Systematically scrutinizing the impact of substitution sites on thermostability and detergent tolerance for *Bacillus subtilis* lipase A

Christina Nutschel[1,2], Alexander Fulton[3], Olav Zimmermann[2], Ulrich Schwaneberg[5,6], Karl-Erich Jaeger[3,4], Holger Gohlke[1,2,7,*]

[1] John von Neumann Institute for Computing (NIC) and Institute for Complex Systems - Structural Biochemistry (ICS-6), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
[2] Jülich Supercomputing Centre (JSC), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
[3] Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 52425 Jülich, Germany
[4] Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
[5] Institute of Biotechnology, RWTH Aachen University, 52074 Aachen, Germany
[6] DWI-Leibniz-Institute for Interactive Materials, 52056 Aachen, Germany
[7] Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

Running title: Impact of substitution sites on thermostability and detergent tolerance

*Corresponding author:
John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), and
Institute for Complex Systems - Structural Biochemistry (ICS-6)
Forschungszentrum Jülich GmbH
Wilhelm-Johnen-Straße
52425 Jülich
Germany
Email: gohlke@uni-duesseldorf.de or h.gohlke@fz-juelich.de

Abstract

Improving an enzyme's (thermo-)stability or tolerance against solvents and detergents is highly relevant in protein engineering and biotechnology. Recent developments have tended towards data-driven approaches, where available knowledge about the protein is used to identify substitution sites with high potential to yield protein variants with improved stability and, subsequently, substitutions are engineered by site directed or site saturation (SSM) mutagenesis. However, the development and validation of algorithms for data-driven approaches has been hampered by the lack of availability of large-scale data measured in a uniform way and being unbiased with respect to substitution types and locations. Here, we extend our knowledge on guidelines for protein engineering following a data-driven approach by scrutinizing the impact of substitution sites on thermostability or / and detergent tolerance for *Bacillus subtilis* lipase A (*Bs*LipA) at very large-scale. We systematically analyze a complete experimental SSM library of *Bs*LipA containing all 3439 possible single variants, which was evaluated as to thermostability and tolerances against four detergents under respectively uniform conditions. Our results provide systematic and unbiased reference data at unprecedented scale for a biotechnologically important protein, identify consistently defined *hot spot* types for evaluating the performance of data-driven protein engineering approaches, and show that the rigidity theory and ensemble-based approach Constraint Network Analysis yields (CNA) *hot spot* predictions with an up to 9-fold gain in precision over *random classification*.

1. Introduction

Improving a protein's (thermo-)stability [1-8] or tolerance against solvents [9-16] and detergents [17-19] has become of utmost importance in protein engineering: Considering that enzymes are predominantly used as detergent additives [20] and that the global industrial enzyme market has been forecast to reach $7.0 billion by 2023 from $5.5 billion in 2018 makes clear that an increasing demand exists for enzymes that are adapted to harsh temperature, solvent, and detergent conditions [20-22].

Modifying protein stability based on rational approaches has a long history [23, 24] and a number of, usually, structure-based algorithms have been developed that estimate the effect of a substitution on the stability of a protein [25-28]. However, despite successful applications in single cases (e.g., see Table II in ref. [20]), the general reliability of these approaches is still unsatisfactory [25, 29-32]. One reason is that multiple attempts to identify key features in protein sequences and/or structures associated with protein stability have failed to paint a clear picture, which makes it difficult to define rules of universal validity and general applicability [20, 33]. Another reason lies in the data used in the design and evaluation of rational design algorithms. The ProTherm database [34, 35], which has been most frequently used for such endeavors, contains on average ~12 single, ~12 double, and ~1 multiple substitution for each of the ~1000 proteins stored [33]. Thus, while overall exhaustive, the data may not include a sufficient number of variants per protein to compensate for outliers and, therefore, may not allow a stratification of the data to derive a generally applicable set of rules. As such data, furthermore, originates from different experimental methods, it is not surprising that different changes in protein stability have been found associated with the same variant [36]. In addition, the data is strongly biased towards substitutions to alanine, whereas it is very limited for some other substitutions [37]. Recently, comprehensive mutagenesis data on a domain level associated with protein stabilities against a denaturating agent have been reported as a means to overcome these limitations [38].

Following the principles of natural evolution, albeit on a reduced timescale, protein engineering by directed evolution has emerged as an attractive strategy to improve stability through iterative cycles of mutagenesis and screening or selection [20, 39]. However, the highly labor-intensive method can become technically challenging if beneficial mutations need to be accumulated over generations of mutagenesis and screening or selection to reach a desired effect [40]. After all, evolution is not good for problems that require multiple, simultaneous, low-probability events [41]. To successfully

investigate the then necessary large protein libraries, powerful automated techniques for rapid high-throughput screenings were established [20, 39].

As an intermediate, third route recent developments have tended towards data-driven approaches [42], where available knowledge about the protein is used to, first, identify a substitution site with high potential to yield protein variants with improved stability and, second, substitutions are engineered by site directed (SDM) or site saturation (SSM) mutagenesis [33]. The "knowledge" can arise from sequence information [42, 43], structure information [44-46] or computational techniques [2, 4, 7, 8, 47, 48]. By such data-driven approaches, the challenge of accurately predicting the effect of a substitution on protein stability is circumvented, and substitution efforts are guided to a few, distinguished sequence positions, making subsequent combinations feasible. However, even with high-throughput screening techniques it is difficult to handle all variants based on combinations of the 20 proteinogenic AAs at more than six substitution sites (i.e., more than $20^6 = 6.4 * 10^7$ variants) [20, 39, 49, 50].

Here, to extend our knowledge on guidelines for time- and cost-efficient protein engineering following a data-driven approach, we scrutinize the impact of substitution sites on thermostability or / and detergent tolerance for one protein at very large-scale. To do so, we systematically analyze a complete experimental site saturation mutagenesis (SSM) library of *Bs*LipA produced by us [15, 16, 19], which contains all 3439 theoretically possible single variants (181 substitution sites of *Bs*LipA x 19 naturally occurring AAs) and was evaluated as to different protein stabilities under respectively uniform conditions. Previously, the SSM library has been characterized regarding solvent and detergent tolerance (*D*) data [15, 16, 19]. Here, we characterize the SSM library for the first time regarding thermostability (T_{50}) as well as combined T_{50} and *D* data. *Bs*LipA is a particularly interesting protein for such analysis, because a high-resolution X-ray crystal structure (PDB ID: 1ISP, 1.3 Å) is known [51], which provides valuable insights in atomic details. Furthermore, the protein has considerable biotechnological importance [52, 53], possesses an α/β-hydrolase fold [54] such that the impact of substitution sites at α-helices, β-strands, and other secondary structure elements can be tested, and has been used frequently as a model system in related experimental and computational small-scale studies [7, 8].

Our systematic large-scale analysis focusses on the following five aspects: **I)** We determined the likelihoods to find substitution sites showing significantly increased T_{50} or *D* and investigated the frequencies and magnitudes of effects caused by single AA substitutions. **II)** We analyzed at which substitution sites variants result with increased T_{50} or / and *D* across the protein and compared the

findings to random mutagenesis. **III)** From these results, we defined *hot spot* classes, i.e. classes of substitution sites particularly promising to increase T_{50} or / and D. **IV)** We probed to what extent *hot spots* can be predicted based on structure or sequence characteristics. **V)** We tested the predictive power of the rigidity theory-based approach Constraint Network Analysis (CNA) previously applied in related scenarios [2, 4-8], i.e., how accurately *hot spots* can be predicted as structural *weak spots* identified in a thermal unfolding simulation of the protein.

The main outcomes from our analyses are that we provide systematic and unbiased reference data at large scale for thermostability measured as T_{50} values and detergent tolerance measured as D for a biotechnologically important protein, we identify and consistently define *hot spot* types for evaluating the performance of data-driven protein engineering approaches, and we show that CNA-based *hot spot* prediction can yield a gain in precision over *random classification* up to 9-fold.

2. Materials and Methods

2.1. Generation and screening of the *Bs*LipA SSM library towards changes in T_{50} or *D*

The *Bs*LipA library was constructed by site saturation mutagenesis (SSM) and site directed mutagenesis (SDM) as described by Frauenkron-Machedjou *et al.* [15, 16] and Fulton *et al.* [19]. In the present study, we defined all 3439 single variants (181 substitution sites of *Bs*LipA x 19 naturally occurring AAs) generated with SSM and SDM as "SSM library".

Previously, the SSM library has been screened towards its tolerance against four different classes of detergents: anionic (sodium dodecyl sulfate, SDS), cationic (cetyltrimethylammonium bromide, CTAB), zwitterionic (3-[hexadecyl(dimethyl)azaniumyl]propane-1-sulfonate, SB3-16), and non-ionic (Polyoxyethylenesorbitan monooleate, Tween 80) by Fulton *et al.* [19]. Residual activities of the variants after incubation in the presence of the respective detergent (*D*) were obtained as described in ref. [19].

As to the screening procedure regarding thermostability, the screening cultures were incubated as described in ref. [19]. The culture supernatant was collected by centrifugation (1500 g, 40 min) and diluted 2.5-fold with Sørensen buffer (42.5 ml Na_2HPO_4 (8.9 g l^{-1}), 2.5 ml KH_2PO_4 (6.8 g l^{-1})) before screening. The protein containing supernatant was incubated in a 0.2 ml PCR microtiter plate (MTP) in a programmable thermal cycler (Eppendorf Mastercycler Thermal Cycler PCR). The supernatant samples were incubated at temperatures between 40°C to 60°C for 20 min. A dry block incubator (MRK 23 Cooling-ThermoMixer, DITABIS) was equipped with a "15 ml and 50 ml falcon tube adaptor" (BT 03, DITABIS). Three falcon tubes with 19.8 ml of *para* nitrophenyl palmitate (*p*NPP) solution A (19.8 ml Sørensen buffer, 45.54 mg sodium deoxycholate, 22 mg gum arabic) were inserted into the falcon tube incubator. All dry block incubators were set to 40°C, 30 min prior to the beginning of the experiment. 20 s before the end of the incubation, 2.2 ml of *p*NPP solution B (48 mg *p*NPP in 8 ml 2-propanol) was added into prewarmed *p*NPP solution A and briefly mixed. The substrate mixture was applied to the wells of the MTPs in 50 µl aliquots to start the measurement of thermostability and measured in a MTP reader (Molecular Devices Spectramax). The enzymatic activity in each sample was measured by the rate of increase in absorption at O.D.410 nm. The residual activity in each sample was calculated from the slope of the change in absorption at O.D.410 nm relative to the slope of the sample heated to 40°C during a measurement time of 3 min. From that, T_{50} was obtained from the inflection point of a sigmoid curve fit. Control experiments with just *p*NPP, or *p*NPP in the presence of *Bs*LipA at temperatures

up to 60.6°C, that way leading to denaturation of *Bs*LipA, show no change in the *para* nitrophenolate (*p*NP) absorption over time, demonstrating that *p*NP is only produced in the presence of a functional enzyme **(Figure S1)**. The T_{50} values are provided as an Excel sheet in the Supporting Information.

2.2. Global characterization of *Bs*LipA variants' changes in T_{50} or D

For analyzing the changes in T_{50} **(Eq. 1)** or D **(Eq. 2)** of *Bs*LipA variants, the values of wt*Bs*LipA were used as references, i.e., the differences between the values of the variants and those of wt*Bs*LipA were calculated. Positive (negative) Δ-values indicate variants with increased (decreased) T_{50} or D.

$$\Delta T_{50} = T_{50}(\text{variant}) - T_{50}(\text{wt}Bs\text{LipA}) \quad \textbf{Eq. 1}$$

$$\Delta D = D\ (\text{variant}) - D\ (\text{wt}Bs\text{LipA}) \quad \textbf{Eq. 2}$$

For the large-scale analysis, only ΔT_{50} of variants higher (lower) than the experimental uncertainty, taken as the standard deviation σ_T for the respective variant determined from three screenings of T_{50}, were considered significantly increased (decreased) in T_{50} compared to wt*Bs*LipA. Furthermore, only ΔD of variants higher (lower) than two times the experimental standard deviation ($2\sigma_D$) of wt*Bs*LipA determined from screenings of 2997 wt*Bs*LipA replicates [19] towards the respective detergent were considered significantly increased (decreased) in D compared to wt*Bs*LipA. Here, σ_D of wt*Bs*LipA was used as significance criterion, as the experimental standard deviation for each variant was not available. $2\sigma_D$ was chosen because it corresponds to a *p*-value below 0.05.

2.3. Definitions of classes of *Bs*LipA substitution sites

The different classes of substitution sites regarding significantly increased T_{50} or / and D were defined based on the set theory. Therefore, the following binary operations on sets were applied:
The *union* of the sets A and B is the set of elements which are in A, in B, or in both A and B **(Eq. 3)** [55].

$$(A \cup B) = \{x : x \in A \vee x \in B\} \quad \textbf{Eq. 3}$$

The *intersection* of the sets *A* and *B* is the set of elements which are in *A* and *B* **(Eq. 4)** [55].

$$(A \cap B) = \{x : x \in A \wedge x \in B\} \quad \textbf{Eq. 4}$$

Finally, the Jaccard index (*J*) was used to compare the similarity of two sets *A* and *B*, i.e., the cardinal number of the respective *intersection* divided by the cardinal number of the respective *union* **(Eq. 5)** [56, 57]. The range of *J* is [0, 1], with one indicating identical sets *A* and *B*.

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} \quad \textbf{Eq. 5}$$

Based on the different classes of substitution sites, we defined *hot spots*, which are substitution sites particularly promising to yield significantly increased T_{50} or / and *D*.

2.4. Structural determinants of *Bs*LipA *hot spots*

Hot spots were assigned to groups according to their location in secondary structure elements (yielding 20 subgroups), solvent accessible surface areas (SASAs) (yielding five subgroups), and physicochemical properties (yielding five subgroups). The secondary structure elements of the wt*Bs*LipA crystal structure (PDB ID: 1ISP with highest resolution of 1.3 Å [51]) were identified with the DSSP program [58]. Additionally, the SASAs of the wt*Bs*LipA were analyzed with the DSSP program [58]. The fractional solvent accessible surface areas (fSASAs) were calculated with respect to the maximum solvent accessible surface area of each *hot spot* (maxSASA) **(Eq. 6)** [59].

$$\text{fSASA} = 100 \cdot \frac{\text{SASA}}{\text{maxSASA}} \quad \textbf{Eq. 6}$$

As the screening studies were performed at pH 8 [19], *hot spots* were subgrouped by their physicochemical properties as follows: aliphatic (Ile, Ala, Val, Leu, Gly), aromatic (Phe, Tyr, Trp), neutral (Cys, Pro, Met, Ser, Thr, Asn, Gln), positively charged (His, Lys, Arg), and negatively charged (Asp, Glu).

2.5. Conservation of wt*Bs*LipA residues within bacterial lipases

Apart from the catalytic triad (S77, D133, and H156), also variants at conserved sequence positions were considered because the SSM library revealed significantly increased T_{50} or / and D at such positions. The conservation of wt*Bs*LipA residues within the bacterial lipases was calculated using the available sequences from the Pfam database [60] for the lipase class 2 (PF01674). The sequences were limited to the bacterial sources, which contain 1138 sequences from 603 bacterial species. All sequences were aligned using Clustal Omega [61, 62]. For the alignment, the full-length sequence of wt*Bs*LipA (UniProt ID: P37957) was used [63]. The conservation was calculated using AACon Calculations [64] through Jalview [65]. The conservation range is [0, 10] with 0 (10) showing no (high) conservation.

2.6. Constraint Network Analysis (CNA)

The Constraint Network Analysis (CNA) aims at linking structural rigidity and flexibility to the biomolecule's structure, (thermo)stability, and function [66-68]. The CNA software acts as front- and back-end to the graph theory-based rigidity analysis software Floppy Inclusions and Rigid Substructure Topography (FIRST) [69]. In CNA, proteins are modelled as constraint networks in a *body-and-bar* representation, which has been described in detail by Hesphenheide *et al* [70]. Based on the modelled constraint network of the protein structure, a *pebble game algorithm* decomposes the network into flexible and rigid subparts [71, 72]. In order to monitor the decay of network rigidity and to identify the *rigidity percolation threshold*, CNA performs thermal unfolding simulations by consecutively removing non-covalent constraints (hydrogen bonds, including salt bridges) from a network in increasing order of their strength [73]. For this, a hydrogen bond energy E_{HB} is computed by a modified version of the potential by Mayo *et al*. [73]. During the thermal unfolding simulations, phase transitions can be identified where the network switches from overall rigid to flexible states. For a given network state $\sigma = f(T)$, hydrogen bonds with an energy $E_{HB} > E_{cut}(\sigma)$ are removed from the network at temperature T. In this study, the thermal unfolding simulation was carried out 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}. E_{cut} can be converted to a temperature T using the linear equation introduced by Radestock *et al*. **(Eq. 7)** [2, 4]. The range of E_{cut} is equivalent to increasing the temperature 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 [74, 75], the number of hydrophobic tethers was kept unchanged during the thermal unfolding simulation, as done previously [7, 8].

$$T = \frac{-20\ \mathrm{K}}{\mathrm{kcal} \cdot \mathrm{mol}^{-1}} E_{\mathrm{cut}} + 300\ \mathrm{K} \qquad \textbf{Eq. 7}$$

The CNA software is available under academic licenses from http://cpclab.uni-duesseldorf.de/software, and the CNA web server is accessible at http://cpclab.uni-duesseldorf.de/cna/.

2.7. Generation of a structural ensemble of wt*Bs*LipA

MD simulations of wt*Bs*LipA were carried out with the GPU-accelerated version of PMEMD [76] of the AMBER14 suite of programs [77] together with the ff14SB force field [78]. As a starting structure, the X-ray crystal structure of wt*Bs*LipA (PDB ID: 1ISP) was used [51]. Hydrogens were added and side-chain orientations ("flips") of Asn, Gln, and His were optimized by the REDUCE program [79] based on suitable hydrogen-bonding geometries and avoiding potential steric clashes. This was done to take into account that O *versus* N or N *versus* C are difficult to distinguish in X-ray crystallography experiments [79]. For neutralization of the system, sodium counter-ions were added. Subsequently, the system was solvated by a truncated octahedral box of TIP3P water [80] such that a layer of water molecules of at least 11 Å widths covers the protein surface. The particle mesh Ewald method [81] was used with a direct-space non-bonded cutoff of 8 Å. Bond lengths involving hydrogen atoms were constrained using the SHAKE algorithm [82], and the time step for the simulation was 2 fs. As done before [8], a trajectory of 100 ns length was generated after thermalization and adjustment of the pressure, simulating in the canonical (NVT) ensemble at T = 300 K, with conformations extracted every 40 ps from the last 80 ns, resulting in a structural ensemble of 2000 conformations. We assessed the statistical independence of the extracted conformations by calculating the autocorrelation function of the cluster configuration entropy H_{type2}, the measure used to identify phase transitions in the constraint networks **(see section 2.9 below)** (**Figure S2**). Because fluctuations of H_{type2} decorrelate already within the first two snapshots, the snapshots used for CNA, which were extracted at time intervals of 40 ps, are considered independent.

2.8. Thermal unfolding simulation of wt*Bs*LipA

For analyzing the rigid cluster decomposition of wt*Bs*LipA, a thermal unfolding simulation was performed by CNA on an ensemble of network topologies (ENT^{MD}) generated from a molecular dynamics (MD) trajectory. The ensemble-based CNA was pursued to increase the robustness of the rigidity analyses [83, 84]Subsequently, the unfolding trajectory was visually inspected by VisualCNA [85] for identifying secondary structure elements that segregate from the largest rigid cluster at each major phase transition. VisualCNA is an easy-to-use PyMOL plug-in that allows setting up CNA runs and analyzing CNA results linking data plots with molecular graphics representations [85]. VisualCNA is available under an academic license from http://cpclab.uni-duesseldorf.de/software.

2.9. Local and global indices for analyzing structural rigidity of wt*Bs*LipA

From the thermal unfolding simulation, CNA computes a comprehensive set of indices to quantify biologically relevant characteristics of the biomolecule's stability [86]. *Global* indices are used for determining the flexibility and rigidity at a macroscopic level. *Local* indices determine the flexibility and rigidity at a microscopic level of bonds.

The cluster configuration entropy H_{type2} is a *global* index, which has been introduced by Radestock and Gohlke [2]. H_{type2} is used to identify the phase transition temperature T_p at which a biomolecule switches from a rigid to a floppy state and the largest rigid cluster stops to dominate the whole protein network. As long as the largest rigid cluster dominates the whole protein network, H_{type2} is low because of the limited number of possible ways to configure a system with a very large cluster. When the largest rigid cluster starts to decay or stops to dominate the network, H_{type2} jumps. There, the network is in a partially flexible state with many ways to configure a system consisting of many small clusters. The percolation behavior of protein networks is usually complex, and multiple phase transitions can be observed [2, 4, 5, 7, 8]. In order to identify T_p, a double sigmoid fit was applied to an H_{type2} *versus* $T(E_{cut})$ curve as done previously [2, 4, 5, 7, 8], and T_p taken as that T value associated with the largest slope of the fit.

The stability map rc_{ij} is a *local* index, which has been introduced by Radestock and Gohlke [4]. rc_{ij} represents the local stability within a protein structure for all residue pairs at which a rigid contact *rc* between two residues *i* and *j* (represented by their C_α atoms) is lost during the thermal unfolding.

rc exists if *i* and *j* belong to the same rigid cluster *c* of the set of rigid clusters $C^{E_{\text{cut}}}$ [86]. Thus, rc_{ij} contains information cumulated over all network states along the unfolding trajectory as to which parts of the network are (locally) mechanically stable at a given σ, and which are not [7]. This stability information is not only available in a qualitative manner but also quantitatively in that each rc_{ij} has been associated with E_{cut} at which the rigid contact is lost. The sum over all entries in rc_{ij} represents the chemical potential energy due to noncovalent bonding, obtained from the coarse-grained, residue-wise network representation of the underlying protein structure. To focus only on the stability of *rc* between structurally close residues, rc_{ij} was filtered such that only rigid contacts between two residues that are at most 5 Å apart from each other were considered (neighbor stability map $rc_{ij,neighbor}$).

Finally, CNA predicts *unfolding nuclei* as structural features from which macroscopic (in)stability originates [87]. *Unfolding nuclei* are represented by residues that percolate from the largest rigid cluster at the latest phase transition. If such residues become flexible, it will have a detrimental effect on protein stability. Fringe residues of the *unfolding nuclei* percolate from the largest rigid cluster during earlier steps of the thermal unfolding. We follow the hypothesis that the more structurally stable the fringes of *unfolding nuclei* are, the more structurally stable will be those *unfolding nuclei* [87]. Therefore, if such fringe residues (termed *weak spots*) are targeted by substitutions, the likelihood to stabilize the rigid core of a protein should be high. If two *unfolding nuclei* were only separated by one residue, this residue was also considered a *weak spot*. This procedure of identifying *weak spots* is in agreement with a previous study of us [87].

2.10. Statistical evaluation of CNA as a binary classifier

The performance of CNA was investigated as a binary classifier with the following possible outcomes: true positives (TP) are predicted *weak spots* that are *hot spots*, whereas false positives (FP) are predicted *weak spots* that are non-*hot spots*. In turn, true negatives (TN) are predicted non-*weak spots* that are non-*hot spots*, whereas false negatives (FN) are predicted non-*weak spots* that are *hot spots*. Different metrics were then applied to evaluate CNA.

The *recall* (*r*) answers the question how many *hot spots* were predicted as *weak spots* **(Eq. 8)** [88].

$$r = \frac{TP}{TP + FN} = \frac{No.\,of\ predicted\ weak\ spots\ that\ are\ hot\ spots}{No.\,of\ hot\ spots} \qquad \textbf{Eq. 8}$$

The *precision* (*p*) evaluates how many predicted *weak spots* are actually *hot spots* **(Eq. 9)** [88].

$$p = \frac{TP}{TP + FP} = \frac{No.of\ predicted\ weak\ spots\ that\ are\ hot\ spots}{No.of\ weak\ spots} \quad \textbf{Eq. 9}$$

The *precision* in *random classification* (p_{random}) indicates how many of the 181 *Bs*LipA residues are actually *hot spots* **(Eq. 10)** [88].

$$p_{random} = \frac{TP + FN}{TP + FP + TN + FN} = \frac{No.of\ hot\ spots}{181\ residues\ of\ BsLipA} \quad \textbf{Eq. 10}$$

The *gain in precision* over *random classification* (*gip*) represents how many predicted *weak spots* are actually *hot spots* in comparison to *random classification* **(Eq. 11)** [88]. The *gip* range is [0, ∞], with values < 1 indicating a lower precision than obtained by *random classification*.

$$gip = \frac{p}{p_{random}} \quad \textbf{Eq. 11}$$

The *F_1-score* (F_1) is a measure of the test's accuracy. It represents the harmonic mean of *p* and *r*, i.e., if there is an uneven class distribution, it is used to seek a balance between *p* and *r* **(Eq. 12)** [89]. The F_1 range is [0, 1], with one indicating perfect *r* and *p*.

$$F_1 = 2 \cdot \frac{p \cdot r}{p + r} \quad \textbf{Eq. 12}$$

2.11. Markov Chain Monte Carlo-based unfolding simulations of wt*Bs*LipA

As an independent method to assess the order of unfolding of wt*Bs*LipA, we used a Markov Chain Monte Carlo (MCMC) simulation with an all-atom model restricted to dihedral degrees of freedom [90]. This method has been successfully used for protein folding simulations [91] and has been shown to reproduce the order of melting temperatures for a set of protein variants [92]. In this MCMC model, implemented in the open source tool ProFASi (Protein Folding and Aggregation Simulator), the protein conformation is modified by changing one or few dihedral angles in each step. A step is accepted according to the Metropolis criterion, i.e., with a probability that depends on the absolute temperature and the resulting change of energy of the system. In ProFASi, the energy is calculated by an all-atom implicit solvent force field [92, 93]. While MCMC simulations allow arbitrarily large

changes to the molecule, the unfolding simulations for this study have been restricted to side chain dihedral updates and small, locally correlated updates of main chain dihedral angles [94]. To ensure adequate sampling, 96 MCMC simulations at 330 K were performed with a total of $3.05 * 10^{10}$ elementary updates.

3. Results

3.1. About one-tenth of all variants in the complete SSM library show significantly increased T_{50} or D towards at least one detergent, and such variants were found at two-thirds of all substitution sites

The *Bs*LipA SSM library contained T_{50} as well as D data towards the four detergents SDS, CTAB, SB3-16, and Tween 80 for all 3439 single variants (181 substitution sites of *Bs*LipA x 19 naturally occurring AAs), including also inactive variants **(see section 2.1)**. Initially, the results of both experimental screening studies of the SSM library with respect to changes in T_{50} (ΔT_{50}) or D towards one detergent (ΔD) were assessed in terms of the variance of the data and its significance **(see section 2.2)**.

As to the T_{50} data, only variants with ΔT_{50} higher (lower) than the experimental uncertainty, taken as the standard deviation σ_T for the respective variant determined from three screenings of T_{50}, were considered significantly increased (decreased) in T_{50} compared to wt*Bs*LipA (ΔT_{50} = 0 K) **(Eq. 1)**. The average σ_T is 0.44 K. In total, 1856 variants with significantly increased T_{50} were obtained, of which 214 (~12%) show an increase and 1642 (~88%) a decrease **(Figure 1A, Table S1)**. This proportion represents what one would obtain in the case of *random mutagenesis*. The distribution of ΔT_{50} is left-skewed, with extreme ΔT_{50} values of -8.3 K and +7.7 K, and the most frequent ΔT_{50} range being -2 K to -1.5 K (~12% out of 1856 variants), followed by ΔT_{50} between -1.5 K and -1 K (~10% out of 1856 variants) **(Figure 1A)**. In turn, for each of 69 substitution sites (~38% out of 181 substitution sites) at least one variant with significantly increased T_{50} was found. These substitution sites are summarized in class **I** (**I** = {Substitution site$_x$ | $1 \leq x \leq 181$, $T_{50}(x)$ is significantly increased}) **(Tables 1 and S2)**.

Likewise, only variants with ΔD higher (lower) than two times the experimental standard deviation ($2\sigma_D$) of wt*Bs*LipA determined from screenings of 2997 wt*Bs*LipA replicates [19] towards the respective detergent were considered significantly increased (decreased) in D compared to wt*Bs*LipA (ΔD = 0) **(Eq. 2)**. The screening revealed the highest σ_D in the presence of SB3-16, followed by Tween 80, CTAB, and SDS **(Table S1)** [19]. This may be related to the fact that SB3-16 and Tween 80 were tested above the critical micelle concentration (CMC), while CTAB and SDS were tested below it [19, 95]. The respective detergent concentration had been chosen based on the inactivation of purified wt*Bs*LipA **(Table S1)** [19]. On average, 900 variants with significantly

increased *D* were obtained, of which 126 (~14%) show an increase and 774 (~86%) a decrease, on average across each detergents **(Figures 1B – E, Table S1)**. This proportion represents what one would obtain in the case of *random mutagenesis*. The distribution of Δ*D* is left-skewed. The magnitude of the increase (decrease) in Δ*D* is between 1.6-fold and 2.4-fold (0.6-fold and 2.9-fold) of the residual activity of wt*Bs*LipA. Furthermore, variants tested against SDS and SB3-16 showed an up to two times higher Δ*D* than against CTAB and Tween 80 **(Figures 1B – E)**. This may be related to the different classes of the detergents [19, 95]. In turn, for each of 74, 42, 46, or 34 substitution sites at least one variant with significantly increased *D* towards SDS, CTAB, SB3-16, or Tween 80 (~41, 23, 25 or 19% out of 181 substitution sites) was found. These substitution sites are summarized in classes **II – V** (**II – V** = {Substitution site$_x$ | $1 \leq x \leq 181$, $D_{\text{SDS / CTAB / SB3-16 / Tween 80}}(x)$ is significantly increased}) **(Tables 1 and S2)**. The *union* of **II – V** contains 109 substitution sites (~60% out of 181 substitution sites) and is represented by class **VI** (**VI** = **II** ∪ **III** ∪ **IV** ∪ **V**) **(Tables 1 and S2, Eq. 3)**. For each of these substitution sites at least one variant shows significantly increased *D* towards at least one detergent**.**

Finally, 124 substitution sites are summarized in the *union* of **I** and **VI** (~69% out of 181 substitution sites) (**VII** = **I** ∪ **VI**) **(Tables 1 and S2, Eq. 3)**. Thus, only for two-thirds of all substitution sites at least one variant with significantly increased T_{50} or *D* towards at least one detergent were obtained.

To conclude, for the first time, we performed a systematic large-scale analysis of a complete experimental SSM library towards two types of stabilities of one protein containing all single variants. The likelihoods to generate variants with significantly increased T_{50} (~12%) or *D* towards one detergent (~14% on average across all detergents) by *random mutagenesis* (**I – V**) are similar. Variants with significantly increased T_{50} or *D* towards at least one detergent were obtained at only two-thirds of all substitution sites (**VII**), and this value falls to about one third or below if T_{50} and *D* towards one detergent are considered separately (**I - V**). Hence, such substitution sites are not uniformly distributed across the protein. For the following analyses, only substitution sites with at least one variant yielding significantly increased T_{50} or *D* towards at least one detergent were considered.

3.2. The higher the frequency of substitution occurrences is that lead to significantly increased T_{50} or D towards one detergent, the more pronounced is the highest effect, and *vice versa*

Next, we investigated the *Bs*LipA SSM library regarding the respective frequency of substitution occurrences at substitution sites that lead to significantly increased T_{50} ($N_{BsLipA;\ T}$) or D ($N_{BsLipA;\ D}$) towards one detergent. Additionally, we analyzed the respective highest effects in significantly increased T_{50} ($\Delta T_{50;\ max}$) or D (ΔD_{max}) towards one detergent at substitution sites. Finally, we address the question if the frequency of substitution occurrences and the highest effects per substitution site are related to each other.

The highest $N_{BsLipA;\ T}$ of **I** was 12 (F17) **(Figure 2A)**, whereas the highest $N_{BsLipA;\ D}$ of **II** – **V** were 14 (E65), 6 (I135 and D144), 11 (G46), and 5 (V99) **(Figure 2B, Table S14)**, respectively, indicating that up to ~60% and more of the variants for some substitution sites yield significantly increased T_{50} or D towards one detergent. Correlations between $N_{BsLipA;\ T}$ of **I** and $N_{BsLipA;\ D}$ of **II** – **V** yielded, on average, $R^2 = 0.03$; $p > 0.1$ **(Figure 2C, Table S3)**. The highest correlation was found between $N_{BsLipA;\ T}$ of **I** and $N_{BsLipA;\ D}$ of **II** ($R^2 = 0.07$, $p < 0.001$). With respect to $N_{BsLipA;\ D}$ of **II** – **V**, overall very weak to weak but mostly significant correlations were obtained (on average: $R^2 = 0.11$, $p < 0.01$) **(Figure 2C, Table S3)**. The highest correlation was observed between $N_{BsLipA;\ D}$ of **III** and **IV** ($R^2 = 0.26$, $p < 0.001$).

The highest $\Delta T_{50;\ max}$ of **I** was 7.7 K (M137), whereas the highest ΔD_{max} of **II** – **V** were 1.49 (M137), 1.63 (T110), 2.41 (G46), and 2.29 (S127), respectively **(Table S9)**, indicating that specific single AA substitutions have a great impact on the magnitudes of the effects. Correlations between $\Delta T_{50;\ max}$ of **I** and ΔD_{max} of **II** – **V** showed, on average, $R^2 = 0.06$; $p > 0.1$ **(Figure 2D, Table S4)**. The highest correlation was observed between $\Delta T_{50;\ max}$ of **I** and ΔD_{max} of **IV** ($R^2 = 0.13$, $p < 0.1$). With respect to ΔD_{max} of **II** – **V**, overall very weak to weak and mostly insignificant correlations were obtained (on average: $R^2 = 0.08$, $p > 0.1$) **(Figure 2D, Table S4)**. The highest correlations were observed between ΔD_{max} of **II** and **V** ($R^2 = 0.24$, $p < 0.05$) as well as ΔD_{max} of **III** and **IV** ($R^2 = 0.13$, $p < 0.1$).

Finally, mostly good to fair and significant correlations between $N_{BsLipA;\ T}$ and $\Delta T_{50;\ max}$ of **I** as well as $N_{BsLipA;\ D}$ and ΔD_{max} of **II** – **V** were found (on average for increase: $R^2 = 0.27$, $p < 0.01$) **(Figure 2E, Table S5)**.

To conclude, these findings indicate that the relation "the higher the frequency of substitution occurrences are that lead to significantly increased T_{50} or D towards one detergent, the more pronounced is the highest effect, and *vice versa*" holds for substitution sites at which at least one variant shows significantly increased T_{50} or D towards one detergent (**I** – **V**). Together with the results from the previous chapter, this result suggests that identifying *a priori* substitution sites with a high likelihood for significantly increased T_{50} or D towards one detergent will also be beneficial with respect to the magnitude of effects that can be achieved there by substitutions.

3.3. Eleven substitution sites yield a ~4.6-fold higher likelihood to find for each detergent variants with significantly increased *D* than *random mutagenesis*

Next, we focused on pairwise *intersections* of **II** – **V** to investigate if there are substitution sites at which for two detergents at least one variant shows significantly increased D, regardless of the magnitude of the single effect **(see section 2.3)**. We compared the pairwise similarities between **II** – **V** by calculating the Jaccard index (J), i.e., the cardinal number of the respective *intersection* divided by the cardinal number of the respective *union* **(Table S6, Eq. 5)** [56, 57]. The highest similarity was found between **III** and **IV** with J(**III**, **IV**) = 0.47, whereas the lowest similarity was observed between **II** and **V** with J(**II**, **V**) = 0.23. This may be related to the different classes of the detergents [19, 95]

Encouraged by the findings of overlapping **II** – **V**, we also looked at the overall *intersection* of **II** – **V** (**VIII** = **II** ∩ **III** ∩ **IV** ∩ **V**), i.e., substitution sites at which for each detergent at least one variant shows significantly increased D, regardless of the magnitude of the single effect **(Tables 1 and S2, Eq. 4)**. **VIII** contains the eleven substitution sites E2, G13, D43, T45, Y49, N51, V54, E65, N98, M134, and M137 (~6% out of 181 substitution sites) **(Tables 1, S2, and S14)**. These substitution sites are associated with 50 variants causing a significant change in D, of which 32 (~64%) show a significant increase, on average across all detergents **(Table S7)**. Thus, this likelihood is ~4.6-fold higher in comparison to *random mutagenesis*. The most promising substitution sites of **VIII** are M134, N51, and T45 with variants showing increased ΔD_{max} of 2.25, 2.10, and 1.90, respectively.

To conclude, a dramatically reduced number of eleven substitution sites (**VIII**) yield a ~4.6-fold higher likelihood to find for each detergent variants with significantly increased D compared to

random mutagenesis. These findings indicate that if a protein engineering study aims at identifying variants showing significantly increased *D* towards each detergent, such substitution sites **(VIII)** should be identified prior to SDM.

3.4. Seven substitution sites yield a ~3.4-fold higher likelihood to find variants with significantly increased T_{50} and a ~4.7-fold higher likelihood to find for each detergent variants with significantly increased *D* than *random mutagenesis*

The same analyses were repeated for *intersections* of **I** and **II** – **V**, respectively, regarding substitution sites at which at least one variant shows significantly increased T_{50} and for one detergent significantly increased *D*, regardless of the magnitude of the single effect **(see section 2.3)**. We compared the pairwise similarities between **I** and **II** – **V**, respectively, by calculating *J* **(Table S6, Eq. 5)**. The highest similarity was found between **I** and **II** with $J(\mathbf{I}, \mathbf{II}) = 0.42$, whereas the lowest similarity was observed between **I** and **V** with $J(\mathbf{I}, \mathbf{V}) = 0.16$.

Encouraged by the findings of overlapping **I** and **II** – **V**, respectively, we also looked at the overall *intersection* of **I** and **II** – **V** (**IX** = **I** ∩ **VIII**), i.e., substitution sites at which at least one variant shows significantly increased T_{50} and for each detergent significantly increased *D*, regardless of the magnitude of the single effect **(Tables 1 and S2, Eq. 4)**. **IX** contains the seven substitution sites E2, G13, T45, Y49, V54, M134, and M137 (~4% out of 181 substitution sites) **(Tables 1, S2, and S14)**. Associated with these are 86 variants causing a significant change in T_{50}, of which 35 (~41%) show a significant increase **(Table S8)**. Thus, this likelihood is ~3.4-fold higher in comparison to *random mutagenesis*. The most promising substitution sites of **IX** are M137, M134, and Y49 with variants showing increased $\Delta T_{50;\ max}$ of 7.7, 5.6, and 1.6 K, respectively. Furthermore, associated with substitution sites of **IX** are 29 variants causing a significant change in *D*, of which 19 (~66%) show a significant increase, on average across all detergents **(Table S8)**. Thus, this likelihood is ~4.7-fold higher in comparison to *random mutagenesis*. The most promising substitution sites of **IX** are M134, T45, and M137 with variants showing increased ΔD_{max} of 2.25, 1.90, and 1.67, respectively.

To conclude, a dramatically reduced number of seven substitution sites (**IX**) yield a ~3.4-fold higher likelihood to find variants with significantly increased T_{50} and a ~4.7-fold higher likelihood to find for each detergent variants with significantly increased *D* compared to *random mutagenesis*.

These findings indicate that if a protein engineering study aims at identifying variants showing significantly increased T_{50} and D towards each detergent, such substitution sites **(IX)** should be identified prior to SDM.

3.5. Six substitution sites with highest $\Delta T_{50;\,max}$ (ΔD_{max}) yield a ~5.3-fold (~4.5-fold) higher likelihood to find variants with significantly increased T_{50} (D) than *random mutagenesis*

The above analyses focused on substitution sites at which significantly increased T_{50} or D towards one detergent (**I** – **V**), significantly increased D towards each detergent (**VIII**), as well as significantly increased T_{50} and D towards each detergent **(IX)** were observed, regardless of the magnitude of the effect. Now, we identified those six substitution sites for which the respective highest effects ($\Delta T_{50;\,max}$ or ΔD_{max}) were found. The number of six is motivated by the current technical limitation to screen more than 20^6 variants [20, 39, 49, 50].

The six substitution sites M137, M134, G155, F17, I157, and Y139 yield variants with the highest $\Delta T_{50;\,max}$ of 7.7, 5.6, 4.5, 3.8, 3.6, and 3.2 K, respectively, and constitute class **X** (**X** = {Substitution sites $_x$ | $1 \leq x \leq 181$, six highest effects in significantly increased $T_{50}(x)$}) **(Tables 1, S2, and S9)**. The substitution sites of **X** are associated with 68 variants causing a significant change in T_{50}, of which 43 (~63%) yield a significantly increased T_{50} **(Table S10)**. Thus, this likelihood is ~5.3-fold higher in comparison to *random mutagenesis*.

The most promising substitution sites exhibiting variants with the highest ΔD_{max} towards one detergent (**XI** – **XIV** = {Substitution sites $_x$ | $1 \leq x \leq 181$, six highest effects in significantly increased $D_{SDS\,/\,CTAB\,/\,SB3\text{-}16\,/\,Tween\,80}(x)$}) are M137 (**XI**), T110 (**XII**), G46 (**XIII**), and S127 (**XIV**) with variants showing highest ΔD_{max} of 1.49, 1.63, 2.41, and 2.29, respectively **(Tables 1, S2, and S9)**. With these substitution sites, 43 variants are associated causing a significant change in D, of which 27 (~63%) cause significantly increased D, on average across all detergents **(Table S10)**. Thus, this likelihood is ~4.5-fold higher in comparison to *random mutagenesis*.

Furthermore, we determined the *union* of **XI** – **XIV**, the set of 20 substitution sites (~11% out of 181 substitution sites) that yield variants showing the respective highest ΔD_{max} towards at least one detergent (**XV** = **XI** ∪ **XII** ∪ **XIII** ∪ **XIV**) **(Tables 1 and S2, Eq. 3)**. Additionally, the *union* of **X** and **XV** was defined as the set of 24 substitution sites (~13% out of 181 substitution sites), which

exhibit variants showing the respective highest $\Delta T_{50;\,max}$ or ΔD_{max} towards at least one detergent (**XVI** = **X** ∪ **XV**) **(Tables 1 and S2, Eq. 3)**.

The *intersection* between **XI** – **XIV** (**XVII** = **XI** ∩ **XII** ∩ **XIII** ∩ **XIV**) is empty, i.e., there are no common substitution sites among those six at which for each detergent variants with highest ΔD_{max} were found **(Tables 1 and S2, Eq. 4)**. The *intersection* between **X** and **XVII** (**XVIII** = **X** ∩ **XVII**) is necessarily empty, too, i.e., there are no common substitution sites among those six at which variants with highest $\Delta T_{50;\,max}$ and ΔD_{max} for each detergent were found **(Tables 1 and S2, Eq. 4)**. Thus, **XVII** and **XVIII** were not considered for the following analyses.

Additionally, we compared the pairwise similarities between **X** – **XIV** by calculating J **(Eq. 5)**. Regarding the highest ΔD_{max}, only **XII** and **XIII** overlap to some extent (J(**XII**, **XIII**) = 0.2) **Table S6)**. Regarding the highest $\Delta T_{50;\,max}$ and ΔD_{max}, only **X** and **XI**, **XII**, or X**III**, respectively, slightly overlap (J(**X**, **XI**) ≈ J(**X**, **XII**) ≈ J(**X**, **XIII**) = 0.1) **(Table S6)**.

To conclude, a highest $\Delta T_{50;\,max}$ of 7.7 K and a highest ΔD_{max} of 2.41 were found. The six substitution sites with highest $\Delta T_{50;\,max}$ yield a ~5.3-fold higher likelihood to find variants with significantly increased T_{50} (**X**); the six substitution sites with highest ΔD_{max} yield a ~4.5-fold higher likelihood to find variants with significantly increased D (**XI** – **XIV**). There are no common substitution sites among those six at which for each detergent variants with highest ΔD_{max} were found (**XVII**). Neither are there common substitution sites among those six at which variants with highest $\Delta T_{50;\,max}$ and ΔD_{max} for each detergent were found (**XVIII**).

3.6. Definition of *hot spots*

Based on these results, we defined seven types of *hot spots*, i.e. substitution sites particularly promising to cause a significant increase in T_{50} or / and D. First, the respective six substitution sites of **X** – **XIV** are considered *hot spots* because variants yield the respective highest $\Delta T_{50;\,max}$ or ΔD_{max} towards one detergent for these substitution sites **(Tables 1, S2, and S9)**. Furthermore, we showed that there is a correlation between the magnitude of an effect found at a substitution site and the frequency of substitution occurrences that lead to significantly increased T_{50} or D towards one detergent **(see section 3.2)**. Finally, generating and evaluating variants based on combinations of all

20 AAs at six substitution sites is still manageable with current protein engineering techniques [20, 39, 49, 50].

As shown above, **XVII** and **XVIII**, which would constitute the substitution sites with the broadest impact on ΔD_{max}, or $\Delta T_{50;\ max}$ and ΔD_{max}, are empty **(see section 3.5)**. Hence, we resorted to defining, second, the eleven substitution sites of **VIII** showing significantly increased *D* towards each detergent, regardless of the magnitude of the single effect **(see section 3.3)** and, third, the seven substitution sites of **IX** showing significantly increased T_{50} and *D* towards each detergent, regardless of the magnitude of the single effect **(see section 3.4)** as *hot spots* **(Tables 1 and S2)**. With eleven and seven substitution sites, these classes are also the smallest besides **X – XIV**.

3.7. *Hot spots* are diverse in terms of localization in secondary structure elements, degree of burial, and sequence-based characteristics of the substituted AAs

Ideally, one would identify such *hot spots* based on structural or sequence characteristics of the protein **(see section 2.4 and 2.5)** prior to performing experiments. Suitable structure-based characteristics are localization in secondary structure elements **(Table S11)** [19, 96-98] and the degree of burial as measured by fSASAs **(Table S12, Eq. 6)** [19, 99, 100].

As to localization in secondary structure elements **(Table S11)**, *hot spots* are rarely found in 3_{10}-helices and β-strands. Exceptions are *hot spots* of class **XIV**, that are enriched in strand β7. With respect to α-helices, at least one and at most four *hot spot*(*s*) of each class is (are) found in that secondary structure class, mainly in helices αB and αE. However, without further information, one would not know which particular secondary structure element to choose for *hot spot* prediction. Hence, if all sites of a certain secondary structure class were chosen as *hot spots*, in the best case, a *gain in precision* (*gip*, **Eq. 11**) over random classification of 4.71 is found for β-strands, albeit at the expense of predicting 32 substitution sites (~18% of 181 AAs), far more than the six sought. As to bridges, turns, loops, and bends defined by DSSP [58], no *hot spot* is found in the first secondary structure type. At most three *hot spots* are found in any of the other three types, but only for *hot spots* of class **XI** and **VIII**. These cases are related to a maximal *gip* of 1.93, albeit at the expense of predicting 47 substitution sites (~26% of 181 AAs). Thus, in our study, identifying *hot spots* based on this secondary structure type results in a low precision.

As to the degree of burial **(Table S12)**, the least *hot spots* are associated with substitution sites that are mostly solvent-exposed ($0.8 < fSASA \leq 1.0$). By contrast, the most *hot spots* are associated with substitution sites that are partially solvent-exposed ($0.6 < fSASA \leq 0.8$), although this statement does not hold for *hot spots* of class **XIV**. This case is related to a maximal *gip* of 6.70, albeit at the expense of predicting 18 substitution sites (~10% of 181 AAs).

Suitable sequence-based characteristics are physicochemical properties of the substituted AAs **(Table S13)** [19, 101-103] and the degree of AA conservation **(Table S14)** [19, 104, 105]. As to the physicochemical properties of the substituted AAs **(Table S13)**, the distribution of *hot spots* over the classes is generally broad. Exceptions are *hot spots* of classes **XIII** and **XIV** (in both cases preferentially found at aliphatic and neutral AAs **(Table S15)**) and class **X** (preferentially found at aliphatic, aromatic, and neutral AAs **(Table S15)**). Therefore, *gip* values are generally low, with the largest one being 4.02 for the case of *hot spots* of class **X** at aromatic AAs, albeit at the expense of predicting 15 substitution sites (~8% of 181 AAs). As to the degree of AA conservation, *hot spots* are located at non-conserved and semi-conserved positions (conservation in the range of 0 - 6) **(Table S14)**. The highest conservations were found for I128 (conservation = 6) and V99, T126, and I128 (conservation = 5).

To conclude, while predicting *hot spots* based on structural characteristics can lead to marked *gip* values, usually many predicted *hot spots* result, which would require considerable experimental efforts. Still, if a higher number of predicted *hot spots* is acceptable, partially solvent-exposed residues are good *hot spot* candidates. Applying sequence-based characteristics, substituting aliphatic and neutral residues should more likely improve T_{50} or / and D. Additionally, non-conserved and semi-conserved regions preferentially contain *hot spots*.

3.8. Rigidity theory-based (CNA) and Markov chain Monte Carlo simulation-based (ProFASi) approaches predict similar thermal unfolding pathways of wt*Bs*LipA

We intend to test if *hot spots* can be predicted as structural *weak spots* by our rigidity theory-based approach CNA [66, 106, 107] **(see section 2.6)**. As a prerequisite, information on the hierarchy of rigid and flexible regions in a protein structure is required. Therefore, a thermal unfolding simulation of

wt*Bs*LipA was carried out with CNA as done previously [7, 8] to predict major phase transitions at which the network switches from overall rigid to flexible states **(see sections 2.7, 2.8, and 2.9)**.

From the thermal unfolding pathway of wt*Bs*LipA, five major phase transitions, T1 – T5, were predicted based on the *global* index H_{type2} **(Figure 3A)**. Depending on the energy cut-off E_{cut}, the phase transitions were characterized as either *early* (T1 – T2; with -0.8 kcal mol^{-1} ≥ E_{cut} ≥ -0.9 kcal mol^{-1}) or *late* (T3 – T5; with -1.7 kcal mol^{-1} ≥ E_{cut} ≥ -1.9 kcal mol^{-1}). E_{cut} can be converted to a temperature T using a linear equation **(Eq. 7)** [87], according to which the ranges of E_{cut} in this study are equivalent to 316 K ≤ T ≤ 318 K for T1 - T2, and 334 K ≤ T ≤ 338 K for T3 – T5. During the early phase transitions αA, 3_{10}-1, αF, and 3_{10}-5 segregate from the largest rigid cluster. αD, αE, αB, αC, and β-strands segregate from the largest rigid cluster during the late phase transitions. Afterwards, the β-sheet becomes sequentially flexible, beginning with β4 and β8, followed by β3, β7, β5, and β6. For the analysis, ~3 h of computational time on a single GPU is required to generate a 100 ns long MD trajectory as well as ~4 h of computational time on a single core for the thermal unfolding simulation.

Since the percolation behavior of a protein network is complex due to the protein's structural hierarchy and composition of different modules, it is often challenging to assign a phase transition with H_{type2} [86]. Thus, in addition to using H_{type2}, we also characterized the hierarchy of rigid and flexible regions of wt*Bs*LipA at a *local* level by computing $rc_{ij,neighbor}$ **(lower triangle in Figure 3B)** based on rc_{ij} **(upper triangle in Figure 3B)**. $rc_{ij,neighbor}$ showed that residue pairs at the *N*-terminus revealed higher E_{cut} values than residue pairs at the *C*-terminus. Thus, $rc_{ij,neighbor}$ demonstrates that the rigid contacts between neighboring residues are stronger at the *N*-terminus than at the *C*-terminus along the thermal unfolding simulation, i.e., the *C*-terminus of wt*Bs*LipA starts to unfold first.

As an independent approach to assess the order of unfolding of wt*Bs*LipA, we used the Markov Chain Monte Carlo (MCMC) simulation software ProFASi (Protein Folding and Aggregation Simulator) **(see section 2.11)** [90]. The results of the simulation were represented in a contact map **(upper triangle in Figure 3C)**. They reveal that the contacts between the residue pairs of the *N*-terminus have a longer lifetime (in terms of MC sweeps) than the contacts of the residue pairs of the *C*-terminus compared to the initial structure. Thus, although methodologically different, ProFASi predicts a very similar unfolding pathway of wt*Bs*LipA with respect to CNA.

To conclude, five major phase transitions, T1 – T5, were predicted by thermal unfolding simulations using CNA at which first the different helices and, finally, the β-strands segregate from the largest rigid cluster during thermal unfolding simulations of wt*Bs*LipA by CNA. Structural rigidity is initially lost at the *C*-terminus, which is uniformly revealed by the *global* index H_{type2} and the *local* index $rc_{ij,neighbor}$. Finally, the two independent approaches CNA and ProFASi predict very similar unfolding pathways of wt*Bs*LipA. The results suggest that the loss of rigidity predicted by CNA along the thermal unfolding simulation closely mimics the temperature-induced unfolding of wt*Bs*LipA.

3.9. *Unfolding nuclei* and major phase transitions are predictive markers of structural *weak spots*

We next probed to what extent structural *weak spots* predicted by CNA agree with the above defined *hot spots*. Following previous work [87], *weak spots* are fringe residues of unfolding nuclei that percolate from the largest rigid cluster during earlier steps of the thermal unfolding **(see section 2.9)**. In total, we predicted ten *weak spots* (~6% out of 181 substitution sites), i.e., I12, G13, G46, G52, P53, T66, M134, I135, V136, and H152 **(Figure 4A, Tables 1, 2, and S2)**. Three *weak spots* each segregate from the largest rigid cluster at T1 or T2, and four from the largest rigid cluster at T4 **(Table 2)**.

The performance of predicting *hot spots* as *weak spots* by CNA was evaluated in terms of a binary classification, considering predicted *weak spots* at *hot spots* true positives (TP) and predicted *weak spots* at not-*hot spots* false positives (FP) **(see section 2.10)**. In particular, the gain in precision over *random classification* (*gip*) **(Eq. 11)** and the *F1*-score (*F1*) **(Eq. 12)**, a measure of a classifier's accuracy, were used as performance measures. Over all seven classes of *hot spots*, between one and three of the predicted *weak spots* are *hot spots* (except for **XIV**, where no *weak spot* was met), resulting in *gip* values between 3.02 and 9.05 **(Tables 1 and S2)**. Note that these results are associated with only ten predicted *weak spots*, about half as many predictions than in the case of identifying *hot spots* as partially solvent-exposed residues **(Table S12)**. As the numbers of *hot spots* in **VIII** to **XIV** are of a very similar magnitude, the CNA predictions are also associated with similar recall (*r*) **(Eq. 8)** and precision (*p*) values **(Eq. 9)** in each case **(Table S2)**, indicating a well-balanced classifier. In the case of **XII**, the CNA predictions yield an *F1*-score of 0.38, higher than

any *F1*-score associated with *hot spot* predictions based on structure or sequence characteristics **(Tables S2, S11, S12, S13, and S14)**, and the *F1*-score for **IX** is 0.24, generally higher than *F1*-scores associated with structure- or sequence-based predictions for this class and on par with the result obtained for identifying these *hot spots* as partially solvent-exposed residues **(Tables S2, S11, S12, S13, and S14)**.

To conclude, predicting *hot spots* as *weak spots* by CNA results in several cases in very good to good *gip* values and good to fair accuracies, and is associated with a very low number of predicted *weak spots*, such that also only little experimental efforts are required later. Considering the low computing time required to perform a CNA analysis, these results indicate that applying CNA-based *weak spot* prediction before experimental engineering is beneficial, in particular if the number of substitution sites that can be dealt with in experiment is low.

4. Discussion

In this study, for the first time, we performed a systematic large-scale analysis of a complete experimental SSM library of a biotechnologically highly relevant protein, *Bs*LipA [52, 53], with respect to two types of protein stability. The library covers all 181 residues of *Bs*LipA and results in 3439 variants, each with a single AA substitution as confirmed by DNA sequencing. Considering the screening results of the library towards thermostability and detergent tolerance together is unique compared to related studies [2, 4-8, 17-19] and important in view of the challenges of multi-dimensional property optimization of modern biocatalysts [108-110]. The measured T_{50} and D values provide valuable reference data for future analyses because, in contrast to other data sources [34-37], the different protein stabilities were measured under respectively uniform conditions, and there is no bias towards any particular substitution type or site. Note, though, that other factors than protein stability may influence T_{50} or D values measured here [52], including that substitutions can directly impact *Bs*LipA function, e.g., when occurring in the vicinity of the active site [8]. Moreover, the measured T_{50} and D values may be influenced by thermodynamic or kinetic factors [7, 8]. Therefore, in our analysis, we focused on scrutinizing the impact of substitution sites on thermostability or / and detergent tolerance to gain generally applicable rules for data-driven protein engineering. The following results stand out:

First, across the library, the likelihoods to find variants with significantly increased T_{50} (~12%) or D towards one detergent (~14%) are almost identical and small. The finding that the overwhelming number of single AA substitutions introduced by *random mutagenesis* causes a destabilizing effect is in agreement with previous studies [33, 111-114]. This finding becomes even more statistically relevant if multiple mutations need to be accumulated over generations to reach a desired effect because frequently, a single, yet rather likely, destabilizing mutation is sufficient to annihilate the effect of several stabilizing ones [20]. The proportions of variants with increased T_{50} or D found here are in line with the composition of databases such as ProTherm [30] but markedly larger than the success rate of ~2% used as a reference to evaluate the performance of FoldX [115]. Hence, beyond the single T_{50} and D data, due to the completeness of our library and the model character of our protein, our results also constitute unbiased reference data as to what efficiency can be expected for a protein system when optimizing thermostability or detergent tolerance by *random mutagenesis*. In turn, largest increases in T_{50} of 7.7 K and D of 2.4 found demonstrate that considerable improvements of protein stability can already be achieved by single AA substitutions. In that

respect, previous studies on *Bs*LipA either applying directed evolution [116] or rational design [7, 8] already yielded close-to-optimal results in terms of increased thermostability.

Second, in the context of data-driven protein engineering, we identified substitution sites for which variants yield significantly increased T_{50} or / and D. Not considering the magnitude of the increase, only about one third or below of all *Bs*LipA residues constitute such favorable substitution sites if T_{50} and D are considered separately, demonstrating that the location of a residue within a protein structure matters with respect to a substitution effect. This result corroborates previous studies [5, 7, 8]. In addition, our complete SSM library allowed us to reveal for such substitution sites a significant and fair correlation between the frequency of T_{50} or / and D-increasing substitutions and the magnitude of the maximum effect. Together, these results show that addressing all substitution sites in an unbiased manner by *random mutagenesis* results in a considerable experimental effort coupled to low efficiency. In turn, approaches that can identify substitution sites with a high likelihood for significantly increased T_{50} or D prior to doing experiments will be of great value in protein engineering studies.

Third, notably, the conclusions from the last paragraph also hold if more than one protein property is considered at a time. As such, we showed that at eleven substitution sites a ~4.6-fold higher likelihood to find for each detergent variants with significantly increased D compared to *random mutagenesis* is found. Additionally, seven substitution sites yield a ~3.4-fold higher likelihood to find significantly increased T_{50} and a ~4.7-fold higher likelihood to find for each detergent variants with significantly increased D compared to *random mutagenesis*. The latter finding suggests that approaches that can identify substitution sites with a high likelihood for significantly increased T_{50} should also be beneficial for identifying substitution sites with a high likelihood for significantly increased D, or *vice versa*. This is an important finding for practical applications as many more algorithms have been devised that address thermostability than detergent tolerance.

Fourth, as another set of reference data, we defined *hot spot* types together with the associated substitution sites to provide benchmark data for evaluating the performance of data-driven approaches. The first five classes follow the strict criterion that only the six substitution sites with the respective highest $\Delta T_{50;\,max}$ or ΔD_{max} are considered, according to that all combinations of the 20 proteinogenic AAs at such sites could still be experimentally investigated [20, 39, 49, 50]. The intersections comprising the substitution sites with the broadest impact on ΔD_{max}, or $\Delta T_{50;\,max}$ and

ΔD_{max}, are empty. Thus, we resorted to defining two further classes with the somewhat relaxed criterion that the comprised substitution sites show significantly increased D towards each detergent, or significantly increased T_{50} and D towards each detergent, regardless of the magnitude of the single effect.

Fifth, we used the complete, unbiased, and uniformly generated T_{50} and D data to probe if universal rules for modulating thermostability or detergent tolerance can be identified. We thereby focused on "one-dimensional" descriptors in terms of location in secondary structure elements, degree of burial, and physicochemical properties and conservation degree of substituted AA. Such descriptors have been widely analyzed before [117, 118] and play a role in data-driven consensus approaches [119, 120]. Analysing "two- or higher dimensional" descriptors in terms of residue-residue interactions, entropic contributions or other collective phenomena, or cross-correlations of "one-dimensional" descriptors [33] remains for future work. Notably, considering our descriptors, many (up to 98 substitution sites) predicted *hot spots* result, which would require considerable experimental efforts particularly if beneficial substitutions need to be accumulated to reach a desired effect. This finding demonstrates on a single protein level that, with these descriptors, no universal and sufficiently discriminating rule(s) can be identified, a finding that is mirrored in studies across protein families [121, 122] and with respect to low successes in assessing thermostabilities [117]. Still, if a higher number of predicted *hot spots* is acceptable, partially solvent-exposed residues are good *hot spot* candidates. This result differs from previous experimental studies showing that especially surface remodeling emerged as an effective strategy to improve protein stability [123, 124]. Furthermore, loop positions, which have elsewhere been identified to preferentially carry favorable substitution sites [125, 126], show mostly destabilizing effects. Finally, and likely surprisingly, *hot spots* were preferentially found at non-conserved and semi-conserved position, a finding that may help refine future consensus concepts where multiple sequence alignments are used to substitute non-consensus residues by consensus ones [42, 127].

Sixth, we made use of the reference data to unequivocally benchmark our ensemble- and rigidity theory-based CNA approach with respect to predicting *hot spots* as structural *weak spots* of the protein. In contrast to previous studies on much smaller data sets [2, 4, 5, 8] the present work allows to systematically assess the quality of our predictions. To do so, and in contrast to other assessments of protein stability predictors [29, 30], we apply recall and precision as basic statistical measures, rather than sensitivity and specificity, because we are interested in the accuracy of predicting *hot spots*,

and not not-*hot spots*, the latter of which furthermore clearly dominate the dataset in terms of occurrence frequency. Methodologically, CNA differs from other state-of-the art methods that do not consider ensemble representations of the protein [128-133]. Furthermore, CNA does not require system-specific weighting or fitting parameters [128, 131, 134, 135]. This should make the results obtained here with CNA transferable to other protein systems. *Weak spot* prediction by CNA relies on a realistic modeling of the thermal unfolding of a protein [66, 106, 107]. The predicted major phase transitions and the order of the segregating secondary structure elements are in agreement with previous computational studies and experimental observations on other proteins with an α/β hydrolase fold [136, 137]. Furthermore, we confirmed the unfolding pathway of wt*Bs*LipA predicted by CNA with the independent MCMC-based ProFASi approach. From a practical point of view, it is relevant that CNA predicted only ten *weak spots*, allowing to focus subsequent substitution efforts on only ~6% of the protein residues. Furthermore, the gain in precision over random classification is between ~3 and ~9, depending on the *hot spot* class. Considering the properties of the majority of predicted *weak spots*, i.e., a location in a loop, turn, or bend and a neutral or aliphatic amino acid type (**Table 2**), the notion may arise that these two properties, when correlated, characterize *hotspots*. The gain in precision over random classification is only between ~0.7 and ~2.1, however, depending on the *hot spot* class (**Table S16**), and, hence, more than 4-fold lower than when *hot spots* are predicted as *weak spots* by CNA (**Table 1**). Together with the low computational demand on the order of hours only, these results lead to the strong recommendation to apply CNA-based *weak spot* prediction for data-driven protein engineering towards increased T_{50} or / and D.

In summary, we provide systematic and unbiased reference data at large scale for thermostability measured as T_{50} values and detergent tolerance measured as D for a biotechnologically important protein, identified consistently defined *hot spot* types for evaluating the performance of data-driven protein engineering approaches, and showed that CNA-based *hot spot* prediction can yield a gain in precision over random classification up to 9-fold.

5. Acknowledgements

CN is funded through a grant ("Vernetzungsdoktorand") provided by the Forschungszentrum Jülich. Parts of the study were supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through funding no. INST 208/704-1 FUGG to HG, INST 208/654-1 FUGG to KEJ. US and KEJ additionally received funding within the DFG research training group 1166 "Biocatalysis using Non-Conventional Media – BioNoCo". HG is grateful for computational support and infrastructure provided by the "Zentrum für Informations- und Medientechnologie" (ZIM) at the Heinrich Heine University Düsseldorf. HG gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JUWELS at Jülich Supercomputing Centre (JSC) (user IDs: HKF7; protil (project ID: 15956)) [138, 139].

6. Authors contributions

HG, KEJ, and US conceived the study. CN analyzed the data, performed MD simulations and CNA computations, analyzed the results, and wrote the manuscript together with HG. AF performed experimental work. OZ performed ProFASi simulations and analyzed the results. HG supervised and managed the project. All authors reviewed and approved the manuscript.

7. Conflict of interest

The authors declare no conflict of interest.

8. References

1. Salazar, O.; Cirino, P. C.; Arnold, F. H., Thermostabilization of a cytochrome P450 peroxygenase. *Chembiochem* **2003**, 4, 891-893.
2. Radestock, S.; Gohlke, H., Exploiting the link between protein rigidity and thermostability for data-driven protein engineering. *Eng Life Sci* **2008**, 8, 507-522.
3. Rader, A., Thermostability in rubredoxin and its relationship to mechanical rigidity. *Phys Biol* **2009**, 7, 016002.
4. Radestock, S.; Gohlke, H., Protein rigidity and thermophilic adaptation. *Proteins: Structure, Function, and Bioinformatics* **2011**, 79, 1089-1108.
5. Rathi, P. C.; Radestock, S.; Gohlke, H., Thermostabilizing mutations preferentially occur at structural weak spots with a high mutation ratio. *J Biotechnol* **2012**, 159, 135-144.
6. Dick, M.; Weiergräber, O. H.; Classen, T.; Bisterfeld, C.; Bramski, J.; Gohlke, H.; Pietruszka, J., Trading off stability against activity in extremophilic aldolases. *Sci Rep* **2016**, 6, 17908.
7. Rathi, P. C.; Jaeger, K.-E.; Gohlke, H., Structural rigidity and protein thermostability in variants of lipase A from Bacillus subtilis. *PLoS One* **2015**, 10, e0130289.
8. Rathi, P. C.; Fulton, A.; Jaeger, K.-E.; Gohlke, H., Application of rigidity theory to the thermostabilization of lipase A from Bacillus subtilis. *PLoS Comput Biol* **2016**, 12, e1004754.
9. Pottkämper, J.; Barthen, P.; Ilmberger, N.; Schwaneberg, U.; Schenk, A.; Schulte, M.; Ignatiev, N.; Streit, W. R., Applying metagenomics for the identification of bacterial cellulases that are stable in ionic liquids. *Green Chem* **2009**, 11, 957-965.
10. Liu, H.; Zhu, L.; Bocola, M.; Chen, N.; Spiess, A. C.; Schwaneberg, U., Directed laccase evolution for improved ionic liquid resistance. *Green Chem* **2013**, 15, 1348-1355.
11. Carter, J. L.; Bekhouche, M.; Noiriel, A.; Blum, L. J.; Doumèche, B., Directed evolution of a formate dehydrogenase for increased tolerance to ionic liquids reveals a new site for increasing the stability. *Chembiochem* **2014**, 15, 2710-2718.
12. Chen, Z.; Pereira, J. H.; Liu, H.; Tran, H. M.; Hsu, N. S.; Dibble, D.; Singh, S.; Adams, P. D.; Sapra, R.; Hadi, M. Z., Improved activity of a thermophilic cellulase, Cel5A, from Thermotoga maritima on ionic liquid pretreated switchgrass. *PloS One* **2013**, 8, e79725.
13. Lehmann, C.; Bocola, M.; Streit, W. R.; Martinez, R.; Schwaneberg, U., Ionic liquid and deep eutectic solvent-activated CelA2 variants generated by directed evolution. *Appl Microbiol and Biotechnol* **2014**, 98, 5775-5785.
14. Nordwald, E. M.; Armstrong, G. S.; Kaar, J. L., NMR-guided rational engineering of an ionic-liquid-tolerant lipase. *ACS Catal* **2014**, 4, 4057-4064.
15. Frauenkron-Machedjou, V. J.; Fulton, A.; Zhu, L.; Anker, C.; Bocola, M.; Jaeger, K. E.; Schwaneberg, U., Towards understanding directed evolution: more than half of all amino acid positions contribute to ionic liquid resistance of Bacillus subtilis lipase A. *Chembiochem* **2015**, 16, 937-945.
16. Zhao, J.; Frauenkron-Machedjou, V. J.; Fulton, A.; Zhu, L.; Davari, M. D.; Jaeger, K.-E.; Schwaneberg, U.; Bocola, M., Unraveling the effects of amino acid substitutions

enhancing lipase resistance to an ionic liquid: a molecular dynamics study. *Phys Chem Chem Phys* **2018**, 20, 9600-9609.
17. Brissos, V.; Eggert, T.; Cabral, J.; Jaeger, K.-E., Improving activity and stability of cutinase towards the anionic detergent AOT by complete saturation mutagenesis. *Protein Eng Des Sel* **2008**, 21, 387-393.
18. Akbulut, N.; Öztürk, M. T.; Pijning, T.; Öztürk, S. İ.; Gümüşel, F., Improved activity and thermostability of Bacillus pumilus lipase by directed evolution. *J Biotechnol* **2013**, 164, 123-129.
19. Fulton, A.; Frauenkron-Machedjou, V. J.; Skoczinski, P.; Wilhelm, S.; Zhu, L.; Schwaneberg, U.; Jaeger, K. E., Exploring the protein stability landscape: Bacillus subtilis lipase A as a model for detergent tolerance. *Chembiochem* **2015**, 16, 930-936.
20. Rigoldi, F.; Donini, S.; Redaelli, A.; Parisini, E.; Gautieri, A., Engineering of thermostable enzymes for industrial applications. *APL Bioeng* **2018**, 2, 011501.
21. Littlechild, J. A., Enzymes from extreme environments and their industrial applications. *Front Bioeng Biotechnol* **2015**, 3, 161.
22. https://www.bccresearch.com/market-research/biotechnology/global-markets-for-enzymes-in-industrial-applications-bio030k.html.
23. Lehmann, M.; Wyss, M., Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution. *Curr Opin Biotechnol* **2001**, 12, 371-375.
24. Eijsink, V. G.; Bjørk, A.; Gåseidnes, S.; Sirevåg, R.; Synstad, B.; van den Burg, B.; Vriend, G., Rational engineering of enzyme stability. *J Biotechnol* **2004**, 113, 105-120.
25. Thiltgen, G.; Goldstein, R. A., Assessing predictors of changes in protein stability upon mutation using self-consistency. *PloS One* **2012**, 7, e46084.
26. Johnston, M. A.; Søndergaard, C. R.; Nielsen, J. E., Integrated prediction of the effect of mutations on multiple protein characteristics. *Proteins: Structure, Function, and Bioinformatics* **2011**, 79, 165-178.
27. Li, Y.; Zhang, J.; Tai, D.; Russell Middaugh, C.; Zhang, Y.; Fang, J., Prots: A fragment based protein thermo-stability potential. *Proteins: Structure, Function, and Bioinformatics* **2012**, 80, 81-92.
28. Capriotti, E.; Fariselli, P.; Rossi, I.; Casadio, R., A three-state prediction of single point mutations on protein stability changes. *BMC Bioinf* **2008**, 9, S6.
29. Potapov, V.; Cohen, M.; Schreiber, G., Assessing computational methods for predicting protein stability upon mutation: good on average but not in the details. *Protein Eng Des Sel* **2009**, 22, 553-560.
30. Khan, S.; Vihinen, M., Performance of protein stability predictors. *Hum Mutat* **2010**, 31, 675-684.
31. Usmanova, D. R.; Bogatyreva, N. S.; Ariño Bernad, J.; Eremina, A. A.; Gorshkova, A. A.; Kanevskiy, G. M.; Lonishin, L. R.; Meister, A. V.; Yakupova, A. G.; Kondrashov, F. A., Self-consistency test reveals systematic bias in programs for prediction change of stability upon mutation. *Bioinformatics* **2018**, 34, 3653-3658.
32. Pucci, F.; Bernaerts, K. V.; Kwasigroch, J. M.; Rooman, M., Quantification of biases in predictions of protein stability changes upon mutations. *Bioinformatics* **2018**, 34, 3659-3665.

33. Modarres, H. P.; Mofrad, M.; Sanati-Nezhad, A., Protein thermostability engineering. *RSC Advances* **2016**, 6, 115252-115270.
34. Bava, K. A.; Gromiha, M. M.; Uedaira, H.; Kitajima, K.; Sarai, A., ProTherm, version 4.0: thermodynamic database for proteins and mutants. *Nucleic Acids Res* **2004**, 32, D120-D121.
35. Kumar, M. S.; Bava, K. A.; Gromiha, M. M.; Prabakaran, P.; Kitajima, K.; Uedaira, H.; Sarai, A., ProTherm and ProNIT: thermodynamic databases for proteins and protein–nucleic acid interactions. *Nucleic Acids Res* **2006**, 34, D204-D206.
36. Zhang, Z.; Wang, L.; Gao, Y.; Zhang, J.; Zhenirovskyy, M.; Alexov, E., Predicting folding free energy changes upon single point mutations. *Bioinformatics* **2012**, 28, 664-671.
37. Kang, S.; Chen, G.; Xiao, G., Robust prediction of mutation-induced protein stability change by property encoding of amino acids. *Protein Eng Des Sel* **2008**, 22, 75-83.
38. Nisthal, A.; Wang, C. Y.; Ary, M. L.; Mayo, S. L., Protein stability engineering insights revealed by domain-wide comprehensive mutagenesis. *PNAS* **2019**, 116, 16367-16377.
39. Zeymer, C.; Hilvert, D., Directed evolution of protein catalysts. *Annual review of biochemistry* **2018**, 87, 131-157.
40. Arnold, F. H., Directed evolution: bringing new chemistry to life. *Angew Chem Int Ed Engl* **2018**, 57, 4143-4148.
41. Romero, P. A.; Arnold, F. H., Exploring protein fitness landscapes by directed evolution. *Nat Rev Mol Cell Biol* **2009**, 10, 866.
42. Chaparro-Riggers, J. F.; Polizzi, K. M.; Bommarius, A. S., Better library design: data-driven protein engineering. *Biotechnol J: Healthcare Nutrition Technology* **2007**, 2, 180-191.
43. Wijma, H. J.; Floor, R. J.; Janssen, D. B., Structure-and sequence-analysis inspired engineering of proteins for enhanced thermostability. *Curr Opin Struct Biol* **2013**, 23, 588-594.
44. Reetz, M. T.; Carballeira, J. D.; Vogel, A., Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. *Angew Chem Int Ed* **2006**, 45, 7745-7751.
45. Huang, X.; Gao, D.; Zhan, C.-G., Computational design of a thermostable mutant of cocaine esterase via molecular dynamics simulations. *Org Biomol Chem* **2011**, 9, 4138-4143.
46. Badieyan, S.; Bevan, D. R.; Zhang, C., Study and design of stability in GH5 cellulases. *Biotechnol Bioeng* **2012**, 109, 31-44.
47. Brooks, B. R.; Brooks III, C. L.; Mackerell Jr, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S., CHARMM: the biomolecular simulation program. *J Comput Chem* **2009**, 30, 1545-1614.
48. Craig, D. B.; Dombkowski, A. A., Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC bioinformatics* **2013**, 14, 346.
49. Singh, J.; Ator, M. A.; Jaeger, E. P.; Allen, M. P.; Whipple, D. A.; Soloweij, J. E.; Chowdhary, S.; Treasurywala, A. M., Application of genetic algorithms to combinatorial synthesis: A computational approach to lead identification and lead optimization. *J Am Chem Soc* **1996**, 118, 1669-1676.

50. Moore, K. W.; Pechen, A.; Feng, X.-J.; Dominy, J.; Beltrani, V. J.; Rabitz, H., Why is chemical synthesis and property optimization easier than expected? *Phys Chem Chem Phys* **2011**, 13, 10048-10070.
51. Kawasaki, K.; Kondo, H.; Suzuki, M.; Ohgiya, S.; Tsuda, S., Alternate conformations observed in catalytic serine of Bacillus subtilis lipase determined at 1.3 Å resolution. *Acta Crystallogr D* **2002**, 58, 1168-1174.
52. Skoczinski, P.; Volkenborn, K.; Fulton, A.; Bhadauriya, A.; Nutschel, C.; Gohlke, H.; Knapp, A.; Jaeger, K.-E., Contribution of single amino acid and codon substitutions to the production and secretion of a lipase by Bacillus subtilis. *Microb Cell Fact* **2017**, 16, 160.
53. Schallmey, M.; Singh, A.; Ward, O. P., Developments in the use of Bacillus species for industrial production. *Can J Microbiol* **2004**, 50, 1-17.
54. Van Pouderoyen, G.; Eggert, T.; Jaeger, K.-E.; Dijkstra, B. W., The crystal structure of Bacillus subtili lipase: a minimal α/β hydrolase fold enzyme1. *J Mol Biol* **2001**, 309, 215-226.
55. Jech, T., *Set theory*. Springer Science & Business Media, Berlin/Heidelberg, **2013**.
56. Hamers, L., Similarity measures in scientometric research: The Jaccard index versus Salton's cosine formula. *Information Processing and Management* **1989**, 25, 315-18.
57. Leydesdorff, L., On the normalization and visualization of author co-citation data: Salton's Cosine versus the Jaccard index. *Journal of the American Society for Information Science and Technology* **2008**, 59, 77-85.
58. Kabsch, W.; Sander, C., Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers: Original Research on Biomolecules* **1983**, 22, 2577-2637.
59. Creighton, T. E., *Proteins: Structures and Molecular Properties*. Freeman: 1992.
60. Punta, M.; Coggill, P. C.; Eberhardt, R. Y.; Mistry, J.; Tate, J.; Boursnell, C.; Pang, N.; Forslund, K.; Ceric, G.; Clements, J., The Pfam protein families database. *Nucleic Acids Res* **2011**, 40, D290-D301.
61. Sievers, F.; Higgins, D. G. Clustal Omega, accurate alignment of very large numbers of sequences. In *Multiple sequence alignment methods*; Springer: **2014**, pp 105-116.
62. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **2011**, 7, 539.
63. Dartois, V.; Baulard, A.; Schanck, K.; Colson, C., Cloning, nucleotide sequence and expression in Escherichia coli of a lipase gene from Bacillus subtilis 168. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* **1992**, 1131, 253-260.
64. Golicz, A.; Troshin, P. V.; Madeira, F.; Martin, D. M. A.; Procter, J. B.; Barton, G. J., AACon: A Fast Amino Acid Conservation Calculation Service. *Submitted.* **2018**.
65. Waterhouse, A. M.; Procter, J. B.; Martin, D. M.; Clamp, M.; Barton, G. J., Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, 25, 1189-1191.
66. Krüger, D. M.; Rathi, P. C.; Pfleger, C.; Gohlke, H., CNA web server: rigidity theory-based thermal unfolding simulations of proteins for linking structure,(thermo-) stability, and function. *Nucleic Acids Res* **2013**, 41, W340-W348.

67. Pfleger, C.; Rathi, P. C.; Klein, D. L.; 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, 1007-15.
68. Hermans, S. M.; Pfleger, C.; Nutschel, C.; Hanke, C. A.; Gohlke, H., Rigidity theory for biomolecules: concepts, software, and applications. *Wiley Interdiscip Rev Comput Mol Sci* **2017**, 7, e1311.
69. Jacobs, D. J.; Rader, A. J.; Kuhn, L. A.; Thorpe, M. F., Protein flexibility predictions using graph theory. *Proteins: Structure, Function, and Bioinformatics* **2001**, 44, 150-165.
70. Hespenheide, B.; Jacobs, D.; Thorpe, M., Structural rigidity in the capsid assembly of cowpea chlorotic mottle virus. *J Phys: Condens Matter* **2004**, 16, S5055.
71. Jacobs, D. J.; Thorpe, M. F., Generic rigidity percolation: the pebble game. *Phys Rev Lett* **1995**, 75, 4051.
72. Jacobs, D. J., Generic rigidity in three-dimensional bond-bending networks. *J Phys A: Math Gen* **1998**, 31, 6653.
73. Dahiyat, B. I.; Benjamin Gordon, D.; Mayo, S. L., Automated design of the surface positions of protein helices. *Protein Sci* **1997**, 6, 1333-1337.
74. 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, 119-127.
75. Privalov, P. L.; Gill, S. J. Stability of protein structure and hydrophobic interaction. In *Adv Protein Chem*; Elsevier: 1988; Vol. 39, pp 191-234.
76. Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C., Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *J Chem Theory Comput* **2013**, 9, 3878-3888.
77. Case, D. A.; Babin, V.; Berryman, J.; Betz, R.; Cai, Q.; Cerutti, D.; Cheatham Iii, T.; Darden, T.; Duke, R.; Gohlke, H., Amber 14. **2014**.
78. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput* **2015**, 11, 3696-3713.
79. Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C., Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation1. *J Mol Biol* **1999**, 285, 1735-1747.
80. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **1983**, 79, 926-935.
81. Darden, T.; York, D.; Pedersen, L., Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. *J Chem Phys* **1993**, 98, 10089-10092.
82. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* **1977**, 23, 327-341.
83. Rathi, P. C.; Radestock, S.; Gohlke, H., Thermostabilizing mutations preferentially occur at structural weak spots with a high mutation ratio. *Journal of biotechnology* **2012**, 159, 135-144.

84. Pfleger, C.; Gohlke, H., Efficient and robust analysis of biomacromolecular flexibility using ensembles of network topologies based on fuzzy noncovalent constraints. *Structure* **2013**, 21, 1725-1734.
85. Rathi, P. C.; Mulnaes, D.; Gohlke, H., VisualCNA: a GUI for interactive constraint network analysis and protein engineering for improving thermostability. *Bioinformatics* **2015**, 31, 2394-2396.
86. Pfleger, C.; Radestock, S.; Schmidt, E.; Gohlke, H., Global and local indices for characterizing biomolecular flexibility and rigidity. *J. Comput. Chem.* **2013**, 34, 220-33.
87. Radestock, S.; Gohlke, H., Exploiting the link between protein rigidity and thermostability for data-driven protein engineering. *Engineering in Life Sciences* **2008**, 8, 507-522.
88. Buckland, M.; Gey, F., The relationship between recall and precision. *Journal of the American society for information science* **1994**, 45, 12-19.
89. Goutte, C.; Gaussier, E. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In European Conference on Information Retrieval, 2005; Springer: 2005; pp 345-359.
90. Irbäck, A.; Mohanty, S., PROFASI: a Monte Carlo simulation package for protein folding and aggregation. *J Comput Chem* **2006**, 27, 1548-1555.
91. Mohanty, S.; Meinke, J. H.; Zimmermann, O., Folding of Top7 in unbiased all-atom Monte Carlo simulations. *Proteins: Structure, Function, and Bioinformatics* **2013**, 81, 1446-1456.
92. Irbäck, A.; Mitternacht, S.; Mohanty, S., An effective all-atom potential for proteins. *PMC Biophys* **2009**, 2, 2.
93. Irbäck, A.; Samuelsson, B.; Sjunnesson, F.; Wallin, S., Thermodynamics of α-and β-structure formation in proteins. *Biophys J* **2003**, 85, 1466-1473.
94. Favrin, G.; Irbäck, A.; Sjunnesson, F., Monte Carlo update for chain molecules: biased Gaussian steps in torsional space. *J Chem Phys* **2001**, 114, 8154-8158.
95. Neugebauer, J., Guide to the Properties and Uses of Detergents in Biology and Biochemistry. **1987**.
96. Villegas, V.; Viguera, A. R.; Avilés, F. X.; Serrano, L., Stabilization of proteins by rational design of α-helix stability using helix/coil transition theory. *Fold Des* **1996**, 1, 29-34.
97. Taddei, N.; Chiti, F.; Fiaschi, T.; Bucciantini, M.; Capanni, C.; Stefani, M.; Serrano, L.; Dobson, C. M.; Ramponi, G., Stabilisation of α-helices by site-directed mutagenesis reveals the importance of secondary structure in the transition state for acylphosphatase folding. *J Mol Biol* **2000**, 300, 633-647.
98. Yu, H.; Yan, Y.; Zhang, C.; Dalby, P. A., Two strategies to engineer flexible loops for improved enzyme thermostability. *Sci Rep* **2017**, 7, 41212.
99. Chen, H.; Zhou, H.-X., Prediction of solvent accessibility and sites of deleterious mutations from protein sequence. *Nucleic Acids Res* **2005**, 33, 3193-3199.
100. Yokota, K.; Satou, K.; Ohki, S.-y., Comparative analysis of protein thermostability: Differences in amino acid content and substitution at the surfaces and in the core regions of thermophilic and mesophilic proteins. *Sci Technol Adv Mater* **2006**, 7, 255.

101. Watanabe, K.; Masuda, T.; Ohashi, H.; Mihara, H.; Suzuki, Y., Multiple Proline Substitutions Cumulatively Thermostabilize Bacillus Cereus ATCC7064 Oligo-1, 6-Glucosidase: Irrefragable Proof Supporting the Proline Rule. *Eur J Biochem* **1994**, 226, 277-283.
102. Kumar, S.; Tsai, C.-J.; Nussinov, R., Factors enhancing protein thermostability. *Nucleic Acids Res* **2000**, 13, 179-191.
103. Ikai, A., Thermostability and aliphatic index of globular proteins. *J Biochem* **1980**, 88, 1895-1898.
104. Ashenberg, O.; Gong, L. I.; Bloom, J. D., Mutational effects on stability are largely conserved during protein evolution. *PNAS* **2013**, 110, 21071-21076.
105. Greene, L. H.; Chrysina, E. D.; Irons, L. I.; Papageorgiou, A. C.; Acharya, K. R.; Brew, K., Role of conserved residues in structure and stability: Tryptophans of human serum retinol-binding protein, a model for the lipocalin superfamily. *Protein Sci* **2001**, 10, 2301-2316.
106. Hermans, S. M.; Pfleger, C.; Nutschel, C.; Hanke, C. A.; Gohlke, H., Rigidity theory for biomolecules: concepts, software, and applications. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2017**, 7, e1311.
107. Pfleger, C.; Rathi, P. C.; Klein, D. L.; Radestock, S.; Gohlke, H., In; ACS Publications: 2013.
108. Hafizah, N. F.; Teh, A.-H.; Furusawa, G., Biochemical Characterization of Thermostable and Detergent-Tolerant β-Agarase, PdAgaC, from Persicobacter sp. CCB-QB2. *Appl Biochem Biotechnol* **2018**, 1-12.
109. Lu, M.; Dukunde, A.; Daniel, R., Biochemical profiles of two thermostable and organic solvent–tolerant esterases derived from a compost metagenome. *Appl Microbiol Biotechnol* **2019**, 1-17.
110. Annamalai, N.; Rajeswari, M. V.; Balasubramanian, T., Extraction, purification and application of thermostable and halostable alkaline protease from Bacillus alveayuensis CAS 5 using marine wastes. *Food Bioprod Process* **2014**, 92, 335-342.
111. Araya, C. L.; Fowler, D. M.; Chen, W.; Muniez, I.; Kelly, J. W.; Fields, S., A fundamental protein property, thermodynamic stability, revealed solely from large-scale measurements of protein function. *Proceedings of the National Academy of Sciences* **2012**, 109, 16858-16863.
112. Foit, L.; Morgan, G. J.; Kern, M. J.; Steimer, L. R.; von Hacht, A. A.; Titchmarsh, J.; Warriner, S. L.; Radford, S. E.; Bardwell, J. C., Optimizing protein stability in vivo. *Molecular cell* **2009**, 36, 861-871.
113. Deng, Z.; Huang, W.; Bakkalbasi, E.; Brown, N. G.; Adamski, C. J.; Rice, K.; Muzny, D.; Gibbs, R. A.; Palzkill, T., Deep sequencing of systematic combinatorial libraries reveals β-lactamase sequence constraints at high resolution. *Journal of molecular biology* **2012**, 424, 150-167.
114. Klesmith, J. R.; Bacik, J.-P.; Wrenbeck, E. E.; Michalczyk, R.; Whitehead, T. A., Trade-offs between enzyme fitness and solubility illuminated by deep mutational scanning. *Proceedings of the National Academy of Sciences* **2017**, 114, 2265-2270.
115. Buß, O.; Rudat, J.; Ochsenreither, K., FoldX as protein engineering tool: better than random based approaches? *Comput Struct Biotechnol J* **2018**, 16, 25-33.

116. Reetz, M. T.; Carballeira, J. D.; Vogel, A., Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. *Angewandte Chemie International Edition* **2006**, 45, 7745-7751.
117. Pack, S. P.; Kang, T. J.; Yoo, Y. J., Protein thermostabilizing factors: high relative occurrence of amino acids, residual properties, and secondary structure type in different residual state. *Appl Biochem Biotechnol* **2013**, 171, 1212-1226.
118. Guerois, R.; Nielsen, J. E.; Serrano, L., Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J Mol Biol* **2002**, 320, 369-387.
119. Polizzi, K. M.; Chaparro-Riggers, J. F.; Vazquez-Figueroa, E.; Bommarius, A. S., Structure-guided consensus approach to create a more thermostable penicillin G acylase. *Biotechnology Journal: Healthcare Nutrition Technology* **2006**, 1, 531-536.
120. Porebski, B. T.; Buckle, A. M., Consensus protein design. *Protein Eng Des Sel* **2016**, 29, 245-251.
121. Vogt, G.; Woell, S.; Argos, P., Protein thermal stability, hydrogen bonds, and ion pairs. *J Mol Biol* **1997**, 269, 631-643.
122. Chakravarty, S.; Varadarajan, R., Elucidation of determinants of protein stability through genome sequence analysis. *FEBS Lett* **2000**, 470, 65-69.
123. Perl, D.; Mueller, U.; Heinemann, U.; Schmid, F. X., Two exposed amino acid residues confer thermostability on a cold shock protein. *Nat Struct Mol Biol* **2000**, 7, 380.
124. Pace, C. N., Single surface stabilizer. *Nat Struct Mol Biol* **2000**, 7, 345.
125. Pokkuluri, P.; Raffen, R.; Dieckman, L.; Boogaard, C.; Stevens, F.; Schiffer, M., Increasing protein stability by polar surface residues: domain-wide consequences of interactions within a loop. *Biophys J* **2002**, 82, 391-398.
126. Jung, S.; Plückthun, A., Improving in vivo folding and stability of a single-chain Fv antibody fragment by loop grafting. *Protein Eng* **1997**, 10, 959-966.
127. Pantoliano, M. W.; Whitlow, M.; Wood, J. F.; Dodd, S. W.; Hardman, K. D.; Rollence, M. L.; Bryan, P. N., Large increases in general stability for subtilisin BPN'through incremental changes in the free energy of unfolding. *Biochemistry* **1989**, 28, 7205-7213.
128. Guerois, R.; Nielsen, J. E.; Serrano, L., Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *Journal of molecular biology* **2002**, 320, 369-387.
129. Worth, C. L.; Preissner, R.; Blundell, T. L., SDM--a server for predicting effects of mutations on protein stability and malfunction. *Nucleic Acids Res* **2011**, 39, W215-22.
130. Pokala, N.; Handel, T. M., Energy functions for protein design: adjustment with protein-protein complex affinities, models for the unfolded state, and negative design of solubility and specificity. *J Mol Biol* **2005**, 347, 203-27.
131. Capriotti, E.; Fariselli, P.; Casadio, R., I-Mutant2. 0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic acids research* **2005**, 33, W306-W310.
132. Rohl, C. A.; Strauss, C. E.; Misura, K. M.; Baker, D. Protein structure prediction using Rosetta. In *Methods in enzymology*; Elsevier: 2004; Vol. 383, pp 66-93.

133. Potapov, V.; Cohen, M.; Inbar, Y.; Schreiber, G., Protein structure modelling and evaluation based on a 4-distance description of side-chain interactions. *Bmc Bioinformatics* **2010**, 11, 374.
134. Benedix, A.; Becker, C. M.; de Groot, B. L.; Caflisch, A.; Böckmann, R. A., Predicting free energy changes using structural ensembles. *Nature methods* **2009**, 6, 3.
135. Bordner, A.; Abagyan, R., Large-scale prediction of protein geometry and stability changes for arbitrary single point mutations. *Proteins: Structure, Function, and Bioinformatics* **2004**, 57, 400-413.
136. Beermann, B.; Guddorf, J.; Boehm, K.; Albers, A.; Kolkenbrock, S.; Fetzner, S.; Hinz, H.-J., Stability, Unfolding, and Structural Changes of Cofactor-Free 1 H-3-Hydroxy-4-oxoquinaldine 2, 4-Dioxygenase. *Biochemistry* **2007**, 46, 4241-4249.
137. Hung, H.-C.; Chang, G.-G., Multiple unfolding intermediates of human placental alkaline phosphatase in equilibrium urea denaturation. *Biophysical journal* **2001**, 81, 3456-3471.
138. Krause, D., JUWELS: Modular Tier-0/1 Supercomputer at the Jülich Supercomputing Centre. *JLSRF* **2019**, 5, A135.
139. Krause, D.; Thörnig, P., JURECA: modular supercomputer at jülich supercomputing centre. *Journal of large-scale research facilities JLSRF* **2018**, 4, 132.

9. Figures

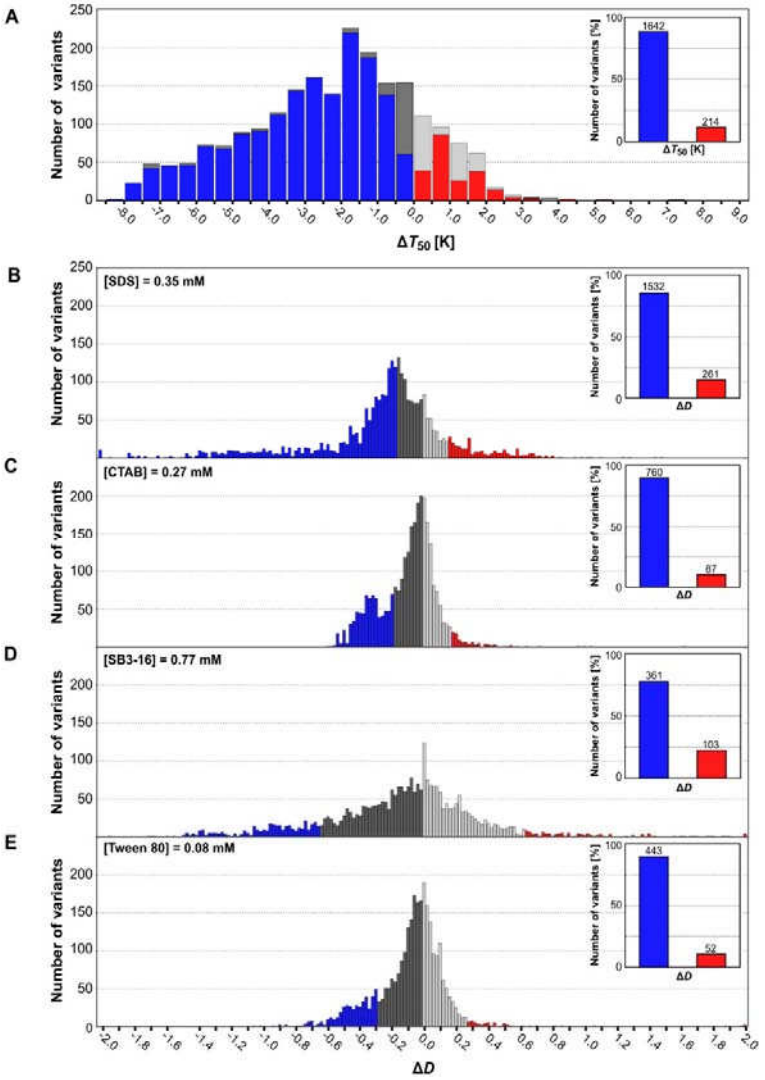


Figure 1: Distribution of *Bs*LipA variants' changes in T_{50} or D towards one detergent. Distribution of *Bs*LipA variants' changes in **(A)** T_{50} (ΔT_{50}) or D (ΔD) with respect to **(B)** SDS, **(C)** CTAB, **(D)** SB3-16, and **(E)** Tween 80 at the indicated concentrations compared to wt*Bs*LipA (ΔT_{50} / $\Delta D = 0$). **(A)** Variants with ΔT_{50} lower than the experimental uncertainty (standard deviation σ_T for the respective variant) were excluded from further analyses (grey). **(B-E)** Variants within $2\sigma_D$ of ΔD of wt*Bs*LipA determined from screenings of 2997 wt*Bs*LipA replicates towards the respective detergent were excluded from further analyses (grey). The insets show the numbers of variants which cause a significant in- or decrease in T_{50} or D towards one detergent. A red (blue) color indicates a significantly increased (decreased) T_{50} or D towards one detergent.

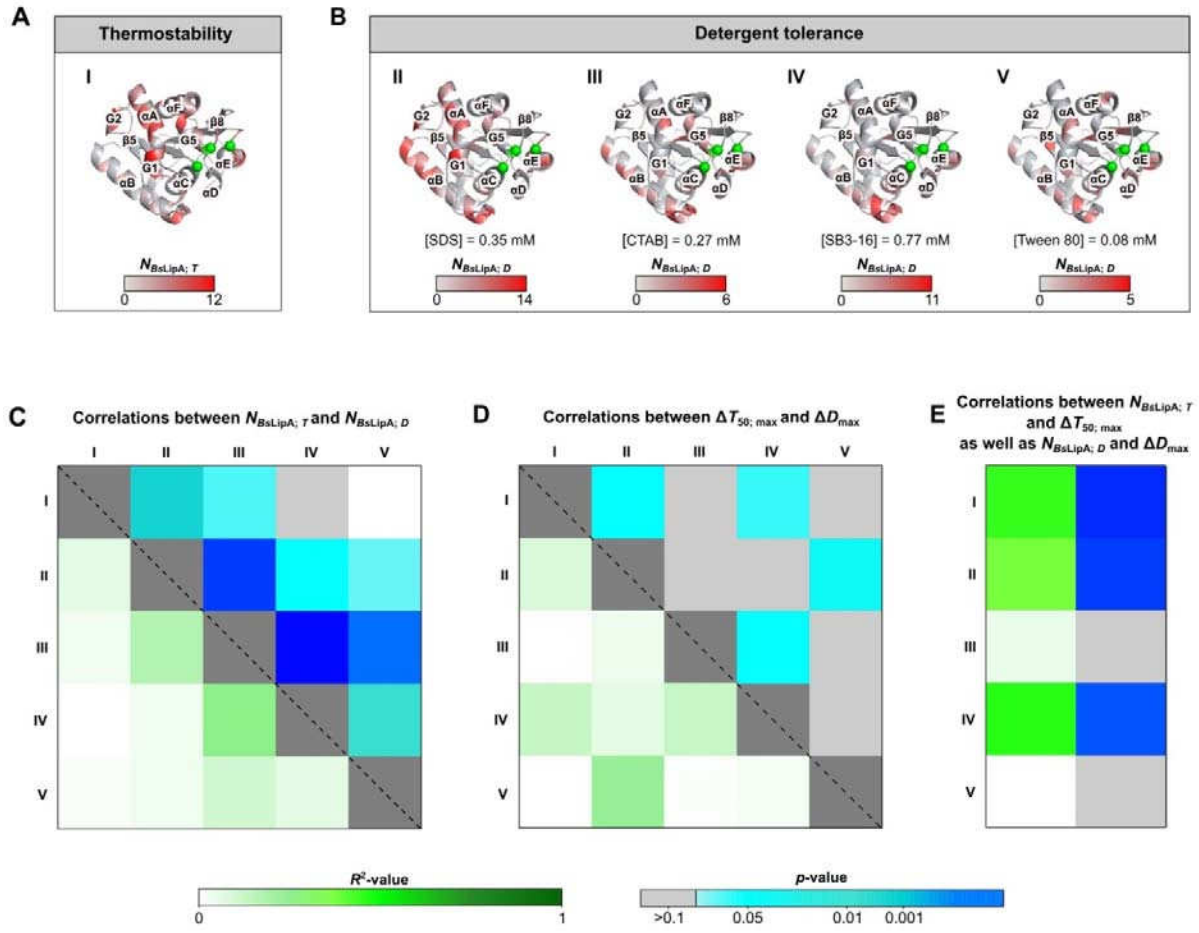


Figure 2: Localization of *Bs*LipA variants as to the frequency of substitution occurrences and highest effects regarding significantly increased T_{50} or D towards one detergent. (A) The maximum number of substitutions that cause significantly increased **(A)** T_{50} ($N_{BsLipA;\ T}$) of **I** (**I** = {Substitution site$_x$ | $1 \leq x \leq 181$, $T_{50}(x)$ is significantly increased}) or **(B)** D ($N_{BsLipA;\ D}$) of **II** – **V** (**II** – **V** = {Substitution site$_x$ | $1 \leq x \leq 181$, $D_{SDS\ /\ CTAB\ /\ SB3\text{-}16\ /\ Tween\ 80}(x)$ is significantly increased}), are mapped onto wt*Bs*LipA (PDB ID: 1ISP). C_α atoms of the catalytic triad S77 / D133 / H156 are shown as green spheres. A red (grey) color indicates a high (low) $N_{BsLipA;\ T}$ of **I** or $N_{BsLipA;\ D}$ of **II** – **V**. **(C)** R^2- and p-values for correlations between $N_{BsLipA;\ T}$ of **I** or $N_{BsLipA;\ D}$ of **II** – **V**. **(D)** Additionally, an analysis of the respective highest effects in significantly increased T_{50} ($\Delta T_{50;\ max}$) of **I** or D (ΔD_{max}) of **II** – **V** were performed. Here, R^2- and p-values for correlations between $\Delta T_{50;\ max}$ of **I** or ΔD_{max} of **II** – **V** are shown. **(E)** R^2- and p-values for correlations between $N_{BsLipA;\ T}$ and $\Delta T_{50;\ max}$ of **I** or $N_{BsLipA;\ D}$ and ΔD_{max} of **II** – **V**.

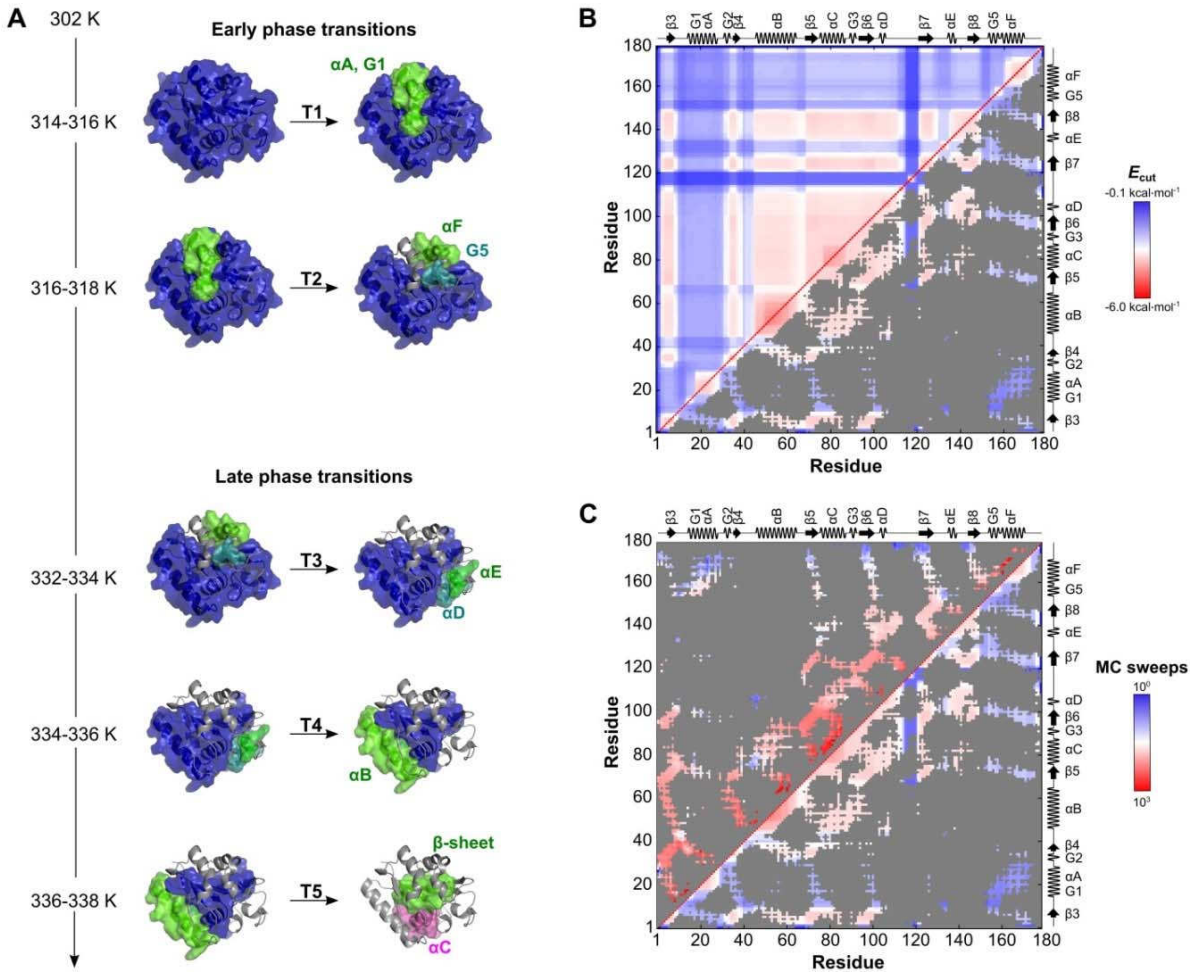


Figure 3: Prediction of the thermal unfolding pathway of wt*Bs*LipA. (A) Thermal unfolding pathway of wt*Bs*LipA (PDB ID: 1ISP) showing the early (T1 – T2) and late (T3 – T5) phase transitions. Rigid clusters are represented as uniformly colored blue, green, magenta, and cyan bodies in the descending order of their sizes. **(B)** For wt*Bs*LipA the stability map rc_{ij} including E_{cut} values at which a rigid contact between two residues is lost for all residue pairs during the thermal unfolding simulation **(upper triangle)**; the neighbor stability map $rc_{ij,neighbor}$ considering only the rigid contacts between two residues that are at most 5 Å apart from each other, with values for all other residue pairs colored gray **(lower triangle)**. The E_{cut} values are calculated with CNA based on a structural ensemble (ENT^{MD}). A red (blue) color indicates that contacts between residue pairs are more (less) rigid. **(C)** The aforementioned $rc_{ij,neighbor}$ **(lower triangle)** was compared with a contact map simulated by ProFASi **(upper triangle)**. A red (blue) color indicates contacts between residue pairs that have a longer (shorter) lifetime (in MC sweeps) than the contacts of the residue pairs of the initial protein structure. 3_{10}-helices are represented as G-helices.

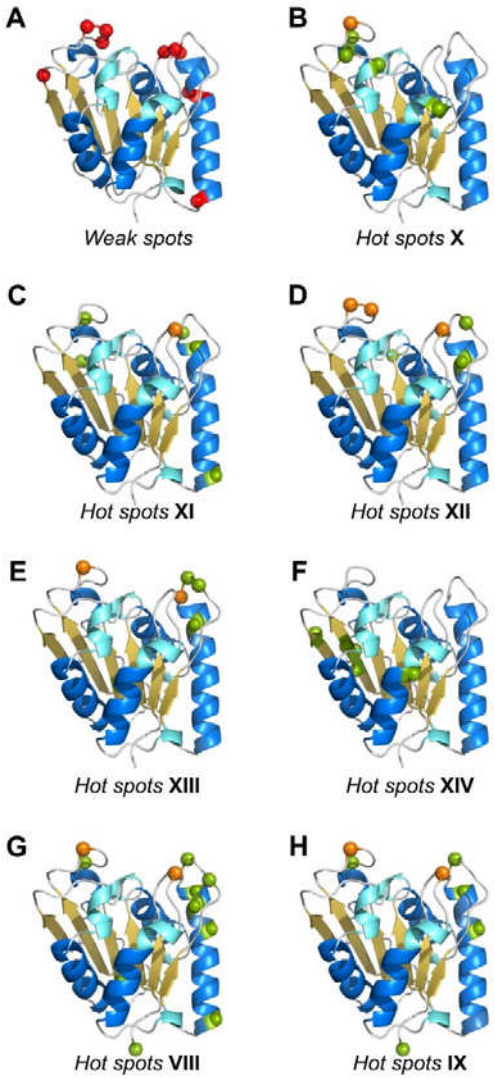
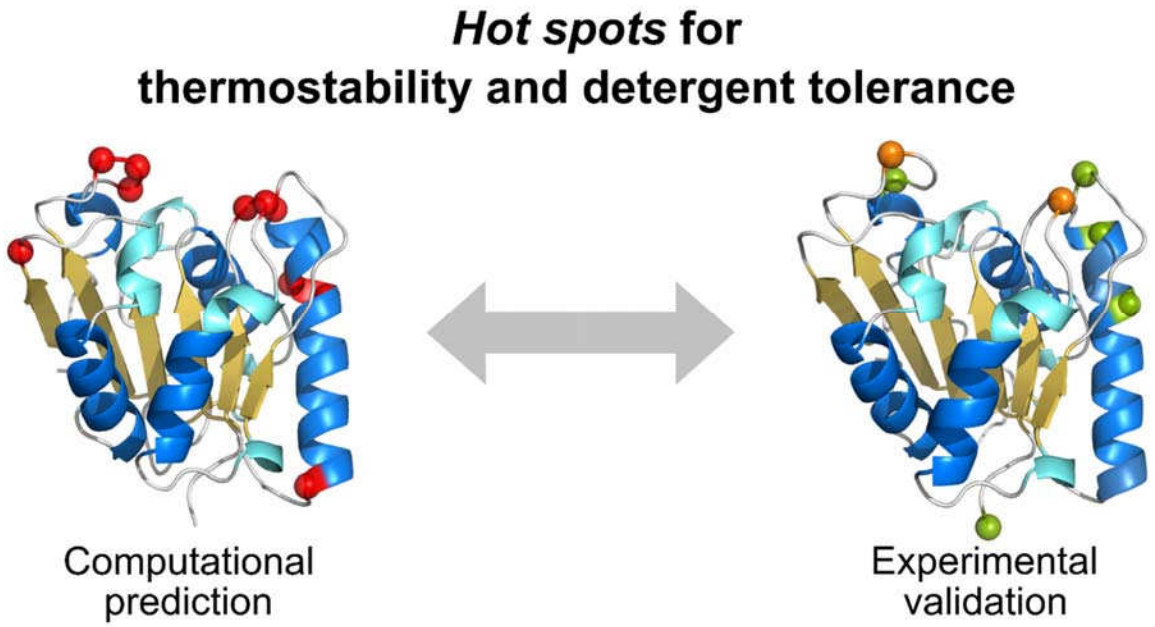


Figure 4: Localization of CNA-predicted *weak spots* and experimental *hot spots* of *Bs*LipA. (A) *Weak spots* and **(B)** *hot spots* of **X**, **(C) XI**, **(D) XII**, **(E) XIII**, **(F) XIV**, **(G) VIII**, and **(H) IV** are mapped onto wt*Bs*LipA (PDB ID: 1ISP). **(A)** Ten *weak spots*, i.e., I12, G13, G46, G52, P53, T66, M134, I135, V136, and H152, were predicted by CNA (red spheres). **(B-F)** The respective six substitution sites of **X** – **XIV** are considered *hot spots* as variants yield the respective six highest $\Delta T_{50;\,max}$ or ΔD_{max} towards one detergent for these substitution sites. **(G)** The eleven substitution sites of **VIII** showing significantly increased *D* towards each detergent, regardless of the magnitude of the single effect, and **(H)** the seven substitution sites of **IX** showing significantly increased T_{50} and *D* towards each detergent, regardless of the magnitude of the single effect, are considered *hot spots.* A green sphere represents a *hot spot*, and an orange sphere indicates a *hot spot* that was correctly predicted as *weak spot*.

TOC graphic


Hot spots for
thermostability and detergent tolerance
Computational
prediction
Experimental
validation

10. Tables

Table 1: Identified classes of substitution sites

Class[a]	Definition	No. of substitution sites	No. of *weak spots*[b]	*gip*[c]
I	{Substitution site$_x$ | $1 \leq x \leq 181$, $T_{50}(x)$ is significantly increased}	69	nd[d]	nd[d]
II	{Substitution site$_x$ | $1 \leq x \leq 181$, $D_{SDS}(x)$ is significantly increased}	74	nd[d]	nd[d]
III	{Substitution site$_x$ | $1 \leq x \leq 181$, $D_{CTAB}(x)$ is significantly increased}	42	nd[d]	nd[d]
IV	{Substitution site$_x$ | $1 \leq x \leq 181$, $D_{SB3\text{-}16}(x)$ is significantly increased}	46	nd[d]	nd[d]
V	{Substitution site$_x$ | $1 \leq x \leq 181$, $D_{Tween\ 80}(x)$ is significantly increased}	34	nd[d]	nd[d]
VI	**II ∪ III ∪ IV ∪ V**	109	nd[d]	nd[d]
VII	**I ∪ VI**	124	nd[d]	nd[d]
VIII	**II ∩ III ∩ IV ∩ V**	11	2	3.30
IX	**I ∩ VIII**	7	2	5.17
X	{Substitution site$_x$ | $1 \leq x \leq 181$, six highest effects in significantly increased $T_{50}(x)$}	6	1	3.02
XI	{Substitution site$_x$ | $1 \leq x \leq 181$, six highest effects in significantly increased $D_{SDS}(x)$}	6	1	3.02
XII	{Substitution site$_x$ | $1 \leq x \leq 181$, six highest effects in significantly increased $D_{CTAB}(x)$}	6	3	9.05
XIII	{Substitution site$_x$ | $1 \leq x \leq 181$, six highest effects in significantly increased $D_{SB3\text{-}16}(x)$}	6	2	6.03
XIV	{Substitution site$_x$ | $1 \leq x \leq 181$, six highest effects in significantly increased $D_{Tween\ 80}(x)$}	6		/
XV	**XI ∪ XII ∪ XIII ∪ XIV**	20	nd[d]	nd[d]
XVI	**X ∪ XV**	24	nd[d]	nd[d]
XVII	**XI ∩ XII ∩ XIII ∩ XIV**	0	nd[d]	nd[d]
XVIII	**X ∩ XVII**	0	nd[d]	nd[d]

[a] Class of substitution sites; underlined classes represent *hot spots*.

[b] Numbers of *hot spots* that are predicted as *weak spots*.

[c] Gain in precision over random classification (**Eq. 11**).

[d] Not determined.

Table 2: CNA-predicted *weak spots* of *Bs*LipA.

Weak spot	**Location at secondary structure elements**	**Phase transition**
I12	Turn	T1
G13	Turn	T1
G46	Loop	T4
G52	αB	T4
P53	αB	T4
T66	αB	T4
M134	Bend	T2
I135	Bend	T2
V136	Bend	T2
H152	Bend	T1