratory analyses and searched another figure in Google Images about PSA (query equation: "accidents de trottinette") and found this one, which seems in favor of our initial idea about the subject, so we performed a graph extraction procedure using Windows' "Ctrl-C Ctrl-V" command [9] (Fig. 2).

As suggested by our reviewer, we could highlighted that due to the home confinement suggested by the WHO for the Covid 19 contingency, a decrease in the circulation of push-scooters can be inferred at the time of 2020 compared to that of 2019, which may be a factor to take into account to say that there is a decrease in accidents.

Adherence of azithromycin was not studied.

3.2 Study 3 – Use of HCQ + AZT in Prophylaxis against PSA

In the treatment group, only 1 of the six participants died during the experiment (though there were a total of three non-deadly accidents). However, death was suspicious as it intervened before participants went down the slope, at the moment he received instructions from the experimenter. According to the coroner, he might have died from a heart problem caused by the treatment. However, as he was very young, we ruled this out as improbable, because hydroxychloroquine is a really safe medication [12]. So, we decided not to count him as a death and rather excluded him from analysis; we counted him as having voluntarily decided to discontinue the treatment.

In the control group, two participants out of four died from an accident (a whopping 50%, the two other participants survived without accident). After several attempts, we were able to find a test that returned a significant effect – paired one-tailed Student t-test: $t(3) = 1.73$, $p < .10$. This suggests that HCQ + AZT was effective and had a stupefying effect in preventing push-scooters accidents. However, because a bunch of petainists methodologists refused to consider our result as significant because it did not reach their completely arbitrary significance threshold, we decided to perform further analyses. Indeed, these are not the international standards, but they are the new standards in France. We realized that oxygen saturation might be a good clinical indicator of mortality (after all, dead people typically do not breathe) – maybe even better than death itself. We thus went back to our participants to measure their oxygen saturation levels (after taking down their surgical mask, to avoid confounds). We did not receive official authorization to dig up the corpses of dead

participants; we did try to dig them up anyway, but the cemetery keeper could no longer find the register of graves, and unfortunately we did not have the necessary material for several blind desecrations. So, we decided to just assume that their transcutaneous oxygen saturation levels were 0%. This analysis returned a "significant" effect – chi-square test: $\chi^{square} = 4.5$, $p < .04$.

4. DISCUSSION

Given the alarmist comments of some people who have not looked at the level of the ice pack on Google Earth by themselves, there is a risk that they will continue to abandon real means of transport and continue to use push-scooters. So, we can certainly expect an increase of the number of push-scooter users, of deaths and injuries [7].

Iodized salt (table salt mixed with iodine) is a well-known preventive method against iodine deficiency (major cause of thyroid affections and intellectual and developmental disabilities) [13], and we suggest that hydroxychloroquine could in the same way be added to table salt in order to

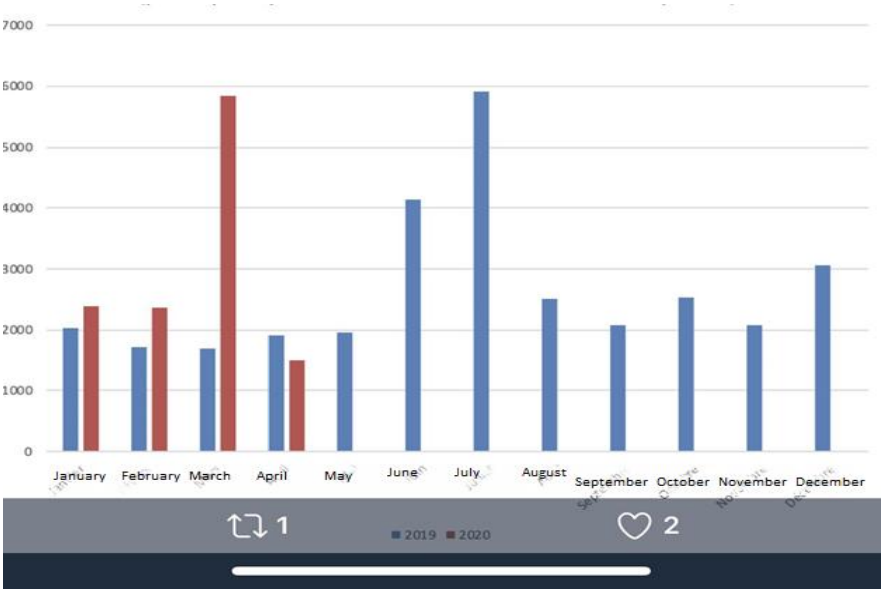


Fig. 1. Number of boxes of hydroxychloroquine reimbursed in the country of Voltaire and Molière (non-exhaustive list of the French elite) in 2019 and 2020

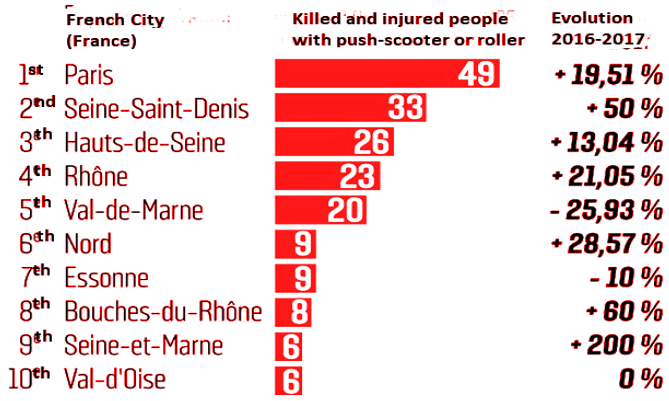


Fig. 2. PSA in France (deaths plus injuries). Note the low level of accidents in Bouches-du-Rhône, where hydroxychloroquine was always widely used for malaria control (Courtesy of Google Images).Translation by authors due to editorial exigencies

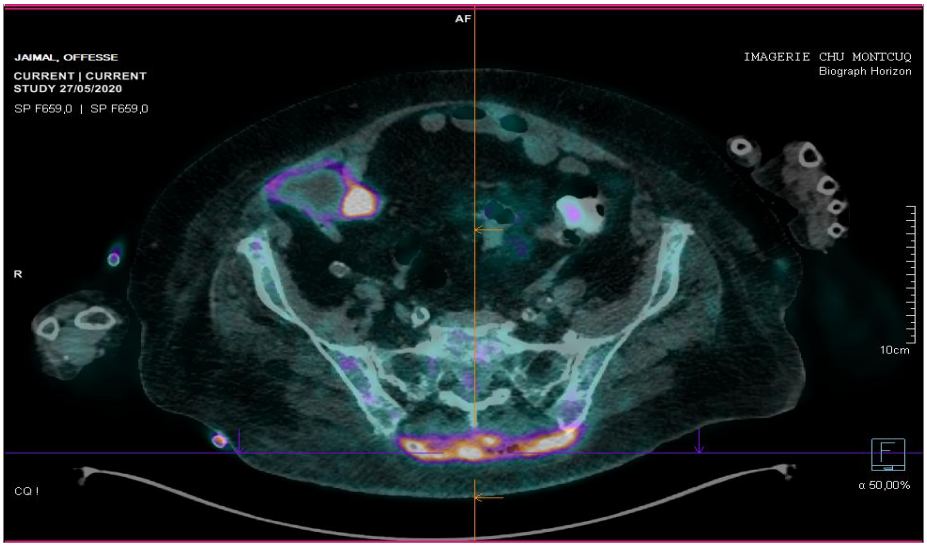


Fig. 3. H type sacral fracture in a 40-years-old female after a push-scooter accident, during COVID-19 pandemia, in Vigneulles-lès-Hattonchâtel, Meuse, France (PET/CT acquisition two days after the trauma). HCQ + AZT was unfortunately not used in this patient, due to location (not Marseille)

prevent both COVID-19 and push-scooter accidents. The pathophysiology of the protective effect of hydroxychloroquine remains to be clarified, but the urgency seems first to prescribe widely. That will save millions of lifes in the world according to our study and other studies [14]. We received the President of the Republic to discuss the results in preview and he was thrilled; following our interview, in his Churchillian televised speech, he insisted that all avenues for preventing scooter accidents be explored.

5. CONCLUSION

In our study, hydroxychloroquine was associated with lower odds of push-scooters accidents. It is urgent to prescribe hydroxychloroquine for all push-scooters users.

Can we publish anything right now? I think that the question, it is quickly answered, and peer-review has never been a scientific method anyway [15].

Further research is needed, especially randomized controlled trials as it was performed for parachutes in aircraft injuries [16]. Obviously, a preprint of an observational study combined with an underpowered CRT will be enough to provide guidelines until that.

Hydroxychloroquine is a cheap, devilishly effective molecule with a higher level of safety than many other drugs. We need to use it more, everywhere, all the time, all around the world. Because hydroxychloroquine is the hero the world deserves, but not the one it needs right now. So the detractors will hunt it. Because hydroxychloroquine can take it. Because it's not our hero. It's a silent guardian, a watchful protector. A dark knight.

Otherwise, the adjunction of zinc [17], ivermectin [18] or any other drug [19] to the association HCQ + AZT should be considered. As the great French scientist Jean-Claude Dusse once said: "you never know, on a misunderstanding, it might work" [20].

DISCLAIMER

The products used for this research are commonly and predominantly use products in our area of research and country. There is absolutely no conflict of interest between the authors and producers of the products because we do not intend to use these products as an avenue for any litigation but for the advancement of knowledge. Also, the research was not funded by the producing company rather it was funded by personal efforts of the authors.

CONSENT

It is not applicable.

ETHICAL APPROVAL

The authors want to mention that the ethics committee that gave retrospective approval for this study was independant; the authors and their colleagues usually participate, but in accordance with the Helsinki rules, they did not take part in the processing of the present submission.

ACKNOWLEDGEMENTS

Hervé de Maisonneuve for his fight against predatory magazines and his daily struggle for good French science. That's probably the only serious sentence in all this (not counting this one, we don't want to risk a paradox).

Violaine Guérin, Martine Wonner and their team, who paid 55 bucks to be published here, thereby allowing us to discover new scientific journals, and to have a lot of fun.

Joachim Son-Forget, Member of Parliament, who taught us that linear regression starts from 3 points; we soon hope to push the limits and reach the purity of linear regression at 1 point.

Idriss Aberkane for his disruptive contribution to world epistemology.

Obviously, Didier Raoult, with or without whom we would be nothing.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Raoult D. Coronavirus: fewer deaths than from push-scooter accidents. International IHU Channel of YouTube; 2020. Available:https://www.mediterranee-infection.com/coronavirus-moins-de-morts-que-par-accident-de-trottinette/
2. Raoult D. Coronavirus in China: should we be concerned? International IHU Channel of YouTube; 2020. Available on: https://www.youtube.com/watch?v=qoBoryHuZ6E
3. Raoult D. This whole thing's gonna end up like contaminated blood. International J of L'Obs; 2020. Available:https://www.nouvelobs.com/coronavirus-de-wuhan/20200430.OBS28205/didier-raoult-toute-cette-histoire-va-finir-comme-le-sang-contamine.html
4. Bègue L, Bushman BJ, Zerhouni O, Subra B, Ourabah M. 'Beauty is in the eye of the beer holder': people who think they are drunk also think they are attractive. Br J Psychol. 2013;104(2):225-34.
5. Mitchell MA, Wartinger DD. Validation of a Functional Pyelocalyceal Renal Model for the Evaluation of Renal Calculi Passage While Riding a Roller Coaster. J Am Osteopath Assoc. 2016;116(10):647-52.
6. Gautret P, Lagier JC, Parola P, et al. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. Int J Antimicrob Agents. 2020:105949.
7. Rochoy M, Lembrouille D, Trottinetta S, Hantome ÖF, Oodendijk W, Macron N, Ruggeri V, Javanica M, Cova F. Picsou doit-il être intégré dans le canon Disney ? Picsou Mag. 2019;7(11):12-12.
8. Sigaps Max. I love the smell of hydroxychloroquine in the morning. Inter J of Secret Agents. 1979;4(8):1516-2342
9. Roberts TL. Text editors. Handbook of human-computer interaction. 1988;655-672.
10. Zaresakhvidi MJ, Zaresakhvidi F, Mehrparvar AH, Foraster M, Dadvand P. Association between noise exposure and diabetes: A systematic review and meta-analysis. Environ Res. 2018;166:647-657.
11. Raoult D. Trust me, I'm the elite and I know the rules of statistics to make your life easier. Useless Commission of National Assembly; 2020. Available:https://www.liberation.fr/checknews/2020/06/28/plus-l-echantillon-d-un-essai-clinique-est-faible-plus-c-est-significatif-que-veut-dire-didier-raoul_1792554
12. Guérin V. et al. Retrospective study in 88 subjects with 3 different therapeutic approaches. International Dropbox of the Authors;2020.
13. Available:https://www.lequotidiendumedecin.fr/cdn/ff/0M-fxVP3K4HEwlQRjWQ3eEiqAiOm7vC4Sf2Wllf3lGA/1588579686/public/2020-05/COVID_19_RAPPORT_ETUDE_RETROSPECTIVE_CLINIQUE_ET_THERAPEUTIQUE_200430[72626].pdf
14. Gallus S, Bosetti C, Negri E, et al. Does pizza protect against cancer? Int J Cancer. 2003;107(2):283-4.

15. Guérin V, Lévy P, Thomas JL, Lardenois T, Lacrosse P, Sarrazin E, Andreis NR-de, Wonner M. Azithromycin and Hydroxychloroquine Accelerate Recovery of Outpatients with Mild/Moderate COVID-19. Asian Journal of Medicine and Health. 2020;18(7):45-55.
16. Aberkane I. Raoult est encore un héros. Interregional Journal of Noopolitics; 2020. Available:https://www.youtube.com/watch?v=s4lF9ExiR8s
17. Parachute use to prevent death and major trauma when jumping from aircraft: randomized controlled trial. BMJ. 2018;363:k5343.
18. Shittu MO, Afolami OI. Improving the efficacy of Chloroquine and Hydroxychloroquine against SARS-CoV-2 may require Zinc additives - A better synergy for future COVID-19 clinical trials. Infez Med. 2020;28(2):192-197.
19. Cova F, Javanica M, Oodendijk W, Rochoy M, Macron N, Ruggeri V, Lembrouille D, Trottinetta S, Hantome ÖF. An apple a day keeps the doctor away. Pommed'Api. 1985;2(22):225-32.
20. Rizzo E. Ivermectin, antiviral properties and COVID-19: a possible new mechanism of action. NaunynSchmiedebergs Arch Pharmacol. 2020;393(7):1153-1156.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/60013

CHAPTER 15

Open Letter: Scientists stand up to protect academic whistleblowers and post-publication peer review

Open Letter: Scientists stand up to protect academic whistleblowers and post-publication peer review.

Lonni Besançon, Alexander Samuel, Thibault G. Sana, Mathieu E. Rebeaud, Anthony Guihur, Marc Robinson-Rechavi, Nicolas Le Berre, Matthieu Mulot, Gideon Meyerowitz-Katz, Hervé Maisonneuve, Brian A. Nosek.

I contributed to this work conceptually and in the manuscript writing.

Open Letter: Scientists stand up to protect academic whistleblowers and post-publication peer review.

Lonni Besançon,[1] Alexander Samuel,[2] Thibault G. Sana,[3] Mathieu E. Rebeaud,[4] Anthony Guihur,[4] Marc Robinson-Rechavi,[5] Nicolas Le Berre,[6] Matthieu Mulot,[7] Gideon Meyerowitz-Katz,[8] Hervé Maisonneuve,[9] Brian A. Nosek.[10,11]

[1]Faculty of Information and Technology, Monash University, Australia. lonni.besancon@gmail.com

[2]Association de Toxicologie Chimie de Paris, France. alexander.samuel@laposte.net

[3]LIPME, Université de Toulouse, INRAE, CNRS, 31320 Castanet-Tolosan, France. thibault.sana@inrae.fr

[4]Department of Plant Molecular Biology, Faculty of Biology and Medicine, University of Lausanne, Switzerland. anthony.guihur@gmail.com rebeaud.mat@gmail.com

[5]Department of Ecology and Evolution, Faculty of Biology and Medicine, University of Lausanne, Switzerland. marc.robinson-rechavi@unil.ch

[6]Intensive Care Unit, Metz-Thionville Regional Hospital, France. n.leberre84@gmail.com

[7]Laboratory of Soil Biodiversity, University of Neuchâtel, 2000 Neuchâtel Switzerland. matthieu.mulot@unine.ch

[8]School of health and society, University of Wollongong, Australia. gidmk.healthnerd@gmail.com

[9]Research Integrity Officer, Schools of Medicine, University of Paris, & Consultant, Paris France. hervemaisonneuve@gmail.com

[10]Professor, Department of Psychology, University of Virginia, Charlottesville, Virginia USA.
[11]Executive Director, Center for Open Science, Charlottesville, Virginia USA. nosek@virginia.edu

For years, Dr. Elisabeth Bik has contributed immensely to scientific research through her much needed and difficult work on detecting errors in published scientific communications. She has detected and reported over 4000 image duplications and manipulations, data concerns, plagiarism cases, and human and animal ethics breaches through comments on the post-publication review platform *PubPeer*, social media, and her website *Science Integrity Digest*. She has also contacted hundreds of editorial boards of journals to report on errors or potential scientific fraud. Nature credited Dr. Bik as the "public face of image sleuthing" [1] and her work has been awarded the *2021 Peter Wildy Prize Lecture* from the *UK Microbiology Society*.

Dr. Bik's work is instrumental to ethical, sound, and reproducible research, but it also introduces her to personal risk as a whistleblower. In one case, she identified image duplication and potential ethical issues in 62 published articles from Prof. Didier Raoult and his institute some of which have also been echoed in published scientific communication [2,3,4]. For several months, Prof. Raoult and some other members of his institute have responded by insulting her on national television, disclosing her personal address on social media, and threatening legal action for harassment and defamation. Prof. Raoult and his team's behaviour toward Dr. Bik and others have been pointed out by many international media outlets such as Times Higher Education or Scientific American. This strategy of harassment and threats creates a chilling effect for whistleblowers and for scholarly criticism more generally [5,6,7]. Scholarly criticism, particularly on issues of research integrity, is fraught with challenges. This makes it particularly important to maintain focus on addressing such critique with scientific evidence and not attacks on people's appearance, character, or person.

The authors and the co-signatories of this open letter are scientists dedicated to transparency and integrity of research. We support the work needed to investigate potential errors and possible misconduct and believe the scientific community can do more to protect whistleblowers against harassment and threats. Individual researchers can provide vocal support for whistleblowers and against harassment to shift norms. Journals, funders, policymakers, and institutions can make explicit policies that protect whistleblowers and establish fair, judicious, and transparent processes for addressing potential misconduct that protect all participants. With this letter, we show our support for post-publication peer-review, to Elisabeth Bik and her work, and to all the whistleblowers that help maintain quality, honesty, integrity and trustworthiness of scientific advances.

References:

[1] Shen, H. (2020). Meet this super-spotter of duplicated images in science papers. *Nature*, *581*(7807), 132-136. doi: https://doi.org/10.1038/d41586-020-01363-z

[2] Marcus, A. A. (2020, May 22). Hydroxychloroquine-COVID-19 study did not meet publishing society's "expected standard." Retraction Watch. https://retractionwatch.com/2020/04/06/hydroxychlorine-covid-19-study-did-not-meet-publishing-societys-expected-standard/

[3] Rosendaal FR. Review of: "Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial" Gautret et al. 2020. https://doi.org/10.1016/j.ijantimicag.2020.105949. Int J Antimicrob Agents. 2020, https://doi.org/10.1016/j.ijantimicag.2020.106063

[4] Lonni Besançon, Nathan Peiffer-Smadja, Corentin Segalas, Haiting Jiang, Paola Masuzzo, Cooper Smout, Eric Billy, Maxime Deforet, Clémence a. Open Science Saves Lives: Lessons from the COVID-19 PandemicbioRxiv 2020.08.13.249847; doi: https://doi.org/10.1101/2020.08.13.249847

[5] Samer, C., Lacombe, K., & Calmy, A. (2021). Cyber harassment of female scientists will not be the new norm. The Lancet Infectious Diseases, 21(4), 457-458. Doi: https://doi.org/10.1016/s1473-3099(20)30944-0

[6] Peiffer-Smadja, N., Rebeaud, M. E., Guihur, A., Mahamat-Saleh, Y., & Fiolet, T. (2020). Hydroxychloroquine and COVID-19: a tale of populism and obscurantism. *The Lancet Infectious Diseases*. Doi: https://doi.org/10.1016/S1473-3099(20)30866-5

[7] Ektorp, E. (2020). Death threats after a trial on chloroquine for COVID-19. *The Lancet. Infectious Diseases*, *20*(6), 661. Doi: https://doi.org/10.1016/S1473-3099(20)30383-2

General Discussion

Perspectives

General Discussion

Three distinct parts are presented in this thesis. Together, they provide new insights into the role of an experimentalist scientist in today's society. This Thesis demonstrated the need to work more broadly on emerging public-health policies, in the context of threats from global warming and pandemics and it emphasized the urgent need to simplify and explain science to the lay audience, journalists, and policymakers. A new plague and a major challenge for the 21 century's scientists came out of the pandemic in the form of world-scale spreading of scientific misinformation, shameless claims of "alternative truths" by a minority of corrupt scientists, supported by predatory journals, populist politicians, divulgated unrestrictedly by the social media. It is necessary to develop ways to present scientific advances to the public, which generally lacks basic scientific knowledge, in a clear, honest, and unbiased manner. This requires undertaking, often risky shortcuts, and over-simplifications by using analogies from other fields better known to all, such as sociology. For example, in 1987, John Ellis (*1*) choose a social term - chaperone - to describe the presumed function of a new class of heat-shock proteins. This prompted the Goloubinoff laboratory to publish in 2007 a book chapter entitled "Molecular crime and cellular punishment" (*2*) in which, parallels were drawn between societal and cellular levels. Therein, the function of the molecular chaperones in stressed cells was compared to that of the police and judicial systems in human societies, in peace and stressful war times. Whereas some parallels (see table 1 in (*2*)) generated amused reactions, old-school scientists generally discarded such comparisons, regarding them as useless, plainly wrong, or even dangerous. Yet, it is a fact that most students, journalists, and politicians, while often lacking basic scientific knowledge, have a profound knowledge, which they have painfully acquired in childhood and teenagerhood, on how to behave in a society of humans. Thus, drawing parallels between molecular chaperones averting the formation of toxic protein aggregates in stressed cells, and police forces averting the formation of toxic criminals in stressed human societies, can simply explain very obscure scientific concepts to scientifically uninformed members of the society, and prevent them from seeking false unverified explanations from the internet.

This thesis is peculiar in the sense that beyond having to investigate molecular chaperone mechanisms in the context of a global pandemic, part of my academic endeavors had to be redirected to address questions of public health. While writing articles on COVID-19 and public health issues, and specific articles concerning the much-publicized treatments with hydroxychloroquine, several unexpected negative societal aspects became apparent: as a consequence of publishing scientific articles on subjects related to the COVID-19 pandemic, I became exposed to harassment, including death threats thru various social networks. How can a young scientist maintain an objective of producing high-

quality unbiased science in an atmosphere of scientific populism, while being threatened by non-scientists and publicly slandered by a minority of fame-seeking corrupted scientists and journalists? While these unexpected fallouts of my research took a negative toll on my mental health, they strengthened my resolution to become a scientist involved in the public debates denouncing the pernicious effects of predatory journals. When necessary, I participated in efforts to neutralize dangerous fake news, by debunking particular publications in predatory journals, preprints, and TV and YouTube interviews. Noticeably, although chaperone biochemistry is the most interesting core of my thesis, it appeared to be the least interesting for the lay audience. Nevertheless, my central interest in chaperone biochemistry is the direction I would like to further investigate in the future.

In addition to the several published articles and preprints presented in the section on chaperone protein biochemistry and the heat-shock response in plants and animals, several additional chaperone-related projects have been initiated during this thesis. Although some have generated important promising results, they have not yet matured into written articles and therefore are not included in this thesis.

How did Hsp70 become an Hsp110 co-disaggregase?

One of the most advanced projects, which has generated significant results and deserves to be included in my future perspectives for this work, focuses on the structure/function analysis of the co-disaggregase Sse1 (HSP110) and its collaboration with the main yeast disaggregase: the Ssa1-Sis1 machinery (HSP70-Hsp40). The project, which is a continuation of chapter 3 (*3*) describes various new hybrid Sse1-Ssa1 mutants that I have generated to address the mechanism by which Hsp110 and Hsp70 collaborate in the disaggregation process of stably misfolded proteins. It aims to understand what the biochemical properties of the first ancestor of all HSP110s were, while it stemmed from a branch of the HSP70s 2 billion years ago, during the formation of the first eukaryotes. This project aims at understanding how in early eukaryotisation, a *bona fide* HSP70 has evolved into being an Hsp110:

1) by apparently losing its ability to bind JDPs (*4*)
2) by apparently losing its ability to close its protein binding domain (*3*)
3) without losing its intrinsic ability to hydrolyze ATP (*3*)
4) without apparently losing its ability to bind proteins (*5*)
5) by gaining lengthy variable extensions in the protein binding domain (*6*)
6) by gaining the ability to transiently interact with the nucleotide-binding domain of HSP70s (*7, 8*).

The questions are how different a NEF HSP110 is, compared to FES1 (*9*) and to the M-domain of HSP104 (*10*), in its ability to accelerate ADP-release and ATP-rebinding in Hsp70, and how, at variance with Fes1, Hsp110 can effectively ameliorate the stand-alone pre-existing disaggregase activity of Hsp70-40?

How by undergoing all these minor changes did the first HSP110 become a new, effective co-disaggregase of the already-existing HSP70-JDP (and HSP70-HSP100-JDP) disaggregating machineries? The mechanism by which sub-stoichiometric amounts of HSP110s that in the yeast cytosol and ER, are ~7 x less abundant than the Hsp70s they act upon, iteratively bind and dissociate from HSP70s in an ATP-hydrolysis-dependent manner, and by doing so, somehow, greatly accelerate the intrinsic disaggregase activity of HSP70-JDP?

In an attempt to address these questions, we have generated several SSE1 mutants, by selecting specific regions in the Sse1 protein and swapping them with topologically equivalent regions from Ssa1, or by deleting them, specifically: 1) Linker mutants, in which Sse1 and Ssa1 linkers were swapped (PFKFED by DLLLLD) (*6*); 2) A loop out mutant, where an extension in the substrate-binding domain (SBD) of Sse1 was removed (*6*); 3) a helix mutant, where the helix of the lid of Sse1 was replaced with that from Ssa1; 4) a C-terminal ΔEEVD Ssa1 and 5) a C-terminal ΔEEVD Ssa1 with the linker of Sse1. We also used several pre-existing mutants, such as the Sse1 K69M (*11*) to test our hypothesis.

With these mutants, we have generated a large number of significant new results which by and large answered the above-listed questions, and we expect to write an article paper soon.

General perspectives

Chaperones are well known to carry out different roles in cellular proteostasis, such as promoting *de novo* folding, protein maturation, assembly and disassembly of protein complexes, protein disaggregation (*12*), and refolding and degradation (*13*). Are over-expressed chaperones such as during a priming heat shock, "good" or "bad" for cells? One answer is that over-expressed chaperones can be bad for metazoans with cancers since chaperones can prevent the beneficial apoptosis of cancer cells and can increase the resistance of cancer cells to various thermo-, radio- and chemotherapies (*14, 15*). For instance, in the case of cancerous tumors, members of the HSP70/40/90 and 110 (like HSP90α and β, HSC70, HOP, HP110, AHA1, and CDC37) are reshaped in stable multiprotein complexes, known as the epichaperome, facilitating tumor survival (*16*). Another answer is that over-expressed chaperones can be good for cells since they can prevent, and even potentially cure, degenerative diseases caused by toxic protein aggregations, as in the case of Parkinson's (*17, 18*) and Alzheimer's (*19, 20*) diseases, Amyotrophic Lateral Sclerosis (*21, 22*), Huntington's (*20, 23, 24*) and aging (*25*) in general.

A central question remains: How come rapidly developing and differentiating embryonic tissues (and Embryonic Stem Cells (ESC)), are immortal like cancer cells, and also massively over-express chaperones, like cancer cells (*26*), in general, can undergo effective beneficial apoptosis as part of embryonic organ development and ultimately turn into fully differentiated healthy organisms, rather than become unhealthy cancerous tissues, invading and ultimately killing the organisms (*27-32*)? The answer likely lies in the details: For optimal healthy protein homeostasis in cells, both the quantity

and quality of the chaperones counts. Chaperones and co-chaperones diversification into a larger array of orthologues acting on specific subsets of polypeptide substrates may make the whole difference between healthy developing embryos that overexpress chaperones and unhealthy invading cancers also over-expressing chaperones.

In this regard, the role of specific co-chaperones, in particular, JDPs (*33*) and NEFs (*34*) (nucleotide exchange factor, like HSP110) in the fine-tuning of the HSP70-core chaperone network could potentially differentiate between health-promoting chaperone-mediated protein homeostasis, as in embryo growth and differentiation and the curing of degenerative diseases caused by protein aggregations, in aging and protein misfolding diseases, compared to promoting unhealthy chaperone-mediated survival of cancer cells. This demonstrates the need for in-depth studies of these families of chaperones and co-chaperones in specific growing and non-growing tissues. Research is needed to address the specific structure and function of chaperones and co-chaperones, and the role of key amino acid changes that may alter overall behavior (*35*), as well as their beneficial or detrimental action in various diseases. Looking at the different tissues of the human body, the landscape formed by molecular chaperones reveals that there is an architecture of different layers of variable chaperones while others form an invariable core. These core chaperones are expressed uniformly across the tissues and are more abundant and important for cell survival than the variable ones (*36*). Recently, a paper (*37*) showed that an intrinsically disordered region present in metazoan HSP110 can suppress aggregate and amyloid formation. Several genetic disorders and diseases arise from mutations in JDPs, known as chaperonopathies (*38, 39*). The majority of these diseases are neurological, pointing to the importance of JDPs to the fitness of neurons (*40*).

From a less anthropocentric point of view, as discussed in Chapters 5 and 6, plants and their adaptation to climate change in the context of global warming, is another area where chaperones need to be studied and considered as part of the potential attempts to improve plant's resistance to heat-damages (*41*). The ongoing and necessary work of quantifying and characterizing chaperones orthologues and new co-chaperones, and particularly JDPs and specific types of NEFs like HSP110s, which exponentially diversified since the first eukaryotisation (*33, 42*), therefore seems of the greatest interest, both for future medical and plant agricultural research.

References

1. J. Ellis, Proteins as molecular chaperones. *Nature* **328**, 378-379 (1987).

2. M. P. Hinault, P. Goloubinoff, Molecular crime and cellular punishment: active detoxification of misfolded and aggregated proteins in the cell by the chaperone and protease networks. *Adv Exp Med Biol* **594**, 47-54 (2007).
3. V. Kumar *et al.*, Interdomain communication suppressing high intrinsic ATPase activity of Sse1 is essential for its co-disaggregase activity with Ssa1. *FEBS J* **287**, 671-694 (2020).
4. J. Shorter, The mammalian disaggregase machinery: Hsp110 synergizes with Hsp70 and Hsp40 to catalyze protein disaggregation and reactivation in a cell-free system. *PLoS One* **6**, e26319 (2011).
5. X. Xu *et al.*, Unique peptide substrate binding properties of 110-kDa heat-shock protein (Hsp110) determine its distinct chaperone activity. *The Journal of biological chemistry* **287**, 5661-5672 (2012).
6. Q. Liu, W. A. Hendrickson, Insights into Hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1. *Cell* **131**, 106-120 (2007).
7. S. Polier, Z. Dragovic, F. U. Hartl, A. Bracher, Structural basis for the cooperation of Hsp70 and Hsp110 chaperones in protein folding. *Cell* **133**, 1068-1079 (2008).
8. C. Andreasson, J. Fiaux, H. Rampelt, S. Druffel-Augustin, B. Bukau, Insights into the structural dynamics of the Hsp110-Hsp70 interaction reveal the mechanism for nucleotide exchange activity. *Proc Natl Acad Sci U S A* **105**, 16519-16524 (2008).
9. N. K. Gowda, G. Kandasamy, M. S. Froehlich, R. J. Dohmen, C. Andreasson, Hsp70 nucleotide exchange factor Fes1 is essential for ubiquitin-dependent degradation of misfolded cytosolic proteins. *Proc Natl Acad Sci U S A* **110**, 5975-5980 (2013).
10. E. C. Duran, C. L. Weaver, A. L. Lucius, Comparative Analysis of the Structure and Function of AAA+ Motors ClpA, ClpB, and Hsp104: Common Threads and Disparate Functions. *Front Mol Biosci* **4**, 54 (2017).
11. H. Raviol, H. Sadlish, F. Rodriguez, M. P. Mayer, B. Bukau, Chaperone network in the yeast cytosol: Hsp110 is revealed as an Hsp70 nucleotide exchange factor. *EMBO J* **25**, 2510-2518 (2006).
12. N. B. Nillegoda, A. S. Wentink, B. Bukau, Protein Disaggregation in Multicellular Organisms. *Trends Biochem Sci* **43**, 285-300 (2018).
13. B. Fauvet *et al.*, Bacterial Hsp90 Facilitates the Degradation of Aggregation-Prone Hsp70-Hsp40 Substrates. *Front Mol Biosci* **8**, 653073 (2021).
14. B. Wawrzynow, A. Zylicz, M. Zylicz, Chaperoning the guardian of the genome. The two-faced role of molecular chaperones in p53 tumor suppressor action. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1869**, 161-174 (2018).
15. P. Goloubinoff, Mechanisms of protein homeostasis in health, aging, and disease. *Swiss Med Wkly* **146**, w14306 (2016).
16. A. Rodina *et al.*, The epichaperome is an integrated chaperome network that facilitates tumour survival. *Nature* **538**, 397-401 (2016).
17. S. Hu *et al.*, Molecular chaperones and Parkinson's disease. *Neurobiology of Disease* **160**, 105527 (2021).
18. J. Tao *et al.*, Hsp70 chaperone blocks α-synuclein oligomer formation via a novel engagement mechanism. *Journal of Biological Chemistry* **296**, 100613 (2021).
19. R. E. Lackie *et al.*, The Hsp70/Hsp90 Chaperone Machinery in Neurodegenerative Diseases. *Front Neurosci* **11**, 254 (2017).
20. C. Zarouchlioti, D. A. Parfitt, W. Li, L. M. Gittings, M. E. Cheetham, DNAJ Proteins in neurodegeneration: essential and protective factors. *Philosophical Transactions of the Royal Society B: Biological Sciences* **373**, 20160534 (2018).
21. S. K. Park *et al.*, Overexpression of the essential Sis1 chaperone reduces TDP-43 effects on toxicity and proteolysis. *PLoS Genet* **13**, e1006805 (2017).
22. J. Shorter, Designer protein disaggregases to counter neurodegenerative disease. *Curr Opin Genet Dev* **44**, 1-8 (2017).

23. Y. Kuo, S. Ren, U. Lao, B. A. Edgar, T. Wang, Suppression of polyglutamine protein toxicity by co-expression of a heat-shock protein 40 and a heat-shock protein 110. *Cell Death & Disease* **4**, e833-e833 (2013).
24. K. Rozales *et al.*, Differential roles for DNAJ isoforms in HTT-polyQ and FUS aggregation modulation revealed by chaperone screens. *Nature Communications* **13**, 516 (2022).
25. M. S. Hipp, P. Kasturi, F. U. Hartl, The proteostasis network and its decline in ageing. *Nature Reviews Molecular Cell Biology* **20**, 421-435 (2019).
26. H. J. Lee, R. Gutierrez-Garcia, D. Vilchez, Embryonic stem cells: a novel paradigm to study proteostasis? *The FEBS Journal* **284**, 391-398 (2017).
27. M. A. Vostakolaei *et al.*, Hsp70 in cancer: A double agent in the battle between survival and death. *Journal of Cellular Physiology* **236**, 3420-3444 (2021).
28. L. Whitesell, S. L. Lindquist, HSP90 and the chaperoning of cancer. *Nature Reviews Cancer* **5**, 761-772 (2005).
29. A. Mitra, L. A. Shevde, R. S. Samant, Multi-faceted role of HSP40 in cancer. *Clinical & Experimental Metastasis* **26**, 559-567 (2009).
30. J. Wu *et al.*, Heat Shock Proteins and Cancer. *Trends in Pharmacological Sciences* **38**, 226-256 (2017).
31. D. D. Mosser, R. I. Morimoto, Molecular chaperones and the stress of oncogenesis. *Oncogene* **23**, 2907-2918 (2004).
32. C. Jolly, R. I. Morimoto, Role of the Heat Shock Response and Molecular Chaperones in Oncogenesis and Cell Death. *JNCI: Journal of the National Cancer Institute* **92**, 1564-1572 (2000).
33. H. H. Kampinga *et al.*, Function, evolution, and structure of J-domain proteins. *Cell Stress Chaperones* **24**, 7-15 (2019).
34. H. Rampelt, M. P. Mayer, B. Bukau, Nucleotide Exchange Factors for Hsp70 Chaperones. *Methods Mol Biol* **1709**, 179-188 (2018).
35. H. Y. Yu *et al.*, Roles of intramolecular and intermolecular interactions in functional regulation of the Hsp70 J-protein co-chaperone Sis1. *J Mol Biol* **427**, 1632-1643 (2015).
36. N. Shemesh *et al.*, The landscape of molecular chaperones across human tissues reveals a layered architecture of core and variable chaperones. *Nature Communications* **12**, 2180 (2021).
37. U. M. Yakubu, K. A. Morano, Suppression of aggregate and amyloid formation by a novel intrinsically disordered region in metazoan Hsp110 chaperones. *Journal of Biological Chemistry* **296**, (2021).
38. A. J. L. Macario, E. Conway de Macario, Chaperonopathies and chaperonotherapy. *FEBS Letters* **581**, 3681-3688 (2007).
39. K. Nisaa, A. Ben-Zvi, Chaperone networks are shaped by cellular differentiation and identity. *Trends Cell Biol*, (2021).
40. H. L. Smith, W. Li, M. E. Cheetham, Molecular chaperones and neuronal proteostasis. *Seminars in Cell & Developmental Biology* **40**, 142-152 (2015).
41. P. Jacob, H. Hirt, A. Bendahmane, The heat-shock protein/chaperone network and multiple stress resistance. *Plant Biotechnol J* **15**, 405-414 (2017).
42. D. P. Easton, Y. Kaneko, J. R. Subjeck, The hsp110 and Grp1 70 stress proteins: newly recognized relatives of the Hsp70s. *Cell stress & chaperones* **5**, 276-290 (2000).