

Fine tuning rigid body docking results using the Dreiding force field: A computational study of 36 known nanobody-protein complexes

Aysima Hacısuleyman¹  | Burak Erman² 

¹Department of Computational Biology,
University of Lausanne, Lausanne, Switzerland

²Chemical and Biological Engineering, Koc
University, Istanbul, Turkey

Correspondence

Burak Erman, Chemical and Biological
Engineering, Koc University, Istanbul, Turkey.
Email: berman@ku.edu.tr

Abstract

This paper aims to understand the binding strategies of a nanobody-protein pair by studying known complexes. Rigid body protein–ligand docking programs produce several complexes, called decoys, which are good candidates with high scores of shape complementarity, electrostatic interactions, desolvation, buried surface area, and Lennard–Jones potentials. However, the decoy that corresponds to the native structure is not known. We studied 36 nanobody-protein complexes from the single domain antibody database, sd-Ab DB, <http://www.sdab-db.ca/>. For each structure, a large number of decoys are generated using the Fast Fourier Transform algorithm of the software ZDOCK. The decoys were ranked according to their target protein–nanobody interaction energies, calculated by using the Dreiding Force Field, with rank 1 having the lowest interaction energy. Out of 36 protein data bank (PDB) structures, 25 true structures were predicted as rank 1. Eleven of the remaining structures required Ångstrom size rigid body translations of the nanobody relative to the protein to match the given PDB structure. After the translation, the Dreiding interaction (DI) energies of all complexes decreased and became rank 1. In one case, rigid body rotations as well as translations of the nanobody were required for matching the crystal structure. We used a Monte Carlo algorithm that randomly translates and rotates the nanobody of a decoy and calculates the DI energy. Results show that rigid body translations and the DI energy are sufficient for determining the correct binding location and pose of ZDOCK created decoys. A survey of the sd-Ab DB showed that each nanobody makes at least one salt bridge with its partner protein, indicating that salt bridge formation is an essential strategy in nanobody-protein recognition. Based on the analysis of the 36 crystal structures and evidence from existing literature, we propose a set of principles that could be used in the design of nanobodies.

KEYWORDS

case-study, Dreiding energy, nanobody, salt-bridge, ZDOCK

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Proteins: Structure, Function, and Bioinformatics* published by Wiley Periodicals LLC.

1 | INTRODUCTION

Nanobodies form a class of proteins produced by the immune system of camelids and sharks and are the antigen-binding domains of heavy chain-only antibodies also denoted as VHH antibodies. These are small proteins, 12–15 kDa, that have three loops which are the three complementarity determining regions, CDR's, that play a significant role in the binding of the nanobody with its target proteins.¹ Although nanobodies have evolved to eliminate harmful molecules in the systems of their parenting species, they can be used in applications of protein crystallization, targeting ligands, biosensing, and diagnostics.^{2,3} Progress in the use of nanobodies in areas concerning human health is impressive.^{4–10} Nanobody technology is rapidly becoming a cornerstone of immunoinformatic as evidenced by the recent integrated database INDI.⁵ Computationally, the three basic challenges of nanobody design are to (1) choose the best binding nanobody from a given library that will bind to a given protein, (2) perform mutations in the CDR regions of a given nanobody to change its binding capacity to a protein, (3) find the binding location and strength of a given nanobody to a given protein. Computationally, these challenges are most commonly investigated by algorithms, the majority of which are based on the assumption of rigid body docking. The advances in rigid body docking algorithms are remarkable.^{11–23} For a given protein and nanobody, the fast fourier transform algorithm (FFT) generates large numbers of their bound forms, called decoys, which are good binders obtained by scoring functions based on shape complementarity, electrostatic interactions, desolvation, buried surface area, Lennard–Jones potentials, and so forth. The search algorithm common to all these studies searches exhaustively the entire rotational and translational space of the ligand with the receptor being fixed. In rotational search, the three Euler angles for the ligand are changed by small increments, ca 5–15°, and for every rotation, the translational space defined by a set of grids on the surface of the protein, with grid spacing ca 0.5–1.2 Å, is explored using FFT. As a result of this technique, a large number of candidate complexes, decoys, are generated, where the nanobody is bound to different suitable locations on the protein and ranked according to the interaction energy criterion mentioned. The main question to be answered is “which one of these decoys is the true solution?” that is, which one matches the crystal structure? In the present study, by examining 36 known nanobody-protein crystal structures, we see that the dominant metric that serves as best distinguishing feature is the interaction energy calculated by the Dreiding force field²⁴ which emerges as an excellent criterion for identifying the decoy that is closest to the known crystal structure. We generate the decoys by the algorithm of ZDOCK.¹¹ These structures are already good candidates that passed the ZDOCK criteria of docking. We then calculate their interaction energies using the Dreiding potential and rank these candidates with respect to their Dreiding interaction (DI) energies, the decoy with the lowest interaction energy being of rank 1. Then, we find the rank of the decoy whose structure is closest to that of the known structure. The known crystal structures are taken from the data bank single domain antibody database²⁵ (sd-Ab DB), <http://www.sdab-db.ca/>. As will be shown in detail as follows, the prediction of the correct nanobody-protein complex using the DI

energy as the force field is extremely accurate and reliable. In several cases, the method predicts the lowest DI energy decoy, that is, rank 1 structure, as the known crystal structure. In some other cases, the decoy that matches the crystal structure is not of rank 1, but small, sub-Ångstrom, rigid body displacements superpose them on the crystal structure. The decoy becomes rank 1 after these rigid body translations. Only for one case, both rigid body translation and rotation are required for the lowest root mean square deviation (RMSD) yielding decoy to become of rank 1. For this purpose, we developed a Monte Carlo script for randomly translating and rotating a decoy until its DI energy reaches a minimum.

The aim of this paper is to learn the binding strategies of nanobody-protein complex formation, with the hope of using them for predicting unknown nanobody-protein systems in the future. There are already several studies that aim at uncovering the rules of recognition. Notably, Gray et al.²⁶ introduced the concept of binding funnel formation which is essentially an energy funnel into which the ligand is trapped prior to finding its exact location on the protein. The concept has its roots in Hill's theory²⁷ of cooperativity in biochemistry where the ligand anchors itself on the surface of the protein and exhibits fluctuations in the three Euler angles formed between the ligand and the protein. While the Hill model was for rigid bodies, Gray et al.²⁶ added side-chain flexibility to small translations and rotations of the ligand as was also done by Fernández-Recio et al. and²⁸ Palma et al.²⁹ Flexibility was also introduced by relaxing interaction potentials by Chen and Weng.³⁰ Our study shows that once the proper decoy is obtained, small rigid body translations and rotations that lower the interaction energy are sufficient to make it the top-ranking decoy. The implication of this observation is that perfect jigsaw-like matching is the dominant potential for defining the best binding pose. The concept of jigsaw matching has indeed been employed in earlier studies of recognition both at the intramolecular and intermolecular level.^{31–34} A survey of the sd-Ab DB showed that, with the exception of one or two cases, each nanobody makes at least one salt bridge with its partner protein. Thus, the binding strategy that emerges from our study has two simple ingredients: First, the nanobody anchors itself on the protein at a location that leads to a salt bridge. Then the ligand performs small rigid body translations and rotations about this anchor point until the binding energy is minimized. The rapid formation of a salt bridge between the ligand and the protein as a determinant of recognition has already been acknowledged.^{35,36} The role of flexibility as a nanobody-protein recognition strategy has yet to be seen.

2 | METHODS

We used the sd-Ab DB,²⁵ <http://www.sdab-db.ca/>, for obtaining the nanobody-protein crystal structures. Currently, there are 195 protein data bank (PDB) entries in sd-AB database. However, some of these are for single nanobody structures, some are for the same protein with small differences of the nanobodies, and some have missing parts. We eliminated those and chose 36 complexes as our data set.

In the first stage of calculations, we used the FFT algorithm of the software ZDOCK,¹¹ version 2.3.2 to generate the decoys. Each decoy

had a different docking pose. Generally, a set of 100 decoys were sufficient to include the structure closest to the known crystal structure. In only two cases, 4CDG and 4AQ1, more than 100 decoys were necessary as we report as follows. Among the various docking programs, we specifically chose ZDOCK because we found it extremely fast, accurate, and user-friendly. The method of generating these poses is described in detail in the original paper by Chen and Weng.³⁰ The algorithm rotates the ligand around the centrally fixed protein by 15°, around each coordinate axis. Following each rotation, it translates the ligand, using the FFT algorithm, along grids with resolution of 1.2 Å constructed on the surface of the protein. We then calculated the DI energy²⁴ for every decoy, using the software Biovia Discovery Studio program,³⁷ which is obtained as the difference between the energy of the complex and the individual energies of unbound protein and nanobody.

$$\begin{aligned} \text{DI energy} = & \text{Dreiding energy of the complex} & (1) \\ & - (\text{Dreiding energy of the target protein} \\ & + \text{Dreiding energy of the nanobody}). \end{aligned}$$

All reported DI energies are in kJ/mol. The protein and the nanobody are taken as in their crystal PDB structures in the absence of water and hydrogens are removed. The DI energy for each decoy calculated in this manner made it possible for us to rank the output structures. The lowest energy decoy is rank 1, the second lowest is rank 2, and so forth. We then superposed each decoy on the known crystal structure and calculated the RMSD between the two and chose the lowest RMSD as the *best matching complex*. RMSD's are calculated using VMD's *measure RMSD* function.³⁸ The ranks and RMSD's of the best matching complexes are reported in Table 1. We also developed a script that translates and rotates a nanobody around a proposed binding pose and calculates the Dreiding energy. The input data for this script are the PDB structures of the nanobody and the protein, the magnitude of the translation of the nanobody, the location of the nanobody atom around which the nanobody will rotate and the amplitudes of translation, and the three Euler angles. For the Monte Carlo scheme, the amplitudes are then multiplied by a uniform random number, r , in the interval, $-1 \leq r \leq 1$, and the nanobody is translated and rotated by the random amount at each step. As will be discussed in detail in Section 4, once the best fitting complex is found, translating and rotating the nanobody rapidly leads to the conformation of the known crystal structure with the lowest energy. Thus, the translation-rotation process may be applied to several candidate decoys to obtain their true lowest energies.

In addition to energy, RMSD calculations and ranking, we used this dataset to identify the salt bridges in the crystal structures using the electrostatic nonbonded interactions module of Discovery Studio. Identifying the salt bridge which will form between the nanobody and the protein is crucial for simplifying the problem since it reduces the 12-dimensional phase space available to the complex to three dimensions defined by the three Euler angles.

3 | RESULTS

3.1 | Energy rank of the nearest-native structure

The list of structures, the energy rank of the best matching complex and RMSD between the decoy and the corresponding crystal structure are given in columns 2–4 of Table 1. A rank of 1 in column 3 indicates that the best matching complex to the crystal structure, that is, the lowest RMSD, has the lowest DI energy among all the decoys generated. Column 3 in Table 1 shows that 25 structures have rank 1, indicating that the decoy closest to the crystal structure already has the lowest DI energy. The closeness to the crystal structure is given by the RMSD values in column 4. Entries from 26 to 35 in Table 1 show predictions which do not have rank 1, that is, the DI energy is not the lowest for these predictions. However, the three-dimensional structures of these structures are close to those of the corresponding crystal structures and may be made to coincide by pure rigid body translation less than an Ångstrom (see later). In the fifth column of Table 1, we compared the ranks obtained using the scoring function 2.3.2 of the ZDOCK server, which is based on surface complementarity, electrostatic interactions, desolvation, buried surface area, and various intermolecular potentials as mentioned in Section 1. The ZDOCK 3.0.2 returns the decoys in a sorted order from highest score being the decoy having the lowest energy to the lowest score decoys having higher energies. A value of 10> in column 5 indicates that the best matching complex is not among the 10 highest score decoys given by the ZDOCK 2.3.2 scoring system. Column 5 shows that 24 of the 36 complexes are not in the set of first 10 decoys according to the ZDOCK 2.3.2 scoring system. Entry 36 in Table 1 required both translation and rotation for a perfect fit with crystal structure.

The Dreiding energy obtained from the crystal structure is given in column 6. Column 7 shows the Dreiding energy for the best matching complex after rigid body translation Monte Carlo cycles. Column 8 shows the available K_D values. The superscript following each entry in column 8 indicates the reference from which the data is taken. Columns 6 and 7 may be used in determining good binders of unknown complexes as will be discussed in more detail in Section 4.

We now give a few representative results from the analysis of the set of 36 known cases:

1. *The case of 3EBA*: As an example of energy ranking, we present the energy rank plot of 100 decoys of 3EBA in Figure 1. Ordinate values represent the DI energies obtained; data points are for 100 decoys where the filled circle shows the energy of the complex that best fits the crystal structure of 3EBA. The empty circles represent results for the remaining decoys.

We see that the best matching complex has the lowest DI energy, meaning its rank is 1. The intermolecular hydrogen bond strength calculated by using Discovery Studio for the decoy with rank 1 is the strongest among all other decoys. However, this is not a general trait, and we cannot conclude that all rank 1 structures have the strongest hydrogen bonding.

TABLE 1 DI energy, RMSD and experimental K_d values of 36 selected protein-nanobody complexes.

	PDB id	DI energy rank of best matching complex	RMSD (Å)	Rank according to ZDOCK	DI energy of the crystal structure (kJ/mol)	DI energy of the best matching complex ^a (kJ/mol)	Experimental K_d s (nM)
1	3EBA	1	0.84/0.16	>10	-117	-107	2.36 ³⁹
2	3K1K	1	3.12/0.38	>10	-146	-133	0.59 ± 0.11 ⁴⁰
3	3 K74	1	0.79/0.27	>10	-107	-105	34 ± 17 ⁴¹
4	3P0G	1	0.33/0.05	>10	-92	-92	0.56 ⁴²
5	4DK3	1	1.07/0.21	1	-64	-62	NA
6	4GFT	1	1.17/0.39	>10	-116	-101	NA
7	4I0C	1	0.55/0.17	>10	-56	-49	460 ⁴³
8	4KML	1	0.81/0.103	>10	-67	-67	9.4 ⁴⁴
9	4KRL	1	1.32/0.09	1	-111	-111	219 ± 20 ⁴⁵
10	4W6X	1	0.95/0.10	>10	-76	-50	5.25 ± 3.08 ⁴⁶
11	5BOP	1	1.03/0.04	1	-120	-120	NA
12	5C2U	1	0.66/0.02	1	-110	-174	NA
13	5F1K	1	0.85/0.09	1	-81	-117	NA
14	5IMM	1	1.00/0.07	>10	-145	-143	3.5 ⁴⁷
15	5O02	1	0.38/0.38	>10	-123	-93	NA
16	6B20	1	0.37/0.11	1	-138	-136	53 ⁴⁸
17	6EQI	1	0.65/0.11	>10	-167	-201	NA
18	6FV0	1	0.59/0.20	>10	-64	-62	1720 ⁴⁹
19	4Z9K	1	2.85/0.84	1	-130	-74	2.2 ⁵⁰
20	5F21	1	0.41/0.03	1	0135	-135	NA
21	5GXB	1	0.16/0.23	>10	+38	-51	15000 ⁵¹
22	5IMK	1	0.78/0.05	>10	-126	-121	850 ⁴⁷
23	5IMO	1	0.58/0.06	>10	-198	-165	3.5 ⁴⁷
24	4EIG	1	1.10/0.13	3	-115	-114	1 ⁵²
25	5G5R	1	0.77/0.19	>10	-94	-94	NA
26	4C57	2/1 ^b	0.74/0.1	10	-124	-124	59.9 ⁵³
27	3CFI	3/1 ^b	2.03/0.11	>10	-97	-97	NA
28	4Y8D	3/1 ^b	0.75/0.18	>10	-93	-96	NA
29	4CDG	6/1 ^b	0.72/0.06	>10	-76	-76	NA
30	4C58	13/1 ^b	0.53/0.06	>10	-83	-83	10.3 ⁵³
31	5BOZ	8/1 ^b	0.67/0.22	1	-103	-146	0.7 ⁵⁴
32	4I1N	8/1 ^b	1.03/0.17	>10	-72	-74	NA
33	5HVF	13/1 ^b	0.37/0.04	4	-137	-140	0.03 ⁵⁵
34	5F7K	20/1 ^b	0.79/0.09	>10	-100	-98	NA
35	4W6Y	7/1 ^b	1.06/0.43	1	-153	-120	1.58 ⁴⁶
36	4AQ1	37 ^c	7.13/1.64	>10	-25	-49	NA

Note: Cases 29 and 36, 4CDG and 4AQ1, required generating more than 100 decoys. The decoy numbers used for translation-rotation cycles are 612 and 239, respectively.

Abbreviations: DI, Dredging interaction; NA, Not available; PDB, protein data bank; RMSD, root mean square deviation.

^aRigid body translation Monte Carlo cycles were applied to the best matching complexes before calculating the DI energies. Translation and rotation Monte Carlo cycles were applied only for sample 36.

^bA rigid body translation in which the centroid of ligand moves in a sphere of radius less than 1 Å matches the structure with the known crystal structure.

^cBoth rotation and translation are required to match the predicted structure with the known crystal structure.

2. *The case of 3CFI*: As another example, we present the energy rank plot of 3CFI decoys in Figure 2. Column 3 for 3CFI entry in Table 1, 3/1^b, indicates that the prediction closest to crystal

structure had energy rank 3 but small rigid body translations resulted in almost perfect superposition with crystal structure and improved the DI energy to rank 1. The entry 2.03/0.46 in column

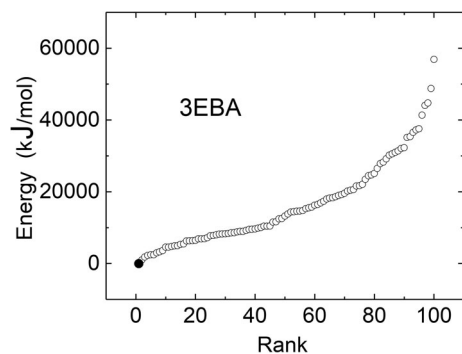


FIGURE 1 DI energy of 100 decoys from ZDOCK for the protein-nanobody complex, 3EBA, is plotted as a function of the rank, where the filled circle shows the DI energy of the best matching complex, meaning the decoy having the lowest RMSD from the crystal structure. DI, Dreding interaction; RMSD, root mean square deviation.

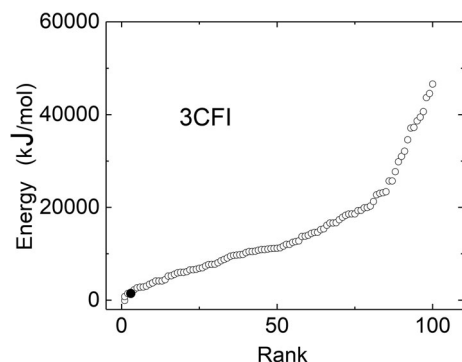


FIGURE 2 DI energy of 100 decoys from ZDOCK for the protein-nanobody complex, 3CFI, is plotted as a function of the rank, where the filled circle shows the DI energy of the best matching complex, meaning the decoy having the lowest RMSD from the crystal structure. DI, Dreding interaction; RMSD, root mean square deviation.

4 of Table 1 for 3CFI indicates that the RMSD of rank 3 decoy, the filled circle, was 2.03 Å and dropped to 0.46 Å after a rigid body translation and the value of its DI energy decreased to -99 kJ/mol upon the rigid body translation. Rigid body translation was applied to several decoys of low energy of Figure 2, but none went below -99 kJ/mol. The next lowest binding energy was -42 kJ/mol. Thus, rigid body translation decreases the energy, but none goes below that of the best matching structure. It should be added that this energy, although negative, does not represent the true energy because the contribution of water and hydrogens are missing.

The 10 decoys indicated by a variable “b” in column 3 of Table 1, have relatively low RMSD's as shown by the values to the left of the slashes in column 4 and can be aligned by small rigid body translation of the centroid within a sphere of radius 1 Å. As a result of this operation, the DI energy then becomes rank 1 for each, with RMSD's shown on the right side of the slashes. The need to translate is because of the size of the grid resolution in generating the decoys,



FIGURE 3 Rank 3 decoy for 3CFI superposed on its crystal structure. Nanobody of the decoy is in green. Crystal structure nanobody is in blue and the target protein is in orange.

which is 1.2 Å in the software we used. The translation operation that is needed to overcome the grid resolution effect is simple and straightforward and may be programmed such that the DI energy is recorded for each small step translation and the structure showing the lowest DI energy is accepted. This is the procedure we followed for the entries between 26 and 35 in Table 1. The conformation of the decoy of rank 3 is presented in Figure 3 where the nanobody of the decoy is highlighted in yellow. Translating the nanobody by 1.42 Å as a rigid body result in overlap with crystal structure. A more detailed explanation of the translation operation is explained in Section 4.

3. *The case of 4AQ1:* The minimum RMSD decoy for 4AQ1 is shown in the left panel of Figure 4 where the nanobody is colored in blue. Due to the size and complexity of the protein, 100 decoys were not sufficient to make the prediction and a set of 300 decoys were used. The decoy that is closest to the crystal structure is the decoy 239 out of 300, with a RMSD of 7.13 Å. When the 300 decoys are ordered with respect to their Dreding energies, this decoy is of rank 37 and cannot be superposed on the crystal structure by pure translation only, and rotation is also required. There is a salt bridge between GLU560 and ARG45 in decoy 239 where the distance between alpha carbons is 5.38 Å. In the crystal structure, there is a salt bridge between LYS565 of the protein and GLU44 of the nanobody where the distance between the alpha carbons is 9.73 Å. The two neighboring pairs LYS565-GLU44 and GLU560-ARG45 may be seen as nuclei of binding according to which a salt bridge between the nanobody and the protein is established first and the nanobody rotates and translates about the salt bridge until its energy is lowered. We first used Monte Carlo translation-rotation cycles on decoy 239 around GLU560 and ARG45. The structure shown on the right panel in Figure 4 is the structure obtained by rigid body translation and rotation about the nitrogen of ARG45 until its DI energy does not decrease further. This operation requires about 1000 Monte Carlo steps of translation and 1000 steps of rotation. The DI energy at the end of these steps is -49 kJ/mol.

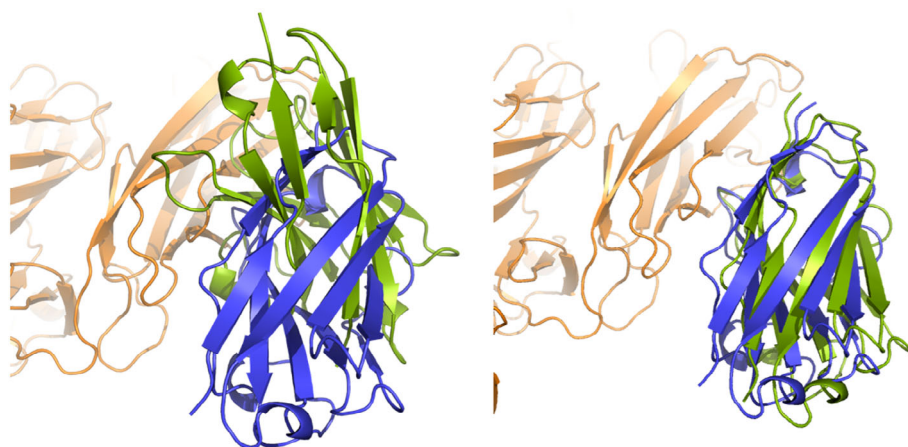


FIGURE 4 (A) Decoy 239 for 4AQI is superposed on 4AQI crystal structure, nanobody of the decoy is highlighted in green and the nanobody from the crystal structure is colored in blue. (B) Decoy 239 is superposed on the crystal structure after performing 1000 Monte Carlo steps of translation and 1000 Monte Carlo steps for rotation about the nitrogen atom of ARG45 of the nanobody, where the nanobody is highlighted in green. (See legend to Figure 3 for color scheme).

Second, we manually translated the decoy 239 such that GLU44 is in the vicinity of LYS565 of the protein. This translation can be done easily using Discovery Studio such that the side chain oxygen atom of GLU44 is approximately at a distance of 3 Å from the side chain nitrogen atom of LYS565. We then applied translation–rotation cycles, 2000 Monte Carlo steps each, until a minimum binding energy of -52 kJ/mol was obtained. However, this minimum energy structure deviated from the crystal structure by a RMSD of 1.64 Å shown in Figure 5. The DI energy of the nanobody in the crystal structure is -24 kJ/mol which is higher than the Monte Carlo translated rotated structure. This may not reflect the real picture and (1) the presence of water and/or (2) deformation of side chains may make the crystal structure more favorable. These two factors have not been thoroughly investigated in the literature. Yet, the conformational deviations due to these two effects are small as verified by Figure 5.

3.2 | Optimizing the DI energy with respect to translations and rotations using Delaunay triangulation

The Monte Carlo search algorithm used in obtaining Figure 5 is quite expensive time-wise due to the ruggedness of the energy surface. Alternatively, using gradient based minimization can be a difficult task also, as the presence of multiple local minima and high-energy barriers may force the technique to become trapped in a local minima. This can prevent the algorithm from reaching the global minima and can result in an inaccurate or incomplete energy surface representation. Also, the gradient descent algorithm relies on the slope of the energy surface to update the position of the search, and the gradients can become small or nonexistent in regions of the energy surface that are nearly flat or have many local minima. In this section, we adopt a simplified visual technique for rapidly identifying the true pose of the decoy 239 shown in Figure 5. The first simplification that we adopt is to apply translations and rotations independently, first translations followed by rotations. If a residue of the nanobody makes a salt bridge with a residue of the protein, we can manually perform translations such that the positively charged nitrogen atom of a basic amino acid residue (such as lysine or arginine, for example) in the protein and the



FIGURE 5 The nanobody of the decoy 239 is superposed on the nanobody crystal structure of 4AQI, after performing a second Monte Carlo translation–rotation cycle of 2000 steps each about the oxygen atom of GLU44 of the nanobody, where the decoy nanobody is colored in green and the nanobody from the crystal structure is colored in blue.

negatively charged oxygen atom of an acidic group (such as carboxylate) in the ligand are separated by a distance of approximately 3–4 Å, and form an electrostatic interaction that stabilizes the protein–nanobody complex. Finding the three Euler angles that lead to the true pose is more challenging. To confine the search to a small number of energy calculations, we use Delaunay triangulation which is a method that can be used to visualize the energy landscape of a system with a rugged energy surface and identify local energy minima. It is a computational geometry technique that constructs a triangulation of a set of points such that no point is inside the circumcircle of any triangle in the triangulation. This triangulation can be used to construct contours of constant energy and locate local energy minima.

The step-by-step procedure for using Delaunay triangulation to find energy minima is as follows:

1. Generate a set of candidate random configurations for the system defined by the values of the three Euler coordinates.

2. Compute the DI energy of each configuration.
3. Construct the Delaunay triangulation of the configuration space using the set of candidate configurations as input. The vertices of the triangles correspond to the candidate configurations, and the edges represent connections between the configurations.
4. Assign the energy value of each configuration to the corresponding vertex in the triangulation.
5. Construct a contour plot of the energy landscape using the triangulation. This involves identifying the triangles in the triangulation that intersect with a given energy level and plotting the corresponding vertices as a contour line.
6. Locate the local energy minima by identifying the vertices of the triangulation that have the lowest energy values. These correspond to the energy minima of the system.

We use the Contour Plot module of the software OriginPro,⁵⁶ which draws contours of DI values for each given pair of Euler angles theta-phi, theta-psi, and phi-psi. The contour lines identify, approximately, the locations of the minima in the corresponding diagrams.

We applied this method to obtain the minimum energy configurations of decoy 239 of 4AQI where LYS565 of the protein and GLU44 of the nanobody are known to make a salt bridge. In the first step, we manually translated the nanobody such that the nitrogen of LYS565 and the oxygen of GLU44 are separated by 3.5 Å. We then generated 432 conformations of the nanobody around the nitrogen of LYS565 by randomly changing the three Euler angles theta, phi, and psi and calculated the corresponding DI energies. The contour lines of

Delaunay triangulation are presented in three panels in Figure 6. The triangulation was applied separately for each pair of the Euler angles. The top panel in Figure 6 is for the theta-phi pair. The lower left and right panels are for the theta-psi and phi-psi pairs, respectively. The black regions are the regions of minima. The three white circles denote the values of the Euler angle pairs for the minimum. These are: theta = 1.23 radians, phi = 1.22 radians, and psi = -0.60 radians. These three Euler angles give the true pose of the nanobody shown in Figure 5. We note that there are several minima in Figure 6.

The challenges identified in the Monte Carlo and energy gradient methods also apply to the Delaunay triangulation method. Figure 6 shows that several sets of triple Euler angles, similar to the three white circles, can be identified in the relatively large minima regions. To determine which of these sets has the lowest binding energy, it is necessary to construct the three-dimensional structure of each nanobody-protein complex using the given three Euler angles and then evaluate the DI values for these configurations.

4 | DISCUSSION

The study of 36 nanobody-protein complexes from the PDB showed that the Dreiding force field is a good metric in the prediction of the crystal structure from among a large number of decoys. For each complex, 100, (300 for one case), decoys were generated by ZDOCK which have favorable shape complementarity, electrostatic interactions, desolvation, and buried surface area scores according to the ZDOCK scoring

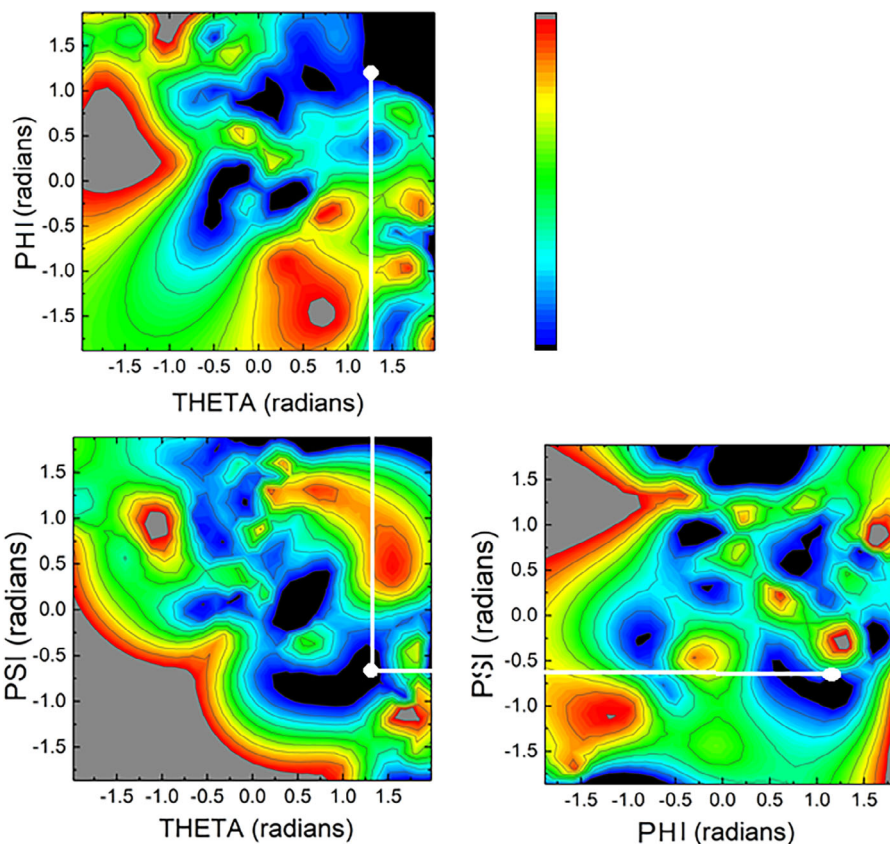


FIGURE 6 The contour of Dreiding interaction (DI) energy values for each given pair of Euler angles theta-phi (upper left panel), theta-psi (lower left panel), and phi-psi (lower right panel), which were generated by random set of translation and rotations applied on 4AQI around the salt bridge forming residues LYS565 of the protein and GLU44 of the nanobody. The three white circles show theta-phi-psi pairs which show a minimum DI energy. These minima angles are theta = 1.23 radians, phi = 1.22 radians, and psi = -0.60 radians, respectively. Color scale common to all three panels is also shown.

system. Out of 36 PDB structures, 25 of them were predicted as rank 1 when ranked according to the Dreiding energy. Eleven of the remaining structures required Ångstrom size rigid body displacements of the nanobody relative to the protein to match the given PDB structure. After the translation the complexes all became rank 1. In one case, neither rigid body rotations nor translations of the nanobody resulted in matching with the crystal structure. This was partly due to the presence of two minima of approximately the same energies. We used a Monte Carlo algorithm that randomly translates and rotates the nanobody of a decoy and calculates the Dreiding energy and the unbiased Monte Carlo simulation did not prefer the crystal structure. This script was applied to the cases that were not rank 1 and sub-Ångstrom size rigid body translation was sufficient for matching.

The observation that all nanobody-protein complexes exhibit at least one salt bridge suggests that salt bridges play essential role on the binding process. Statistical mechanical arguments given earlier²⁷ have suggested that binding is initiated by anchoring of the ligand at a suitable site on the protein followed by rotational rearrangements of the ligand around the fixed anchor point. The identification of the salt bridging residue pairs by the ligand constitutes the onset of the binding funnel²⁶ in which the nanobody exhibits small translational and rotational rearrangements until it locks into the protein. This idea forms the basis of the Monte Carlo script that we used in the present study. Among the 98 nanobody-protein complexes whose crystal structures are listed in sd-Ab DB and listed in the Table S1, all but five complexes exhibit at least one salt bridge. The reason why five complexes do not have salt bridges is because there are no charged residues in the three CDR loops of the nanobodies. Most of the bonding in these five exceptional cases is through polar residues. The distribution of salt bridge pairs with respect to residue type is presented in Table 2.

The C^α-C^α distances of salt bridge forming residues lie in the range 8–12 Å. More importantly, the distance between the donor and acceptor atoms on the side chains is approximately 3 Å. The ZDOCK decoys all exhibit salt bridges, which point to the anchor point around which the nanobody will fluctuate. Fixing the anchor point of the ligand reduces the phase space significantly. Thus, selecting the salt bridge at the beginning of the docking and using Monte Carlo translations and rotations may be a useful simplifying strategy. The total number, N_{SB} , of possible salt bridges between a protein and a nanobody is

$$N_{SB} = (L_R + L_K)(P_D + P_E) + (L_D + L_E)(P_R + P_K), \quad (2)$$

where, L and R denote the ligand and the protein and the subscripts R , K , D , and E denote Arg, Lys, Asp, and Glu, respectively, on the right hand side of Equation (1). Of course, the ZDOCK decoy predictions followed by DI energy calculations eliminate most of the N_{SB}

TABLE 2 Number of salt bridge forming residue pairs from 98 nanobody-protein complexes listed in sd-Ab DB.

	ASP	GLU
ARG	57	71
LYS	33	16

structures and the prediction algorithm may reduce to only few Monte Carlo simulations around each of the designated candidate decoy.

The DI energies of the complexes with available experimental K_D values, column 7 of Table 1, are plotted as a function of available K_D 's, column 8. The solid line in Figure 7 is the linear fit through the points given by the equation

$$\Delta E_D = (-129.3 \pm 11.1) \text{Ln} k_D + 8.8, \quad (3)$$

where ΔE_D is the DI binding energy. Solving for K_D leads to

$$k_D = e^{\left(\frac{\Delta E_D + 129.3 \pm 11.1}{8.8}\right)}. \quad (4)$$

Equation (3) is plotted in Figure 8 where the ordinate is in logarithmic scale and the abscissa values span the range $-50 \leq \Delta E_D \leq -200$ kJ/mol which covers the scale from poor to excellent binding. This figure may be used as a master chart to assess the quality of binding in

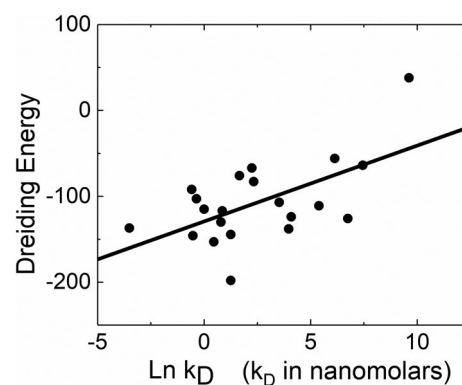


FIGURE 7 Dreiding interaction energy (kJ/mol) of complexes with available experimental K_D values are plotted against their experimental $\text{Ln}K_D$ (nM) values.

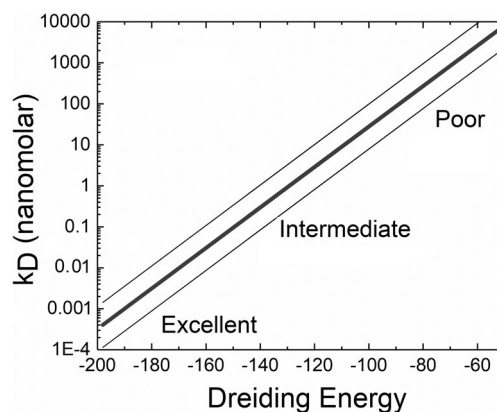


FIGURE 8 The K_D values are plotted against the Dreiding interaction (DI) energy using Equation (4). Ordinate scale is logarithmic. K_D of a given protein-nanobody complex whose DI energy is known can be obtained from this figure.

unknown nanobody-protein pairs. The DI energy obtained from the Monte Carlo translation-rotation operation on low-rank decoys will give the corresponding K_D values.

The findings of the present work may be used as design principles in nanobody engineering. We now discuss two cases:

Case 1. Crystal structures of the nanobody and the protein are known but the structure of the complex is not known. We run a ZDOCK simulation and obtain several decoys, one of which is possibly the true solution. We determine the DI energies of all the decoys and rank them. Starting from the lowest rank, we perform Monte Carlo translation-rotation simulations and obtain the lowest DI energy decoy. This will possibly be the structure closest to the true structure.

Case 2. The structure of the complex is known, and modifications are required to increase the binding affinity: We repeat the steps of Case 1 for each mutation and obtain the binding pose and the DI energy of the structure closest to the true structure.

The bottleneck in the above workflow is the Monte Carlo step. In an ordinary laptop, 2000 step Monte Carlo cycle of diverse structures and the determination of the DI energy is in the order of 1 h.

AUTHOR CONTRIBUTIONS

Aysima Hacisuleyman: Conceptualization; investigation; writing – original draft; software; formal analysis; methodology; writing – review and editing. **Burak Erman:** Conceptualization; investigation; writing – original draft; software; formal analysis; methodology; writing – review and editing; project administration.

PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/prot.26529>.

ORCID

Aysima Hacisuleyman  <https://orcid.org/0000-0003-2471-0296>

Burak Erman  <https://orcid.org/0000-0002-2496-6059>

REFERENCES

- Muyldermans S. A guide to: generation and design of nanobodies. *FEBS J.* 2021;288(7):2084-2102.
- Holliger P, Hudson PJ. Engineered antibody fragments and the rise of single domains. *Nat Biotechnol.* 2005;23(9):1126-1136.
- Muyldermans S. Nanobodies: natural single-domain antibodies. *Annu Rev Biochem.* 2013;82:775-797.
- Karvouni M, Vidal-Manrique M, Lundqvist A, Alici E. Engineered NK cells against cancer and their potential applications beyond. *Front Immunol.* 2022;13:825979.
- Deszyński P, Młokosiewicz J, Volanakis A, et al. INDI—integrated nanobody database for immunoinformatics. *Nucleic Acids Res.* 2022; 50(D1):D1273-D1281.
- Poustforoosh A, Faramarz S, Negahdaripour M, Hashemipour H. Modeling and affinity maturation of an anti-CD20 nanobody: a comprehensive in-silico investigation. *Sci Rep.* 2023;13(1):1-11.
- Kok BH, Lim HT, Lim CP, Lai NS, Leow CY, Leow CH. Dengue virus infection—a review of pathogenesis, vaccines, diagnosis and therapy. *Virus Res.* 2022;324:199018.
- Bessalah S, Jebahi S, Mejri N, Salhi I, Khorchani T, Hammadi M. Perspective on therapeutic and diagnostic potential of camel nanobodies for coronavirus disease-19 (COVID-19). *3 Biotech.* 2021;11(2):1-14.
- Zare H, Aghamollaei H, Hosseindokht M, Heiat M, Razei A, Bakherad H. Nanobodies, the potent agents to detect and treat the coronavirus infections: a systematic review. *Mol Cell Probes.* 2021;55: 101692.
- Sasisekharan R. Preparing for the future—nanobodies for covid-19? *N Engl J Med.* 2021;384(16):1568-1571.
- Pierce BG, Wiehe K, Hwang H, Kim B-H, Vreven T, Weng Z. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics.* 2014;30(12):1771-1773.
- Comeau SR, Gatchell DW, Vajda S, Camacho CJ. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics.* 2004;20(1):45-50.
- Gabb HA, Jackson RM, Sternberg MJ. Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J Mol Biol.* 1997;272(1):106-120.
- Roberts VA, Thompson EE, Pique ME, Perez MS, Ten Eyck L. DOT2: macromolecular docking with improved biophysical models. *J Comput Chem.* 2013;34(20):1743-1758.
- Matsuzaki Y, Uchikoga N, Ohue M, et al. MEGADOCK 3.0: a high-performance protein-protein interaction prediction software using hybrid parallel computing for petascale supercomputing environments. *Source Code Biol Med.* 2013;8(1):1-8.
- Zhang C, Lai L. SDOCK: a global protein-protein docking program using stepwise force-field potentials. *J Comput Chem.* 2011;32(12): 2598-2612.
- Tovchigrechko A, Vakser IA. Development and testing of an automated approach to protein docking. *Proteins.* 2005;60(2):296-301.
- Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proc Natl Acad Sci.* 1992;89(6):2195-2199.
- Bajaj CL, Chowdhury R, Siddahanavalli V. F^2 dock: fast fourier protein-protein docking. *IEEE/ACM Trans Comput Biol Bioinform.* 2009;8(1):45-58.
- Mandell JG, Roberts VA, Pique ME, et al. Protein docking using continuum electrostatics and geometric fit. *Protein Eng.* 2001;14(2):105-113.
- Li L, Guo D, Huang Y, Liu S, Xiao Y. ASPDock: protein-protein docking algorithm using atomic solvation parameters model. *BMC Bioinformatics.* 2011;12(1):1-9.
- Ritchie DW, Kemp GJ. Protein docking using spherical polar Fourier correlations. *Proteins.* 2000;39(2):178-194.
- Garzon JI, López-Blanco JR, Pons C, et al. FRODOCK: a new approach for fast rotational protein-protein docking. *Bioinformatics.* 2009;25(19):2544-2551.
- Mayo SL, Olafson BD, Goddard WA. DREIDING: a generic force field for molecular simulations. *J Phys Chem.* 1990;94(26):8897-8909.
- Wilton EE, Opyr MP, Kailasam S, Kothe RF, Wieden H-J. *sdAb-DB: The Single Domain Antibody Database.* ACS Publications; 2018.
- Gray JJ, Moughon S, Wang C, et al. Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J Mol Biol.* 2003;331(1):281-299.
- Hill TL. *Cooperativity Theory in Biochemistry: Steady-State and Equilibrium Systems.* Springer Science & Business Media; 2013.
- Fernández-Recio J, Totrov M, Abagyan R. Soft protein-protein docking in internal coordinates. *Protein Sci.* 2002;11(2):280-291.
- Palma PN, Krippahl L, Wampler JE, Moura JJ. BiGGER: a new (soft) docking algorithm for predicting protein interactions. *Proteins.* 2000; 39(4):372-384.
- Chen R, Weng Z. Docking unbound proteins using shape complementarity, desolvation, and electrostatics. *Proteins.* 2002;47(3):281-294.

31. Bates PA, Kelley LA, MacCallum RM, Sternberg MJ. Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. *Proteins*. 2001;45(S5):39-46.
32. Kutikhin AG, Yuzhalin AE, Tsitko EA, Brusina EB. Pattern recognition receptors and DNA repair: starting to put a jigsaw puzzle together. *Front Immunol*. 2014;5:343.
33. Huber EM, Groll M. The 19S cap puzzle: a new jigsaw piece. *Structure*. 2012;20(3):387-388.
34. Koh IY, Eylich VA, Marti-Renom MA, et al. EVA: evaluation of protein structure prediction servers. *Nucleic Acids Res*. 2003;31(13):3311-3315.
35. Kumar S, Nussinov R. Close-range electrostatic interactions in proteins. *ChemBiochem*. 2002;3(7):604-617.
36. Schreiber G, Fersht AR. Rapid, electrostatically assisted association of proteins. *Nat Struct Biol*. 1996;3(5):427-431.
37. Dassault Systèmes BIOVIA, D. S. M. E. *Release 2017*. Dassault Systèmes; 2016.
38. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996;14(1):33-38.
39. Vincke C, Loris R, Saerens D, Martinez-Rodriguez S, Muyldermans S, Conrath K. General strategy to humanize a camelid single-domain antibody and identification of a universal humanized nanobody scaffold. *J Biol Chem*. 2009;284(5):3273-3284.
40. Kirchofer A, Helma J, Schmidhals K, et al. Modulation of protein properties in living cells using nanobodies. *Nat Struct Mol Biol*. 2010;17(1):133-138.
41. Oyen D, Srinivasan V, Steyaert J, Barlow JN. Constraining enzyme conformational change by an antibody leads to hyperbolic inhibition. *J Mol Biol*. 2011;407(1):138-148.
42. Rasmussen SG, Choi H-J, Fung JJ, et al. Structure of a nanobody-stabilized active state of the β_2 adrenoceptor. *Nature*. 2011;469(7329):175-180.
43. De Genst E, Chan P-H, Pardon E, et al. A nanobody binding to non-amyloidogenic regions of the protein human lysozyme enhances partial unfolding but inhibits amyloid fibril formation. *J Phys Chem B*. 2013;117(42):13245-13258.
44. Abskharon RN, Giachin G, Wohlkonig A, et al. Probing the N-terminal β -sheet conversion in the crystal structure of the human prion protein bound to a nanobody. *J Am Chem Soc*. 2014;136(3):937-944.
45. Schmitz KR, Bagchi A, Roovers RC, en Henegouwen PMVB, Ferguson KM. Structural evaluation of EGFR inhibition mechanisms for nanobodies/VHH domains. *Structure*. 2013;21(7):1214-1224.
46. Moonens K, De Kerpel M, Coddens A, et al. Nanobody mediated inhibition of attachment of F18 fimbriae expressing *Escherichia coli*. *PLoS One*. 2014;9(12):e114691.
47. Wen Y, Ouyang Z, Schoonoghe S, et al. Structural evaluation of a nanobody targeting complement receptor Vsig4 and its cross reactivity. *Immunobiology*. 2017;222(6):807-813.
48. Gulati S, Jin H, Masuho I, et al. Targeting G protein-coupled receptor signaling at the G protein level with a selective nanobody inhibitor. *Nat Commun*. 2018;9(1):1-15.
49. Pernigo S, Chegkazi MS, Yip YY, et al. Structural basis for isoform-specific kinesin-1 recognition of Y-acidic cargo adaptors. *elife*. 2018;7:e38362.
50. Rudolph MJ, Vance DJ, Cassidy MS, Rong Y, Shoemaker CB, Mantis NJ. Structural analysis of nested neutralizing and non-neutralizing B cell epitopes on ricin toxin's enzymatic subunit. *Proteins*. 2016;84(8):1162-1172.
51. Jiang X, Smirnova I, Kasho V, et al. Crystal structure of a LacY-nanobody complex in a periplasmic-open conformation. *Proc Natl Acad Sci*. 2016;113(44):12420-12425.
52. Oyen D, Wechselberger R, Srinivasan V, Steyaert J, Barlow JN. Mechanistic analysis of allosteric and non-allosteric effects arising from nanobody binding to two epitopes of the dihydrofolate reductase of *Escherichia coli*. *Biochim Biophys Acta*. 2013;1834(10):2147-2157.
53. Chaikwad A, Keates T, Vincke C, et al. Structure of cyclin G-associated kinase (GAK) trapped in different conformations using nanobodies. *Biochem J*. 2014;459(1):59-69.
54. Rudolph MJ, Vance DJ, Cassidy MS, Rong Y, Mantis NJ. Structural analysis of single domain antibodies bound to a second neutralizing hot spot on ricin toxin's enzymatic subunit. *J Biol Chem*. 2017;292(3):872-883.
55. Zhou X, Weeks S, Ameloot P, Callewaert N, Strelkov S, Declerck P. Elucidation of the molecular mechanisms of two nanobodies that inhibit thrombin-activatable fibrinolysis inhibitor activation and activated thrombin-activatable fibrinolysis inhibitor activity. *J Thromb Haemost*. 2016;14(8):1629-1638.
56. Seifert E. OriginPro 9.1: scientific data analysis and graphing software—software review. *J Chem Inf Model*. 2014;54(5):1552.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Hacisuleyman A, Erman B. Fine tuning rigid body docking results using the Dreiding force field: A computational study of 36 known nanobody-protein complexes. *Proteins*. 2023;91(10):1417-1426. doi:[10.1002/prot.26529](https://doi.org/10.1002/prot.26529)