Title: Discovery of catalases in members of the Chlamydiales order.
Authors: Rusconi B, Greub G
Journal: Journal of bacteriology
Year: 2013 Aug
Volume: 195
Issue: 16
Pages: 3543-51
DOI: 10.1128/JB.00563-13

Discovery of Catalases in Members of the *Chlamydiales* Order

Brigida Rusconi, Gilbert Greub*

Institute of Microbiology, University of Lausanne and University Hospital Center, Lausanne, 1011, Switzerland

*Corresponding author:

Mailing address: Gilbert Greub, Institute of Microbiology, University of Lausanne and University Hospital Center, Bugnon 48, 1011 Lausanne, Switzerland.

Phone: +41 21 314 49 79 Fax: +41 21 314 40 60

E-mail: Gilbert.Greub@chuv.ch

Running Title: Catalases of the *Chlamydiales* Order

Abstract

Catalase is an important virulence factor for survival in macrophages and other phagocytic cells. In *Chlamydiaceae* no catalase was described so far. With the sequencing and annotation of the full genomes of *Chlamydia*-related bacteria, presence of different catalase-encoding genes has been documented. However, the distribution in the *Chlamydiales* order and the functionality of these catalases remain unknown. Phylogeny of chlamydial catalases was inferred using Mr. Bayes, Maximum Likelihood, and Maximum Parsimony algorithms, allowing the description of three clade 3 and two clade 2 catalases. Only monofunctional catalases were found (no catalase-peroxidase nor Mn-catalase). All presented a conserved catalytic domain and tertiary structure. Enzymatic activity of cloned chlamydial catalases was assessed by measuring hydrogen peroxide degradation. The catalases are enzymatically active with different efficiencies. The catalase of *Parachlamydia acanthamoebae* is the least efficient of all (its catalytic activity was 2 logs lower than *Pseudomonas aeruginosa*). Based on the phylogenetic analysis, we hypothesize that an ancestral class 2 catalase was probably present in the common ancestor of all current *Chlamydiales* but was only retained in *Criblamydia sequanensis* and *Neochlamydia hartmannellae*. The catalases of class 3, present in *Estrella lausannensis* and *Parachlamydia acanthamoebae*, were probably acquired by lateral gene transfer from *Rhizobiales*, whereas for *Waddlia chondrophila* likely originated from *Legionellales* or *Actinomycetales.* The acquisition of catalases at several occasions in the *Chlamydiales* suggests the importance of this enzyme for the bacteria in their host environment.

Catalases / *Chlamydiales* / Catalytic activity / Phylogeny

Introduction

Macrophages are often a target of intracellular bacteria (1). The bacteria can be obligate intracellular bacteria like *Rickettsia* spp. (2) or facultative intracellular such as *Legionella pneumophila* (3), *Brucella abortus* (4), and *Mycobacterium tuberculosis* (5). Among members of the *Chlamydiales* order*, Chlamydia trachomatis* (serovarK9) do not resist macrophage microbicidal effectors (6) whereas *Waddlia chondrophila* are able to replicate very efficiently in macrophages (7, 8) and *Parachlamydia acanthamoebae* is able to replicate to a lower extend inducing rapidly the apoptotic death of the macrophage (9-11).

The ability to grow in a professional phagocyte offers several advantages to the invading pathogen. First, macrophages represent an interesting cell target due to their presence in almost all tissues. Moreover, infecting macrophages will give the bacteria an opportunity to hamper macrophage activation, therefore delaying the development of an effective cytotoxic T cell response. It is therefore crucial to determine, which factors define the resistance of macrophages to certain bacteria, especially when macrophages exhibited a different permissivity to bacteria belonging to the same order.

One of the first lines of defense of macrophages is the rapid degradation of the bacteria by acidic pH, lysosomal hydrolases and various other microbicidal effectors, including reactive oxygen species (ROS) produced by the transmembranous NADPH oxidase complex (NOX2). Catalases are bacterial enzymes that degrade ROS (H_2O_2). Catalases belong to a very diverse functional class of proteins that can be classified in four main groups: first the heme-containing monofunctional catalases, second the heme-containing bifunctional catalase-peroxidases, third non-heme catalases, and fourth unclassified catalases (12). Mutations in certain components of the NOX2 complex cause an immune disease called chronic granulomatous disease (CGD) (13). This genetic disease is associated with recurrent

bacterial and fungal infections. Interestingly, pathogens that infect CGD patients are expressing a ROS degrading enzyme catalase (14), belonging to the heme-containing mono- or bifunctional catalases group.

Catalase-peroxidases are thought to have been acquired through lateral gene transfer from archea (15), while monofunctional heme-containing catalases are believed to be very ancient. The latter can be further subdivided into three main clades.

During annotation of the genome of *Waddlia chondrophila* in our group, a catalase encoding gene was identified (16). Since none of the members of the *Chlamydiaceae* family encode for a catalase we investigated the presence of these genes in other members of the *Chlamydiales* order, including *Waddliaceae*, *Parachlamydiaceae*, *Simkaniaceae*, and *Criblamydiaceae* families. The functional properties and evolutionary history of identified catalases was then assessed.

Results

Genetic and phylogenetic analyses of chlamydial catalases

A prototypic small subunit catalase from *Staphylococcus aureus* was used to perform a BLASTP on all *Chlamydiales* genomes available including unpublished in-house ongoing genomes of *Criblamydia sequanensis*, *Estrella lausannensis*, *Protochlamydia naegleriophila,* and *Neochlamydia hartmannellae*. Catalases were found in the two members of the *Criblamydiaceae* family (*C. sequanensis* (KatE) and *E. lausannensis* (KatA)) and in two members of the *Parachlamydiaceae* family (*P. acanthamoebae* (KatA) and *N. hartmannellae* (KatE)), as well as the previously annotated KatA of *Waddlia chondrophila*. When using the catalases sequences of all these members of the *Chlamydiales* order to perform additional

BLASTP and PSIBLAST searches, no additional catalases were identified in any of the available genomes of *Chlamydiaceae*, *Simkaniaceae* or *Protochlamydia* spp..

The amino acid sequence identity of the *Chlamydia*-related bacteria proteins compared to *Staphylococcus aureus* and to the prototypical small subunit *Pseudomonas aeruginosa* KatA ranged between 35 to 60 % and 38 to 64%, respectively. When compared to catalases from clade 2, the identity of KatE of *C. sequanensis* reached 45 to 68%, except with KatE of *N. hartmanellae* (75.9%) (Figure S1A). Identity among members of the clade 3 reached around 54% of sequence similarity. Between KatA of *E. lausannensis* and KatA of *P. acanthamoebae* (77.6%) and KatA of *W. chondrophila* and *L. longbeacheae* (KatA) (77%) (Figure S1B). These differences in identity led us to perform a detailed phylogenetic analysis of the chlamydial catalase proteins.

In addition to reference sequences used by Klotz et al., 2003 and Zamocky et al., 2010 we added the monofunctional catalases found in other amoeba-resisting bacteria (*Legionellales*, *Mycobacteria, Bradyrhizobiales*) (12, 17, 18). Moreover, catalases of amoebal origin were added to determine a possible lateral gene transfer from the host. Sequences were directly retrieved by BLASTP from amoebal genomes available on NCBI (*Dictyosteliida, Acanthamoeba, Tetrahymena, Naegleria*). For the amoebae belonging to the *Jakobidae* family the sequences were reconstructed from expressed sequence tags (ESTs) retrieved by tBLASTn, since no genome was available at the time. Other amoebal sequences found in the EST database did not cover the whole protein sequence and were therfore excluded. Moreover, the three highest BLAST hits for each chlamydial catalase was added to the alignment as well. Phylogeny of the bacterial catalases was performed with maximum likelihood (ML), maximum parsimony (MP), and MrBayes (MB), all resulting in clear separation of the three catalases categories (i.e. clade 1, 2, and 3) (Figure1 , Figure S2).

Sampled trees from MP analyses showed constant posterior probabilities, therefore excluding any problem with convergence (Figure S1C).

Tree topology was conserved between ML and MB in clade1 and clade2, although with lower confidence in some nodes of ML (Figure S2A). For MP the *Bacillales* branch of clade 2 catalases clustered with the *Mycobacteria* instead of branching off after *E. coli* like in the two other analyses. However, phylogenetic relationship of *Bacillales* could not be accurately determined since the bootstrap support was lower than 0.5 (Figure S2B). The catalases present in *C. sequanensis* and *N. hartmannellae* were assigned to the clade 2 catalases that comprise large tetrameric catalases. As previously observed, the negibacteria catalases are divided in 2 branches (12). KatE of *C. sequanensis* and KatE of *N. hartmanellae* clustered with the branch comprising the posibacterial lineage (Figure 1A).

Amoebal clade 3 catalases all clustered together with a deep rooting compared to other bacterial clade 3 catalases. Between the three methods there were small differences in the node placement. This was mostly due to the inability of MP and ML to infer the phylogenetic relationship of a few sequences as can be seen by the low bootstrap values (Figure S2A-B). The chlamydial clade 3 catalases separate into two distinct branches in all three methods. Within the *W. chondrophila* KatA branch five nodes hade very low values for both MP and ML (Figure S2A-B). However, with MB the phylogenetic relationship was determined with high posterior probabilites (Figure 1A). *E. lausannensis* and *P. acanthamoebae* cluster in the same branch, but not together.

To better understand the origin of these different catalases in *Chlamydiales*, we then analyzed their genetic environment and we observed seven transposase and integrase elements immediately upstream of *katA* of *W. chondrophila* (Figure 1B) as well as seven

transposase and integrase elements 15 genes downstream of waddlial *katA*. None of the genes located between these transposases exhibited a BLASTP hit with a *L. longbeacheae* gene; precluding confirmation that *katA* was acquired from *L. longbeacheae*. However, the presence of these mobile elements is supporting the hypothesis of a horizontal acquisition of *katA* by *W. chondrophila*. Around the other catalases, no transposases were identified. Moreover, the genetic environment of the KatA encoding-gene from *E. lausannensis* and *P. acanthamoebae* was not conserved, despite these proteins exhibited a sequence similarity of 78% and clustered together in phylogenetic trees. However, the synteny is low between members of different families of the *Chlamydiales* order (19), explaining the absence of conserved genetic environment despite a likely common origin for *E. lausannensis* and *P. acanthamoebae katA*.

Conservation of domains and motifs

The heme binding sites are conserved among all clades of monofunctional catalases (17). We therefore analyzed the sequences of the chlamydial catalases to determine “motif conservation”. Differences in prevalence of a given amino acid within the motifs were observed depending on the catalase clade. Amino acid variants that were mostly found in clade 2 catalases are highlighted in bold letters in Figure 2. The proximal heme-binding site was very conserved and was detected in all the chlamydial catalases (Figure 2E). The active sites in both sites were conserved (Figure 2D-E, red boxes). For the distal heme binding site (Figure 2D), analysis by PROSITE did not give any hit for KatA of *P. acanthamoebae* and KatA of *E. lausannensis*. This was due to two mutations at positions F44Q and S59W, respectively (Figure 2D). The phenylalanine is very conserved and its substitution with glutamine could strongly affect the binding site, especially its topology. The second

substitution is also quite problematic, since tryptophan is more reactive and bulkier than serine.

Moreover, the sequences were analyzed for the conservation of surface epitopes determined in a previous study (20). Since the consensus sequence for these epitopes was only determined with four sequences from small subunit catalases we further confirmed them by analyzing the MUSCLE alignment of all bacterial catalases obtained from http://peroxibase.toulouse.inra.fr/ (Figure S3). All three epitopes were conserved in all clades of catalases, with minor differences between small and large subunit catalases, except for the second epitope. The latter one had a more conserved sequence in clade 2 catalases than clade 3. Moreover, the consensus sequences differed substantially between the two clades (Figure S3, orange, green, and blue boxes). The first epitope (Figure 2A) was conserved in all *Chlamydiales* catalases. The last surface epitope (Figure 2C) was well conserved in all *Chlamydiae* except KatA of *W. chondrophila* that presented two mutations in the more conserved sites (N304D, F306H). Moreover, at position 305 KatA of *W. chondrophila* had a phenylalanine, like large subunit catalases.

Enzymatic activity

Since *Chlamydiales* present two developmental stages, the activity of the catalase would be required during the early steps of infection, when the bacteria are still in their elementary body form. Amount of bacteria persent in each vial was determined by qPCR with the pan-*Chlamydiales* primers. 10^7 purified elementary bodies of *P. acanthamoebae*, *E. lausannensis*, *C. sequanensis*, *N. hartmannellae* and *W. chondrophila* all encoding for a catalase were exposed to 0.2M of hydrogen peroxide (Figure 3A). All of them produced oxygen bubbles to a similar extent as the *P. aeruginosa* positive control. However, *N. hartmannellae* and *W.*

chondrophila displayed a reduced oxygen bubble formation. A mock negative control of the amoebal co-culture was also tested and proved to be negative. , *Simkania negevensis* was used as second negative control was used since it does not encode a catalase and therefore, as expected, only produced one small oxygen bubble.. This could be due to a slight contamination of purified elementary bodies by *Acanthamoebae castellanii* catalase. However, the contamination is insignificant compared to the catalase activity displayed by the different bacteria. The clade 3 catalase of *A. castellanii* was also blocked by 0.1M azide. The activity of the catalase was blocked in all bacteria with 0.1M azide, a non-competitive inhibitor of catalases (21). Only when the concentration was reduced to 1mM did we observed again oxygen formation for *P. aeruginosa* bacteria, while *E. lausannesis* showed no catalase activity even at 0.1mM azide (3A).

Once we observed this oxygen production, we then determined the catalytic activity of these catalases at various pH in order to define whether the activity is correlating with host range and phagolysosome survival. As a positive control we cloned KatA from *P. aeruginosa.* To control for eventual contamination by the *Escherichia coli* catalase we used as a negative control the Hsp60 from *W. chondrophila* purified in the same conditions as the catalases. The enzymatic activity was assessed by measuring the absorption of H_2O_2. For KatA of *P. acanthamoebae* and KatA of *W. chondrophila* we had to increase the amount of protein from 200ng/ml to 2μg/ml and 800ng/ml respectively to detect a sustained degradation of hydrogen peroxide (Figure 3B). The catalytic units (Figure 3C) derived from the assay differed quite substantially depending on the species. The KatA of *P. acanthamoebae* had the lowest ability to degrade hydrogen peroxide. KatA of *E. lausannensis* was more efficient at pH 7.0 whereas KatA of *W. chondrophila* appeared to be less sensitive to changes in pH.

194 *In silico modeling of chlamydial catalases*

195 Several crystal structures exist for bacterial clade 2 and 3 catalases. We therefore used this
196 information to build an *in silico* model of the chlamydial catalases using 3DJIGSAW (22-24)
197 and SWISS-MODEL Workspace (25-28). The tetrameric structure of the clade 3 catalases
198 was built on the crystal structure of *Enterobacter faecalis* (1SI8) (Figure S4A) for all three
199 chlamydial catalases. The organization of the tetramer was retained for all chlamydial small
200 subunit catalases (Figure S4A). For the chlamydial large subunit catalases the crystal
201 structure of *E. coli* HPII (1IPH) was used (Figure S4B). The organization of the tetramer of
202 KatE of *N. hartmanellae* catalase was more similar to HPII from *E. coli* than KatE of *C.*
203 *sequanensis* (Figure S4B).

204 To determine the conservation of the tertiary structure of the tetramer subunits the RMS
205 values for each chlamydial catalase were determined. Overall, the tertiary structures were
206 well conserved with only limited deviations in the loop regions (Figure 4A-C, Figure S4C).
207 KatE of *C. sequanensis* presented one helix with more than 5Å deviation from the crystal
208 structure. However, this part of the subunit is neither involved in the contact site of the
209 tetramer nor in the catalytic domain (Figure S4C left). Moreover, the quality of the tertiary
210 structure was further determined by looking at residues with bad backbone conformation and
211 side chains without hydrogen bonds. There were 24 residues in the C-terminal with clashes in
212 the backbone. Of these 24, fifteen are mutated compared to *E. coli* HPII and four were not
213 aligned. The remaining residues were close to mutated residues that caused the distortion of
214 the backbone (Figure S5A). Moreover, in the C-terminal of *KatE of C. sequanensis* nine
215 residues were mutated to prolines that strongly influenced the backbone orientation. Only
216 three prolines (P383, P522, P543) were not in an optimal conformation. As for *KatE of C.*

sequanensis most of the residues (11) that clashed with the backbone were located in the C-terminal part of the protein.

Chlamydial small subunit catalases had no side chains that lacked hydrogen bonds. In KatA of *P. acanthamoebae*, KatA of *E. lausannensis*, and KatA of *W. chondrophila* there were only five residues of which 3 mutated compared to *Helicobacter pylori* (2IQF) catalase sequence (Figure S4B). They were not all located at the same residues. Moreover, KatA of *P. acanthamoebae* catalase had one proline mutation (Q379P), KatA of *W. chondrophila* two (V6P, V55P), and KatA of *E. lausannensis* one (V6P) that affected the backbone organization.

The water molecules for both catalases were added to the catalytic site residues according to Diaz *et al*., 2012 (29) (Figure 4A). For the small subunit catalase the residues orientation is almost completely conserved. For KatA of *W. chondrophila* R53 the side chain is in another orientation that replaces the hydrogen bond with the heme with the water molecule (W2) (Figure 4B). The rotation of the N128 in KatA of *P. acanthamoebae* and KatA of *E. lausannensis* caused a weakening of the interaction with the water molecule (3.2 Å) of the nitrogen of the histidine ring and the oxygen of the asparagine side chain (Figure 4C). Moreover, the interaction most likely occurs with the oxygen and not the nitrogen of the asparagine side chain like in *H. pylori*. In KatA of *P. acanthamoebae* the slight tilt of the H53 further increased the distance to the water molecule weakening the interaction. All other hydrogen bonds were in the range of moderate interactions (2.5-3.2 Å). The flipping of the H198 side chain in KatA of *P. acanthamoebae* disrupted the hydrogen bond chain between the catalytic residues Y338, R334, H198, and N328. Indeed, the distances passing 4 Å did not allow for a hydrogen bond formation. This could affect the stability of the catalytic site. For chlamydial clade 2 catalases, the catalytic site was also conserved and all the hydrogen

bonds could be formed (Figures S4D). Only Y415 in *KatE of C. sequanensis* had a 14° shift of the aromatic ring lengthening the distance between the tyrosine and the iron of the heme d from 2.11 Å to 2.35 Å. However, the tyrosine was still close enough to interact with the iron.

Discussion

In this work, we identified catalase-encoding genes in the genomes of five different chlamydia-related bacteria and we could demonstrate their enzymatic activity and a significant conservation of their active sites. These catalases belonged to clade 2 (KatE of *N. hartmanellae* and KatE of *C. sequanensis*), and clade 3 (KatA of *W. chondrophila*, KatA of *E. lausannensis*, and KatA of *P. acanthamoebae*) catalases. Moreover, waddlial catalase was branching far from the other two clade 3 catalases, respectively encoded by *P. acanthamoebae* and *E. lausannensis*, suggesting a complex evolutionary history of these proteins. The limited amount of sequences available for chlamydial catalases did not allow us to determine definitely which one is most likely to be the ancestral *Chlamydiales* catalase (Figure 5A). However, according to phylogenetic analyses performed by Klotz *et al.,* 2003 the first catalase was likely a large subunit catalase that underwent gene duplication with loss of the C-terminal domain and partial loss at the N-terminal domain (12). It is thus likely that the common ancestor of the *Chlamydiales* order had a large subunit catalase that was lost by most of the species. Moreover, lateral gene transfer of large subunit catalases was so far only observed from bacteria to archea and not among bacteria (12).

The clade 3 catalase shared by *E. lausannensis* and *P. acanthamoebae* was probably acquired later, since this enzymatic clade has originated later in posibacteria, and was then distributed on other species by lateral gene transfer (12). This transfer in *Chlamydiales* has likely taken place after the separation of the *Criblamydiaceae / Parachlamydiaceae* ancestor

from the *Waddliaceae,* since they encode for clade 3 catalases of different origins (Figure 5A). Since *E. lausannensis* and *P. acanthamoebae* were clustering in the same branch, but not together it is more likely that they both received the catalase from a *Rhizobiales* at different occasion than an internal transfer from on to the other after lateral gene transfer from a non chlamydial bacteria (Figure 5B). The transfer of the catalase present in *E. lausannensis* and *P. acanthamoebae* probably occurred from a *Rhizobiales* precursor, since the majority of the first 30 BLASTP hits belong to this order. For *W. chondrophila* the origin of the lateral gene transfer is probably an *Actinomycetales* or *Legionellales* rather than the *Planctomycetes* best BLAST hits. This hypothesis is based on the fact that *Legionellales* and several members of the *Actinomycetales* are able to survive within amoebae (18). Interestingly, host-associated genera like *Brucella* and *Bordetella* lost the large subunit catalase upon acquisition of the clade 3 catalase. Such a loss could explain the absence of the clade 2 catalase in the chlamydia-related bacteria that encode a clade 3 catalase. As proposed by Klotz et al., this might be due to the presence of clade 3 catalases in eukaryotes, that are forcing the bacteria to adapt to the physiological selective pressure present within the host (12). When looking at the most common human bacterial pathogens the majority encode for a catalase or a catalase-peroxidase. As seen for other anaerobic bacteria *Streptococcaceae* do not possess a catalase, since their exposure to reactive oxygen species is much lower. Only *Mycoplasma, Rickettsia rickettsii, Borrelia burgdorferi, Treponema pallidum* and *Chlamydiaceae*. These bacteria all underwent strong genome size reduction and adaptation to a specific niche, making the presence of a catalase redundant. Since the reservoir of *Chlamydia*-related bacteria is much broader these bacteria need a more diverse pannel of virulence factors to adapt to each host. For *W. chondrophila* and *E. lausannensis* the

presence of a catalase could indeed prove useful during the infection of humans, when encountering professional phagocytes.

Despite significant differences at the amino acid sequence level, tertiary structure modeling showed significant conservation especially around the catalytic sites. The histidine 198 of KatA of *P. acanthamoebae* probably destabilizes the interaction between the heme and the catalytic tyrosine due to a disruption of the hydrogen bond chain. This correlates with the reduced enzymatic activity of KatA of *P. acanthamoebae* observed *in vitro.* Such lower activity of KatA of *P. acanthamoebae* compared to KatA of *W. chondrophila* may impact the ability of these bacteria to grow within macrophages. Indeed *W. chondrophila* has a productive and rapid growth cycle in macrophages (7, 8), whereas *P. acanthamoebae* grow only poorly in macrophages(10, 11). However, the presence of a catalase is not essential for growth in free-living amoebae, since *Protochlamydia* spp. do not encode for any catalases and are nevertheless growing successfully in these protists (30, 31). As revealed by phylogenetic analyses, the horizontal transfer of clade 3 catalases probably occurred on several occasions and after the divergence of the different families within the *Chlamydia*-related bacterial branch. Moreover, within the *Parachlamydiaceae,* the Protochlamydia spp. that do not encode for any catalase diverged prior to acquisition of catalases by *Parachlamydia* spp..

Materials and Methods

Phylogenetic analysis

BLASTP was performed on the NCBI BLAST platform (32). Protein sequences used for search were retrieved from Pubmed and Uniprot. Sequences were aligned with MUSCLE (33)

312 in Geneious v6. Domains were determined with PROSITE (34, 35). Consensus sequences
313 display was created with WebLogo (36).
314 Prior to phylogenetic analyses the amino acid substitution model was assessed with ProtTest
315 v.3.2 and modelgenerator (37). Phylogenetic analysis was done with the following models:
316 Maximum Likelihood was performed with the PhyML 3.0 platform using the LG substitution
317 matrix, with invariant gamma distribution (4 categories) and 500 bootstraps (38). The
318 Maximum Parsimony tree was obtained using the Close-Neighbor-Interchange algorithm (39)
319 with search level 1 in which the initial trees were obtained with the random addition of
320 sequences (10 replicates) and 500 bootstraps. Evolutionary analyses were conducted in
321 MEGA5 (40). A third tree was constructed with MrBayes using LG and 1'000'000 generations
322 with invariant gamma distribution (4 categories) and frequency (41). Quality of the bayesian
323 phylogentic tree was assessed by AWTY (42). Trees were visualized with FigTree v1.3.1.
324 Newly submitted chlamydial proteins and all catalase sequences used are listed in appendix
325 tables A1 and A2.
326 *Protein modeling*
327 Structure was modeled with 3DJIGSAW (22-24) and SWISS-MODEL Workspace (25-28).
328 Quality of the models obtained with SWISS-MODEL Workspace were assessed with
329 QMEAN4 (43, 44). The model was then further analyzed in Deepview (26). Other clade 3
330 catalases from *P. mirabilis* (PDB: 1M85) (45) and *E. faecalis* (PDB: 1SI8) (46) were used as
331 controls. RMS, threading energy, and conformation clashes were determined with Deepview
332 (26). For clade 2 catalases *E. coli* HPII PDB: 1GGE (47) and PDB: 1IPH (48) were used for
333 modeling.
334
335 *Catalase cloning and expression*

Genomic DNA of *W. chondrophila*, *P. acanthamoebae*, *E. lausannensis*, and *P. aeruginosa* ATCC 27853 have been extracted with Promega Wizard SV genomic DNA kit (Promega Corporation, Madison, USA) following manufacturer's instructions. The following primers were used *W. chondrophila* catalase: Wch_KatA_F: 5'-CAC CAA AAG AGA TCG CCC AGC CACT-3', Wch_KatA_R: 5'-TTA GGA GGA ACA GCC TGC TGC TTT TTT GAT TCGC-3', *P. acanthamoebae* Pac_katA_F: 5'-CAC CGA GAA TAA AGA TAC GCT GAC CAC CA-3', Pac_KatA_R: 5'-TCA GTT TTT ACG AGA GAG TAG GGCA-3', E. lausannensis Ela_KatA_F: 5'-CAC CAC AGA TAA GCC CCC CCT AT-3', Ela_KatA_R: 5'-CTA TTT TTT TCT CTT ATC CAG CGC TT-3', P. aeruginosa Pae_KatA_F: 5'-CAC CGA AGA GAA GAC CCG CCT-3', Pae_KatA_R: 5'-TCA GTC CAG CTT CAG GCC GAG-3'. The following PCR conditions were used for gene amplifications: 98°C 30s, 35 cycles of 98°C 10s, 68°C 30s, 72°C 90s, additional extension of 10min at 72°C. Annealing temperature was lowered to 61°C for *E. lausannensis* and *P. acanthamoebae* catalase and to 60°C for *P. aeruginosa.* High-fidelity Phusion (Fermentas) amplified genes were cloned in pET200 TOPO vector (Invitrogen). Expression of protein was done in BL21Star (Invitrogen). Purification of protein was performed in non-denaturing conditions with MagneHis (Promega). Purified proteins were concentrated in PBS with Amicon columns according to manufacturer's instructions (Milipore). Protein concentration was determined by fluorescence with Qubit Protein Assay following manufacturer's instructions (Invitrogen).

Enzymatic activity

Catalase activity was assessed with minor modifications by decrease in absorption of H_2O_2 at 240nm due to degradation as described in Li & Schellhorn, 2007 (49). Briefly, hydrogen peroxide was used at a concentration of 10 mM. Purified proteins were used at a

concentration of 50ng, 200ng, or 500ng depending on the enzymatic activity. PureGrade 96-well plates (Brand, England) were used in microplate reader SynergyH1 (BioTek, USA) at 240nm.

Elementary bodies of *P. acanthamoebae*, *E. lausannensis*, *W. chondrophila*, and *S. negevensis* grown in amoebae were purified according to Bertelli et al., 2010 (16), with a minor modification. Prior to gastrographin gradient ultracentrifugation the bacteria were treated for 1 hour at 37°C with Dnase to reduce amoebal protein contamination. An aliquot of frozen bacteria were then washed twice with PBS before exposure to hydrogen peroxide. Bacterial concentration was determined by qPCR with the panChlamydiales primer pair and probe (50). *P. aeruginosa* (ATCC 27853) over night culture was washed twice with PBS and pelleted for 10min at 8000g at room temperature before exposure to hydrogen peroxide. Bacterial concentration was assessed by CFU. *A. castellanii* liquid axenic culture were counted in a kovar chamber and 10^6 amoebae were pelleted at 500g for 5min and washed twice with PBS prior to exposure to hydrogen peroxide.

Acknowledgements

This work was supported by the Swiss National Science Foundation (project n° PDFMP3-127302). Brigida Rusconi is supported by the Swiss National Science Foundation within the PRODOC program "Infection and Immunity". The authors declare that they have no conflict of interest.

We thank C. Bertelli and T. Pillonel for helpful comments.

References

1. **Duclos, S., and M. Desjardins.** 2000. Subversion of a young phagosome: the survival strategies of intracellular pathogens. Cell Microbiol **2:**365-77.
2. **Radulovic, S., P. W. Price, M. S. Beier, J. Gaywee, J. A. Macaluso, and A. Azad.** 2002. Rickettsia-macrophage interactions: host cell responses to Rickettsia akari and Rickettsia typhi. Infect Immun **70:**2576-82.
3. **Brand, B. C., A. B. Sadosky, and H. A. Shuman.** 1994. The Legionella pneumophila icm locus: a set of genes required for intracellular multiplication in human macrophages. Mol Microbiol **14:**797-808.
4. **Harmon, B. G., L. G. Adams, and M. Frey.** 1988. Survival of rough and smooth strains of Brucella abortus in bovine mammary gland macrophages. Am J Vet Res **49:**1092-7.
5. **Maxwell, K. W., and S. Marcus.** 1968. Phagocytosis and intracellular fate of Mycobacterium tuberculosis: in vitro studies with guinea pig peritoneal and alveolar mononuclear phagocytes. J Immunol **101:**176-82.
6. **Schmitz, E., E. Nettelnbreker, H. Zeidler, M. Hammer, E. Manor, and J. Wollenhaupt.** 1993. Intracellular persistence of chlamydial major outer-membrane protein, lipopolysaccharide and ribosomal RNA after non-productive infection of human monocytes with Chlamydia trachomatis serovar K. J Med Microbiol **38:**278-85.
7. **Goy, G., A. Croxatto, and G. Greub.** 2008. Waddlia chondrophila enters and multiplies within human macrophages. Microbes Infect **10:**556-62.
8. **Croxatto, A., and G. Greub.** 2010. Early intracellular trafficking of Waddlia chondrophila in human macrophages. Microbiology **156:**340-355.
9. **Greub, G., J.-L. Mege, and D. Raoult.** 2003. Parachlamydia acanthamoebae enters and multiplies within human macrophages and induces their apoptosis [corrected]. Infect Immun **71:**5979-85.
10. **Roger, T., N. Casson, A. Croxatto, J. M. Entenza, M. Pusztaszeri, S. Akira, M. K. Reymond, D. Le Roy, T. Calandra, and G. Greub.** 2010. Role of MyD88 and Toll-Like Receptor (TLR) 2 and TLR4 in the Sensing of Parachlamydia acanthamoebae. Infect Immun **78:**5195-5201.
11. **Greub, G., B. Desnues, D. Raoult, and J.-L. Mege.** 2005. Lack of microbicidal response in human macrophages infected with Parachlamydia acanthamoebae. Microbes Infect **7:**714-9.
12. **Klotz, M. G., and P. C. Loewen.** 2003. The molecular evolution of catalatic hydroperoxidases: evidence for multiple lateral transfer of genes between prokaryota and from bacteria into eukaryota. Mol Biol Evol **20:**1098-112.
13. **Curnutte, J. T.** 1993. Chronic granulomatous disease: the solving of a clinical riddle at the molecular level. Clin Immunol Immunopathol **67:**S2-15.
14. **Soler-Palacin, P., C. Margareto, P. Llobet, O. Asensio, M. Hernandez, I. Caragol, and T. Espanol.** 2007. Chronic granulomatous disease in pediatric patients: 25 years of experience. Allergol Immunopathol (Madr) **35:**83-9.
15. **Faguy, D. M., and W. F. Doolittle.** 2000. Horizontal transfer of catalase-peroxidase genes between archaea and pathogenic bacteria. Trends Genet **16:**196-7.
16. **Bertelli, C., F. Collyn, A. Croxatto, C. Rückert, A. Polkinghorne, C. Kebbi-Beghdadi, A. Goesmann, L. Vaughan, and G. Greub.** 2010. The Waddlia genome: a window into chlamydial biology. PLoS ONE **5:**e10890.
17. **Zamocky, M., P. G. Furtmuller, and C. Obinger.** 2008. Evolution of catalases from bacteria to humans. Antioxid Redox Signal **10:**1527-48.
18. **Greub, G., and D. Raoult.** 2004. Microorganisms resistant to free-living amoebae. Clin Microbiol Rev **17:**413-33.
19. **Collingro, A., P. Tischler, T. Weinmaier, T. Penz, E. Heinz, R. C. Brunham, T. D. Read, P. M. Bavoil, K. Sachse, S. Kahane, M. G. Friedman, T. Rattei, G. S. A. Myers, and M. Horn.** 2011. Unity in Variety - the Pan-Genome of the Chlamydiae. Molecular biology and evolution **28:**3253–3270.
20. **Alyamani, E. J., P. Brandt, J. A. Pena, A. M. Major, J. G. Fox, S. Suerbaum, and J. Versalovic.** 2007. Helicobacter hepaticus catalase shares surface-predicted epitopes with mammalian catalases. Microbiology **153:**1006-16.

21. **Ghadermarzi, M., and A. A. Moosavi-Movahedi.** 1999. Influence of different types of effectors on the kinetic parameters of suicide inactivation of catalase by hydrogen peroxide. Biochim Biophys Acta **1431:**30-6.
22. **Bates, P. A., L. A. Kelley, R. M. MacCallum, and M. J. Sternberg.** 2001. Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. Proteins **Suppl 5:**39-46.
23. **Bates, P. A., and M. J. Sternberg.** 1999. Model building by comparison at CASP3: using expert knowledge and computer automation. Proteins **Suppl 3:**47-54.
24. **Contreras-Moreira, B., P. W. Fitzjohn, and P. A. Bates.** 2002. Comparative modelling: an essential methodology for protein structure prediction in the post-genomic era. Appl Bioinformatics **1:**177-90.
25. **Arnold, F. W., J. T. Summersgill, A. S. Lajoie, P. Peyrani, T. J. Marrie, P. Rossi, F. Blasi, P. Fernandez, T. M. File, J. Rello, R. Menendez, L. Marzoratti, C. M. Luna, J. A. Ramirez, and C.-A. P. O. C. Investigators.** 2007. A worldwide perspective of atypical pathogens in community-acquired pneumonia. Am J Respir Crit Care Med **175:**1086-93.
26. **Guex, N., and M. C. Peitsch.** 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis **18:**2714-23.
27. **Kiefer, F., K. Arnold, M. Kunzli, L. Bordoli, and T. Schwede.** 2009. The SWISS-MODEL Repository and associated resources. Nucleic Acids Res **37:**D387-92.
28. **Schwede, T., J. Kopp, N. Guex, and M. C. Peitsch.** 2003. SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res **31:**3381-5.
29. **Diaz, A., P. C. Loewen, I. Fita, and X. Carpena.** Thirty years of heme catalases structural biology. Arch Biochem Biophys **525:**102-10.
30. **Casson, N., R. Michel, K. D. Muller, J. D. Aubert, and G. Greub.** 2008. Protochlamydia naegleriophila as etiologic agent of pneumonia. Emerg Infect Dis **14:**168-72.
31. **Collingro, A., E. R. Toenshoff, M. W. Taylor, T. R. Fritsche, M. Wagner, and M. Horn.** 2005. 'Candidatus Protochlamydia amoebophila', an endosymbiont of Acanthamoeba spp. Int J Syst Evol Microbiol **55:**1863-6.
32. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J Mol Biol **215:**403-10.
33. **Edgar, R. C.** 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics **5:**113.
34. **Hulo, N., A. Bairoch, V. Bulliard, L. Cerutti, E. De Castro, P. S. Langendijk-Genevaux, M. Pagni, and C. J. Sigrist.** 2006. The PROSITE database. Nucleic Acids Res **34:**D227-30.
35. **Sigrist, C. J., L. Cerutti, N. Hulo, A. Gattiker, L. Falquet, M. Pagni, A. Bairoch, and P. Bucher.** 2002. PROSITE: a documented database using patterns and profiles as motif descriptors. Brief Bioinform **3:**265-74.
36. **Crooks, G. E., G. Hon, J. M. Chandonia, and S. E. Brenner.** 2004. WebLogo: a sequence logo generator. Genome Res **14:**1188-90.
37. **Darriba, D., G. L. Taboada, R. Doallo, and D. Posada.** 2011. ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics **27:**1164-5.
38. **Guindon, S., J. F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel.** 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol **59:**307-21.
39. **Nei, M., and S. Kumar.** 2000. Molecular Evolution and Phylogenetics. Oxford University Press, New York.
40. **Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol **28:**2731-9.
41. **Huelsenbeck, J. P., and F. Ronquist.** 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17:**754-5.
42. **Wilgenbusch, J. C., D. L. Warren, and D. L. Swofford.** 2004. AWTY: A system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference.
43. **Benkert, P., M. Biasini, and T. Schwede.** Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics **27:**343-50.
44. **Benkert, P., S. C. Tosatto, and D. Schomburg.** 2008. QMEAN: A comprehensive scoring function for model quality assessment. Proteins **71:**261-77.

45. **Gouet, P., H. M. Jouve, and O. Dideberg.** 1995. Crystal structure of Proteus mirabilis PR catalase with and without bound NADPH. J Mol Biol **249:**933-54.
46. **Hakansson, K. O., M. Brugna, and L. Tasse.** 2004. The three-dimensional structure of catalase from Enterococcus faecalis. Acta Crystallogr D Biol Crystallogr **60:**1374-80.
47. **Melik-Adamyan, W., J. Bravo, X. Carpena, J. Switala, M. J. Mate, I. Fita, and P. C. Loewen.** 2001. Substrate flow in catalases deduced from the crystal structures of active site variants of HPII from Escherichia coli. Proteins **44:**270-81.
48. **Bravo, J., N. Verdaguer, J. Tormo, C. Betzel, J. Switala, P. C. Loewen, and I. Fita.** 1995. Crystal structure of catalase HPII from Escherichia coli. Structure **3:**491-502.
49. **Li, Y., and H. E. Schellhorn.** 2007. Rapid kinetic microassay for catalase activity. J Biomol Tech **18:**185-7.
50. **Lienard, J., A. Croxatto, S. Aeby, K. Jaton, K. Posfay-Barbe, A. Gervaix, and G. Greub.** 2011. Development of a new Chlamydiales-specific real-time PCR and its application to respiratory clinical samples. J Clin Microbiol**:**1-26.

Figure 1. Gene environment and phylogenetic tree of typical bacterial catalases including chlamydial catalases. (A) Representative tree of bacterial catalases obtained from www.peroxibase.toulouse.inra.fr with catalase sequences of *Chlamydiales* and other amoeba-resisting bacteria. The evolutionary history was inferred by using MB (1'000'000 generations) and confirmed by ML (500 bootstraps) and by MP (500 bootstraps). Only posterior probabilities below 1 are marked (●0.99-0.95, ●0.95-0.9, ●0.85-0.8, ●0.77-0.76). Bar indicates 0.1 amino acid substitution per site. (**B**) The gene environment of the catalases present in the different *Chlamydiales* exhibited no homology. Only *W. chondrophila* encodes for transposases and integrases near the catalase location suggesting that this catalase was likely obtained by lateral gene transfer.

Figure 2. Surface epitopes and heme binding domains. Bold font amino acids are differently conserved in large subunit catalases. (**A-C**) Surface epitopes conserved in *Chlamydiales,* previously determined by *in silico* modeling of *H. hepaticus* (Alyamani et al., 2007). Please note N304D and F306H mutations (arrows, *Pseudomonas* numbering). (**D**) Distal site of the prosthetic heme with catalytic histidine (red box). Please note F44Q and S59W mutations (arrows) (**E**) Proximal site of the prosthetic heme with catalytic tyrosine (red box).

Figure 3. Enzymatic activity of chlamydial catalases *in vitro*. (**A**)10^7 elementary bodies (EBs) of *Chlamydiales* and *P. aeruginosa* were exposed to 0.2M H_2O_2. Catalase activity is blocked by 0.1M azide. . (**B**) Degradation of hydrogen peroxide by purified catalase was followed by decrease in absorbance at 240nm. Catalase of *P. acanthamoebae* exhibited the weakest activity. The activity of the enzymes was conserved even when the pH was lower.

Mean of at least six independent experiments, +/- SEM for each condition and protein. Hsp60 of *W. chondrophila* was used as negative control. (**C**) Catalytic units derived from decrease in absorbance are in $M^{-1}*sec^{-1}$.

Figure 4. *In silico* modeling of small subunit chlamydial catalases. Tertiary structure of small subunit chlamydial catalases marked according to RMS values. (**A**) WchKatA tertiary structure is strongly conserved. (**B-C**) KatA of *P. acanthamoebae* and KatA of *E. lausannensis* tertiary structure are conserved except in some loops (arrows). (**D**) Catalytic site of *H. pylori* (white, PDB 2IQF), *P. mirabilis* (salmon, PDB: 1M85), and *E. faecalis* (yellow PDB: 1SI8). Numbering according to *H. pylori.* (**E**) Catalytic site of WchKatA (purple). All residues are conserved except R53 (arrow). (**F**) Catalytic site of KatA of *P. acanthamoebae* (red) and KatA of *E. lausannensis* (white). Orientation of the active site residues is conserved, except for H198 of *P. acanthamoebae* (arrow) and N128 in both bacteria (arrow). Numbering according to KatA of *P. acanthamoebae* sequence. Water molecules are marked W1-3, with hydrogen bonds showed in dotted green lines. The heme of *H. pylori* was introduced into the model structure of the small subunit chlamydial catalases.

Figure 5. Model of catalase acquisition and loss in *Chlamydiales*. (**A**) MrBayes phylogeny of fully sequenced *Chlamydiales* based on concatenation of gyrA, gyrB, rpoA, rpoB, recA, secY, tufA, and topA amino acid sequences. Only nodes with posterior probability inferior to 1 are marked. Tree constructed with ML placed the *Criblamydiaceae* node before the *Waddliaceae.* Families are delimited by colored boxes. Distribution of catalases in the *Chlamydiales* order is depicted in the column. (**B**) Two hypotheses for lateral transfer of clade 3 catalase in *P. acanthamoebae* and *E. lausannensis.* Transfer from a *Rhizobiales* to one of

555 the two chlamydial species and then internal transfer (1) or transfer from a *Rhizobiales* to
556 both species at about the same time (2). (**C**) Catalase of *W. chondrophila* could originate from
557 a lateral transfer from *Legionellales* (1) or an *Actinomycetales* (2).
558 **Figure S1: Phylogenetic trees based on Maximum Likelihood and Maximum Parsimony**.
559 (**A**) Tree constructed with PhyML with LG matrix, invariant gamma distribution (4 categories)
560 and 500 bootstraps. (**B**) Tree constructed with MEGA5 Maximum Parsimony with 500
561 bootstraps.
562 **Figure S2. Identity of catalases used for phylogenetic analysis.** (**A**) Clade 2 catalases
563 used in phylogenetic analyses. Please note that the average identity is of about 45 to 55%.
564 (**B**) Selection of clade 3 catalases used in phylogenetic analyses with average identity of 49
565 to 58%. Alignments were performed with MUSCLE. (**C**) Posterior probabilities of splits over
566 two MCMC runs are constant. (**D**) Pair wise comparison between the posterior probabilities
567 showed no divergence between runs.
568
569 **Figure S3. Surface epitopes consensus sequence.** Surface epitope motifs (grey boxes)
570 were confirmed with protein sequence alignment of clade 3 (**A**) and 2 catalases (**B**) available
571 on http://peroxibase.toulouse.inra.fr/. Amino acids that differ in conservation between clade 2
572 and clade 3 catalases are highlighted in green and blue boxes. (**C**) Overview of consensus
573 sequence of clade 2 catalases. Conservation is low at the N- and C-terminal part of the
574 protein. Alignments were built with MUSCLE.
575
576 **Figure S4. *In silico* modeling of large subunit chlamydial catalases and tetramers.** (**A**)
577 Tetramer model of chlamydial clade 3 catalases with heme group from *E. faecalis* (1SI8). (**B**)
578 Tetramer structure of clade 2 chlamydial catalases with heme group from *E. coli* (1IPH). (**C**)

579 Tertiary structure of *C. sequanensis* (left) and *N. hartmannellae* (right) colored according to
580 RMS values compared to *E. coli* (1IPH) structure. (**D**) Catalytic site of *C. sequanensis*
581 (magenta) and *N. hartmannellae* are conserved compared to *E. coli* (1GGE, white). Residues
582 numbering according to *C. sequanensis*.
583
584 **Figure S5. Clashes of backbone in chlamydial catalases.** *In silico* models of chlamydial
585 catalases were checked for possible residues that clashed with backbone and side chains
586 without proper hydrogen bonds. Prolines, which strained the backbone are highlighted in red.
587 All other residues that clashed with the backbone are depicted in yellow (**A**) For clade 2 no
588 side chains were missing hydrogen bonds. Even in the model of the crystallized *E. coli*
589 subunit some residues are not in an optimal conformation. (**B**) Only very few residues clashed
590 with the backbone in clade 3 catalases and no side chains were missing stabilizing hydrogen
591 bonds.

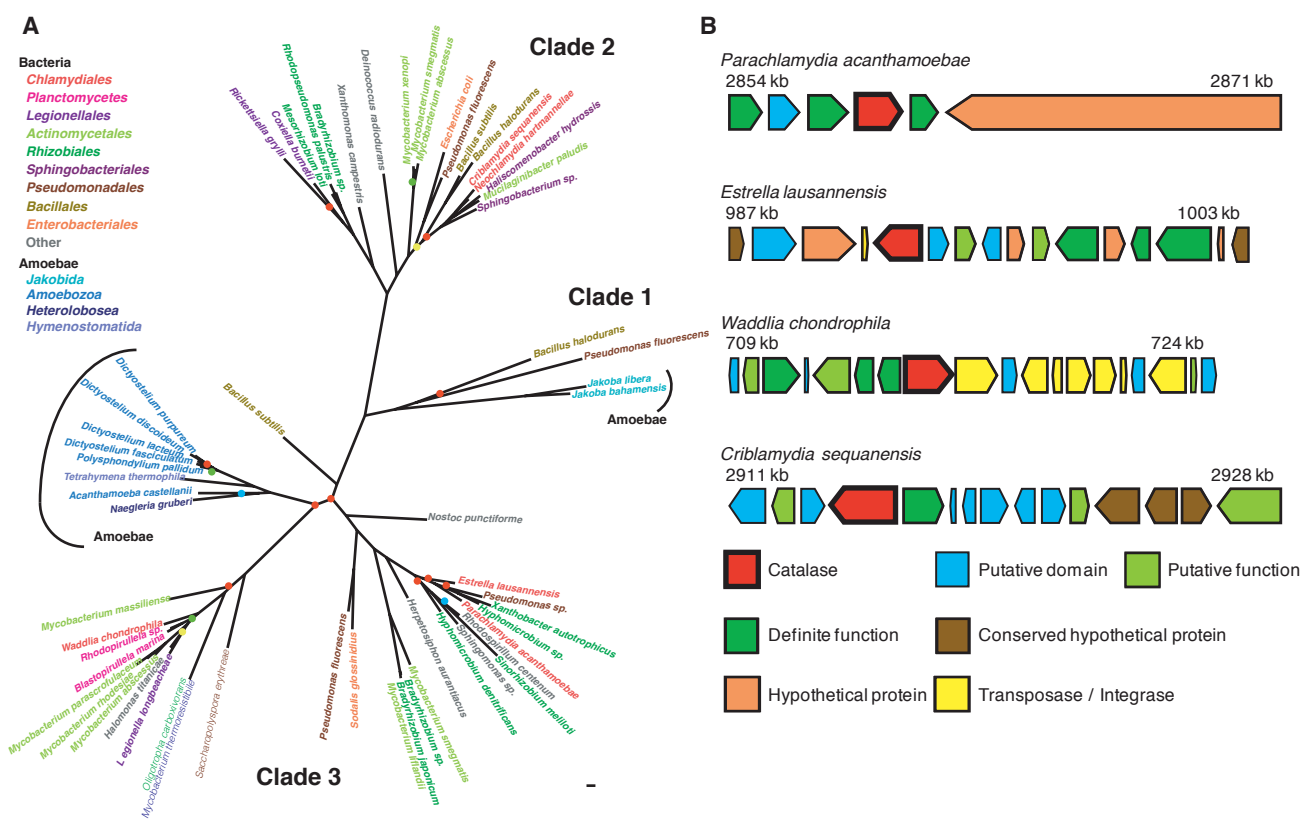


Figure 1. Gene environment and phylogenetic tree of typical bacterial catalases including chlamydial catalases. (**A**) Representative tree of bacterial catalases obtained from www.peroxibase.toulouse.inra.fr with catalase sequences of *Chlamydiales* and other amoeba-resisting bacteria. The evolutionary history was inferred by using MB (1'000'000 generations) and confirmed by ML (500 bootstraps) and by MP (500 bootstraps). Only posterior probabilities below 1 are marked (●0.99-0.95,●0.95-0.9, ●0.85-0.8, ●0.77-0.76). Bar indicates 0.1 amino acid substitution per site. (**B**) The gene environment of the catalases present in the different *Chlamydiales* exhibited no homology. Only *W. chondrophila* encodes for transposases and integrases near the catalase location suggesting that this catalase was likely obtained by lateral gene transfer.

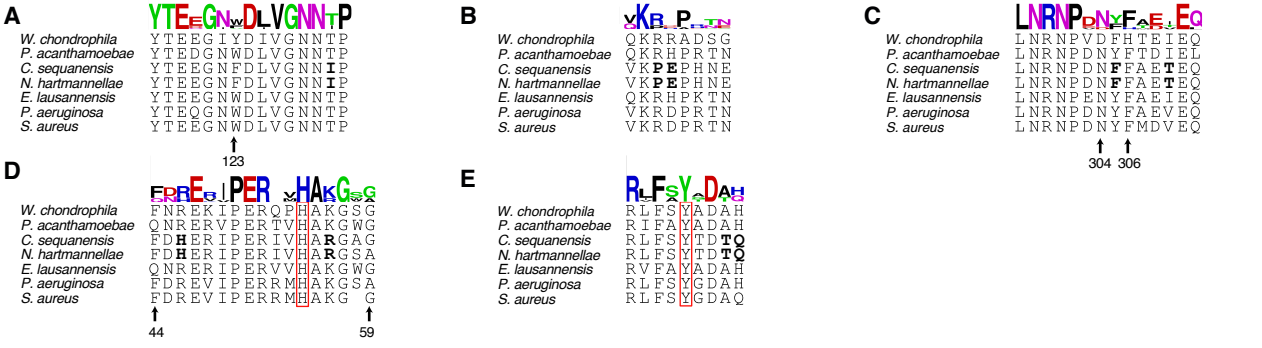
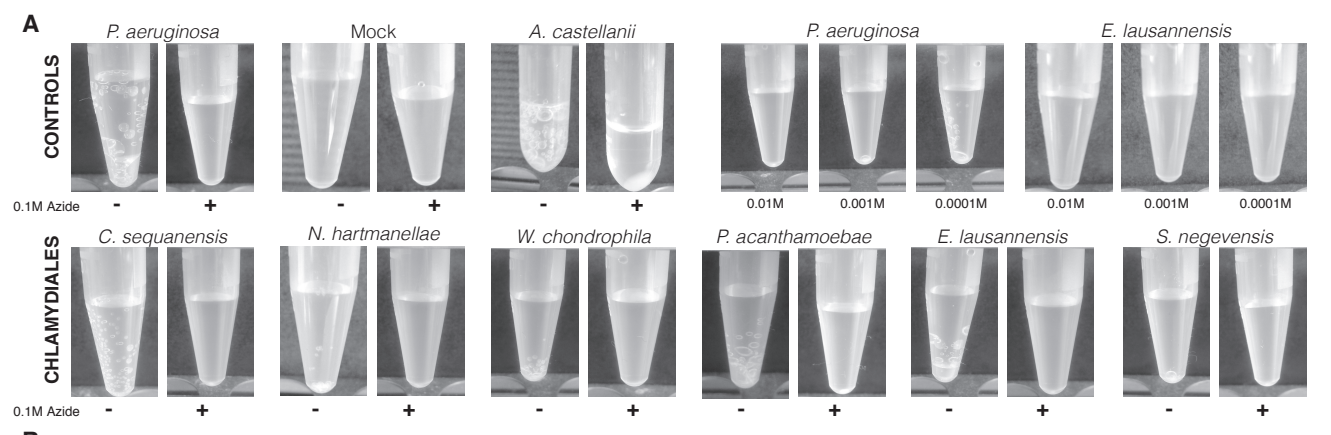
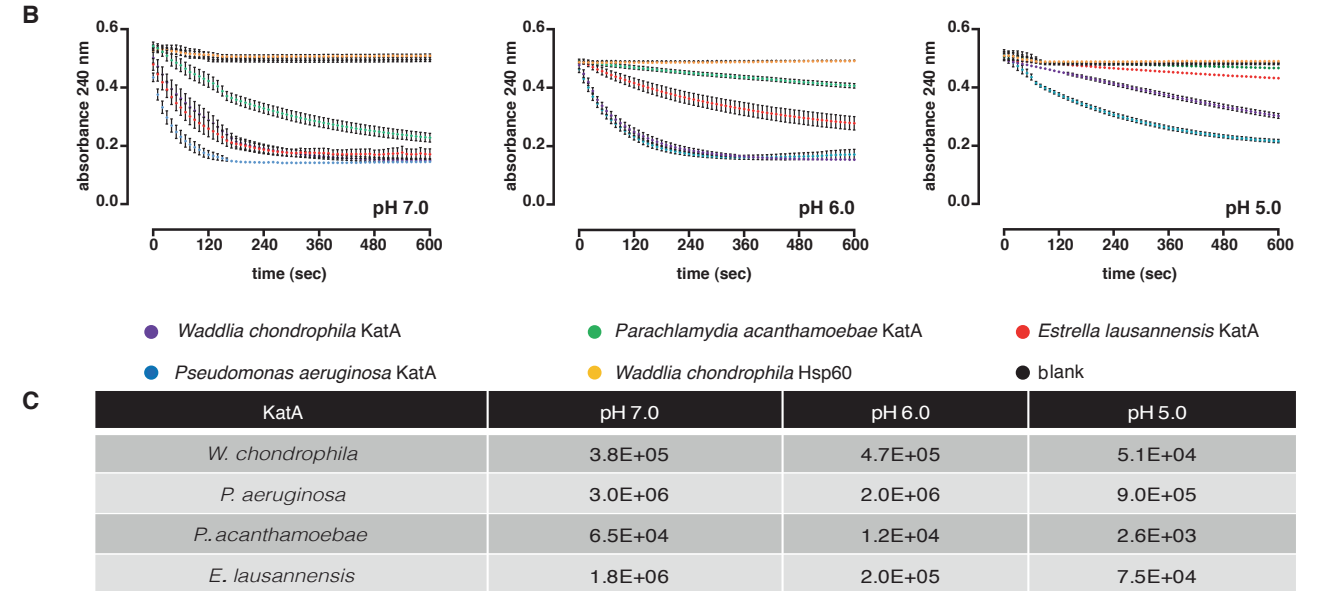


Figure 2. Surface epitopes and heme binding domains. Bold font amino acids are differently conserved in large subunit catalases. (**A-C**) Surface epitopes conserved in *Chlamydiales*, previously determined by in silico modeling of *H. pylori* (Alyamani et al., 2007). Please note N304D and F306H mutations (arrows, *Pseudomonas* numbering). (**D**) Distal site of the prosthetic heme with catalytic histidine (red box). Please note F44Q and S59W mutations (arrows) (**E**) Proximal site of the prosthetic heme with catalytic tyrosine (red box).





C

KatA	pH 7.0	pH 6.0	pH 5.0
W. chondrophila	3.8E+05	4.7E+05	5.1E+04
P. aeruginosa	3.0E+06	2.0E+06	9.0E+05
P. acanthamoebae	6.5E+04	1.2E+04	2.6E+03
E. lausannensis	1.8E+06	2.0E+05	7.5E+04

Figure 3. Enzymatic activity of chlamydial catalases in vitro. (**A**) 10^7 Elementary bodies (EBs) of *Chlamydiales* and *P. aeruginosa* were exposed to 0.2M H_2O_2. Catalase activity is blocked by 0.1M azide. (**B**) Degradation of hydrogen peroxide by purified catalase was followed by decrease in absorbance at 240nm. Catalase of *P. acanthamoebae* exhibited the weakest activity. The activity of the enzymes was conserved even when the pH was lower. Mean of at least six independent experiments, +/- SEM for each condition and protein. Hsp60 of *W. chondrophila* was used as negative control. (**C**) Catalytic units derived from decrease in absorbance are in M-1*sec-1.

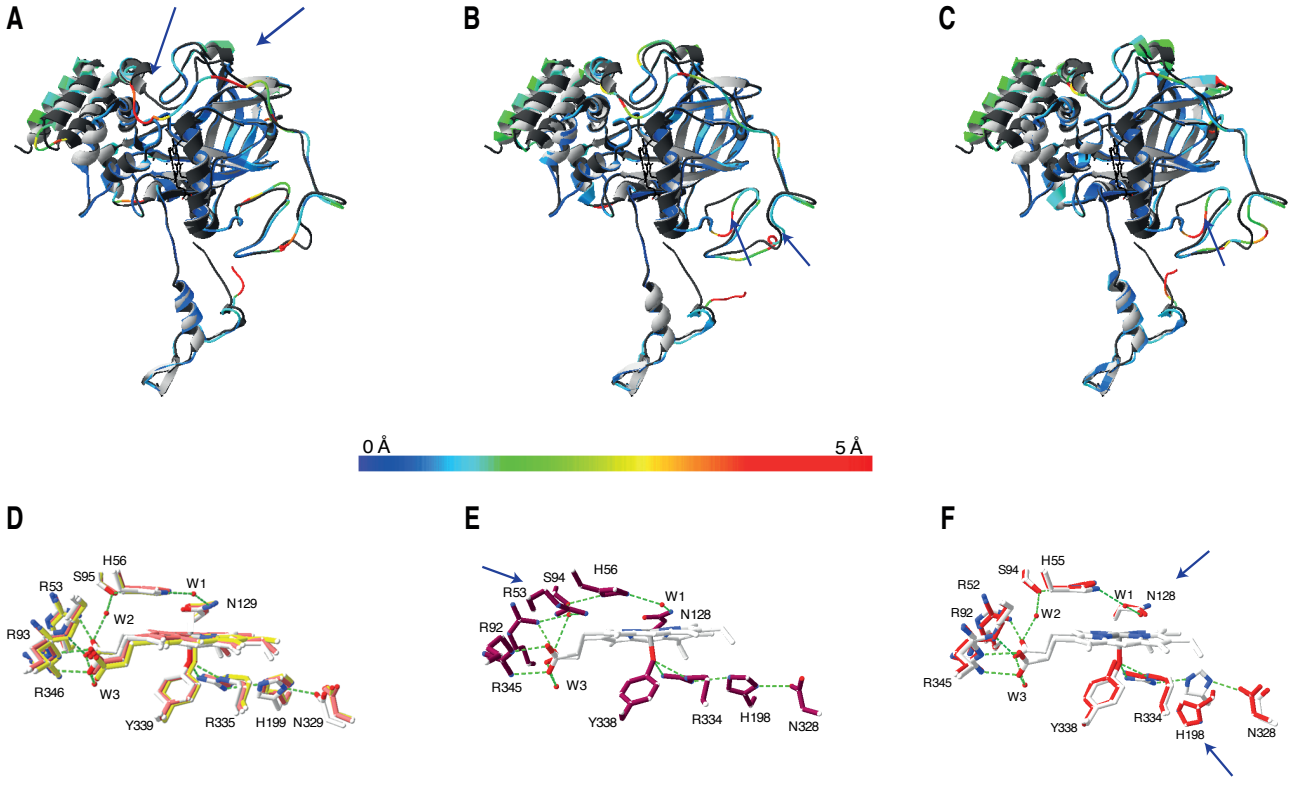


Figure 4. *In silico* modeling of small subunit chlamydial catalases. Tertiary structure of small subunit chlamydial catalases marked according to RMS values. (**A**) *W. chondrophila* KatA tertiary structure is strongly conserved. (**B-C**) *P. acanthamoebae* KatA and *E. lausannensis* KatA tertiary structure are conserved except in some loops (arrows). (**D**) Catalytic site of *H. pylori* (white, PDB 2IQF), *P. mirabilis* (salmon, PDB: 1M85), and *E. faecalis* (yellow PDB: 1SI8). Numbering according to *H. pylori*. (**E**) Catalytic site of *W. chondrophila* KatA (purple). All residues are conserved except R53 (arrow). (**F**) Catalytic site of *P. acanthamoebae* KatA (red) and *E. lausannensis* KatA (white). Orientation of the active site residues is conserved, except for H198 of *P. acanthamoebae* (arrow) and N128 in both bacteria (arrow). Numbering according to *P. acanthamoebae* KatA sequence. Water molecules are marked W1-3, with hydrogen bonds showed in dotted green lines. The heme of *H. pylori* was introduced into the model structure of the small subunit chlamydial catalases.

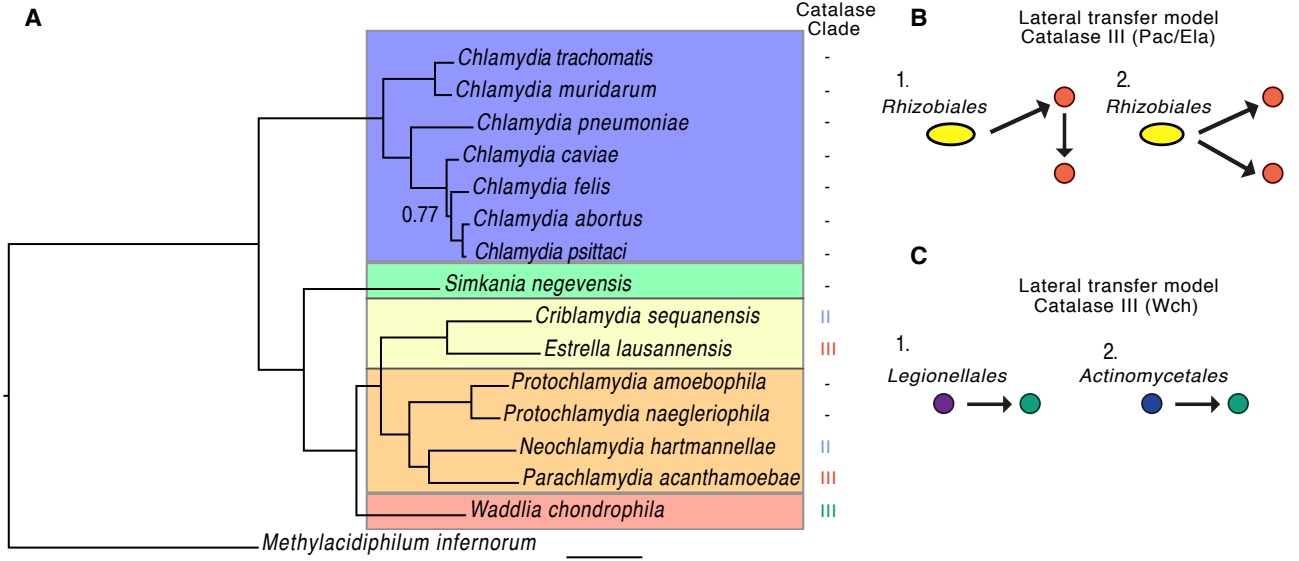


Figure 5. Model of catalase acquisition and loss in *Chlamydiales*. (**A**) MrBayes phylogeny of fully sequenced *Chlamydiales* based on concatenation of gyrA, gyrB, rpoA, rpoB, recA, secY, tufA, and topA amino acid sequences. Only nodes with posterior probability inferior to 1 are marked. Tree constructed with ML placed the *Criblamydiaceae* node before the *Waddliaceae*. Families are delimited by colored boxes. Distribution of catalases in the *Chlamydiales* order is depicted in the column. (B) Two hypotheses for lateral transfer of clade 3 catalase in *P. acanthamoebae* and *E. lausannensis*. Transfer from a *Rhizobiales* to one of the two chlamydial species and then internal transfer (1) or transfer from a *Rhizobiales* to both species at about the same time (2). (C) Catalase of *W. chondrophila* could originate from a lateral transfer from *Legionellales* (1) or an *Actinomycetales* (2).

A

	Cse	Nha	Rgr	Cbu	Mlo	Bsp	Rpa	Bha	Bsu	Eco	Mab	Msm	Mxe	Pfl	Hhy	Mpa	Ssp	Xca	Dra
Criblamydia sequanensis		75.9	44.9	48.5	47.1	48.4	46.9	53.3	53.5	48	53.4	51.5	51	56.9	66.8	68.6	65.9	49.7	46.3
Neochlamydia hartmanellae	75.9		44.7	47.1	46.3	47.8	46.8	54.1	55.9	50.4	53.3	53.1	51.7	57.9	69.7	69.7	68	51.6	45.5
Rickettsiella grylli	44.9	44.7		61.5	59.1	58.6	57.3	42.7	43.1	40	44.3	43.6	42.6	44.9	44	44.9	44.4	51.5	40.4
Coxiella burnetii	48.5	47.1	61.5		65.8	66.8	64.8	44.6	47.2	46	46.4	48.9	46.8	47.8	48.6	47.3	49.2	56.8	46.1
Mesorhizobium loti	47.1	46.3	59.1	65.8		78.6	77.4	43	45.1	43	48.2	48.1	47.7	48.1	47.7	47.6	48.4	60.5	45.2
Bradyrhizobium sp.	48.4	47.8	58.6	66.8	78.6		79.9	44.4	45.4	43.6	49.2	49.1	48.4	47.7	49.2	48	49.2	59.7	46.7
Rhodopseudomonas palustris	46.9	46.8	57.3	64.8	77.4	79.9		44.6	45.5	42.5	49.5	48.6	48.2	47.2	48.7	46.8	48.6	58.4	45.3
Bacillus halodurans	53.3	54.1	42.7	44.6	43	44.4	44.6		66.3	45.3	53.2	53.1	53.6	51.5	52.2	52.4	54.6	45.1	46.4
Bacillus subtilis	53.5	55.9	43.1	47.2	45.1	45.4	45.5	66.3		47.8	55	54.9	55.1	52.9	54.6	55.2	56	47.7	47.2
Escherichia coli	48	50.4	40	46	43	43.6	42.5	45.3	47.8		50.3	47.1	47.8	56.5	47.9	50.5	49	45.6	40.8
Mycobacterium abscessus	53.4	53.3	44.3	46.4	48.2	49.2	49.5	53.2	55	50.3		76.4	77.4	54.4	54.9	53.5	55.9	50.1	48.5
Mycobacterium smegmatis	51.5	53.1	43.6	48.9	48.1	49.1	48.6	53.1	54.9	47.1	76.4		84.2	52.9	54.3	53.4	55.2	46.9	46.2
Mycobacterium xenopi	51	51.7	42.6	46.8	47.7	48.4	48.2	53.6	55.1	47.8	77.4	84.2		54.4	54	52.3	54.5	47	46.3
Pseudomonas fluorescens	56.9	57.9	44.9	47.8	48.1	47.7	47.2	51.5	52.9	56.5	54.4	52.9	54.4		55.5	57.2	56.3	50.6	43.9
Haliscomenobacter hydrossis	66.8	69.7	44	48.6	47.7	49.2	48.7	52.2	54.6	47.9	54.9	54.3	54	55.5		72.1	66.7	50.5	46.4
Mucilaginibacter paludis	68.6	69.7	44.9	47.3	47.6	48	46.8	52.4	55.2	50.5	53.5	53.4	52.3	57.2	72.1		68.4	50.2	46.3
Sphingobacterium sp.	65.9	68	44.4	49.2	48.4	49.2	48.6	54.6	56	49	55.9	55.2	54.5	56.3	66.7	68.4		49.9	45.6
Xanthomonas campestris	49.7	51.6	51.5	56.8	60.5	59.7	58.4	45.1	47.7	45.6	50.1	46.9	47	50.6	50.5	50.2	49.9		46.8
Deinococcus radiodurans	46.3	45.5	40.4	46.1	45.2	46.7	45.3	46.4	47.2	40.8	48.5	46.2	46.3	43.9	46.4	46.3	45.6	46.8	

0-49.9
50-59.9
60-79.9
80-100

B

	Ela	Pac	Wch	Llo	Bma	Oca	Xau	Bsu	Hde	Mab	Mli	Mma	Msm	Ser	Pfl	Hau	Npu	Rce	Aca	Ddi	Ppa	Ngu	Tth
Estrella lausannensis		77.6	51	49.9	48.9	46.3	74.5	58	79.3	48.3	58.8	51.1	57.6	50.2	62	70.8	61.2	75.6	53	50.2	51.2	52.2	47.7
Parachlamydia acanthamoebae	77.6		49.1	48.9	47.6	46.6	76.2	55.1	74.6	47.2	57.6	50.3	56.2	48.8	60.5	70.8	57.9	75.4	51.6	48.4	48.6	50.8	46.4
Waddlia chondrophila	51	49.1		75.8	79.1	58.8	49.2	50	50.2	74.4	49.3	66	46.9	64.7	50.5	51	49.2	50.7	49.8	47.6	48.6	50.4	44.3
Legionella longbeacheae	49.9	48.9	75.8		76.9	58.2	49.4	48.8	51.4	74.8	48.3	65.2	46.7	66.9	50.5	50.4	47.3	50.1	51	48.4	48.4	50.2	45.9
Blastopirellula marina	48.9	47.6	79.1	76.9		60.6	49.6	52.1	50.2	76	48.5	65.6	46.3	66.7	49.7	50.6	48.6	49.9	49.6	46.9	47.6	50	44.3
Oligotropha carboxidovorans	46.3	46.6	58.8	58.2	60.6		45.9	47	45.8	57.2	45.7	58.6	43.7	57.8	46.4	48.6	43.1	47.3	46.8	44.5	44.3	45.2	41.5
Xanthobacter autotrophicus	74.5	76.2	49.2	49.4	49.6	45.9		54.4	74.9	49	56.6	51.1	55.2	51	56.7	69.1	56.8	74.3	51.3	50.3	50.5	52.3	46.3
Bacillus subtilis	58	55.1	50	48.8	52.1	47	54.4		57.6	49.4	52.4	52.5	53.1	53.8	59.5	61.6	59.8	56.6	57.1	59.3	61.7	61.6	51
Hyphomicrobium denitrificans	79.3	74.6	50.2	51.4	50.2	45.8	74.9	57.6		50.1	59.1	50.7	58	53.2	61.6	70.6	61.6	78.6	54.3	51	51	53	47
Mycobacterium abscessus	48.3	47.2	74.4	74.8	76	57.2	49	49.4	50.1		48.3	64.2	46.5	64	49.6	49.7	47.4	50.3	48.1	45.9	46.5	48.4	42.5
Mycobacterium liflandii	58.8	57.6	49.3	48.3	48.5	45.7	56.6	52.4	59.1	48.3		48.2	68.9	46.8	57.9	61	53.4	60.8	50.6	48.5	48.5	49	43.5
Mycobacterium massiliense	51.1	50.3	66	65.2	65.6	58.6	51.1	52.5	50.7	64.2	48.2		45.7	64	50.4	51.9	51.7	51.8	52.1	50.7	51.3	51.2	43.6
Mycobacterium smegmatis	57.6	56.2	46.9	46.7	46.3	43.7	55.2	53.1	58	46.5	68.9	45.7		46.3	56.9	61.9	52.8	59.4	48.8	47.1	48.1	48.7	44
Saccharopolyspora erytraea	50.2	48.8	64.7	66.9	66.7	57.8	51	53.8	53.2	64	46.8	64	46.3		50.2	52	49	52.1	51.9	49.3	49.7	52.9	46.8
Pseudomonas fluorescens	62	60.5	50.5	50.5	49.7	46.4	56.7	59.5	61.6	49.6	57.9	50.4	56.9	50.2		64.9	63.6	61.5	55.5	54.4	55	55.8	51
Herpetosiphon aurantiacus	70.8	70.8	51	50.4	50.6	48.6	69.1	61.6	70.6	49.7	61	51.9	61.9	52	64.9		60.9	72	51.9	54.1	54.3	55.1	48.7
Nostoc punctiforme	61.2	57.9	49.2	47.3	48.6	43.1	56.8	59.8	61.6	47.4	53.4	51.7	52.8	49	63.6	60.9		61.9	57.1	55.8	56.6	56.1	49.3
Rhodospirillum centenum	75.6	75.4	50.7	50.1	49.9	47.3	74.3	56.6	78.6	50.3	60.8	51.8	59.4	52.1	61.5	72	61.9		53.2	51.3	51.9	53.1	48.7
Acanthamoeba castellanii	53	51.6	49.8	51	49.6	46.8	51.3	57.1	54.3	48.1	50.6	52.1	48.8	51.9	55.5	51.9	57.1	53.2		66.7	66.7	68.3	60.1
Dictyostelium discoideum	50.2	48.4	47.6	48.4	46.9	44.5	50.3	59.3	51	45.9	48.5	50.7	47.1	49.3	54.4	54.1	55.8	51.3	66.7		82.3	63.9	59.4
Polysphondylium pallidum	51.2	48.6	48.6	48.4	47.6	44.3	50.5	61.7	51	46.5	48.5	51.3	48.1	49.7	55	54.3	56.6	51.9	66.7	82.3		68.6	60.1
Neagleria gruberi	52.2	50.8	50.4	50.2	50	45.2	52.3	61.6	53	48.4	49	51.2	48.7	52.9	55.8	55.1	56.1	53.1	68.3	63.9	68.6		60.8
Tetrahymena thermophila	47.7	46.4	44.3	45.9	44.3	41.5	46.3	51	47	42.5	43.5	43.6	44	46.8	51	48.7	49.3	48.7	60.1	59.4	60.1	60.8	

C

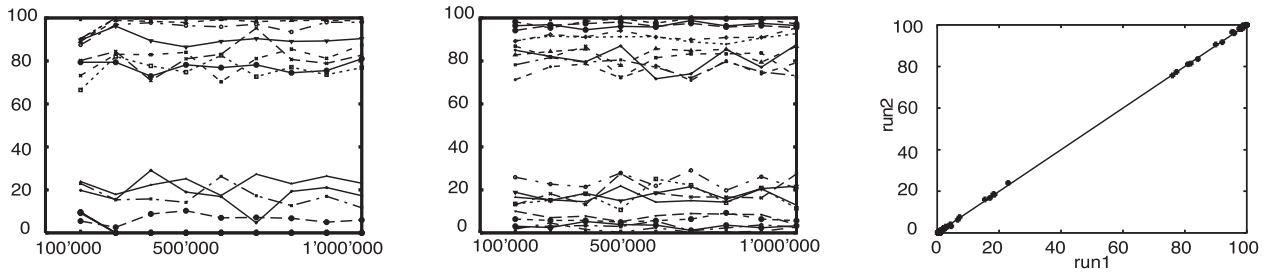


Figure S1. Identity of catalases used for phylogenetic analysis. (**A**) Clade 2 catalases used in phylogenetic analyses. Please note that the average identity is of about 45 to 55%. (**B**) Selection of clade 3 catalases used in phylogenetic analyses with average identity of 49 to 58%. Alignments were performed with MUSCLE. (**C**) Posterior probabilities of splits over two MCMC runs are constant. (**D**) Pairwise comparison between the posterior probabilities of the two runs showed no divergence between runs.

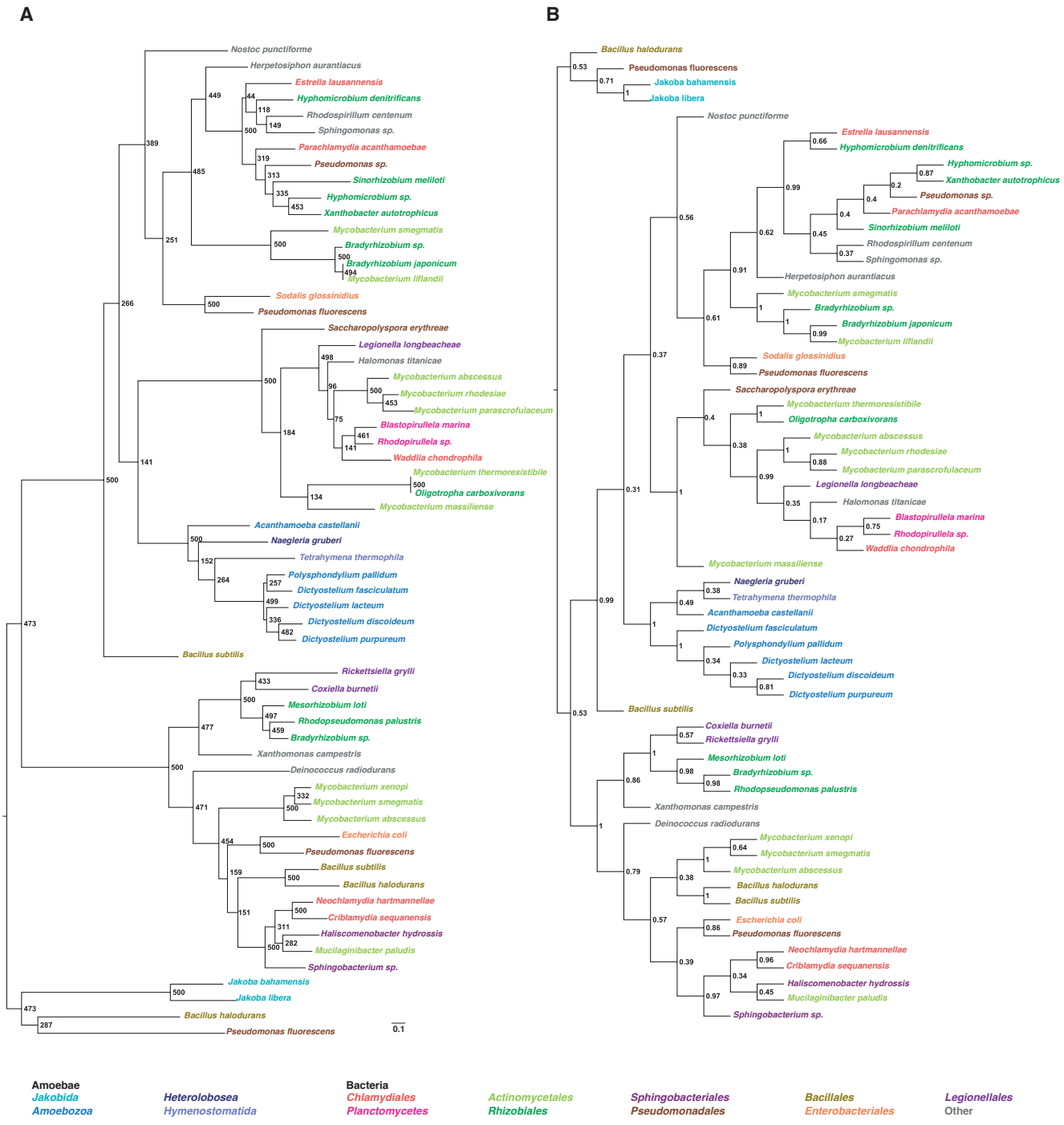


Figure S2: Phylogenetic trees based on Maximum Likelihood and Maximum Parsimony. (**A**) Tree constructed with PhyML with LG matrix, invariant gamma distribution (4 categories) and 500 bootstraps. (**B**) Tree constructed with MEGA5 Maximum Parsimony with 500 bootstraps.

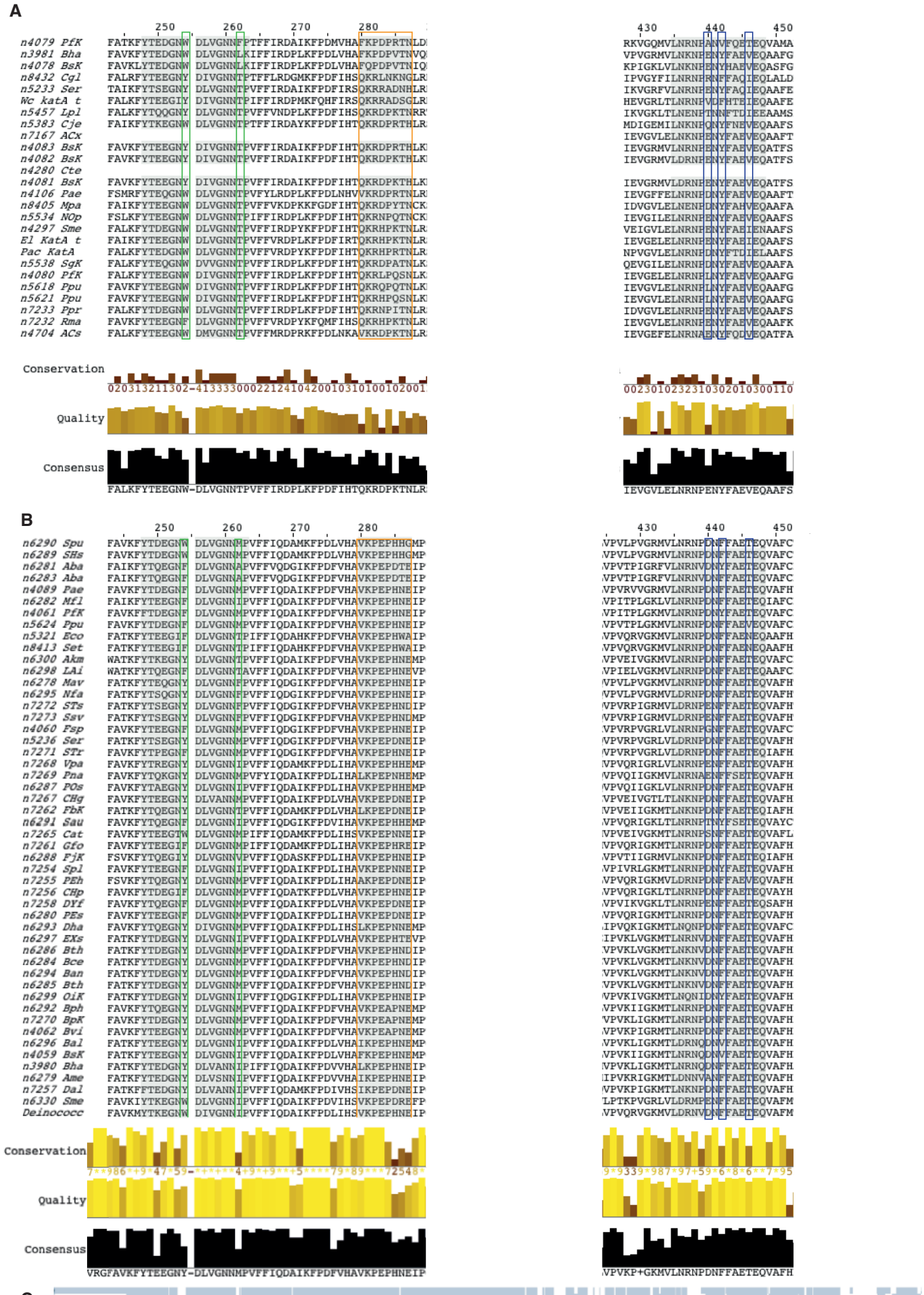


Figure S3: Surface epitopes of consensus sequence. Surface epitope motifs (grey boxes) were confirmed with protein sequence alignment of clade 3 (**A**) and 2 catalases (**B**) available on http://peroxibase.toulouse.inra.fr/. Amino acids that differ in conservation between clade 2 and clade 3 catalases are highlighted in green and blue boxes. (**C**) Overview of consensus sequence of clade 2 catalases. Conservation is low at the N- and C-terminal part of the protein. Alignments were built with MUSCLE.

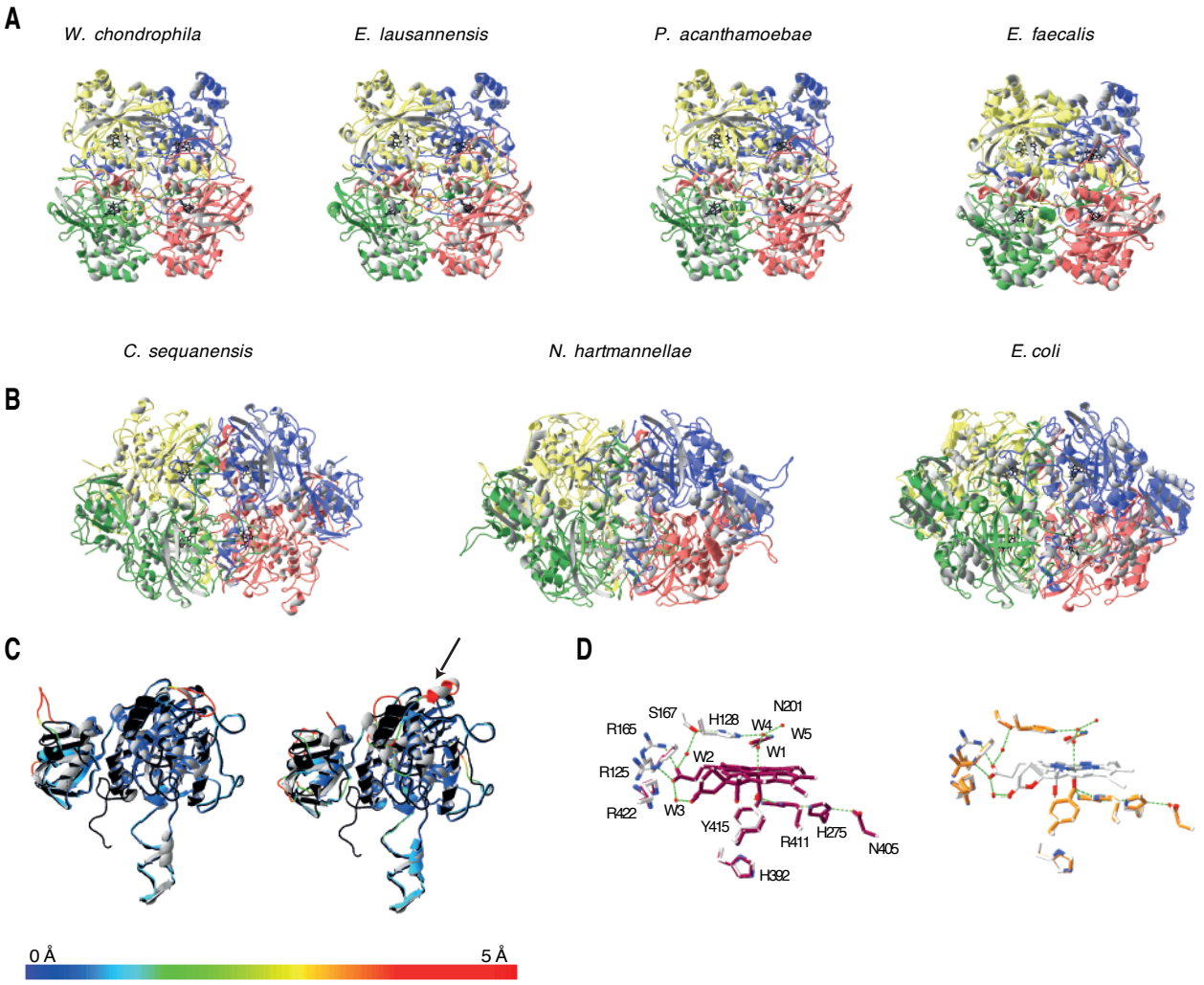


Figure S4. In silico modeling of large subunit chlamydial catalases and tetramers. (**A**) Tetramer model of chlamydial clade 3 catalases with heme group from *E. faecalis* (1SI8). (**B**) Tetramer structure of clade 2 chlamydial catalases with heme group from *E. coli* (1IPH). (**C**) Tertiary structure of *C. sequanensis* (left) and *N. hartmannellae* (right) colored according to RMS values compared to *E. coli* (1IPH) structure. (D) Catalytic site of *C. sequanensis* (magenta) and *N. hartmannellae* are conserved compared to *E. coli* (1GGE, white). Residues numbering according to *C. sequanensis*.

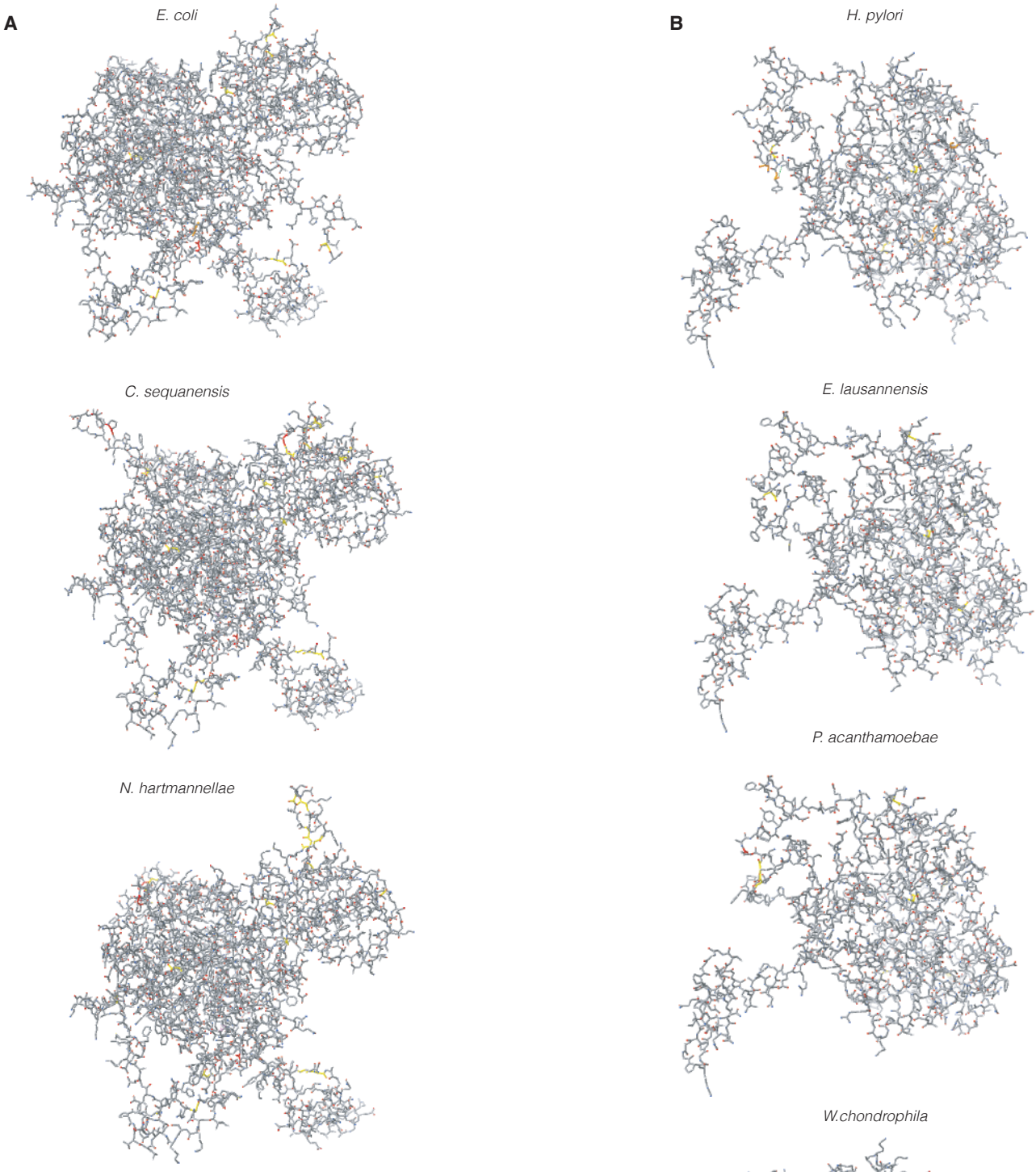


Figure S5. Clashes of backbone in chlamydial catalases. In silico models of chlamydial catalases were checked for possible residues that clashed with backbone and side chains without proper hydrogen bonds. Prolines, which strained the backbone are highlighted in red. All other residues that clashed with the backbone are depicted in yellow. (**A**) For clade 2 no side chains were missing hydrogen bonds. Even in the model of the crystallized *E. coli* subunit some residues are not in an optimal conformation. (**B**) Only very few residues clashed with the backbone in clade 3 catalases and no side chains were missing stabilizing hydrogen bonds.

Table S1: Previously unpublished chlamydial proteins used for phylogeny of *Chlamydiales* order

Chlamydiales	**gyrA**	**gyrB**	**rpoB**	**tuf**	**rpoA**	**recA**	**secY**	**topA**
C. sequanensis				KC514611		KC514612		
N. hartmannellae	KC514613	KC514614	KC514617	KC514618	KC514616	KC514620	KC514615	KC514619
P. naegleriophila	KC514601	KC514602	KC514603	KC514604	KC514605	KC514606	KC514607	KC514608
E. lausannensis				KC514609		KC514610		