PLOS | ONE

RESEARCH ARTICLE

Structural Rigidity and Protein Thermostability in Variants of Lipase A from *Bacillus subtilis*

Prakash Chandra Rathi[1], **Karl-Erich Jaeger**[2,3], **Holger Gohlke**[1]*

1 Institute of Pharmaceutical and Medical Chemistry, Heinrich-Heine-University, Düsseldorf, Germany, **2** Institute of Molecular Enzyme Technology, Heinrich-Heine-University, Düsseldorf, Germany, **3** Institute of Bio- and Geosciences IBG-1: Biotechnology, Research Centre Jülich, Jülich, Germany

* gohlke@uni-duesseldorf.de

Citation: Rathi PC, Jaeger K-E, Gohlke H (2015) Structural Rigidity and Protein Thermostability in Variants of Lipase A from *Bacillus subtilis*. PLoS ONE 10(7): e0130289. doi:10.1371/journal.pone.0130289

Editor: Annalisa Pastore, National Institute for Medical Research, Medical Research Council, London, UNITED KINGDOM

Received: October 1, 2014

Accepted: May 18, 2015

Published: July 6, 2015

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The study was supported by the Ministry of Innovation, Science, and Research of North Rhine-Westphalia and Heinrich-Heine-University Düsseldorf (HHU) by a scholarship to PCR within the CLIB2021 Graduate Cluster Industrial Biotechnology. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Understanding the origin of thermostability is of fundamental importance in protein biochemistry. Opposing views on increased or decreased structural rigidity of the folded state have been put forward in this context. They have been related to differences in the temporal resolution of experiments and computations that probe atomic mobility. Here, we find a significant ($p = 0.004$) and fair ($R^2 = 0.46$) correlation between the structural rigidity of a well-characterized set of 16 mutants of lipase A from *Bacillus subtilis* (*Bs*LipA) and their thermodynamic thermostability. We apply the rigidity theory-based Constraint Network Analysis (CNA) approach, analyzing directly and in a time-independent manner the statics of the *Bs*LipA mutants. We carefully validate the CNA results on macroscopic and microscopic experimental observables and probe for their sensitivity with respect to input structures. Furthermore, we introduce a robust, local stability measure for predicting thermodynamic thermostability. Our results complement work that showed for pairs of homologous proteins that raising the structural stability is the most common way to obtain a higher thermostability. Furthermore, they demonstrate that related series of mutants with only a small number of mutations can be successfully analyzed by CNA, which suggests that CNA can be applied prospectively in rational protein design aimed at higher thermodynamic thermostability.

Introduction

Sufficiently high thermostability of proteins is important for both organisms living in high temperature environments and for biotechnological applications where enzymes are used as biocatalysts under often harsh reaction conditions [1, 2]. From a mechanistic point of view, "protein thermostability" embraces at least two different meanings [3, 4]: (1) *thermodynamic* thermostability describes the folded-unfolded equilibrium of a protein, and (2) *kinetic* thermostability refers to the length of time a protein remains active before undergoing irreversible denaturation at an elevated temperature. Several factors have been frequently attributed to elevated protein thermostability including improved hydrogen bonding [5], ion pair and salt

bridge networks [6], better hydrophobic packing [7], shortened loops [8], and higher secondary structure content [9], in all favoring an increased structural rigidity of the folded state [10–13]. As an opposing view, proteins from thermophilic organisms have been reported to be as flexible as or even more flexible than homologs from mesophilic organisms [14–17].

These different views on the relation between protein thermostability and structural rigidity have been a matter of ongoing discussion [10, 18–23]. In particular, it has been argued that atomic movements, which are the primary mobility data from which information on protein statics (rigidity and flexibility) is derived, cover a wide range of timescales within a protein [15, 24, 25]. Hence, depending on the temporal resolution of the experimental technique or computational analysis used to detect such movements, (parts of) a protein can come out as rigid or flexible [26–32]. Here, we address the question of the relation between protein thermostability and structural rigidity by analyzing *directly* the static properties of a well-characterized set of 16 mutants of lipase A from *Bacillus subtilis* (*Bs*LipA). We do so by applying the rigidity theory-based Constraint Network Analysis (CNA) approach developed by us [33–35], thereby considering the *Bs*LipA variants to be in static equilibrium, hence avoiding that the results depend on the temporal resolution of the approach.

*Bs*LipA is an important member of the lipase class of enzymes and used in diverse biotechnological applications [36, 37]. Owing to its importance, *Bs*LipA has been extensively studied with respect to structure [38–41] and thermostability [42–48]. As to the latter, Reetz *et al.* applied iterative saturation mutagenesis on the most flexible amino acids as identified by crystallographic B-factors, which resulted in *Bs*LipA mutants that were more thermostable than the wild type showing an increase in T_{50}^{60} (the temperature required to reduce the initial enzymatic activity by 50% within 60 min) of $\leq$ 45 K [42]. Subsequent biophysical characterization of the three most thermostable mutants revealed that the improved activity retention resulted from a reduced rate of protein unfolding and a reduced precipitation of the unfolding intermediates, i.e., due to kinetic reasons [49]. In contrast, Rao *et al.* sequentially developed several thermostable *Bs*LipA mutants using directed evolution assisted by structural information. These mutants were shown to be more thermostable than the wild type due to predominantly thermodynamic reasons [44–48, 50]; the most thermostable mutant displayed an increase in the melting temperature T_m of ~22 K.

In the CNA approach, a protein is modeled as a constraint network where bodies (representing atoms) are connected by sets of bars (constraints, representing covalent and noncovalent interactions) [51]. A rigidity analysis performed on the network [52, 53] results in a decomposition into rigid parts and flexible links in between. By analyzing a series of "perturbed" networks in which noncovalent interactions are included in a temperature-dependent manner [11, 13, 54], the loss of rigidity of a protein is simulated, which can be related to thermal unfolding [12, 13, 54]. Results of these analyses can be linked to biologically relevant characteristics of a biomolecular structure by a set of global and local indices [55]. In particular, a phase transition point T_p can be identified during the thermal unfolding simulation at which a largely rigid network becomes almost flexible; this phase transition point has been related to the thermodynamic thermostability of a protein [11–13]. For improving the robustness of the analyses, the rigidity analyses are performed on ensembles of network topologies (ENT^{FNC}) [56]. That way, thermal fluctuations of a protein are considered without actually sampling conformations.

The main outcome of this work is the finding of a significant and good correlation between the structural rigidity of all *Bs*LipA variants and their thermodynamic thermostability. On the way, we carefully probed for the sensitivity of the results with respect to the input structures and developed an approach for detecting outliers based on differences in the pathways of thermal unfolding. We furthermore introduced a local stability measure for predicting

thermodynamic thermostability, which complements the detection of the (global) phase transition point T_p. As the *Bs*LipA variants are sequentially closely related, these results have important implications for applying CNA in a prospective manner in rational protein design aimed at higher thermodynamic thermostability. Finally, we discuss our results in terms of potentially different mechanisms underlying the increased protein thermostabilities of mutants isolated by Reetz *et al.*and Rao *et al.*

Materials and Methods

Data set

The wild type structure of *Bs*LipA with the highest resolution (PDB ID: 1ISP; resolution = 1.3 Å) was obtained from the Protein Data Bank (PDB; www.pdb.org) [57]. For probing the sensitivity of the CNA results on the conformation of the input structures, five additional crystal structures of wild type *Bs*LipA were analyzed (PDB IDs: 1I6W, 1R4Z, 1R50, 2QXT, 2QXU). We included in our study all mutants from Rao *et al.* for which T_m values were determined [44–48]. In addition, we included the three most thermostable mutants developed in the last rounds of iterative saturation mutagenesis by Reetz *et al.* [42]. Models of mutant structures for which crystal structures were not available in the PDB were generated with the SCWRL program [58], using the respective *Bs*LipA structure as a template that is closest in sequence to the mutant. SCWRL constructs mutant models by predicting backbone-dependent side chain conformations with the help of a rotamer library; coordinates of backbone atoms remain unchanged. Conformations of side chains of all residues within 8 Å of a mutated residue were re-predicted in order to allow for a local structural relaxation. For all structures, hydrogen atoms were added using REDUCE [59]; side chains of Asn, Gln, and His were flipped in this stage if necessary to optimize the hydrogen bond network. All water molecules, buffer ions, and crystal solvents were removed from the structures. Finally, all structures were minimized by 5000 steps of conjugate gradient minimization (including an initial steepest descent minimization for 100 steps) or until the root mean-square gradient of the energy was $< 1.0 \cdot 10^{-4}$ kcal mol^{-1} $Å^{-1}$. The energy minimization was carried out with Amber11 [60] using the Cornell *et al.* force field [61] with modifications for proteins (ff99SB) [62] and the GB^{OBC} generalized Born model [63]. All variants of *Bs*LipA used in this study are summarized in Table 1.

Construction of the constraint network and rigidity analysis

As described in the previous section, only the protein part was considered for network construction, i.e., all non-protein molecules including water molecules were discarded. This was done based on previous findings that including water molecules does not significantly change the rigidity analysis results [64, 65]. Proteins were modeled as constraint networks in a *body-and-bar* representation (see section "Body-and-bar networks" in S1 File) [66, 67] using the CNA software [35] that acts as a front- and back-end to the Floppy Inclusion and Rigid Substructure Topography (FIRST) program [51, 68]. Once the constraint network is built, rigidity analysis is carried out, which identifies (rigid) clusters of atoms with no internal motion and flexible links in between, using the pebble game algorithm [52, 53] as implemented in the FIRST software [51].

Thermal unfolding simulation

By sequentially removing non-covalent constraints from a network, one can simulate a loss of structural rigidity due to a temperature rise. Specifically, hydrogen bonds were removed from the network in increasing order of their strength following the idea that stronger hydrogen

Table 1. Summary of *Bs*LipA variants used in the study.

*Bs*LipA variant[a]	PDB ID[b]	Resolution[c]	Mutations	T_m (K)	T_i (K) [d]	T_{50} (K)	$\widetilde{rc}_{ij.neighbor}$ [e]		Reference
							(kcal mol^{-1})	(K)	
Wild type	1ISP	1.3	–	329.15	324.95	321.15[f]	-0.87 (-0.80)[h]	317.4 (315.9) [h]	[39, 42, 44]
IX	1ISP*	–	K112D, M134D, Y139C, I157M	–	318.75	335.95[g]	-0.67	313.5	[42, 49]
X	1ISP*	–	R33Q, D34N, K35D, K112D, M134D, Y139C, I157M	–	321.65	362.15[f]	-0.73	314.5	[42, 49]
XI	1ISP*	–	R33G, K112D, M134D, Y139C, I157M	–	322.45	366.15[f]	-0.74	314.9	[42, 49]
TM	1T2N	1.8	L114P, A132D, N166Y	334.35	–	–	-0.88	317.6	[44]
1-14F5	1T2N*	-	TM + N89Y	336.15	–	–	-0.98	319.5	[44]
1-17A4	3D2A	1.73	TM + I157M	336.55	–	–	-1.00	319.9	[44]
1-8D5	1T2N*	–	TM + F17S	337.55	–	–	-0.81	316.3	[44]
2D9	3D2B	1.95	TM + F17S, N89Y, I157M	340.55	–	–	-0.98	319.7	[44]
3-18G4	3D2B*	–	2D9 + G111D	341.55	–	–	-0.92	318.5	[44]
3-11G1	3D2B*	–	2D9 + A20E	341.75	–	–	-0.98	319.6	[44]
3-3A9	3D2B*	–	2D9 + A15S	341.85	–	–	-0.87	317.5	[44]
4D3	3D2C	2.18	2D9 + A15S, A20E,G111D	344.35	–	–	-1.18	323.6	[44]
5-D	3D2C*	–	4D3 + S163P	345.35	–	–	-1.00	320.0	[45]
5-A	3D2C*	–	4D3 + M134E	346.05	–	–	-0.97	319.4	[45]
5-B	3D2C*	–	4D3 + M137P	347.25	–	–	-1.03	320.5	[45]
6B	3QMM	1.89	4D3 + M134E, M137P, S163P	351.35	–	–	-1.20	324.0	[47]

[a] Names of *Bs*LipA structures are taken from the respective references.

[b] A PDB ID marked with an asterisk indicates that the model of the corresponding variant was built using the structure with that PDB ID as a template.

[c] In Å.

[d] The temperature at which the unfolding transition begins.

[e] Median stability of rigid contacts between residue neighbors computed by applying the ENT^{FNC} approach (see section "Median stability of rigid contacts between residue neighbors as a new measure for predicting thermodynamic thermostability") (left column). Values in the right column were obtained by converting the median stabilities to a temperature scale according to Eq 1.

[f] T_{50}^{60} values, i.e., the temperature required to reduce the initial enzymatic activity by 50% within 60 min.

[g] T_{50}^{15} values, i.e., the temperature required to reduce the initial enzymatic activity by 50% within 15 min.

[h] Average $\widetilde{rc}_{ij.neighbor}$ over six wild type structures (see the main text for details).

doi:10.1371/journal.pone.0130289.t001

bonds break at higher temperatures than weaker ones [69]. As such, only hydrogen bonds with an energy $E_{HB} \leq E_{cut}(\sigma)$ were included in the network of state σ. A thermal unfolding trajectory of 60 network states was generated for each input network by decreasing E_{cut} from −0.1 kcal mol^{-1} to −6.0 kcal mol^{-1} with a step size of 0.1 kcal mol^{-1}. According to the linear relationship between E_{cut} and the temperature T introduced by Radestock and Gohlke on 20 pairs of orthologs from mesophilic and thermophilic organisms, respectively (Eq 1) [12, 13], the range of E_{cut} used in this study is equivalent to increasing the temperature of the system from 302 K to 420 K with a step size of 2 K. Because hydrophobic interactions remain constant or become even stronger as the temperature increases [70, 71], the number of hydrophobic tethers were kept unchanged throughout the thermal unfolding simulation. Rigidity analysis was performed on all such generated network states, and then local and global rigidity characteristics were calculated (see section "Local and global rigidity indices" in S1 File). The setup of the thermal

unfolding simulation and the subsequent rigidity analysis were performed using the CNA software [35], which is available from http://cpclab.uni-duesseldorf.de/software. A web service for performing CNA analysis can be accessed via http://cpclab.uni-duesseldorf.de/cna [34].

$$T = \frac{-20\ \mathrm{K}}{\mathrm{kcal} * \mathrm{mol}^{-1}} E_{cut} + 300\ \mathrm{K} \quad (1)$$

Ensemble of networks generated by using fuzzy noncovalent constraints

For improving the robustness of rigidity analyses, CNA is generally carried out on an ensemble of structures (e.g., generated by molecular dynamics (MD) simulations), and then results are averaged [11, 64]. The preceding MD simulation compromises the efficiency of the rigidity analysis, however. To overcome this drawback, Pfleger *et al.* [56] recently introduced an approach that performs rigidity analyses on an ensemble of network topologies (ENT^{FNC}) generated from a single input structure by using fuzzy noncovalent constraints. Here, the number and distribution of non-covalent constraints (hydrogen bonds and hydrophobic tethers) are modulated by random components within certain ranges as specified in ref. [56], thus simulating thermal fluctuations of a biomacromolecule without actually moving atoms. An ensemble of 2000 network configurations was generated using these definitions of fuzzy noncovalent constraints for all *Bs*LipA variants, respectively. Finally, average local indices were calculated, as were average phase transition temperatures identified by the global index cluster configuration entropy H_{type2}. The index H_{type2} monitors the degree of disorder in the realization of a given network state σ: As long as a network is dominated by a very large rigid cluster, H_{type2} tends to be low because there are only a few configurations of a system with a large rigid cluster possible; H_{type2} increases when larger rigid clusters break down in smaller clusters (see section "Local and global rigidity indices" in S1 File and ref. [55] for details).

Clustering of unfolding pathways

Recently, we showed that curves of the rigidity order parameter, which characterizes the general percolation behavior of a constraint network during thermal unfolding, for mesophilic proteins and their thermophilic counterparts are almost identical except for a shift of the curve of the thermophilic protein to higher temperatures [12]. This finding supported the hypothesis of corresponding states according to which mesophilic and thermophilic enzymes are in corresponding states of similar rigidity and flexibility at their respective optimal temperature [12]. The percolation index p_i is a local analog to the rigidity order parameter. It monitors for each bond when it segregates from the largest rigid cluster present at the beginning of a thermal unfolding simulation (see section "Local and global rigidity indices" in S1 File and ref. [55] for details). That way, a residue-wise p_i profile of a protein, generated by taking the lower of the p_i values of the two backbone bonds for each residue, expresses the hierarchical break-down of the largest rigid cluster during a thermal unfolding simulation.

We thus reasoned that the (dis)similarity of unfolding pathways of *Bs*LipA variants can be measured by Manhattan distances between their respective p_i profiles. We used this distance measure for clustering the network topologies of all *Bs*LipA variants into 10 clusters using the Partitioning Around Medoids algorithm [72] as implemented in the R program (http://www.r-project.org). This optimal number of clusters was chosen based on monitoring the change in the objective function of the clustering (the mean of the dissimilarities of all objects to their nearest medoids) as a function of the number of clusters (Figure A in S1 File) and visual inspection of cluster medoids for their dissimilarity to other medoids (residue-wise p_i profiles for

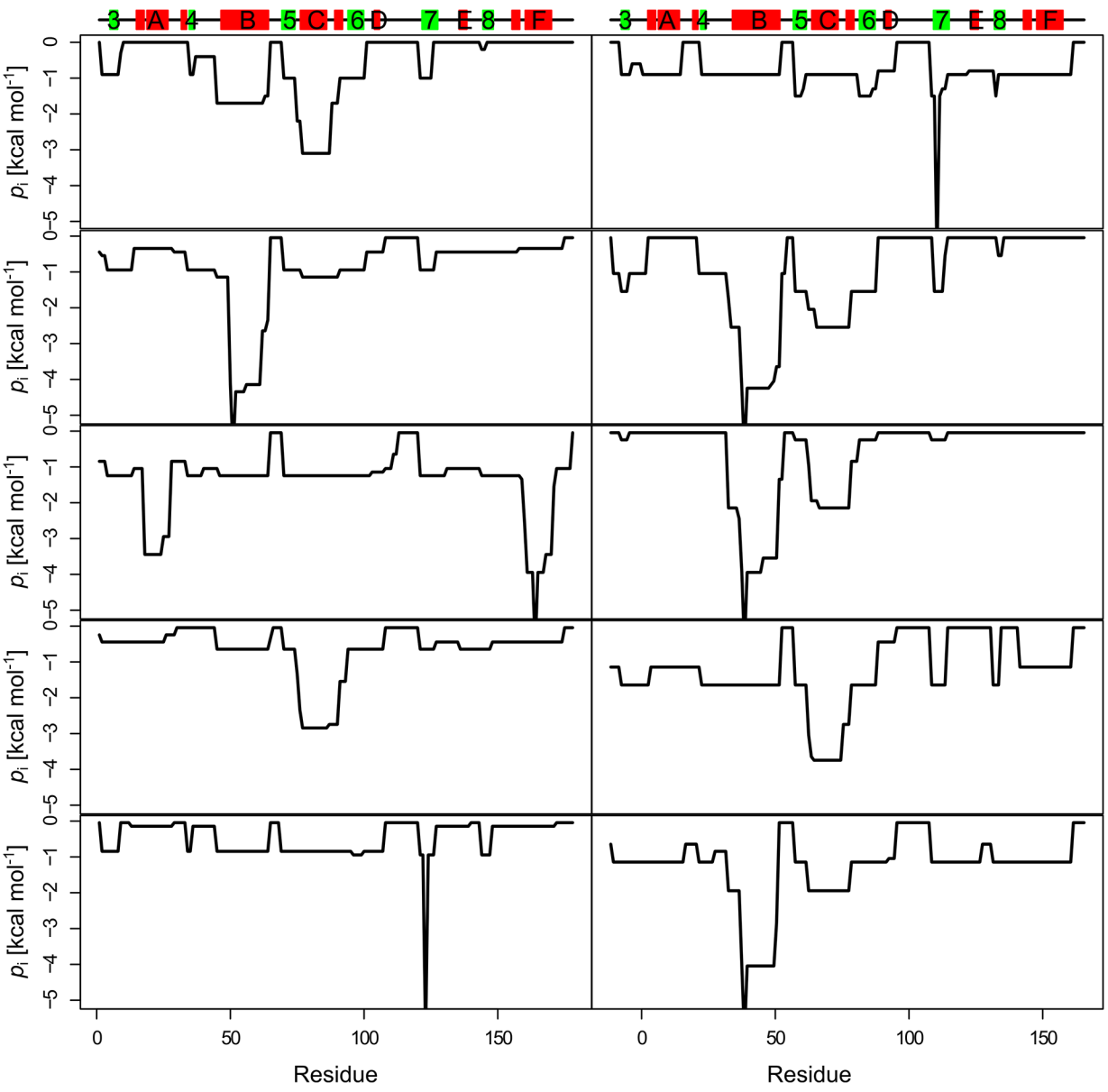

Fig 1. Residue-wise p_i plots for medoids of the 10 clusters. Secondary structure elements as computed by the DSSP program [88, 89] are indicated on the top of the plots and are labeled: α-helix (red rectangle), β-strands (green rectangle), loop (black line).

doi:10.1371/journal.pone.0130289.g001

medoids of the 10 clusters are shown in Fig 1). A clustering in more than 10 clusters essentially created additional clusters that were very similar to other clusters. From this, the cluster distribution (frequencies of network topologies in each of the 10 clusters out of in total 2000 network topologies) for each *Bs*LipA variant was calculated by counting the number of networks that belongs to each of the 10 clusters. A high (low) correlation between cluster distributions for two *Bs*LipA variants then indicates that both variants unfold in a similar (different) manner. Finally, a matrix of all pairwise correlations of cluster distributions of *Bs*LipA variants was generated.

Results

Data set

*Bs*LipA is a protein of 181 amino acids with a minimal α/β hydrolase fold; in this fold, a central parallel β-sheet of six β-strands is surrounded by six α-helices. Ser77, Asp133, and His156 constitute the catalytic triad (Fig 2). Unlike other lipases, the catalytic site in *Bs*LipA is not covered with a lid. Hence, *Bs*LipA does not show interfacial activation [40]. The data set used in this study contains structures of the wild type *Bs*LipA, thirteen mutants from Rao *et al.* [44–48], and three mutants from Reetz *et al.* [42, 49] (Table 1). The mutants differ from the wild type by three to twelve mutations, i.e., the sequence identity is $> 93\%$. Models for the mutants for which X-ray structures were not available were built using the SCWRL program. As the number of mutations in the modeled variants is ≤ 7 with respect to the template structures ($< 4\%$ with respect to the sequence length) (Table 1), an overall similar backbone confirmation can be expected as can be an overall reliable modeling of side chain conformations by SCWRL. This was also evident from a very good structural alignment and low root-mean-square deviations (RMSD) between the wild type and those mutants for which crystal structures were available (C_α atom-based RMSD values between the wild type and the mutants < 0.38 Å). The high structural similarity allows a direct comparison of results from rigidity analyses for these structures [11–13].

The melting temperature T_m of the wild type is 329.15 K. The T_m values of the mutants of Rao *et al.* range from 334.35 to 351.35 K (Table 1). For the mutants of Reetz *et al.* no T_m values are available. Rather, unfolding initiation temperatures T_i were reported, which are lower by

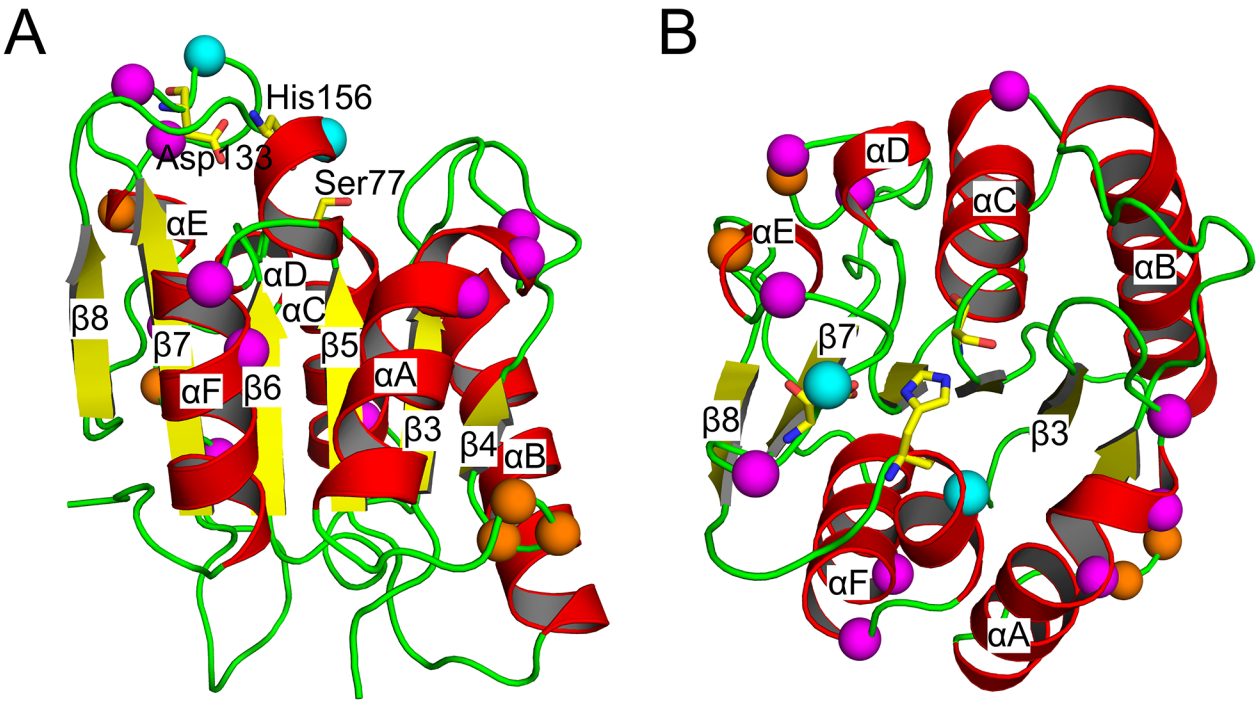

Fig 2. Cartoon representation of wild type *Bs*LipA with mutated residues indicated by spheres of their C_α atoms (mutations from Rao *et al.* [44–48]: magenta; Reetz *et al.* [42, 49]: orange; mutations common in both data sets: cyan). The catalytic triad (Ser77-Asp133-His156) is shown in stick representation with yellow carbons. The protein is colored according to secondary structure (α-helices: red; β-sheets: yellow; loops: green). The right view (**B**) differs from the left (**A**) by an anti-clockwise rotation of ~90° about a horizontal axis. All figures of *Bs*LipA structures were generated with PyMOL (http://www.pymol.org).

doi:10.1371/journal.pone.0130289.g002

2.5 to 6.2 K than that of the wild type. This suggests that mutants of Reetz *et al.* are thermodynamically less thermostable than the wild type [49], in contrast to mutants from Rao *et al.* [44–48]. However, we note that, while T_m reports on the temperature at which 50% of the protein is unfolded and, hence, properly describes the folded-unfolded equilibrium of a protein, T_i only reports on the temperature at which the unfolding transition begins. Therefore, we will only consider relations within mutants of Rao *et al.* and to the wild type and distinguish those from relations within mutants of Reetz *et al.* and to the wild type. Finally, the $T_{50}{}^{t}$ values of the mutants of Reetz *et al.* are higher than that of the wild type (Table 1), showing that these mutants more efficiently refold upon cooling after incubation at high temperatures than does the wild type. The location of mutations in all of the mutants investigated in this study is shown in Fig 2; all mutations are located on the protein surface.

Thermal unfolding pathway of *Bs*LipA

From monitoring the loss in rigidity percolation during thermal unfolding simulations, major phase transitions in the protein can be identified that relate to the unfolding pathway [11–13, 54, 73]. Here, we describe the loss of rigidity percolation of the wild type *Bs*LipA (PDB ID 1ISP) as an example. Similarity or dissimilarity, respectively, of the unfolding pathways across all variants is described below. During the thermal unfolding, a giant rigid cluster that exists at low temperature (equivalent to a high E_{cut}) breaks down in smaller sub-clusters until, finally, the whole protein becomes flexible at a high temperature (Fig 3; see also S1 Video showing the loss of rigidity percolation during the thermal unfolding of the wild type). As such, nearly the entire protein structure constitutes a single giant rigid cluster initially (at 302 K; Fig 3). As the temperature increases, loops segregate first from the giant rigid cluster. Then, at 314 K, α-helix D (αD) and αE segregate to form individual small rigid clusters (Fig 3), as do αA and αF at 318 K. The giant rigid cluster at this temperature is formed by the central β-sheet region and the two helices αB and αC (Fig 3). Next, the β-sheet region becomes sequentially flexible, beginning with β4 and β8 at 320 K (Fig 3). Then, the remaining β-strands become flexible in the order β3, β7, and β5–β6, leading to a completely flexible β-sheet region at 332 K (Fig 3 The immediate next step at which αB and αC become two separate rigid clusters is identified as a phase transition point: Now most of the structure has become flexible. This transition is most prominent with respect to going from a structurally stable wild type *Bs*LipA to an unfolded one (Figure B in S1 File). After this phase transition point, the remaining rigidity is sequentially lost, and the structure finally becomes completely flexible at 374 K (Fig 3).

During the thermal unfolding of *Bs*LipA, helices segregate from the giant rigid cluster as independent small rigid clusters. This is due to two reasons: First, in the *body-and-bar* network representation, a helix with a minimum of seven amino acids is already rigid by itself due to constraints arising from covalent and backbone hydrogen bonds [66]. Second, with the current energy function E_{HB} [69], all backbone hydrogen bonds are assigned a very similar strength, irrespective of their location along a helix. Thus, a helix will persist as an independent rigid cluster during the thermal unfolding simulation until all backbone hydrogen bonds break almost simultaneously at a high temperature, which most likely represents an overstabilization of a helix [74]. Considering this behavior, the unfolding pathway identified for the wild type *Bs*LipA is in good agreement with respect to the early segregation of α-helices with experimental findings on the unfolding of proteins with an α/β hydrolase fold [75, 76]. This indicates that side chain-mediated interactions between amino acids are well represented by the applied definitions of non-covalent constraints in the network. This is important as we want to detect effects of changes in such interactions due to mutations.

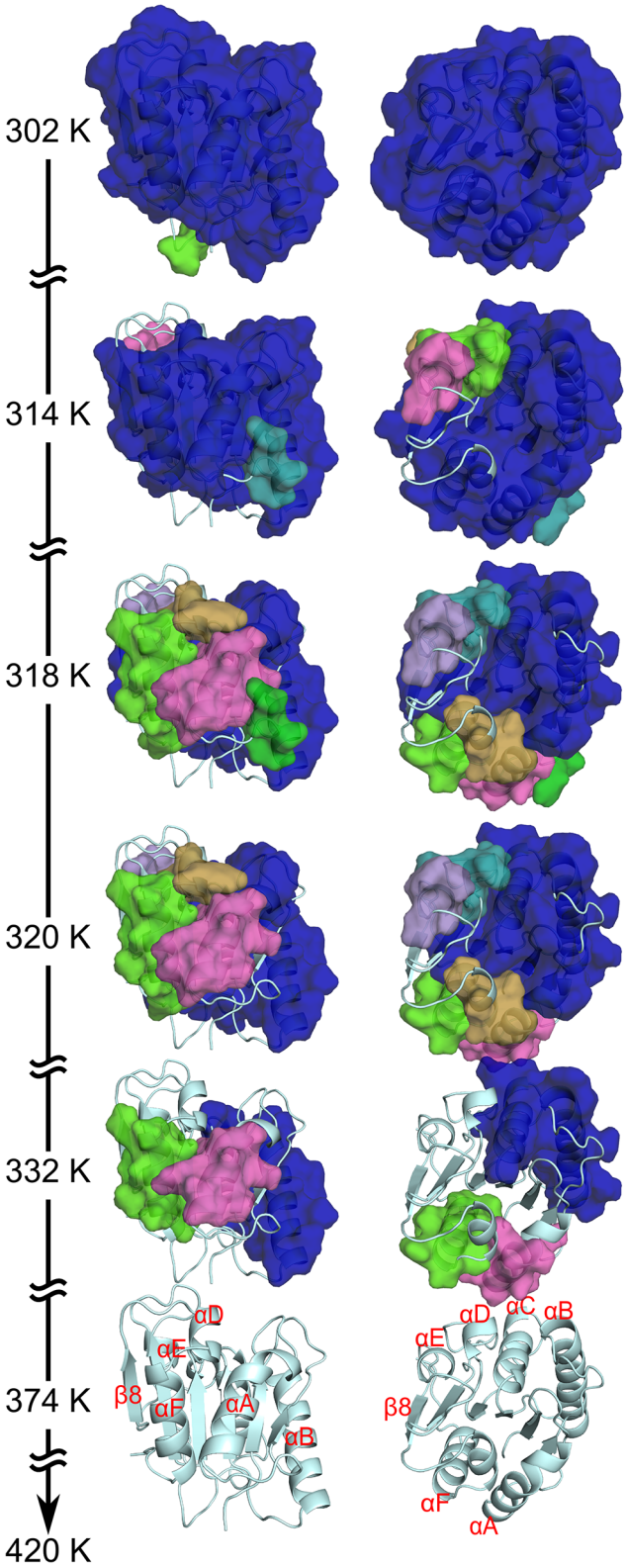

Fig 3. Average loss of structural rigidity of the wild type *Bs*LipA during a thermal unfolding simulation. Rigid clusters are depicted as uniformly colored bodies, with the largest rigid cluster shown in blue and smaller rigid clusters in the order of the colors green, magenta, cyan, orange, and violet. Temperatures are indicated for each depiction of a rigid cluster decomposition. At the beginning of the

thermal unfolding simulation (302 K), almost the complete structure is part of the giant rigid cluster; in contrast, the structure becomes completely flexible at temperatures ≥374 K. The right views differ from the left ones by an anti-clockwise rotation of ~90° about a horizontal axis. Important secondary structure elements are labeled. Note that the unfolding pathway shown here represents an *average loss of rigidity percolation* calculated from a stability map (see section "Local and global rigidity indices" in S1 File) averaged over all unfolding trajectories obtained for the ensemble of 2000 network topologies. Hence, the temperature at the phase transition point identified that way (Figure B in S1 File) cannot be compared to the *average phase transition temperature*, which is obtained from 2000 individual T_p values and used for predicting the thermodynamic thermostability of *Bs*LipA variants (see section "Prediction of thermodynamic thermostability of *BsLipA* variants")

doi:10.1371/journal.pone.0130289.g003

Prediction of thermodynamic thermostability of *Bs*LipA variants based on the global index H_{type2}

From the thermal unfolding simulations, the temperature of the phase transition point T_p was identified as described in the section "Local and global rigidity indices" in S1 File. Note that T_p values determined that way should be considered relative values only, as stated in previous studies [12, 34, 35]. Initially, we calculated phase transition points using single network topologies generated from the input structures of wild type *Bs*LipA and mutants of Rao *et al.*; however, this resulted in a very poor prediction of thermodynamic thermostability with a coefficient of determination (R^2) for a linear fit between experimental T_m and predicted T_p of 0.22 (Figure C in S1 File). We anticipated that this result reflects the high sensitivity of CNA on the conformation of the input structures as also found previously [11, 56, 64, 65]. We thus resorted to averaging T_p values over an ensemble of *Bs*LipA, applying the recently developed ENT^{FNC} approach. This approach generates an ensemble of network topologies from a single input structure and has been shown to yield results of rigidity analyses both at the local and global level that agree almost perfectly with those obtained from MD simulations-generated ensembles of structures [56]. However, this yielded a significant ($p = 0.002$) correlation between T_p and T_m with $R^2 = 0.58$ only if the two structures with the lowest (wild type) and highest (mutant 6B) T_m were considered outliers (Fig 4A; see below for an explanation regarding the outliers; note that removing the two outliers in the case of using single network topologies only marginally improved R^2 from 0.22 to 0.29). The mutants IX, X and XI of Reetz *et al.* were predicted to be slightly less thermostable than the wild type (Fig 4A). This is in line with experimental findings by Reetz *et al.* that suggest that these mutants are thermodynamically less stable than the wild type [49]. In summary, these results suggest that CNA coupled with the ENT^{FNC} approach can sense effects on the thermodynamic thermostability that arise from only a few sequence variations (pairwise sequence identity > 93%; pairwise RMSD < 0.38 Å). However, the false predictions for wild type *Bs*LipA and mutant 6B are dissatisfying.

Difference in unfolding pathways explains outliers

Next, we investigated why the thermostabilities of the wild type and the mutant 6B were predicted falsely. Since the precision of the computations shown in Fig 4A is high (the standard error in the mean is < 0.38 K in all cases), we reasoned that the false prediction must arise from a systematic difference between the wild type and 6B *versus* all other mutants of Rao *et al.* Thus, we mutually compared all unfolding pathways of the systems as described in "Materials and Methods". After partitioning unfolding pathways of *Bs*LipA variants characterized on a residue basis by the percolation index p_i into 10 clusters (see Fig 1 for the p_i profiles of the 10 cluster medoids), we calculated correlation coefficients from the resulting cluster distributions for all pairs of variants (Fig 5; Tables A and B in S1 File).

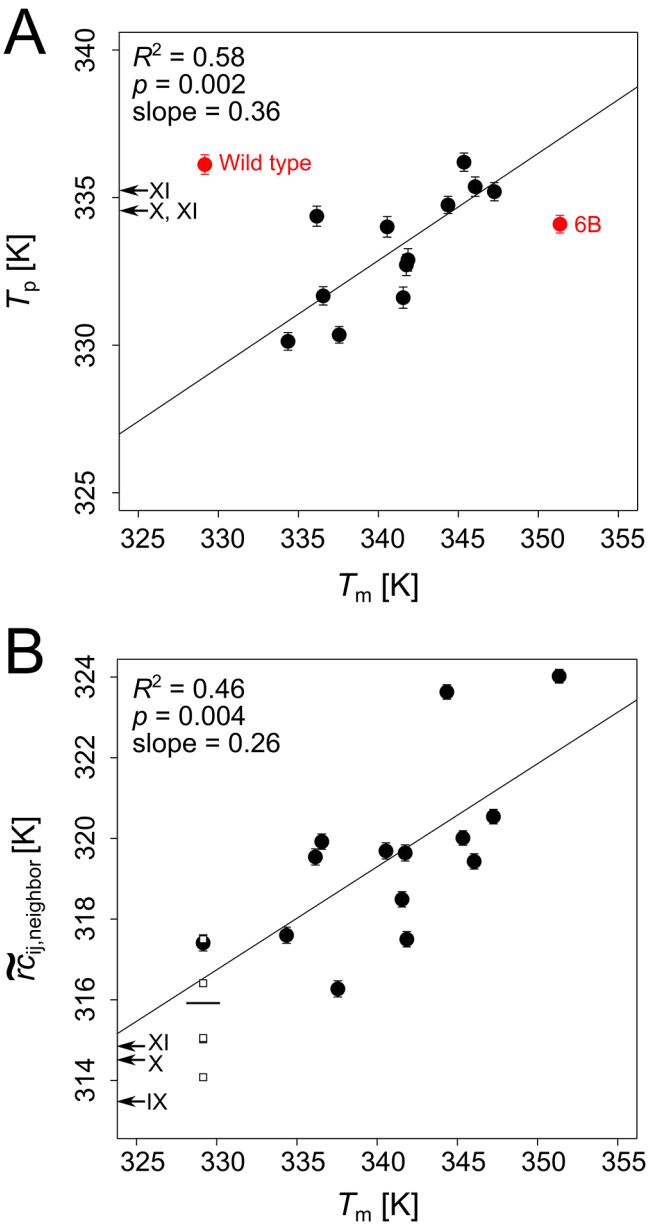

Fig 4. Correlation between predicted and experimental thermostabilities (T_m values) of *Bs*LipA variants; for the predictions, the ENT^{FNC} approach was used. A: Correlation between T_p derived from the global index H_{type2} and T_m values for thermodynamically thermostabilized mutants from Rao *et al*. Data points colored red were considered outliers (see main text for explanation) and excluded when calculating R^2 values and the correlation lines. **B**: Correlation between $\widetilde{rc}_{ij,neighbor}$ and T_m values for thermodynamically thermostabilized mutants from Rao *et al*. Data points shown as empty squares represent $\widetilde{rc}_{ij,neighbor}$ values for five additional wild type crystal structures (see main text for details; two of the squares closely overlap; mean $\widetilde{rc}_{ij,neighbor}$ over all six data points for wild type structures is shown as a small horizontal line: 315.9 ± 0.6 K). A and B: Error bars represent the standard error in the mean. T_p and $\widetilde{rc}_{ij,neighbor}$ values for kinetically thermostabilized mutants from Reetz *et al*. are marked by arrows on the corresponding ordinates.

doi:10.1371/journal.pone.0130289.g004

These results revealed that the wild type enzyme shows an unfolding pathway distribution very distinct from other *Bs*LipA variants from Rao *et al*. with correlation coefficients *r* ranging from −0.69 to 0.54 (Fig 5, Table A in S1 File). The average *r* value for the wild type against all

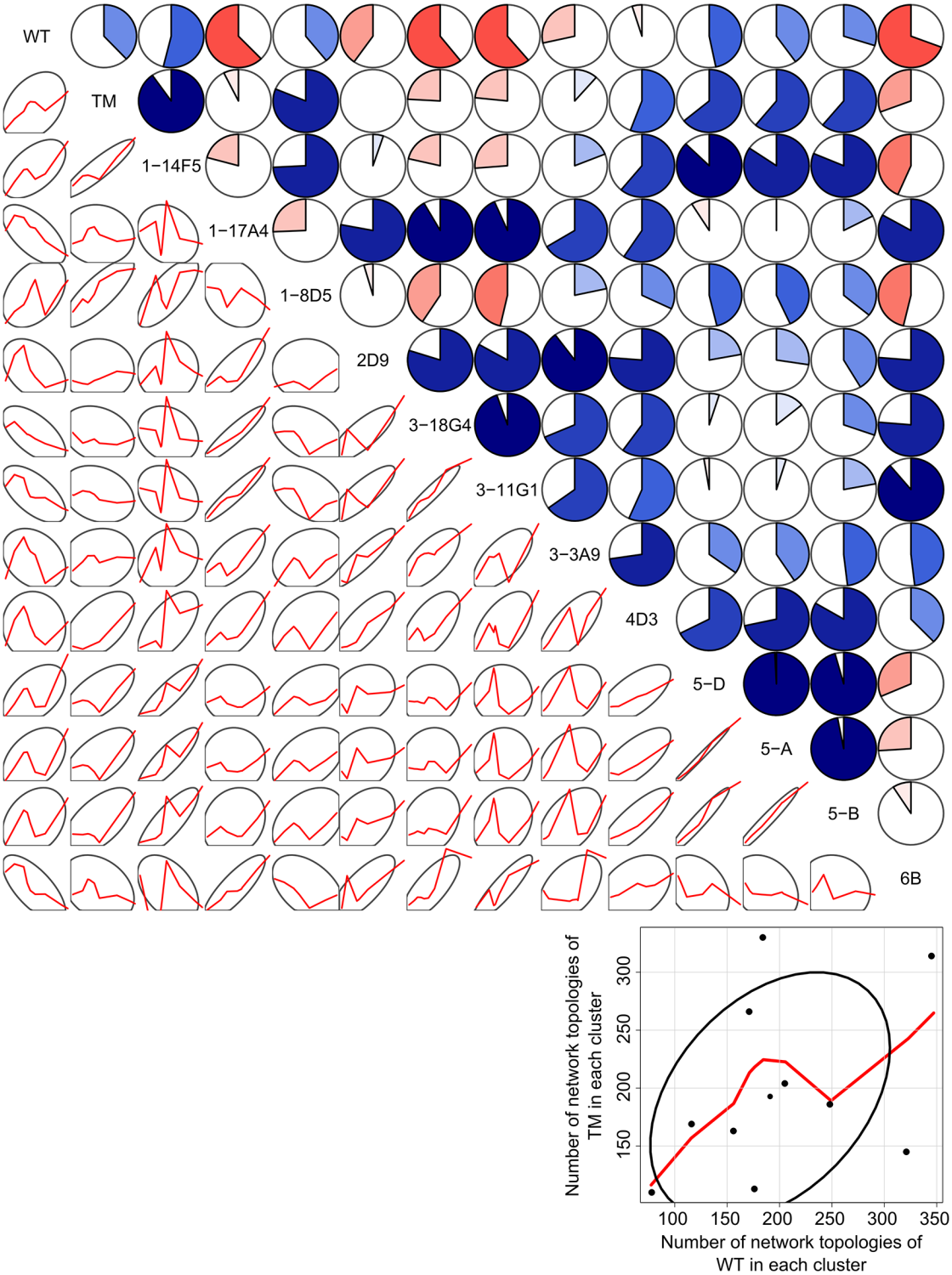

Fig 5. Pairwise correlations of cluster distributions (using 10 clusters) of unfolding pathways of wild type *Bs*LipA and mutants from Rao *et al*. The upper triangle shows pairwise correlation coefficients as dial plots where a filled portion of a pie indicates the magnitude of the correlation (*r*) and blue (red) color indicates a positive (negative) correlation. The lower triangle shows 68% data ellipses (depicting the bivariate mean ± 1 standard deviation) [90] and scatterplots of the respective cluster distributions (frequencies of network topologies in each of the 10 clusters) of the two *Bs*LipA variants as red lines smoothed by locally-weighted polynomial regression [91]. Data points and axes for the plots in the lower triangle are omitted for clarity. The blow-up in the right bottom corner depicts exemplarily for the TM *versus* WT case axes, axes labels, data points, the smoothing line obtained by locally-weighted polynomial regression, and the 68% data ellipsis. The figure was generated using the "corrgram" package [92] of the R program (http://www.r-project.org).

doi:10.1371/journal.pone.0130289.g005

other variants from Rao *et al.* is −0.06 ± 0.14 (mean ± SEM), which is lower than that of the other variants (≥°0.16 except for the outlier 6B) (Table A in S1 File). The second outlier, mutant 6B, has an average *r* value of 0.12 ± 0.16 when comparing its unfolding pathway distribution to those of other variants from Rao *et al.* This average *r* value is lower than the corresponding average *r* values of all other mutants from Rao *et al.* (Table A in S1 File). The thermal unfolding pathway of 6B is shown in Figure D in S1 File. While the overall unfolding pathway of 6B is comparable with that of the wild type *Bs*LipA in that the helices segregate from the giant rigid cluster as individual rigid clusters in the early phase of unfolding, they do so in a different order (αD, αA–αF, αE, αB–αC; Figure D in S1 File). A probability density function (PDF) of *r* values of unfolding pathway distributions of the two outliers wild type and mutant 6B with all other variants shows a bimodal distribution and is shifted towards lower *r* values compared to the PDF of the *r* values of other mutants from Rao *et al.* Furthermore, about half of this distribution is related to negative *r* values (Figure E in S1 File). In all, this suggests that the two outliers have unfolding pathways different from all other mutants from Rao *et al.* for which the prediction of thermodynamic thermostability was successful. Finally, we note that the unfolding pathway distributions of the wild type and the three mutants from Reetz *et al.* are highly similar to each other ($r > 0.79$; $p \leq 0.001$; Table B in S1 File).

These findings have important implications: First, the results strongly suggest that the misprediction of the thermostabilities of the wild type and mutant 6B arises from them showing different unfolding pathways from all of the remaining mutants from Rao *et al.*. Apparently, the present approach of identifying phase transition points by monitoring the *global* index H_{type2} (see section "Local and global rigidity indices" in S1 File) is too sensitive with respect to the details of such pathways. Consequently, alternative methods should be explored (see section "Median stability of rigid contacts between residue neighbors as a new measure for predicting thermodynamic thermostability"). Second, the results suggest that the history of the generation of the *Bs*LipA structures may play a role for the observed differences in the unfolding pathways: generally, the most similar unfolding pathways (Tables A and B in S1 File) (and then the most coherent T_p predictions) are found for those variants that originate from a common structural "ancestor" (Table 1). Third, the results propose to apply the similarity/dissimilarity of unfolding pathway distributions as a measure to judge the reliability of thermostability predictions in future studies: the lower the similarity for two variants, the less confident should one be that relative thermostability predictions are correct. Finally, we cannot exclude at the present stage that thermostabilizing mutations lead to an unfolding pathway that is different from the one of the wild type. Considering that intrinsic and extrinsic modifications in other systems that led to thermostabilization have been shown to influence not just the folded state but the entire (un)folding free energy landscape [77, 78], this possibility also exists for *Bs*LipA mutants [45, 47].

Median stability of rigid contacts between residue neighbors as a new measure for predicting thermodynamic thermostability

The above findings called for predicting the thermodynamic thermostability in a way that is less sensitive to the details of the unfolding pathway than the present approach relying on the *global* index H_{type2}. The sensitivity arises here from the need to accurately identify the phase transition point from the percolation behavior of the constraint network as the most pronounced jump in H_{type2} during the unfolding (Figure B in S1 File). As shown previously, however, the percolation behavior of networks from protein structures is complex [13] (in contrast to that of network glasses [54, 79]), reflecting that a protein structure is hierarchical and

composed of modules. As a consequence, often more than one pronounced jump in H_{type2} is observed, which then makes it difficult to assign a phase transition point (Figure B in S1 File).

As an alternative, we set out to characterize thermodynamic thermostability at the *local* level [55], i.e., by monitoring residue pair-wise descriptors of local stability within a protein structure as a function of the temperature. The most comprehensive information in that direction is provided by stability maps rc_{ij} [12], which depict when a rigid contact *rc* between two residues *i* and *j* ceases to exist along a thermal unfolding trajectory. As such, rc_{ij} contains information cumulated over all states σ of a network along the trajectory as to which parts of the network are (locally) mechanically stable at a given state σ, and which are not [12, 55]. Of note, this stability information is not only available in a qualitative manner (i.e., in terms of local rigidity and flexibility) but also quantitatively in that each rc_{ij} has associated with it the energy E_{cut} at which this rigid contact is lost. Thus, $\sum_{i,j>i} rc_{ij}$ represents the chemical potential energy due to non-covalent bonding, obtained from the coarse-grained, residue-wise network representation of the underlying protein structure. With respect to a reference state where no non-covalent interactions are present anymore (i.e., an unfolded state), $\sum_{i,j>i} rc_{ij}$ can be considered an unfolding energy then. Three modifications were applied to $\sum_{i,j>i} rc_{ij}$ here for technical reasons. I) In order to stress the locality of interactions within a protein, which will later aid in understanding how structural differences relate to thermostability differences (see section "Influence of mutations on local structural rigidity"), we focused on the stability of rigid contacts $rc_{ij,neighbor}$ between structurally close residues only (i.e., those residues where at least one pair of respective atoms is within 5 Å distance). II) To suppress the influence of extreme values in the double summation on the outcome of the unfolding energy, we used the median stability of rigid contacts $\widetilde{rc}_{ij,neighbor}$ instead. Such extreme values can occur in regions that are highly stabilized by interactions to hydrophobic atoms [56]. III) Applying the ENT^{FNC} approach, $\widetilde{rc}_{ij,neighbor}$ were averaged over ensembles of 2000 constraint networks, which has been shown to significantly improve the robustness of rigidity analyses [56]. The $\widetilde{rc}_{ij,neighbor}$ values are given in Table 1. In addition, Table 1 and Fig 4B show these values after converting them to a temperature scale via Eq 1

A significant and fair linear correlation of $\widetilde{rc}_{ij,neighbor}$ with T_m values of the thermodynamically stable mutants from Rao *et al.* is obtained ($R^2 = 0.46$, $p = 0.004$; Fig 4B). No outlier is observed now, indicating that our definition of an average local stability correctly reflects differences in the thermodynamic thermostability. This finding substantiates our above interpretation of $\widetilde{rc}_{ij,neighbor}$ as an approximation to the unfolding energy, because under the condition of a temperature-independent heat capacity the unfolding energy is linearly correlated to the melting temperature, with the heat capacity as the scaling factor [80]. The slope of the correlation line (0.26) in Fig 4B deviates from unity. This indicates that the linear relationship in Eq 1 used for converting $\widetilde{rc}_{ij,neighbor}$ to a temperature scale, which was derived for H_{type2}-based thermostability prediction [12, 13], may need to be reparameterized for application with $\widetilde{rc}_{ij,neighbor}$. In this case, as the heat capacity has been shown to scale linearly with the number of residues for small globular proteins [80, 81], a normalization with respect to protein size needs to be applied. When only considering the six X-ray structures in the dataset of Rao *et al.*, a good correlation of $\widetilde{rc}_{ij,neighbor}$ with T_m values of $R^2 = 0.87$ ($p = 0.007$) is found (Table 1). In contrast, a weaker correlation ($R^2 = 0.33$, $p = 0.07$) is obtained for the eight variants that were modeled using SCWRL (). This suggests an influence of the quality of the input structures on the prediction of thermodynamic thermostability. Finally, as before, the mutants from Reetz *et al.* are found to have a lower thermodynamic thermostability than the wild type, in very good agreement with experimental findings (see above and Table 1) [49].

The $\widetilde{rc}_{ij,neighbor}$-based measure is apparently less sensitive to differences in the unfolding pathway because the wild type and mutant 6B are now much better ranked. However, comparing the prediction of thermostabilities by $\widetilde{rc}_{ij,neighbor}$ and H_{type2}, the latter yields a better correlation with T_m for mutants with similar unfolding pathways. From an application point of view, we thus recommend using H_{type2}-derived T_p values for comparing thermostabilities of variants of a protein unless the underlying unfolding pathways are dissimilar; in that case, we recommend using $\widetilde{rc}_{ij,neighbor}$.

When applied to hen egg white lysozyme the ENT^{FNC} approach has been shown to significantly improve the robustness of rigidity analyses with respect to the conformation of the input structures [56]. To probe if this also holds for *Bs*LipA investigated here, we computed $\widetilde{rc}_{ij,neighbor}$ using the ENT^{FNC} approach for five additional crystal structures of wild type *Bs*LipA (see section "Materials and methods"). The standard error of the mean in $\widetilde{rc}_{ij,neighbor}$ over all six wild type *Bs*LipA structures is 0.57 K (Fig 4B) including PDB ID 1ISP discussed so far. This error is likely within the experimental uncertainty, confirming our previous results of robust rigidity analyses with ENT^{FNC} [56]. Still, if the average $\widetilde{rc}_{ij,neighbor}$ over all six crystal structures (315.9 K; see horizontal line in Fig 4B;Table 1) is considered for the $\widetilde{rc}_{ij,neighbor}$ *versus* T_m correlation, the quality of the correlation improves considerably to $R^2 = 0.55$ ($p = 0.001$) compared to if only $\widetilde{rc}_{ij,neighbor}$ of PDB ID 1ISP is used (see above). This indicates that the use of multiple input structures in connection with the ENT^{FNC} approach further increases the accuracy of thermostability predictions.

Influence of mutations on local structural rigidity

Considering that the average local stability defined above correctly reflects differences in the (macroscopic) thermodynamic thermostability, we analyzed on a residue basis how changes in thermostability relate to changes in local structural stability (rigidity). First, we compared stability maps of variants from Rao *et al*. with distinct thermostabilities to analyze the effect of mutations on the local rigidity. In particular, we compared the wild type to a more thermostable variant 1-14F5 and the most thermostable variant 6B. We averaged stability maps of the six wild type structures (see above and Fig 4B) and used this average for comparison against the thermostable variants of *Bs*LipA. Difference stability maps for 1-14F5/wild type (Fig 6A) and 6B/wild type (Fig 6B) pairs demonstrate that mutations in general improve the strength of rigid contacts to and in between neighboring residues of the mutations (lower triangles in Fig 6A and 6B) but also in between residue pairs not in contact distance (upper triangles in Fig 6A and 6B). This effect is more pronounced for 6B/wild type than 1-14F5/wild type.

In more detail, the four mutations (indicated by arrows in Fig 6A and shown in Fig 6D) on 1-14F5 stabilize contacts of αD with its neighboring helix αC and contacts of αA with αF (Fig 6A and 6D). More importantly, the contacts of helices αA and αF with their neighboring β-strands in the central β-sheet region are stabilized, which delays the early loss of these helices observed during the thermal unfolding of the wild type (Fig 3). Similarly, the contacts between αB and the central β-sheet region also become stronger, which delays the decay of structural stability of the β-sheet during thermal unfolding. On average, contacts between all residue neighbors are ~−0.1 kcal mol^{-1} or ~2 K more stable in 1-14F5 than in the wild type.

Residues mutated in 6B (indicated by arrows in Fig 6B and shown in Fig 6E) include the mutations already found in 1-14F5. This explains a strengthening of inter-helical contacts and of the contacts between α helices and the central β-sheet region as discussed already for 1-14F5 (Fig 6D and 6E). However, the additional mutations in 6B stabilize contacts between other α-helices (αD and αE) and the central β-sheet region and further reinforce those between αA or

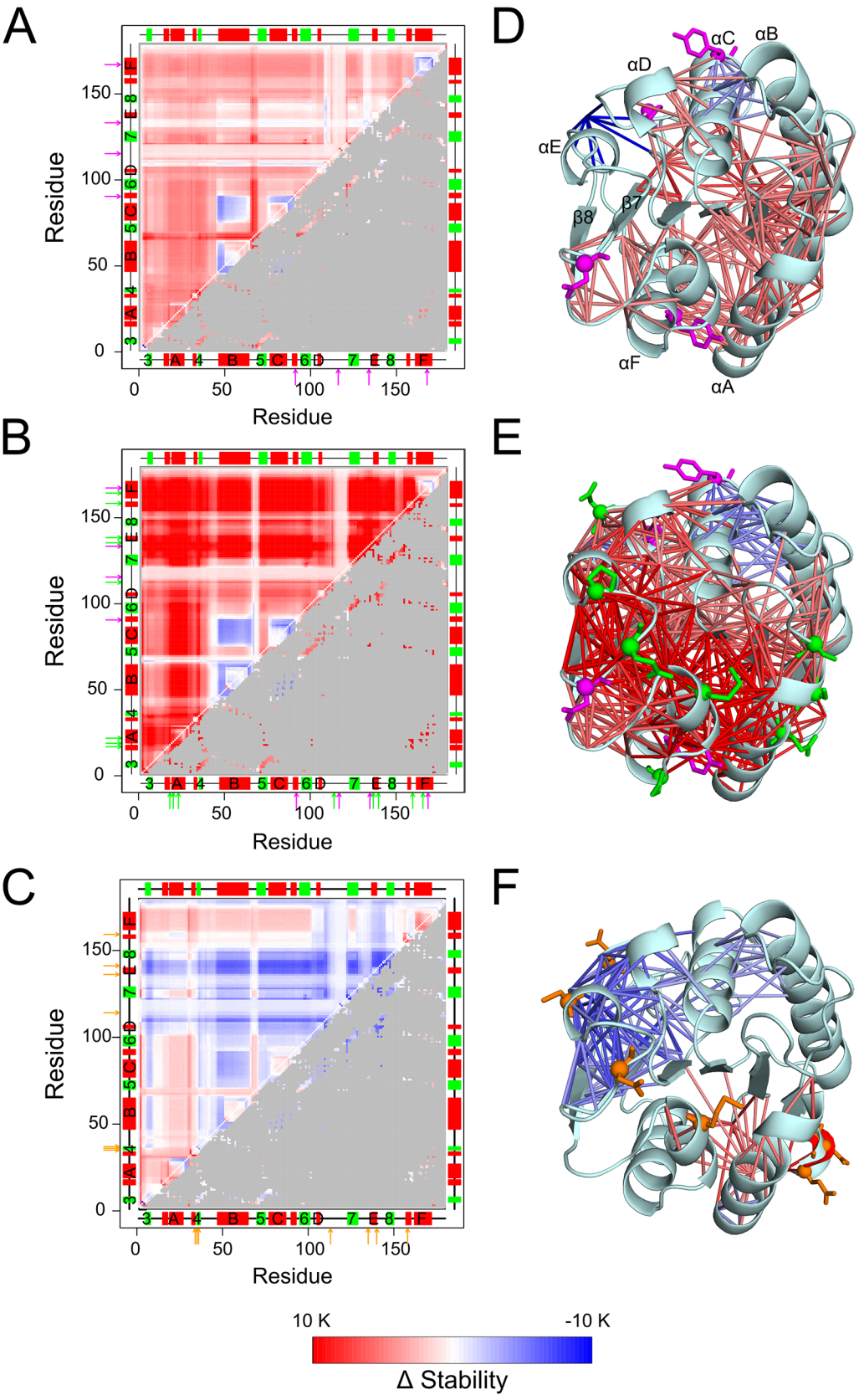
A
B
C
D
E
F
Residue
Residue
0
50
100
150
αA
αB
αC
αD
αE
αF
β7
β8
10 K
-10 K
Δ Stability

Fig 6. Differences in the stability of rigid contacts between wild type and mutants of *Bs*LipA. Maps depict differences between stability maps of the respective mutants and an average stability map of the six wild type structures (see the main text for explanation) for **A**: mutant 1-14F5, **B**: mutant 6B, and **C**: mutant X. A red (blue) color indicates that a rigid contact in the mutant is more (less) stable than in the wild type (see color scale at the bottom). The upper triangles show differences in the stability values for all residue pairs; the lower triangles show differences in the stability values only for residue pairs that are within 5 Å of each other, with values for all other residue pairs colored gray. Secondary structure elements as computed by the DSSP program [88, 89] are indicated on both abscissa and ordinate and are labeled: α-helix (red rectangle), β-strands (green rectangle), loop (black line). Arrows represent the mutation positions with respect to the wild type sequence: Common mutations in 1-14F5 (A) and 6B (B) are shown in magenta, unique mutations in 6B (B) are shown in green, and mutations in X (C) are shown in orange. The differences in the stability of rigid contacts for residue neighbors is also displayed on the structures of the mutants by sticks connecting C_α atoms of residue pairs colored according to the color scale of the maps for **D**: 1-14F5, **E**: 6B, and **F**: X. Only those contacts that are stabilized by $\geq$ 4 K or destabilized by $\geq$ 3 K are shown for clarity; for the same reason, contacts between two residues of the same secondary structure element are not shown. Mutated residues are shown as sticks and a sphere at their C_α atoms (D, E, and F) in the same color used for arrows (A, B, and C).

doi:10.1371/journal.pone.0130289.g006

αF and the β-sheet. On average, contacts between all residue neighbors are ~−0.4 kcal mol^{-1} or ~8 K more stable in 6B than in the wild type (Fig 6E).

Taken together, contacts between peripheral helices and the central β-sheet region are stronger in 6B than in 1-14F5. This delays the loss of α-helices during thermal unfolding (Fig 3) to a larger extent in 6B than in 1-14F5, explaining at a structural level why 6B is more stable than 1-14F5. Remarkably, many of these stabilizations must arise from the long-range aspect of rigidity percolation [52, 64, 82, 83], because almost all mutations in 6B are on the surface, i.e., far from the central β-sheet region. In contrast, inter-helical contacts of the αB/αC helix pair become weaker in the mutants than in wild type (Fig 6D and 6E) indicating that the strengthened stability between these helices and the central β-sheet region is sufficient to keep the structure folded. At last, for all other thermodynamically more thermostable mutants, a similar profile of changes in contact stability between various secondary structure elements was observed (Figure F in S1 File). Not unexpected, the increase in contact stability compared to wild type (Figure F in S1 File) was generally the more pronounced the higher the thermodynamic thermostability is of the mutant (Table 1).

Second, we compared the mutants from Reetz *et al*. to the wild type. Regarding mutant X, seven residues have been mutated (indicated by arrows in Fig 6C and shown in Fig 6G). In strict contrast to what was observed for the thermodynamically thermostabilized mutants, this mutant showed a destabilization of rigid contacts both locally and globally (Fig 6C and 6F; see also Figure G in S1 File, where a similar finding is depicted for mutants IX and XI). For mutant X, the average decrease in stability over all residue neighbors is ~0.06 kcal mol^{-1} or ~1.2 K. The destabilization found on the local scale agrees with results of a lower T_p found when analyzing the mutants globally. Furthermore, the results are in line with experimental findings which suggest that the mutants are thermodynamically less stable than the wild type (Table 1) [49]. Our findings are also in good agreement with results obtained by comparative crystal structure analysis of wild type and variant X [49]: Loop region 14–21, for which lower B-factors in X than in the wild type structure were observed, shows increased contact stabilities with its neighboring residues in X (Fig 6C and 6F; Figure H in the S1 File). Likewise, regions 129–153 and 177–181, for which higher B-factors in X than in the wild type structure were observed, show decreased contact stabilities with their neighboring residues in X (Fig 6C and 6F; Figure H in the S1 File). However, region 60–70 shows increased contact stabilities in X (Fig 6C and 6F and Figure H in the S1 File) despite higher B-factors observed in the comparative crystal structure analysis. The latter may reflect increased motions of a stabilized region as a whole, taking into consideration that B-factors can report on rigid body motions of a structurally stable part [84].

Finally, it would be very satisfying from both the biochemical and structural biology point of view, if the effects of the three to twelve mutations on increased or decreased local rigidity could be immediately related to the changes in specific interactions with neighboring residues. Our above observation for the mutants from Rao *et al*. that many of the stability changes must

arise from the long-range aspect of rigidity percolation [52, 64, 82, 83] speaks against such an endeavor. Nevertheless, we analyzed differences in the per-residue number of hydrogen bonds and hydrophobic tethers of the variants 1_14F5, 6B, and X with respect to the wild type (Figure I in S1 File). While these analyses reveal differences in the number of interactions > 1 indeed only for a small set of residues, between ~10 and ~35% of all residues show differences of 1. This finding is remarkable given the small number of mutations in the variants and the high structural similarity between crystal structures of the wild type and mutants found for backbone atoms above; apparently, these differences arise from subtle changes in the conformations of the side chains due to the mutations. This finding also suggests that an interpretation of the relationship between increased or decreased local rigidity due to changes in specific interactions of a mutated residue and a change in thermostability of a mutant may fall short of the actual complexity underlying this relationship.

Discussion

Understanding the origin of thermostability is of fundamental importance in protein biochemistry. Here, we have probed the relation between protein thermostability and structural rigidity by directly analyzing static properties of a well-characterized set of 16 *Bs*LipA mutants. The main outcome of this work is the finding of a good correlation between the structural rigidity of all *Bs*LipA variants and their thermodynamic thermostability. This finding of a quantitative relation between structural rigidity and thermodynamic thermostability within a series of closely related protein variants complements a previous study that showed for pairs of homologous proteins from thermophilic and mesophilc organisms that raising the structural stability is the most common way (~77% of all cases) to obtain a higher thermostability [85].

Intense discussions are ongoing regarding the question if elevated protein thermostability is related to increased or decreased structural rigidity of the folded state [10, 18–23]. Part of this discussion is related to how information on structural rigidity is derived from information on mobility, in particular with respect to the temporal resolution of the experimental techniques and computational analysis [26–32]. In this context, the finding we describe here is highly relevant. As the rigidity theory-based CNA approach applied characterizes rigidity and flexibility of proteins directly, i.e., without the requirement of information on atomic movements, it does not suffer from such time dependence. Another part of the discussion is related to the fact that changes in the enthalpy, entropy and/or heat capacity can lead to thermodynamic stabilization; these changes can be linked to distinct effects on the structural stability of the folded state [19]. It was thus instructive to observe that the general increase in rigidity in the mutants of Rao *et al.* is accompanied by certain inter-helical contacts becoming weaker than in the wild type; these weakened contacts between the "modular" helices may increase the entropy of the folded state and so may further contribute to the overall stability of the systems [17, 86, 87]. This finding again calls attention to analyzing the origin of thermostability with methods that cover a wide range of temporal and spatial resolution because otherwise one effect may be hidden beneath another.

Our results are backed up with a careful validation of the accuracy and robustness of the CNA approach on the data set both from a macroscopic and microscopic point of view. As to the former, good and statistically significant correlations between experimental melting temperatures (T_m) of mutants of Rao *et al.* and predicted thermodynamic thermostabilities have been found based on two independent measures (H_{type2} and $\widetilde{rc}_{ij,neighbor}$), as was correctly predicted that the thermodynamic thermostability of the mutants of Reetz *et al.* is lower than that of the wild type. Furthermore, $\widetilde{rc}_{ij,neighbor}$-based predictions of the thermodynamic thermostability on six crystal structures of wild type *Bs*LipA revealed a standard error of the mean likely

within experimental error, confirming previous results of robust rigidity analyses when applying the ENT^{FNC} approach [56]. As to the latter, the detailed analysis of the unfolding pathway of wild type *Bs*LipA revealed a good agreement with respect to the early segregation of α-helices with experimental observations on other proteins with an α/β hydrolase fold. These findings are in line with previous successful applications of CNA in predicting melting temperatures and identifying structural weak spots [11–13].

From a methodological point of view, some additional comments are in order. First, in the present study we successfully predicted the thermodynamic thermostability for mutants that differ by as few as three to twelve mutations from the wild type. Compared to previous applications of CNA on either pairs of mesophilic and thermophilic homologues [12, 13] or a series of homologous proteins from different organisms living at varying temperatures [11], this finding considerably broadens the application domain of CNA towards data-driven protein engineering: There, related series of mutants with only a small number of respective mutations will be the major focus of investigations. Second, we introduced a measure for the similarity/dissimilarity of unfolding pathways of mutants and used it for explaining false thermostability predictions. We suggest to use the measure in future studies as a significance criterion to judge the reliability of thermostability predictions from CNA. Third, we introduced the median stability of rigid contacts as a new local measure for predicting thermodynamic thermostability and showed that this measure is less sensitive to details of the unfolding pathway. The measure is thus recommended for comparing thermostabilities of mutants the underlying unfolding pathways of which are dissimilar.

Finally, regarding the subset of mutants of Reetz *et al*., we find a decreased local rigidity compared to wild type, in line with findings of lower unfolding initiation temperatures, yet the mutants are more "thermostable" than the wild type in that they preserve enzymatic activity better after subjecting them to higher temperatures [42]. It would have been tempting to investigate how this relates to a potential kinetic stabilization of the mutants. However, we refrained from doing so due to the lack of direct experimental evidence for such a kinetic stabilization [49]. In turn, this finding draws attention to the fact that the term "protein thermostability" is often used in a non-discriminating sense, i.e., data reported in the literature does not allow to establish whether a protein is thermodynamically or kinetically stable [49]. This adds another layer of complexity to the question of the relation between protein thermostability and structural rigidity as it may be required to decouple observations on "increased *vs*. decreased structural rigidity" from the general description of "protein thermostability" in future studies.

Supporting Information

S1 File. The file contains additional information to the manuscript: Pairwise Pearson correlation coefficients *r* and corresponding *p* values between cluster distributions of *Bs*LipA variants from Rao *et al.* (Table A) and Reetz *et al.* (Table B), objective function of the clustering (Figure A), cluster configuration entropy H_{type2} *vs*. temperature obtained from the *average loss of rigidity percolation* of wild type *Bs*LipA (Figure B), correlation between predicted T_p derived from the global index H_{type2} and experimental thermostabilities (T_m values) of *Bs*LipA variants using single input structures. (Figure C), average loss of structural rigidity of mutant 6B during a thermal unfolding simulation (Figure D), probability density functions (PDFs) of all pairwise Pearson correlation coefficients between cluster distributions of *Bs*LipA variants (Figure E), differences in the stability of rigid contacts between wild type and variants of *Bs*LipA from Rao *et al.* (Figure F) and Reetz *et al.* (Figure G), differences in the stability of rigid contacts between variant X and wild type for selected residue neighbors (Figure H), and differences in the number of hydrogen bonds and hydrophobic tethers between *Bs*LipA mutants and

wild type (Figure I).
(PDF)

S1 Video. The video file shows the average loss of structural rigidity during thermal unfolding of wild type *Bs*LipA and the corresponding global rigidity index H_{type2} *vs*. temperature plot. See captions of Fig 3 in the main text and Figure B in S1 File for details.
(AVI)

Acknowledgments

We are grateful to the Ministry of Innovation, Science, and Research of North Rhine-Westphalia and Heinrich-Heine-University Düsseldorf (HHU) for a scholarship to PCR within the CLIB^{2021} Graduate Cluster Industrial Biotechnology. We acknowledge the „Zentrum fuer Informations- und Medientechnologie"(ZIM) at HHU for computational support. We thank Anuseema Bhadauriya, HHU, for fruitful discussions on the experimental data of the investigated mutants.

Author Contributions

Conceived and designed the experiments: HG. Performed the experiments: PCR. Analyzed the data: PCR KEJ HG. Wrote the paper: PCR KEJ HG.

References

1. Demirjian DC, Moris-Varas F, Cassidy CS. Enzymes from extremophiles. Current Opinion in Chemical Biology. 2001; 5(2):144–51. Epub 2001/04/03. S1367-5931(00)00183-6 [pii]. PMID: 11282340.
2. Van den Burg B. Extremophiles as a source for novel enzymes. Current Opinion in Microbiology. 2003; 6(3):213–8. Epub 2003/07/02. S1369527403000602 [pii]. PMID: 12831896.
3. Ó'Fágáin C. Engineering protein stability. In: Walls D, Loughran ST, editors. Protein Chromatography: Springer; 2011. p. 103–36.
4. Polizzi KM, Bommarius AS, Broering JM, Chaparro-Riggers JF. Stability of biocatalysts. Current Opinion in Chemical Biology. 2007; 11(2):220–5. Epub 2007/02/20. doi: 10.1016/j.cbpa.2007.01.685 PMID: 17307381.
5. Vogt G, Woell S, Argos P. Protein thermal stability, hydrogen bonds, and ion pairs. Journal of Molecular Biology. 1997; 269(4):631–43. Epub 1997/06/20. S0022-2836(97)91042-1 [pii] doi: 10.1006/jmbi.1997.1042 PMID: 9217266.
6. Kumar S, Tsai CJ, Nussinov R. Factors enhancing protein thermostability. Protein Engineering. 2000; 13(3):179–91. Epub 2000/04/25. PMID: 10775659.
7. Gromiha MM, Pathak MC, Saraboji K, Ortlund EA, Gaucher EA. Hydrophobic environment is a key factor for the stability of thermophilic proteins. Proteins. 2013; 81(4):715–21. doi: 10.1002/Prot.24232 PMID: ISI:000316003500015.
8. Russell RJ, Hough DW, Danson MJ, Taylor GL. The crystal structure of citrate synthase from the thermophilic archaeon, Thermoplasma acidophilum. Structure. 1994; 2(12):1157–67. Epub 1994/12/15. PMID: 7704526.
9. Querol E, PerezPons JA, MozoVillarias A. Analysis of protein conformational characteristics related to thermostability. Protein Engineering. 1996; 9(3):265–71. doi: 10.1093/protein/9.3.265 PMID: ISI:A1996UN04900003.
10. Vihinen M. Relationship of protein flexibility to thermostability. Protein Engineering. 1987; 1(6):477–80. Epub 1987/12/01. doi: 10.1093/protein/1.6.477 PMID: 3508295.
11. Rathi PC, Radestock S, Gohlke H. Thermostabilizing mutations preferentially occur at structural weak spots with a high mutation ratio. Journal of biotechnology. 2012; 159(3):135–44. Epub 2012/02/14. doi: 10.1016/j.jbiotec.2012.01.027 PMID: 22326626.
12. Radestock S, Gohlke H. Protein rigidity and thermophilic adaptation. Proteins. 2011; 79(4):1089–108. Epub 2011/01/20. doi: 10.1002/prot.22946 PMID: 21246632.
13. Radestock S, Gohlke H. Exploiting the link between protein rigidity and thermostability for data-driven protein engineering. Eng Life Sci. 2008; 8(5):507–22. doi: 10.1002/elsc.200800043

14. Hernandez G, LeMaster DM. Reduced temperature dependence of collective conformational opening in a hyperthermophile rubredoxin. Biochemistry-Us. 2001; 40(48):14384–91. doi: 10.1021/Bi0112560 PMID: ISI:000172465100009.

15. Hernandez G, Jenney FE, Adams MWW, LeMaster DM. Millisecond time scale conformational flexibility in a hyperthermophile protein at ambient temperature. P Natl Acad Sci USA. 2000; 97(7):3166–70. doi: 10.1073/pnas.040569697 PMID: ISI:000086195200039.

16. Fitter J, Heberle J. Structural equilibrium fluctuations in mesophilic and thermophilic alpha-amylase. Biophys J. 2000; 79(3):1629–36. PMID: ISI:000089060900043.

17. Danciulescu C, Ladenstein R, Nilsson L. Dynamic arrangement of ion pairs and individual contributions to the thermal stability of the cofactor-binding domain of glutamate dehydrogenase from Thermotoga maritima. Biochemistry-Us. 2007; 46(29):8537–49. doi: 10.1021/Bi7004398 PMID: ISI:000248073200005.

18. Jaenicke R. Do ultrastable proteins from hyperthermophiles have high or low conformational rigidity? Proceedings of the National Academy of Sciences. 2000; 97(7):2962–4. doi: 10.1073/pnas.97.7.2962

19. Jaenicke R, Böhm G. The stability of proteins in extreme environments. Current Opinion in Structural Biology. 1998; 8(6):738–48. PMID: ISI:000077614300011.

20. Kalimeri M, Rahaman O, Melchionna S, Sterpone F. How Conformational Flexibility Stabilizes the Hyperthermophilic Elongation Factor G-Domain. J Phys Chem B. 2013; 117(44):13775–85. doi: 10.1021/jp407078z PMID: 24087838

21. Basu S, Sen S. Do Homologous Thermophilic-Mesophilic Proteins Exhibit Similar Structures and Dynamics at Optimal Growth Temperatures? A Molecular Dynamics Simulation Study. J Chem Inf Model. 2013; 53(2):423–34. doi: 10.1021/Ci300474h PMID: ISI:000315478900012.

22. Oyeyemi OA, Sours KM, Lee T, Kohen A, Resing KA, Ahn NG, et al. Comparative Hydrogen-Deuterium Exchange for a Mesophilic vs Thermophilic Dihydrofolate Reductase at 25 degrees C: Identification of a Single Active Site Region with Enhanced Flexibility in the Mesophilic Protein. Biochemistry-Us. 2011; 50(38):8251–60. doi: 10.1021/Bi200640s PMID: ISI:000295058700016.

23. Marcos E, Jimenez A, Crehuet R. Dynamic Fingerprints of Protein Thermostability Revealed by Long Molecular Dynamics. J Chem Theory Comput. 2012; 8(3):1129–42. doi: 10.1021/Ct200877z PMID: ISI:000301396300036.

24. Henzler-Wildman K, Kern D. Dynamic personalities of proteins. Nature. 2007; 450(7172):964–72. Epub 2007/12/14. doi: 10.1038/nature06522 PMID: 18075575.

25. Henzler-Wildman KA, Lei M, Thai V, Kerns SJ, Karplus M, Kern D. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. Nature. 2007; 450(7171):913–6. Epub 2007/11/21. doi: 10.1038/nature06407 PMID: 18026087.

26. Ishima R, Torchia DA. Protein dynamics from NMR. Nature structural biology. 2000; 7(9):740–3. Epub 2000/08/31. doi: 10.1038/78963 PMID: 10966641.

27. Englander SW, Kallenbach NR. Hydrogen exchange and structural dynamics of proteins and nucleic acids. Quarterly reviews of biophysics. 1983; 16(4):521–655. Epub 1983/11/01. PMID: 6204354.

28. Weiss S. Fluorescence spectroscopy of single biomolecules. Science (New York, NY. 1999; 283 (5408):1676–83. doi: 10.1126/science.283.5408.1676 PMID: 10073925.

29. Zhang XJ, Wozniak JA, Matthews BW. Protein flexibility and adaptability seen in 25 crystal forms of T4 lysozymes. Journal of Molecular Biology. 1995; 250(4):527–52. doi: 10.1006/jmbi.1995.0396 PMID: 7616572.

30. Frank J, Agrawal RK. A ratchet-like inter-subunit reorganization of the ribosome during translocation. Nature. 2000; 406(6793):318–22. Epub 2000/08/05. doi: 10.1038/35018597 PMID: 10917535.

31. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. Nature structural biology. 2002; 9(9):646–52. doi: 10.1038/nsb0902-646 PMID: 12198485

32. Case DA. Normal mode analysis of protein dynamics. Current Opinion in Structural Biology. 1994; 4 (2):285–90. doi: 10.1016/S0959-440X(94)90321-2

33. Rathi PC, Pfleger C, Fulle S, Klein DL, Gohlke H. Statics of biomacromolecules. In: Comba P, editor. Molecular Modeling. Weinheim: Wiley-VCH; 2011. p. 281–99.

34. Kruger DM, Rathi PC, Pfleger C, Gohlke H. CNA web server: rigidity theory-based thermal unfolding simulations of proteins for linking structure, (thermo-)stability, and function. Nucleic acids research. 2013; 41(Web Server issue):W340–W8. Epub 2013/04/24. doi: 10.1093/nar/gkt292 PMID: 23609541; PubMed Central PMCID: PMC3692064.

35. Pfleger C, Rathi PC, Klein DL, Radestock S, Gohlke H. Constraint Network Analysis (CNA): A Python software package for efficiently linking biomacromolecular structure, flexibility, (thermo-)stability, and function. J Chem Inf Model. 2013; 53(4):1007–15. Epub 2013/03/23. doi: 10.1021/ci400044m PMID: 23517329.

36. Jaeger KE, Eggert T. Lipases for biotechnology. Curr Opin Biotech. 2002; 13(4):390–7. doi: 10.1016/S0958-1669(02)00341-5 PMID: ISI:000177248000017.

37. Jaeger KE, Ransac S, Dijkstra BW, Colson C, Vanheuvel M, Misset O. Bacterial Lipases. Fems Microbiol Rev. 1994; 15(1):29–63. doi: 10.1111/j.1574-6976.1994.tb00121.x PMID: ISI:A1994PH57500003.

38. Droge MJ, Boersma YL, van Pouderoyen G, Vrenken TE, Ruggeberg CJ, Reetz MT, et al. Directed evolution of Bacillus subtilis lipase A by use of enantiomeric phosphonate inhibitors: crystal structures and phage display selection. Chembiochem: a European journal of chemical biology. 2006; 7(1):149–57. Epub 2005/12/13. doi: 10.1002/cbic.200500308 PMID: 16342303.

39. Kawasaki K, Kondo H, Suzuki M, Ohgiya S, Tsuda S. Alternate conformations observed in catalytic serine of Bacillus subtilis lipase determined at 1.3 A resolution. Acta crystallographica Section D, Biological crystallography. 2002; 58(Pt 7):1168–74. Epub 2002/06/22. PMID: 12077437.

40. van Pouderoyen G, Eggert T, Jaeger KE, Dijkstra BW. The crystal structure of Bacillus subtilis lipase: a minimal alpha/beta hydrolase fold enzyme. Journal of Molecular Biology. 2001; 309(1):215–26. Epub 2001/08/09. PMID: 11491291.

41. Rajakumara E, Acharya P, Ahmad S, Sankaranaryanan R, Rao NM. Structural basis for the remarkable stability of Bacillus subtilis lipase (Lip A) at low pH. Bba-Proteins Proteom. 2008; 1784(2):302–11. doi: 10.1016/j.bbapap.2007.10.012 PMID: ISI:000253353600006.

42. Reetz MT, Carballeira JD, Vogel A. Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. Angewandte Chemie International Edition. 2006; 45 (46):7745–51.

43. Abraham T, Pack SP, Yoo YJ. Stabilization of Bacillus subtilis Lipase A by increasing the residual packing. Biocatal Biotransfor. 2005; 23(3–4):217–24. doi: 10.1080/10242420500193013 PMID: ISI:000233591500012.

44. Ahmad S, Kamal MZ, Sankaranarayanan R, Rao NM. Thermostable Bacillus subtilis lipases: In vitro evolution and structural insight. Journal of Molecular Biology. 2008; 381(2):324–40. doi: 10.1016/j.jmb.2008.05.063 PMID: ISI:000258483600008.

45. Ahmad S, Rao NM. Thermally denatured state determines refolding in lipase: Mutational analysis. Protein Sci. 2009; 18(6):1183–96. doi: 10.1002/Pro.126 PMID: ISI:000267882100007.

46. Acharya P, Rajakumara E, Sankaranarayanan R, Rao NM. Structural basis of selection and thermostability of laboratory evolved Bacillus subtilis lipase. Journal of Molecular Biology. 2004; 341(5):1271–81. doi: 10.1016/j.jmb.2004.06.059 PMID: ISI:000223480000011.

47. Kamal MZ, Ahmad S, Molugu TR, Vijayalakshmi A, Deshmukh MV, Sankaranarayanan R, et al. In Vitro Evolved Non-Aggregating and Thermostable Lipase: Structural and Thermodynamic Investigation. Journal of Molecular Biology. 2011; 413(3):726–41. doi: 10.1016/j.jmb.2011.09.002 PMID: ISI:000296950200018.

48. Kamal MZ, Ahmad S, Yedavalli P, Rao NM. Stability curves of laboratory evolved thermostable mutants of a Bacillus subtilis lipase. Bba-Proteins Proteom. 2010; 1804(9):1850–6. doi: 10.1016/j.bbapap.2010.06.014 PMID: ISI:000280976800018.

49. Augustyniak W, Brzezinska AA, Pijning T, Wienk H, Boelens R, Dijkstra BW, et al. Biophysical characterization of mutants of Bacillus subtilis lipase evolved for thermostability: factors contributing to increased activity retention. Protein Sci. 2012; 21(4):487–97. Epub 2012/01/24. doi: 10.1002/pro.2031 PMID: 22267088; PubMed Central PMCID: PMC3375749.

50. Kamal MZ, Mohammad TAS, Krishnamoorthy G, Rao NM. Role of Active Site Rigidity in Activity: MD Simulation and Fluorescence Study on a Lipase Mutant. Plos One. 2012; 7(4):e35188. ARTN e35188 doi: 10.1371/journal.pone.0035188 PMID: ISI:000305341600126.

51. Jacobs DJ, Rader AJ, Kuhn LA, Thorpe MF. Protein flexibility predictions using graph theory. Proteins. 2001; 44(2):150–65. doi: 10.1002/prot.1081 PMID: 11391777.

52. Jacobs DJ, Thorpe MF. Generic rigidity percolation: the pebble game. Physical review letters. 1995; 75 (22):4051–4. Epub 1995/11/27. doi: 10.1103/PhysRevLett.75.4051 PMID: 10059802.

53. Jacobs DJ, Hendrickson B. An algorithm for two-dimensional rigidity percolation: the pebble game. Journal of Computational Physics. 1997; 137(2):346–65. doi: 10.1006/jcph.1997.5809

54. Rader AJ, Hespenheide BM, Kuhn LA, Thorpe MF. Protein unfolding: rigidity lost. P Natl Acad Sci USA. 2002; 99(6):3540–5. Epub 2002/03/14. doi: 10.1073/pnas.062492699 PMID: 11891336; PubMed Central PMCID: PMC122559.

55. Pfleger C, Radestock S, Schmidt E, Gohlke H. Global and local indices for characterizing biomolecular flexibility and rigidity. Journal of computational chemistry. 2013; 34(3):220–33. Epub 2012/09/26. doi: 10.1002/jcc.23122 PMID: 23007873.

56. Pfleger C, Gohke H. Efficient and robust analysis of biomacromolecular flexibility using ensembles of network topologies based on fuzzy noncovalent constraints. Structure. 2013; 21(10):1725–34. Epub 2013/09/03. doi: 10.1016/j.str.2013.07.012 PMID: 23994009.

57. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. Nucleic acids research. 2000; 28(1):235–42. PMID: 10592235

58. Krivov GG, Shapovalov MV, Dunbrack RL. Improved prediction of protein side-chain conformations with SCWRL4. Proteins. 2009; 77(4):778–95. doi: 10.1002/Prot.22488 PMID: ISI:000271602000003.

59. Word JM, Lovell SC, Richardson JS, Richardson DC. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation1. Journal of Molecular Biology. 1999; 285 (4):1735–47. PMID: 9917408

60. Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, et al. The Amber biomolecular simulation programs. Journal of computational chemistry. 2005; 26(16):1668–88. Epub 2005/10/04. doi: 10.1002/jcc.20290 PMID: 16200636; PubMed Central PMCID: PMC1989667.

61. Wang JM, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? Journal of computational chemistry. 2000; 21(12):1049–74. PMID: ISI:000088366200003.

62. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins. 2006; 65(3):712–25. doi: 10.1002/Prot.21123 PMID: ISI:000241247100017.

63. Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized born model. Proteins. 2004; 55(2):383–94. doi: 10.1002/Prot.20033 PMID: ISI:000220980600016.

64. Gohlke H, Kuhn LA, Case DA. Change in protein flexibility upon complex formation: analysis of Ras-Raf using molecular dynamics and a molecular framework approach. Proteins-Structure Function and Bioinformatics. 2004; 56(2):322–37. doi: 10.1002/prot.20116 PMID: 15211515.

65. Mamonova T, Hespenheide B, Straub R, Thorpe MF, Kurnikova M. Protein flexibility using constraints from molecular dynamics simulations. Phys Biol. 2005; 2(4):S137–S47. doi: 10.1088/1478-3975/2/4/S08 PMID: 16280619.

66. Whiteley W. Counting out to the flexibility of molecules. Phys Biol. 2005; 2(4):S116–S26. Epub 2005/11/11. doi: 10.1088/1478-3975/2/4/S06 PMID: 16280617.

67. Hespenheide BM, Jacobs DJ, Thorpe MF. Structural rigidity in the capsid assembly of cowpea chlorotic mottle virus. Journal of Physics: Condensed Matter. 2004; 16(44):S5055–S64. doi: 10.1088/0953-8984/16/44/003

68. FIRST, a program for analysing flexibility of networks. Available: http://flexweb.asu.edu/. Accessed 2014 January 17).

69. Dahiyat BI, Gordon DB, Mayo SL. Automated design of the surface positions of protein helices. Protein Sci. 1997; 6(6):1333–7. Epub 1997/06/01. doi: 10.1002/pro.5560060622 PMID: 9194194; PubMed Central PMCID: PMC2143725.

70. Folch B, Rooman M, Dehouck Y. Thermostability of salt bridges versus hydrophobic interactions in proteins probed by statistical potentials. J Chem Inf Model. 2008; 48(1):119–27. Epub 2007/12/29. doi: 10.1021/ci700237g PMID: 18161956.

71. Privalov PL, Gill SJ. Stability of protein structure and hydrophobic interaction. Advances in Protein Chemistry. 1988; 39:191–234. Epub 1988/01/01. doi: 10.1016/S0065-3233(08)60377-0 PMID: 3072868.

72. Reynolds AP, Richards G, de la Iglesia B, Rayward-Smith VJ. Clustering rules: a comparison of partitioning and hierarchical clustering algorithms. J Math Model Algorithms. 2006; 5(4):475–504.

73. Hespenheide BM, Rader AJ, Thorpe MF, Kuhn LA. Identifying protein folding cores from the evolution of flexible regions during unfolding. Journal of Molecular Graphics & Modelling. 2002; 21(3):195–207. Epub 2002/12/05. S1093-3263(02)00146-8 [pii]. PMID: 12463638.

74. Graf J, Nguyen PH, Stock G, Schwalbe H. Structure and Dynamics of the Homologous Series of Alanine Peptides: A Joint Molecular Dynamics/NMR Study. Journal of the American Chemical Society. 2007; 129(5):1179–89. doi: 10.1021/ja0660406 PMID: 17263399

75. Beermann B, Guddorf J, Boehm K, Albers A, Kolkenbrock S, Fetzner S, et al. Stability, unfolding, and structural changes of cofactor-free 1H-3-hydroxy-4-oxoquinaldine 2,4-dioxygenase. Biochemistry-Us. 2007; 46(14):4241–9. doi: 10.1021/Bi0622423 PMID: ISI:000245370400004.

76. Hung HC, Chang GG. Multiple unfolding intermediates of human placental alkaline phosphatase in equilibrium urea denaturation. Biophys J. 2001; 81(6):3456–71. PMID: ISI:000172407800038.

77. Sapra KT, Balasubramanian GP, Labudde D, Bowie JU, Muller DJ. Point mutations in membrane proteins reshape energy landscape and populate different unfolding pathways. Journal of Molecular

Biology. 2008; 376(4):1076–90. Epub 2008/01/15. doi: 10.1016/j.jmb.2007.12.027 PMID: 18191146; PubMed Central PMCID: PMC2742699.

78. Kayatekin C, Zitzewitz JA, Matthews CR. Zinc binding modulates the entire folding free energy surface of human Cu,Zn superoxide dismutase. Journal of Molecular Biology. 2008; 384(2):540–55. Epub 2008/10/09. doi: 10.1016/j.jmb.2008.09.045 PMID: 18840448; PubMed Central PMCID: PMC2756654.

79. Thorpe MF, Jacobs DJ, Chubynsky MV, Phillips JC. Self-organization in network glasses. Journal of Non-Crystalline Solids. 2000; 266-269(Part 2):859–66.

80. Rees DC, Robertson AD. Some thermodynamic implications for the thermostability of proteins. Protein Sci. 2001; 10(6):1187–94. Epub 2001/05/23. doi: 10.1110/ps.180101 PMID: 11369857; PubMed Central PMCID: PMC2374017.

81. Robertson AD, Murphy KP. Protein Structure and the Energetics of Protein Stability. Chemical reviews. 1997; 97(5):1251–68. Epub 1997/08/05. PMID: 11851450.

82. Yilmaz LS, Atilgan AR. Identifying the adaptive mechanism in globular proteins: Fluctuations in densely packed regions manipulate flexible parts. The Journal of Chemical Physics. 2000; 113(10):4454–64. http://dx.doi.org/10.1063/1.1288512.

83. Rader AJ, Yennamalli RM, Harter AK, Sen TZ. A rigid network of long-range contacts increases thermostability in a mutant endoglucanase. Journal of biomolecular structure & dynamics. 2012; 30(6):628–37. Epub 2012/06/27. doi: 10.1080/07391102.2012.689696 PMID: 22731517.

84. Fulle S, Gohlke H. Analyzing the flexibility of RNA structures by constraint counting. Biophys J. 2008; 94(11):4202–19. Epub 2008/02/19. S0006-3495(08)70077-8 [pii]doi: 10.1529/biophysj.107.113415 PMID: 18281388; PubMed Central PMCID: PMC2480660.

85. Razvi A, Scholtz JM. Lessons in stability from thermophilic proteins. Protein Sci. 2006; 15(7):1569–78. Epub 2006/07/04. doi: 10.1110/ps.062130306 PMID: 16815912; PubMed Central PMCID: PMC2242557.

86. Seewald MJ, Pichumani K, Stowell C, Tibbals BV, Regan L, Stone MJ. The role of backbone conformational heat capacity in protein stability: Temperature dependent dynamics of the B1 domain of Streptococcal protein G. Protein Sci. 2000; 9(6):1177–93. PMID: ISI:000087856300013.

87. Stone MJ, Gupta S, Snyder N, Regan L. Comparison of protein backbone entropy and beta-sheet stability: NMR-derived dynamics of protein G B1 domain mutants. Journal of the American Chemical Society. 2001; 123(1):185–6. Epub 2001/03/29. PMID: 11273620.

88. Kabsch W, Sander C. Dictionary of Protein Secondary Structure—Pattern-Recognition of Hydrogen-Bonded and Geometrical Features. Biopolymers. 1983; 22(12):2577–637. doi: 10.1002/bip.360221211 PMID: ISI:A1983RV60400010.

89. Joosten RP, Beek TAHT, Krieger E, Hekkelman ML, Hooft RWW, Schneider R, et al. A series of PDB related databases for everyday needs. Nucleic acids research. 2011; 39:D411–D9. doi: 10.1093/Nar/Gkq1105 PMID: ISI:000285831700067.

90. Friendly M, Monette G, Fox J. Elliptical insights: understanding statistical methods through elliptical geometry. Statistical Science. 2013; 28(1):1–39.

91. Cleveland WS. Robust Locally Weighted Regression and Smoothing Scatterplots. J Am Stat Assoc. 1979; 74(368):829–36. doi: 10.2307/2286407 PMID: ISI:A1979JD82100012.

92. Friendly M. Corrgrams: Exploratory displays for correlation matrices. Am Stat. 2002; 56(4):316–24. doi: 10.1198/000313002533 PMID: ISI:000178751400011.