

Supporting Information

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

Christina Nutschel^{1,2}, Alexander Fulton³, Olav Zimmermann², Ulrich Schwaneberg^{5,6},
Karl-Erich Jaeger^{3,4}, Holger Gohlke^{1,2,7,*}

¹ John von Neumann Institute for Computing (NIC) and Institute of Biological Information Processing (IBI-7: Structural Biochemistry), 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

*Corresponding author:

John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), and Institute of Biological Information Processing (IBI-7: Structural Biochemistry)

Forschungszentrum Jülich GmbH

Wilhelm-Johnen-Straße

52425 Jülich

Germany

Email: gohlke@uni-duesseldorf.de or h.gohlke@fz-juelich.de

Table of Contents

Supplemental Tables	S4-S23
Table S1: Variants with significantly changed T_{50} or D towards one detergent by <i>random mutagenesis</i> .	S4
Table S2: Identified classes of substitution sites.	S5
Table S3: Correlations between the frequency of substitution occurrences per substitution site of classes I – V , where variants yield significantly increased T_{50} ($N_{BsLipA; T}$) or D ($N_{BsLipA; D}$) towards one detergent.	S7
Table S4: Correlations between the highest effects per substitution site of class I – V , considering variants with significantly increased T_{50} ($\Delta T_{50; \max}$) or D (ΔD_{\max}) towards one detergent.	S8
Table S5: Correlations between $N_{BsLipA; T}$ and $\Delta T_{50; \max}$ as well as $N_{BsLipA; D}$ and ΔD_{\max} of substitution sites of classes I – V .	S9
Table S6: Jaccard indices (J) for substitution sites of classes I – V or X – XIV .	S10
Table S7: Variants with significantly changed D at substitution sites of class VIII .	S11
Table S8: Variants with significantly changed T_{50} or D at substitution sites of class IX .	S12
Table S9: Substitution sites of classes X – XIV with the respective six highest effects in significantly increased T_{50} or D towards one detergent.	S13
Table S10: Variants with significantly changed T_{50} or D at substitution sites of classes X – XIV .	S14
Table S11: Distribution of <i>hot spots</i> regarding secondary structure elements.	S15
Table S12: Distribution of <i>hot spots</i> regarding fractional solvent accessible surface areas.	S16
Table S13: Distribution of <i>hot spots</i> regarding physicochemical properties.	S17
Table S14: Conservation of wt Bs LipA residues within bacterial lipases.	S18
Table S15: Amino acid types that lead to maximum changes in thermostability or detergent tolerance.	S22
Table S16: <i>Hot spot</i> classes for amino acids located in loops, turns, or bends and being of neutral or aliphatic type.	S23
Supplemental Figures	S24-S25
Figure S1: Control experiments regarding <i>p</i> NP absorption.	S24
Figure S2: Autocorrelation function of the cluster configuration entropy $H_{\text{type}2}$.	S25

Supplemental References

S26

Supplemental Tables

Table S1: Variants with significantly changed T_{50} or D towards one detergent by *random mutagenesis*.

Type of protein stability	No. of variants with $\Delta T_{50} > 0$ K or $\Delta D > 0$ ^[a]	No. of variants with $\Delta T_{50} < 0$ K or $\Delta D < 0$ ^[a]	Total no. of variants	Concentration [mM] ^[b]	σ_D [mM] ^[c]
T_{50}	214 (11.5)	1642 (88.5)	1856	/	/
D_{SDS}	261 (14.6)	1532 (85.4)	1793	0.35	0.08
D_{CTAB}	87 (10.3)	760 (89.7)	847	0.27	0.09
D_{SB3-16}	103 (22.2)	361 (77.8)	464	0.77	0.32
$D_{Tween\ 80}$	52 (10.5)	443 (89.5)	495	0.08	0.14
Mean (D)	126 (14.4)	774 (85.6)	900	/	/

^[a] Number of variants; values in brackets represent the likelihood [%] to find variants with significantly changed T_{50} or D in relation to the total number of variants, respectively.

^[b] Used detergent concentration; CMC values according to published data: SDS (7 mM); CTAB (1 mM); SB3-16 (0.01 mM); Tween 80 (0.012 mM)¹.

^[c] Standard deviations of 2997 wtBsLipA replicates for each concentration ².

Table S2: Identified classes of substitution sites.

Class ^[a]	Definition	No. of substitution sites	Random classification ^[b]	CNA				
				No. of weak spots ^[c]	r ^[d]	p ^[e]	gip ^[f]	FI ^[g]
I	{Substitution site _x 1 ≤ x ≤ 181, $T_{50}(x)$ is significantly increased}	69	0.38	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]
II	{Substitution site _x 1 ≤ x ≤ 181, $D_{SDS}(x)$ is significantly increased}	74	0.41	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]
III	{Substitution site _x 1 ≤ x ≤ 181, $D_{CTAB}(x)$ is significantly increased}	42	0.23	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]
IV	{Substitution site _x 1 ≤ x ≤ 181, $D_{SB3-16}(x)$ is significantly increased}	46	0.25	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]
V	{Substitution site _x 1 ≤ x ≤ 181, $D_{Tween\ 80}(x)$ is significantly increased}	34	0.19	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]
VI	II U III U IV U V	109	0.60	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]
VII	I U VI	124	0.69	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]
VIII	II ∩ III ∩ IV ∩ V	11	0.06	2	0.18	0.20	3.30	0.19
IX	I ∩ VIII	7	0.04	2	0.29	0.20	5.17	0.24
X	{Substitution site _x 1 ≤ x ≤ 181, six highest effects in significantly increased $T_{50}(x)$ }	6	0.03	1	0.17	0.10	3.02	0.13
XI	{Substitution site _x 1 ≤ x ≤ 181, six highest effects in significantly increased $D_{SDS}(x)$ }	6	0.03	1	0.17	0.10	3.02	0.13
XII	{Substitution site _x 1 ≤ x ≤ 181, six highest effects in significantly increased $D_{CTAB}(x)$ }	6	0.03	3	0.50	0.30	9.05	0.38
XIII	{Substitution site _x 1 ≤ x ≤ 181, six highest effects in significantly increased $D_{SB3-16}(x)$ }	6	0.03	2	0.33	0.20	6.03	0.25
XIV	{Substitution site _x 1 ≤ x ≤ 181, six highest effects in significantly increased $D_{Tween\ 80}(x)$ }	6	0.03	0	/	/	/	/

Table S2 continued.

Class ^[a]	Definition	No. of substitution sites	Random classification ^[b]	CNA				
				No. of weak spots ^[c]	r ^[d]	p ^[e]	gip ^[f]	FI ^[g]
XV	XI U XII U XIII U XIV	20	0.11	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]
XVI	X U XV	24	0.13	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]
XVII	XI ∩ XII ∩ XIII ∩ XIV	0	0	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]
XVIII	X ∩ XVII	0	0	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]	nd ^[h]

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

^[b] Likelihood for randomly choosing substitution sites / *hot spots* of the respective class.

^[c] Numbers of *hot spots* predicted as *weak spots*.

^[d] Recall (**Eq. 8 in the main text**).

^[e] Precision (**Eq. 9 in the main text**).

^[f] Gain in precision over *random classification* (**Eq. 11 in the main text**).

^[g] F1-score (**Eq. 12 in the main text**).

^[h] Not determined.

Table S3: Correlations between the frequency of substitution occurrences per substitution site of classes **I – V**, where variants yield significantly increased T_{50} ($N_{Bs\text{LipA}; T}$) or D ($N_{Bs\text{LipA}; D}$) towards one detergent.

		Class of substitution site^[a]				
		I	II	III	IV	V
Class of substitution site^[a]	I	/	< 0.001	< 0.05	> 0.1	< 0.1
	II	0.066	/	< 0.001	< 0.05	< 0.05
	III	0.033	0.176	/	< 0.001	< 0.001
	IV	0.004	0.036	0.263	/	< 0.001
	V	0.015	0.031	0.105	0.059	/

^[a] Upper values are p -values; lower values are R^2 -values.

Table S4: Correlations between the highest effects per substitution site of classes **I – V**, considering variants with significantly increased T_{50} ($\Delta T_{50; \text{max}}$) or D (ΔD_{max}) towards one detergent.

		Class of substitution site ^[a]				
		I	II	III	IV	V
Class of substitution site ^[a]	I	/	< 0.1	> 0.1	< 0.1	> 0.1
	II	0.089	/	> 0.1	> 0.1	< 0.05
	III	0.002	0.041	/	< 0.1	> 0.1
	IV	0.132	0.064	0.128	/	> 0.1
	V	$4 \square 10^{-5}$	0.235	0.007	0.029	/

^[a] Upper values are p -values; lower values are R^2 -values.

Table S5: Correlations between $N_{Bs\text{LipA}; T}$ and $\Delta T_{50; \text{max}}$ as well as $N_{Bs\text{LipA}; D}$ and ΔD_{max} of substitution sites of classes **I – V**.

Class of substitution sites	R^2 -value	p-value
I	0.449	< 0.001
II	0.382	< 0.001
III	0.054	> 0.1
IV	0.464	< 0.001
V	0.008	> 0.1

Table S6: Jaccard indices (J) for substitution sites of classes **I – V** or **X – XIV**.

		Class of substitution site ^[a]				
		I / X	II / XI	III / XII	IV / XIII	V / XIV
Class of substitution site ^[a]	I / X	/	0.091	0.091	0.091	0
	II / XI	0.416	/	0	0	0
	III / XII	0.291	0.333	/	0.200	0
	IV / XIII	0.264	0.250	0.467	/	0
	V / XIV	0.157	0.227	0.310	0.250	/

^[a] Lower values are J of substitution sites for which variants yield significantly increased T_{50} and D towards one detergent as well as significantly increased D towards two detergents (**I – V**); upper values are the J of the six substitution sites for which variants yield the respective highest effects regarding significantly increased T_{50} and D towards one detergent as well as significantly increased D towards two detergents (**X – XIV**).

Table S7: Variants with significantly changed D at substitution sites of class **VIII**.

Type of protein stability	No. of variants with $\Delta D > 0^{[a]}$	No. of variants with $\Delta D < 0^{[a]}$	Total no. of variants
D_{SDS}	63 (54.8)	52 (45.2)	115
D_{CTAB}	19 (76.0)	6 (24.0)	25
D_{SB3-16}	31 (79.5)	8 (20.5)	39
$D_{Tween\ 80}$	16 (84.2)	3 (15.8)	19
Mean (D)	32 (64.0)	17 (36.0)	50

[a] Number of variants; values in brackets represent the likelihood [%] to find variants with significantly changed D in relation to the total number of variants, respectively.

Table S8: Variants with significantly changed T_{50} or D at substitution sites of class **IX**.

Type of protein stability	No. of variants with $\Delta T_{50} > 0$ or $\Delta D > 0^{[a]}$	No. of variants with $\Delta T_{50} < 0$ or $\Delta D < 0^{[a]}$	Total no. of variants
T_{50}	35 (40.7)	51 (59.3)	86
D_{SDS}	39 (56.5)	30 (43.5)	69
D_{CTAB}	10 (79.9)	3 (23.1)	13
D_{SB3-16}	18 (78.3)	5 (21.7)	23
$D_{Tween\ 80}$	10 (83.3)	2 (16.7)	12
Mean (D)	19 (65.5)	10 (34.5)	29

[a] Number of variants; values in brackets represent the likelihood [%] to find variants with significantly changed T_{50} or D in relation to the total number of variants, respectively.

Table S9: Substitution sites of class X – XIV with the respective six highest effects in significantly increased T_{50} or D towards each detergent.

Class of substitution site	Substitution site ^[a]	$\Delta T_{50; \text{max}} [\text{K}]$ or ΔD_{max}	$N_{\text{BsLipA}; T \text{ or }} N_{\text{BsLipA}; D}$	Location ^[b]
X	M137	7.7	10	Loop
	M134	5.6	9	Bend
	G155	4.5	4	Loop
	F17	3.8	12	G1
	I157	3.6	1	G5
	Y139	3.2	7	αE
XI	M137	1.49	6	Loop
	R142	1.45	7	Loop
	T47	1.29	8	Loop
	E65	1.10	14	αB
	G13	1.03	6	Turn
	Y49	0.94	9	αB
XII	T110	1.63	1	Loop
	K44	1.04	4	Turn
	I135	1.01	6	Bend
	G13	0.72	1	Turn
	M134	0.58	4	Bend
	N51	0.55	1	αB
XIII	G46	2.41	11	Loop
	K44	2.26	3	Turn
	M134	2.25	1	Bend
	N51	2.10	7	αB
	T45	1.90	4	Turn
	V99	1.87	5	β6
XIV	S127	2.29	3	β7
	I128	2.00	1	β7
	T126	1.98	1	β7
	L123	1.89	1	β7
	Q150	1.07	2	β8
	A20	0.86	1	αA

^[a] Substitution sites highlighted in bold are predicted as *weak spots* by CNA.

^[b] Location of the substitution site in terms of secondary structure elements.

Table S10: Variants with significantly changed T_{50} or D at substitution sites of classes **X – XIV**.

Type of protein stability	No. of variants with $\Delta T_{50} > 0$ or $\Delta D > 0$ ^[a]	No. of variants with $\Delta T_{50} < 0$ or $\Delta D < 0$ ^[a]	Total no. of variants
T_{50}	43 (63.2)	25 (36.8)	68
D_{SDS}	50 (70.4)	21 (29.6)	71
D_{CTAB}	17 (53.1)	15 (46.9)	32
D_{SB3-16}	31 (79.5)	8 (20.5)	39
$D_{Tween\ 80}$	9 (32.1)	19 (67.9)	28
Mean (D)	27 (62.8)	16 (37.2)	43

^[a] Number of variants; values in brackets represent the likelihood [%] to find variants with significantly increased T_{50} or D in relation to the total number of variants, respectively.

Table S11: Distribution of *hot spots* regarding secondary structure elements.

Secondary structure element	No of substitution sites	Class of hot spots																												
		X				XI				XII				XIII				XIV				VIII								
		#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$				
aA	9	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	1	0.17	0.11	0.13	3.35	/	/	/	/				
aB	19	/	/	/	/	/	2	0.33	0.11	0.16	3.18	1	0.17	0.05	0.08	1.59	1	0.17	0.05	0.08	1.59	/	/	/	/					
aC	12	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/					
aE	4	1	0.17	0.25	0.20	7.54	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/					
aF	11	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/					
a(all)	55	1	0.17	0.02	0.03	0.55	2	0.33	0.04	0.07	1.10	1	0.10	0.02	0.03	0.55	1	0.10	0.02	0.03	0.55	1	0.10	0.02	0.03	0.55				
G1	4	1	0.17	0.25	0.20	7.54	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	4	0.36	0.07	0.12	1.20				
G2	3	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	2	0.29	0.04	0.06	0.94					
G3	3	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/					
G4	3	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/					
G5	5	1	0.17	0.20	0.18	6.03	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/					
G(all)	18	2	0.33	0.11	0.17	3.35	/	/	/	/	/	/	/	/	/	/	/	/	/	/	1	0.1	0.1	0.1	1					
β 3	4	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/					
β 4	3	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/					
β 5	6	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/					
β 6	7	/	/	/	/	/	/	/	/	/	/	/	/	/	1	0.17	0.14	0.15	4.31	/	/	/	/	1	0.09	0.14	0.11	2.35		
β 7	7	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	4	0.67	0.57	0.62	17.24	/	/	/	/	/				
β 8	5	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	1	0.17	0.20	0.18	6.03	/	/	/	/	/				
β (all)	32	/	/	/	/	/	/	/	/	/	/	/	/	1	0.10	0.03	0.05	0.94	5	0.83	0.16	0.26	4.71	1	0.09	0.03	0.05	0.51		
Bridge	3	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/					
Turn	13	/	/	/	/	/	1	0.17	0.08	0.11	2.32	2	0.33	0.15	0.21	4.64	2	0.33	0.15	0.21	4.64	/	/	/	/	2	0.18	0.15	0.17	2.53
Loop	47	2	0.33	0.04	0.08	1.28	3	0.50	0.06	0.11	1.93	1	0.17	0.02	0.04	0.64	1	0.17	0.02	0.04	0.64	/	/	/	/	3	0.27	0.06	0.10	1.05
Bend	13	1	0.17	0.08	0.11	2.32	/	/	/	/	/	2	0.33	0.15	0.21	4.64	1	0.17	0.08	0.11	2.32	/	/	/	/	1	0.09	0.08	0.08	1.27
Rest(all)	76	3	0.50	0.04	0.07	1.19	4	0.67	0.05	0.10	1.59	5	0.83	0.07	0.12	1.98	4	0.67	0.05	0.10	1.59	/	/	/	/	6	0.55	0.08	0.14	1.30
		5	0.71	0.07	0.12	1.70																								

[a] Recall (Eq. 8 in the main text).

[b] Precision (Eq. 9 in the main text).

[c] F1-score (Eq. 12 in the main text).

[d] Gain in precision over random classification (Eq. 11 in the main text).

Table S12: Distribution of *hot spots* regarding fractional solvent accessible surface areas.

fsASA	No of substitution sites	Class of <i>hot spots</i>																								
		X				XI				XII				XIII				XIV				VIII				
		#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$
$\geq 0; \leq 0.2$	98	1	0.17	0.01	0.02	0.31	1	0.17	0.01	0.02	0.31	2	0.33	0.02	0.04	0.62	2	0.3	0.03	0.05	0.76	4	0.70	0.04	0.08	1.23
$> 0.2; \leq 0.4$	31	/	/	/	/	/	1	0.17	0.03	0.05	0.97	/	/	/	/	/	/	/	/	/	3	0.27	0.10	0.14	1.59	
$> 0.4; \leq 0.6$	26	2	0.33	0.08	0.13	2.32	2	0.33	0.08	0.13	2.32	/	/	/	/	/	3	0.5	0.06	0.10	1.68	2	0.30	0.08	0.13	2.32
$> 0.6; \leq 0.8$	18	3	0.50	0.17	0.25	5.03	2	0.33	0.11	0.17	3.35	4	0.67	0.22	0.33	6.70	1	0.2	0.05	0.07	1.44	/	/	/	/	/
$> 0.8; \leq 1.0$	8	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	1	0.09	0.13	0.11	2.06	
																					1	0.14	0.13	0.13	3.23	

[a] Recall (Eq. 8 in the main text).

[b] Precision (Eq. 9 in the main text).

[c] F1-score (Eq. 12 in the main text).

[d] Gain in precision over random classification (Eq. 11 in the main text).

Table S13: Distribution of *hot spots* regarding physicochemical properties.

Physicochemical property	No of substitution sites	Class of hot spots																																		
		X				XI				XII				XIII				XIV				VIII														
		#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$										
Aliphatic	79	2	0.33	0.03	0.05	0.76	1	0.17	0.01	0.02	0.38	2	0.33	0.03	0.05	0.76	3	0.50	0.04	0.07	1.15	3	0.50	0.04	0.07	1.15	2	0.18	0.03	0.04	0.42	2	0.29	0.03	0.05	0.65
Aromatic	15	2	0.33	0.13	0.19	4.02	1	0.17	0.07	0.10	2.01	/	/	/	/	/	/	/	/	/	/	/	1	0.09	0.07	0.08	1.10	1	0.14	0.07	0.09	1.72				
Neutral	54	2	0.33	0.04	0.07	1.12	2	0.33	0.04	0.07	1.12	3	0.5	0.06	0.10	1.68	3	0.50	0.06	0.10	1.68	3	0.50	0.06	0.10	1.68	5	0.45	0.09	0.15	1.52	3	0.43	0.06	0.10	1.44
Charged (+)	21	/	/	/	/	/	1	0.17	0.05	0.07	1.44	1	0.17	0.05	0.07	1.44	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/		
Charged (-)	12	/	/	/	/	/	1	0.17	0.08	0.11	2.51	/	/	/	/	/	/	/	/	/	/	/	3	0.27	0.25	0.26	4.11	1	0.14	0.08	0.11	2.15				

[a] Recall (Eq. 8 in the main text).

[b] Precision (Eq. 9 in the main text).

[c] F1-score (Eq. 12 in the main text).

[d] Gain in precision over *random classification* (Eq. 11 in the main text).

Table S14: Conservation of wtBsLipA residues within bacterial lipases.

Residue ^[a]	Conservation	$N_{Bs\text{LipA}; T}$	$N_{Bs\text{LipA}; D}$ (SDS)	$N_{Bs\text{LipA}; D}$ (CTAB)	$N_{Bs\text{LipA}; D}$ (SB3-16)	$N_{Bs\text{LipA}; D}$ (Tween 80)
A1	2	3	0	0	0	0
<u>E2</u>	2	4	2	1	1	1
H3	2	1	0	1	0	0
N4	2	0	1	0	0	0
P5	10	0	1	0	0	0
V6	9	0	1	0	1	0
V7	9	0	0	0	0	0
M8	7	1	0	0	0	0
V9	7	0	1	0	0	0
H10	3	1	0	0	0	0
G11	10	0	0	0	0	0
I12	5	4	3	1	0	3
G13	2	2	6	1	5	1
G14	4	2	0	0	0	0
A15	5	5	1	1	0	1
S16	0	2	1	0	1	0
<u>F17</u>	4	12	1	0	1	0
N18	0	3	1	0	0	0
F19	9	0	0	0	0	0
<u>A20</u>	3	7	5	0	0	1
G21	0	7	5	2	0	1
I22	0	0	0	0	0	0
K23	0	4	4	0	0	0
S24	0	6	8	0	0	0
Y25	1	0	0	0	0	0
L26	8	0	0	0	0	0
V27	2	4	3	0	1	0
S28	3	2	2	0	0	0
Q29	3	1	0	0	0	0
G30	10	0	0	0	0	0
W31	9	0	0	0	0	0
S32	3	9	1	0	0	0
R33	2	6	0	0	0	0
D34	0	1	12	0	0	0
K35	0	1	0	0	0	0
L36	0	0	0	0	0	0
Y37	0	0	1	0	0	2
A38	0	0	0	0	0	0
V39	0	0	0	0	1	0
D40	4	0	1	0	0	0
F41	7	0	0	0	0	0
W42	4	2	0	0	0	0
<u>D43</u>	2	0	4	1	4	1
<u>K44</u>	0	1	0	4	3	0
<u>T45</u>	0	1	11	1	4	1
G46	0	0	0	3	11	0
<u>T47</u>	0	3	8	5	4	0
N48	5	0	0	0	3	0
<u>Y49</u>	1	8	9	1	3	2
N50	2	1	7	4	6	0
<u>N51</u>	3	0	2	1	7	2
G52	3	1	0	0	3	0

Table S14 continued.

Residue^[a]	Conservation	$N_{Bs\text{LipA}; T}$	$N_{Bs\text{LipA}; D}$ (SDS)	$N_{Bs\text{LipA}; D}$ (CTAB)	$N_{Bs\text{LipA}; D}$ (SB3-16)	$N_{Bs\text{LipA}; D}$ (Tween 80)
P53	2	1	0	0	1	0
V54	3	1	1	1	2	1
L55	7	0	0	0	0	0
S56	3	0	0	0	0	0
R57	2	0	2	2	2	0
F58	3	0	8	0	0	3
V59	9	1	3	0	0	0
Q60	3	2	10	2	1	0
K61	3	0	11	4	0	0
V62	7	0	3	0	0	0
L63	5	0	8	1	0	0
D64	3	3	5	0	0	0
E65	2	0	14	3	1	1
T66	8	0	9	0	0	0
G67	9	1	10	2	0	0
A68	7	0	1	0	0	0
K69	2	1	4	1	0	0
K70	6	0	0	0	0	0
V71	7	0	3	0	0	0
D72	4	0	0	0	0	0
I73	9	0	1	0	0	0
V74	9	0	2	0	0	0
A75	8	0	0	0	0	0
H76	5	0	0	0	0	0
S77	10	0	0	0	0	0
M78	4	0	0	0	0	0
G79	10	0	0	0	0	0
G80	1	0	0	0	0	0
A81	4	0	0	1	1	0
N82	5	0	0	2	4	0
T83	4	0	0	0	1	1
L84	3	0	0	0	0	0
Y85	3	0	0	0	3	0
Y86	6	0	0	0	0	0
I87	7	0	0	0	0	0
K88	5	1	2	0	0	0
N89	3	4	0	0	1	0
L90	0	0	0	0	1	0
D91	0	0	0	0	0	0
G92	0	0	0	0	0	0
G93	0	0	0	0	0	0
N94	0	0	1	0	0	0
K95	0	0	0	0	0	0
V96	9	0	1	0	0	0
A97	2	0	0	0	0	0
<u>N98</u>	2	0	4	4	1	2
<u>V99</u>	5	0	0	5	5	5
V100	7	1	0	0	0	0
T101	0	2	0	0	0	0
L102	1	0	0	1	0	0
G103	1	0	0	0	0	0
G104	1	0	0	0	0	0

Table S14 continued.

Residue^[a]	Conservation	<i>N_{BsLipA; T}</i>	<i>N_{BsLipA; D} (SDS)</i>	<i>N_{BsLipA; D} (CTAB)</i>	<i>N_{BsLipA; D} (SB3-16)</i>	<i>N_{BsLipA; D} (Tween 80)</i>
A105	0	0	0	0	0	1
N106	1	0	3	2	2	0
R107	1	0	0	0	1	1
L108	3	1	0	0	1	0
T109	5	0	1	0	0	0
<u>T110</u>	0	0	0	1	0	0
G111	0	4	1	0	0	0
K112	0	5	1	0	0	0
A113	0	0	0	0	0	1
L114	0	1	1	0	0	0
P115	0	0	0	0	0	1
G116	0	0	0	0	0	0
T117	1	0	0	1	1	1
D118	0	0	0	0	0	0
P119	0	0	0	0	0	0
N120	1	3	0	0	0	1
Q121	0	1	0	0	0	0
K122	1	0	0	0	0	0
<u>L123</u>	0	0	1	0	0	1
L124	2	0	0	0	1	0
Y125	1	0	0	0	0	1
<u>T126</u>	5	0	1	0	0	1
<u>S127</u>	5	0	1	0	0	3
<u>I128</u>	6	0	0	0	0	1
Y129	0	0	0	0	0	0
S130	7	0	0	0	0	0
S131	1	2	1	0	0	0
A132	2	3	5	0	0	0
D133	2	0	0	0	0	0
M134	0	9	4	4	1	3
<u>I135</u>	1	7	1	6	2	0
V136	0	0	0	0	0	0
<u>M137</u>	0	10	6	1	2	1
N138	2	1	0	0	0	0
<u>Y139</u>	0	7	1	1	2	0
L140	0	0	0	1	0	2
S141	0	0	1	0	0	0
<u>R142</u>	0	4	7	1	0	0
L143	0	0	0	0	0	0
D144	0	1	7	6	1	0
G145	0	0	0	0	1	0
A146	2	0	1	0	0	0
R147	1	0	0	0	0	0
N148	2	0	0	0	0	0
V149	1	1	1	0	0	0
<u>Q150</u>	0	0	0	3	1	2
I151	1	0	0	0	0	0
H152	0	1	0	0	0	1
G153	0	1	1	1	1	0
V154	0	0	0	0	0	0
<u>G155</u>	0	4	1	0	0	0
H156	5	0	0	0	0	0

Table S14 continued.

Residue^[a]	Conservation	$N_{Bs\text{LipA}; T}$	$N_{Bs\text{LipA}; D}$ (SDS)	$N_{Bs\text{LipA}; D}$ (CTAB)	$N_{Bs\text{LipA}; D}$ (SB3-16)	$N_{Bs\text{LipA}; D}$ (Tween 80)
I157	0	1	0	1	0	0
G158	0	2	0	1	1	0
L159	0	0	0	0	0	0
L160	0	0	0	0	0	0
Y161	0	7	2	1	0	0
S162	2	2	2	0	0	0
S163	0	1	0	0	0	1
Q164	0	0	1	0	0	0
V165	1	0	0	0	0	0
N166	0	3	0	0	1	0
S167	0	0	0	0	0	0
L168	0	0	0	0	0	0
I169	1	0	0	0	0	0
K170	0	1	0	0	0	0
E171	0	0	0	0	0	0
G172	1	1	0	0	0	0
L173	2	0	0	0	0	0
N174	0	7	0	0	0	0
G175	0	0	0	0	0	0
G176	0	0	0	0	0	0
G177	0	0	0	0	0	0
Q178	0	0	3	0	0	0
N179	0	0	0	0	0	0
T180	0	0	0	0	1	0
N181	0	0	1	0	0	0

^[a] Underlined substitution sites are identified as *hot spots*; substitution sites highlighted in bold are predicted as *weak spots* by CNA

Table S15: Amino acid types that lead to maximum changes in thermostability or detergent tolerance.

AA type	$\Delta T_{50; \text{max}} [\text{K}]$	$\Delta D_{\text{max}; \text{SDS}}$	$\Delta D_{\text{max}; \text{CTAB}}$	$\Delta D_{\text{max}; \text{SB3-16}}$	$\Delta D_{\text{max}; \text{Tween80}}$
I	3.59	0.86	1.01	1.01	2.00
A	2.79	0.58	0.45	0.97	0.86
V	1.19	0.79	0.44	1.87	0.70
L	1.51	0.67	0.22	0.76	0.45
G	4.51	1.03	0.72	2.41	0.48
F	3.78	0.68	-0.19	1.07	0.39
Y	4.22	-1.37	0.35	1.44	0.51
W	1.68	-0.20	-0.19	0.67	-0.38
C	na ^[a]	na ^[a]	na ^[a]	na ^[a]	na ^[a]
P	0.54	0.35	-0.19	0.90	0.30
M	7.67	1.49	0.58	2.25	0.45
S	2.55	0.68	-0.19	0.74	2.29
T	0.94	1.29	1.63	1.90	1.98
N	2.65	0.76	0.55	2.10	0.51
Q	0.66	0.68	0.34	0.87	1.07
H	0.84	0.17	0.25	-0.67	0.31
K	1.87	0.80	1.04	2.26	-0.29
R	2.79	1.45	0.51	1.76	0.67
D	1.92	0.74	0.32	1.10	0.33
E	1.00	1.10	0.23	1.27	0.55

^[a] Not available.

Table S16: *Hot spot* classes for amino acids located in loops, turns, or bends and being of neutral or aliphatic type.

Criterion	No of substitution sites	Classes of <i>hot spot</i>																VIII		IX															
		X				XI				XII				XIII				XIV																	
		#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$	#	$r^{[a]}$	$p^{[b]}$	$FI^{[c]}$	$gip^{[d]}$														
Loop or turn or bend and aliphatic or neutral	43	1	0.17	0.02	0.04	0.70	3	0.50	0.07	0.12	2.10	3	0.50	0.07	0.12	2.10	1	0.17	0.02	0.04	0.70	/	/	/	/	2	0.18	0.05	0.07	0.77	2	0.29	0.05	0.08	1.20

[a] Recall (**Eq. 8 in the main text**).

[b] Precision (**Eq. 9 in the main text**).

[c] F1-score (**Eq. 12 in the main text**).

[d] Gain in precision over *random classification* (**Eq. 11 in the main text**).

Supplemental Figures

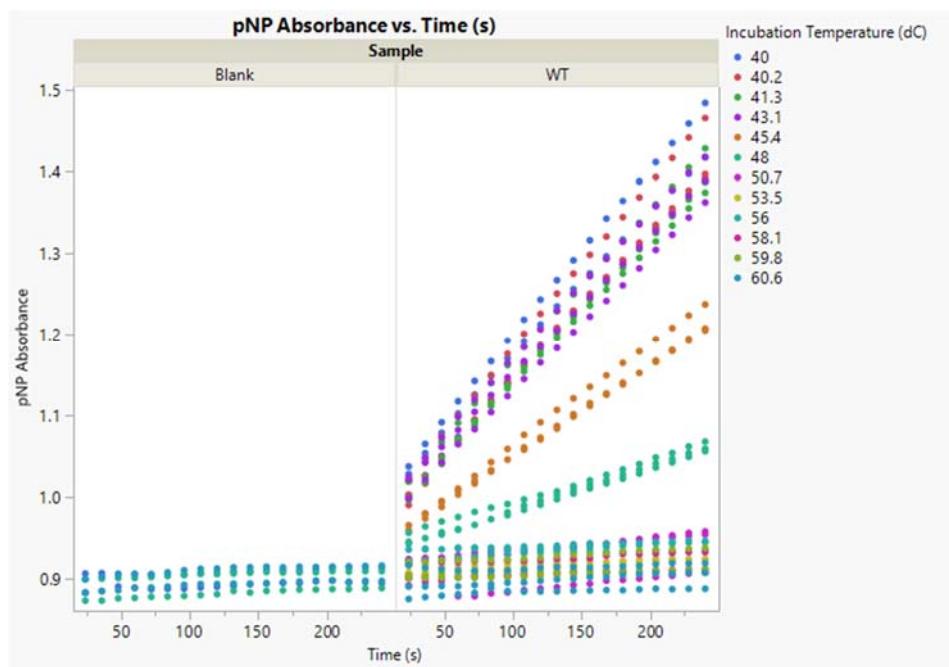


Figure S1: Control experiments regarding *p*NP absorption. The *p*NP absorption over time was measured at different temperatures between 40 and 60.6°C. On the left, the results of the control experiments, i.e., just *p*NPP in solution, are shown for temperatures up until 60.6°C. On the right, the results of *p*NP absorption in the presence of *BsLipA* are shown. At temperatures above 48°C, the protein denatures; no increase in *p*NP absorption over time is observed then.

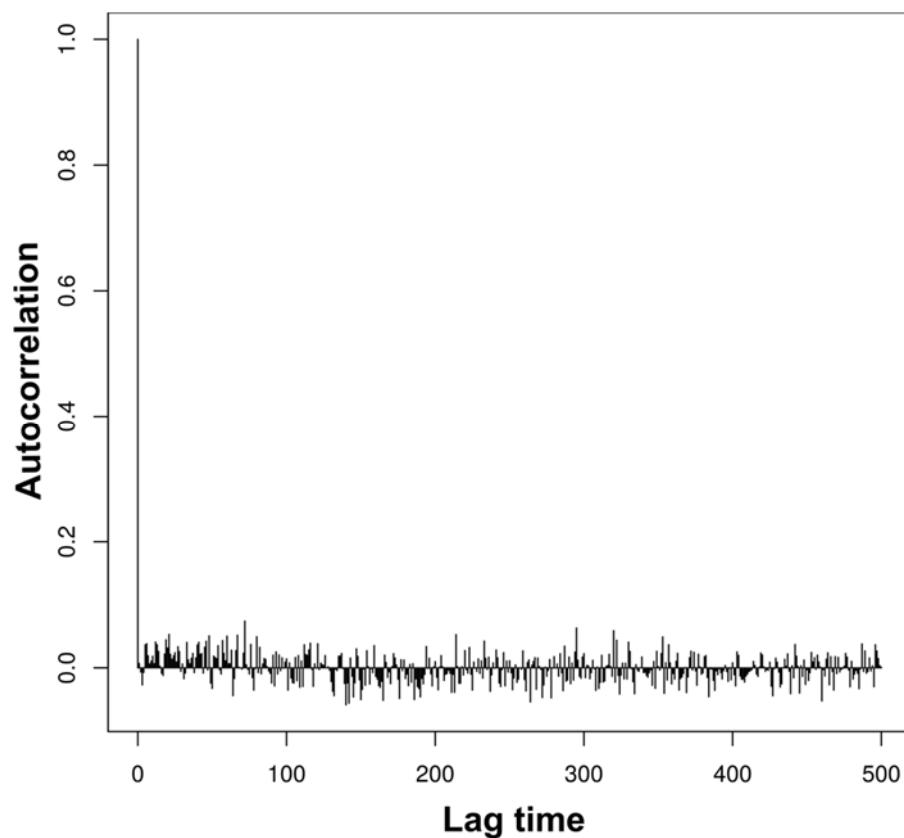


Figure S2: Autocorrelation function of the cluster configuration entropy H_{type2} . The snapshots were extracted at time intervals of 40 ps, and the lag time is in multiples of 40 ns.

Supplemental References

1. Bhairi, S, A Guide to the Properties and Uses of Detergents in Biology and Biochemistry. *Calbiochem-Novobiochem Corporation San Diego, 1997.*
2. 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.