Electronic supplementary information (ESI)

Insights into the role of electrostatics in temperature adaptation: A comparative study of psychrophilic, mesophilic, and thermophilic subtilisin-like serine proteases

Yuan-Ling Xia,‡[a] Jian-Hong Sun,‡[a] Shi-Meng Ai,[b] Yi Li,[a] Xing Du,[a] Peng Sang,[c] Li-Quan Yang,[c] Yun-Xin Fu*[a,d] and Shu-Qun Liu*[a,e]

[a]State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Yunnan University, Kunming, Yunnan, P.R. China

[b]Department of Applied Mathematics, Yunnan Agricultural University, Kunming, Yunnan, P. R. China

[c]College of Agriculture and Biological Science, Dali University, Dali, Yunnan, P. R. China

[d]Human Genetics Center and Division of Biostatistics, School of Public Health, the University of Texas Health Science Center, Houston, Texas, USA

[e]Key Laboratory for Tumor molecular biology of High Education in Yunnan Province, School of Life Sciences, Yunnan University, Kunming, Yunnan, P. R. China

‡ These authors contributed equally to this work

* Corresponding author

Email: Yunxin.Fu@uth.tmc.edu (YXF), shuqunliu@ynu.edu.com (SQL)

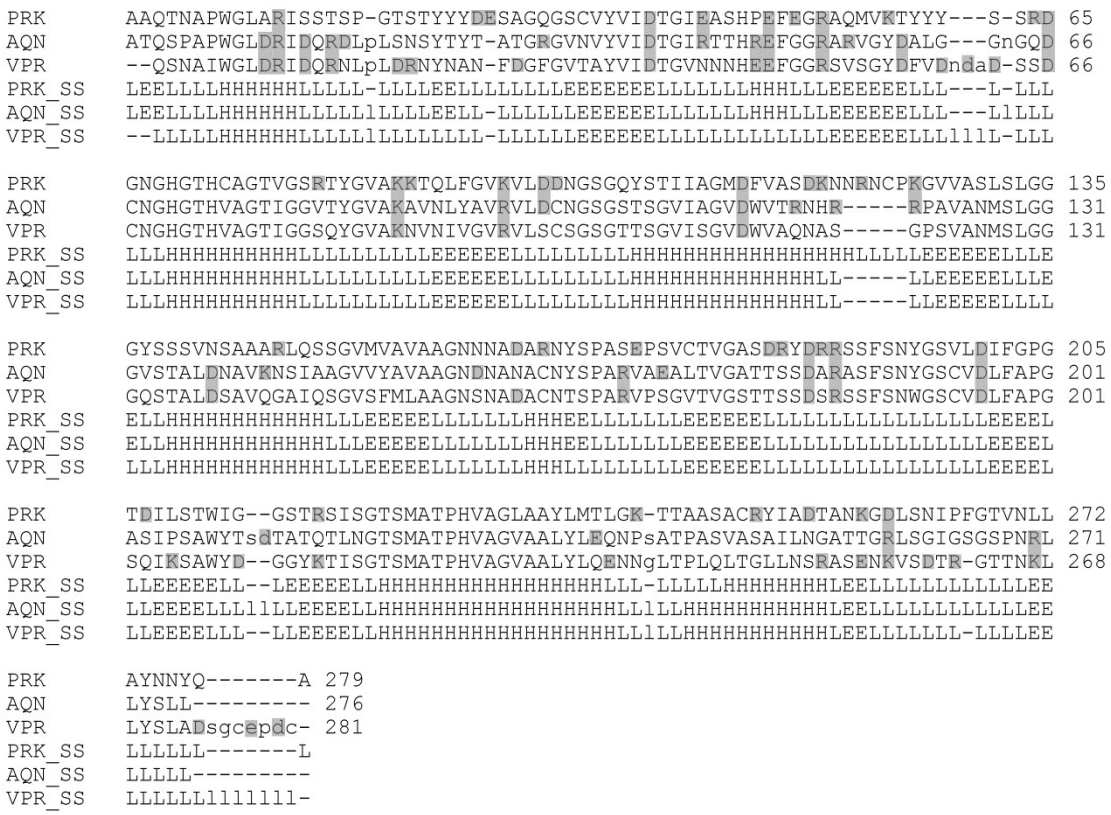

Fig. S1 Structure-based multiple sequence alignment of the psychrophilic VPR, mesophilic PRK, and thermophilic AQN. Protein secondary structure (SS) is shown below the alignment, with H, E, and L/l representing the α-helix (or 3/10 helix), β-strand, and loop, respectively. Residue insertion and deletion are denoted by lowercase single-letter amino acid code and '-', respectively. The charged residues are highlighted in grey.

Table S1 pK_a values of histidines in the three protease structures predicted by DelPhiPKa[†].

Histidines[a]	VPR	PRK	AQN
His46	6.15	5.78	5.08
His69	6.99	6.86	6.82
His72	6.13	6.10	5.73
His118	-	-	5.50
His229	5.47	5.44	5.47

[†]DelPhiPKa default parameters were used except for the force field and salt concentration, which were set to CHARMM and 100 mM, respectively.

[a]Residue numbering is according to the PRK sequence and structurally equivalent residue positions were determined from structure-based multiple sequence alignment (Fig. S1).

Table S2 Comparison of amino acid composition and numbers of salt bridges and salt-bridge networks among the psychrophilic VPR, mesophilic PRK, and thermophilic AQN.

Parameter	VPR	PRK	AQN
PDB ID	1SH7	1IC6	4DZT
Resolution (Å)	1.84	0.98	1.95
Residue number	281	279	276
Charged[a]	38 (13.5%)	39 (14.0%)	33 (12.0%)
Acidic[a]	24 (8.5%)	19 (6.8%)	16 (5.8%)
Basic[a]	14 (5.0%)	20 (7.2%)	17 (6.2%)
Polar uncharged[a]	141 (50.2%)	137 (49.1%)	126 (45.7%)
Hydrophobic[a]	98 (34.9%)	99 (35.5%)	112 (40.6%)
Aromatic[a]	19 (6.8%)	25 (9.0%)	18 (6.5%)
Salt bridge[b]	7 (36.8%)	8 (41.0%)	6 (36.4%)
Salt-bridge network[b]	1 (7.9%)	1 (7.7%)	0

[a]Numbers and percentages (in parentheses) of charged (Asp, Glu, Lys, and Arg), acidic (Asp and Glu), basic (Lys and Arg), polar uncharged (Gly, Ser, Thr, Asn, Gln, Tyr, Cys, and His), hydrophobic (Leu, Ile, Met, Val, Trp, Pro, Ala, and Phe), and aromatic (Phe, Tyr, and Trp) residues in protease sequences.

[b]Numbers of salt bridges and salt-bridge networks in crystal structures. Percentages shown in parentheses are proportion of the charged residues forming salt bridges/salt-bridge networks out of the total number of charged residues. A salt bridge is considered to exist if at least a pair of nitrogen and oxygen atoms in side-chain charged groups of two charged residues is within a 4-Å distance; a salt-bridge network is considered to exist if more than three residues participate in the formation of at least two salt bridges.

Table S3 Salt bridges identified in the crystal structures of the psychrophilic VPR, mesophilic RPK, and thermophilic AQN.

	VPR[a]	PRK[a]	AQN[a]
Common to all[b]	Arg12-Asp187	Arg12-Asp187	Arg12-Asp187
Common to VPR and AQN[c]	Asp58-Arg94		Asp58-Arg94
	Asp142-Arg173		Asp142-Arg173
Unique[d]	Arg16-Asp278	Arg12-Asp260	Asp17-Arg260
	Asp61-Arg94	Glu48-Arg80	Arg31-Glu239
	Arg189-Asp263	Glu50-Arg52	Arg43-Asp214b
	Glu240-Arg255	Lys94-Asp98	
		Asp112-Arg147	
		Asp117-Arg121	
		Asp184-Arg188	

[a]Salt bridges are indicated as their constituent residues and residue numbers. Residue numbering is according to the PRK sequence and structurally equivalent residue positions were determined from structure-based multiple sequence alignment (Fig. S1). Salt bridges highlighted by underline participate in the formation of a salt-bridge network.

[b]Salt bridges observed in all three crystal structures (absolutely conserved).

[c]Salt bridges observed in the crystal structures of VPR and AQN (relatively conserved).

[d]Salt bridges unique to the respective crystal structures (non-conserved).

Table S4 Relative solvent-accessible surface area (R-SASA) and average energy values of individual salt-bridge networks during MD simulations of VPR, PRK, and AQN at the three simulation temperatures.

Salt-bridge network[a]	Prot &Temp[b] (K)	R-SASA[c] (%)	$\Delta\Delta G_{\text{dslv-ntwk}}$[d] (kcal/mol)	$\Delta\Delta G_{\text{brd-ntwk}}$[e] (kcal/mol)	$\Delta\Delta G_{\text{prt-ntwk}}$[f] (kcal/mol)	$\Delta\Delta G_{\text{tot-ntwk}}$[g] (kcal/mol)
Asp58-Arg94-Asp61(c)	VPR 283	21.0	21.2 (1.0)	−6.5 (1.4)	−33.8 (4.2)	−19.2 (3.4)
Asp58-Arg94-Asp61(c)	VPR 300	20.3	22.7 (1.6)	−6.7 (1.1)	−38.1 (2.8)	−22.2 (2.2)
Asp58-Arg94-Asp61(c)	VPR 343	21.7	25.4 (1.7)	−9.2 (2.5)	−40.6 (5.8)	−24.4 (4.5)
Arg189-Asp263-Arg265	VPR 283	43.3	10.1 (2.3)	−2.2 (2.9)	−11.9 (2.5)	−4.0 (1.9)
Arg189-Asp263-Arg265	VPR 300	37.4	13.6 (2.3)	−6.3 (3.2)	−11.6 (1.8)	−4.4 (1.4)
Arg189-Asp263-Arg265	VPR 343	36.5	16.4 (2.3)	−9.2 (3.2)	−13.0 (2.1)	−5.7 (1.8)
Arg52-Glu50-Lys87	PRK 283	35.1	12.1 (2.4)	−6.4 (2.9)	−9.6 (1.1)	−3.9 (1.4)
Arg52-Glu50-Lys87	PRK 300	34.1	13.4 (2.9)	−7.6 (3.6)	−10.4 (1.4)	−4.6 (1.6)
Arg52-Glu50-Lys87	PRK 343	33.1	16.3 (3.2)	−10.0 (4.2)	−12.7 (1.7)	−6.4 (2.0)
Asp65-Lys94-Asp98	PRK 283	12.6	30.1 (1.8)	−10.8 (1.9)	−47.8 (2.5)	−28.5 (2.3)
Asp65-Lys94-Asp98	PRK 300	21.6	24.1 (5.8)	−9.6 (5.8)	−29.5 (7.5)	−15.0 (5.9)
Asp65-Lys94-Asp98	PRK 343	15.4	33.4 (4.1)	−12.3 (3.8)	−51.5 (8.0)	−30.4 (6.0)
Asp187-Arg12-Asp260(c)	PRK 283	19.6	19.4 (2.0)	−7.3 (1.3)	−29.7 (2.4)	−17.5 (2.3)
Asp187-Arg12-Asp260(c)	PRK 300	20.8	19.3 (2.0)	−7.4 (1.4)	−30.2 (2.4)	−18.4 (2.1)
Asp187-Arg12-Asp260(c)	PRK 343	19.8	23.1 (2.9)	−9.1 (1.7)	−37.5 (3.3)	−23.5 (2.5)
Arg43-Asp214b-Arg47	AQN 283	44.1	10.7 (2.0)	−6.7 (2.4)	−4.6 (1.5)	−0.6 (1.4)
Arg43-Asp214b-Arg47	AQN 300	48.9	9.6 (3.1)	−5.7 (3.7)	−4.2 (1.6)	−0.2 (2.0)
Arg43-Asp214b-Arg47	AQN 343	50.4	9.1 (3.3)	−4.0 (3.8)	−5.5 (2.7)	−0.4 (2.5)
Asp58-Arg94-Asp97	AQN 283	12.7	21.5 (4.1)	−7.5 (2.5)	−28.4 (6.2)	−14.5 (4.1)
Asp58-Arg94-Asp97	AQN 300	17.2	23.8 (2.2)	−9.0 (1.0)	−29.5 (6.6)	−14.7 (4.6)
Asp58-Arg94-Asp97	AQN 343	18.6	25.8 (2.7)	−10.9 (1.9)	−29.8 (7.3)	−15.0 (5.0)

[a]Crystal salt-bridge networks retained during MD simulations are followed by 'c' in parentheses.

[b]Protein and MD simulation temperature.

[c]R-SASA was calculated as the average of relative solvent-accessible surface areas of the network-participating residues.

[d-g]Average values of the desolvation energy penalty, bridge energy term, protein energy term, and total electrostatic free energy (or electrostatic strength) of a salt-bridge network, respectively. SD is in parentheses.

Table S5 Average values of solvent-accessible surface area (SASA) and energy terms of individual calcium ions during MD simulations of VPR, PRK, and AQN at the three simulation temperatures.

No.	Prot & Temp[a] (K)	Ca^{2+}-binding site	SASA (Å^2)	$\Delta\Delta G_{dslv\text{-}ca}$[b] (kcal/mol)	$\Delta\Delta G_{prt\text{-}ca}$[c] (kcal/mol)	$\Delta\Delta G_{tot\text{-}ca}$[d] (kcal/mol)
Ca1	VPR 283	P175, G177, D200	10.5 (3.8)	28.6 (3.2)	−32.7 (3.5)	−4.1 (1.9)
Ca1	VPR 300	P175, G177, D200	10.8 (4.1)	30.1 (3.5)	−35.5 (3.8)	−5.4 (2.1)
Ca1	VPR 343	P175, G177, D200	14.2 (8.3)	32.3 (7.2)	−39.6 (7.8)	−7.3 (2.6)
Ca2	VPR 283	D58, D61b, D62	4.9 (3.3)	42.4 (2.9)	−68.5 (5.2)	−26.1 (3.1)
Ca2	VPR 300	D58, D61b, D62	3.7 (2.6)	45.9 (2.2)	−74.5 (4.3)	−28.6 (3.3)
Ca2	VPR 343	D58, D61b, D62	4.7 (3.7)	51.6 (4.8)	−85.6 (8.4)	−34.1 (4.5)
Ca3	VPR 283	D11, D14, Q15, D20,	0.2 (0.8)	59.6 (5.3)	−91.4 (9.1)	−31.8 (4.6)
Ca3	VPR 300	D11, D14, Q15, D20,	0.2 (0.9)	64.7 (6.0)	−100.8 (9.5)	−36.1 (4.7)
Ca3	VPR 343	D11, D14, Q15, D20,	0.5 (1.4)	72.8 (8.3)	−118.1	−45.3 (5.9)
Ca1	PRK 283	P175, V177, D200	11.2 (3.5)	31.9 (2.0)	−35.8 (2.6)	−3.9 (1.7)
Ca1	PRK 300	P175, V177, D200	14.6 (8.2)	30.3 (6.9)	−34.9 (7.0)	−4.6 (1.9)
Ca1	PRK 343	P175, V177, D200	12.4 (7.7)	37.9 (6.8)	−45.2 (7.4)	−7.4 (2.3)
Ca2	PRK 283	T16, D260	33.5 (4.5)	19.5 (1.5)	−27.1 (2.0)	−7.6 (1.3)
Ca2	PRK 300	T16, D260	35.7 (4.9)	19.6 (1.9)	−27.8 (2.1)	−8.2 (1.3)
Ca2	PRK 343	T16, D260	36.3 (5.0)	22.6 (2.1)	−34.0 (2.6)	−11.4 (1.7)
Ca1	AQN 283	D11, D14, Q15, S20, S22	1.3 (1.9)	54.4 (3.0)	−73.0 (4.1)	−18.5 (2.8)
Ca1	AQN 300	D11, D14, Q15, S20, S22	2.1 (2.3)	56.3 (4.5)	−77.0 (5.7)	−20.7 (3.3)
Ca1	AQN 343	D11, D14, Q15, S20, S22	2.5 (3.0)	63.2 (6.4)	−90.7 (8.4)	−27.4 (4.6)
Ca2	AQN 283	V174, A177, T179, D200	13.4 (5.5)	25.0 (3.1)	−29.6 (3.9)	−4.6 (2.1)
Ca2	AQN 300	V174, A177, T179, D200	12.5 (3.9)	28.0 (2.3)	−34.2 (2.9)	−6.3 (1.9)
Ca2	AQN 343	V174, A177, T179, D200	13.8 (5.1)	31.7 (3.6)	−39.5 (5.5)	−7.7 (3.3)

[a]Protein and simulation temperature

[b-d]Average values of the desolvation energy penalty, protein energy term, and total electrostatic free energy (or electrostatic strength) of a calcium ion, respectively. SD is in parentheses.

Table S6 Comparison of electrostatic free energy contributions to protein stability by different types of electrostatic interactions among VPR, PRK, and AQN at the three simulation temperatures.

Prot & Temp[a] (K)	Salt bridge			Salt-bridge network			Calcium ion		
	No.[b]	Average[c] (kcal/mol)	Total[d] (kcal/mol)	No.[b]	Average[c] (kcal/mol)	Total[d] (kcal/mol)	No.[b]	Average[c] (kcal/mol)	Total[d] (kcal/mol)
VPR 283	9	–3.0 (3.7)	–27.0	2	–11.6 (7.6)	–23.2	3	–20.7 (11.9)	–62.0
VPR 300	9	–3.8 (3.9)	–34.2	2	–13.3 (8.9)	–26.6	3	–23.4 (13.1)	–70.1
VPR 343	10	–4.7 (4.2)	–47.0	2	–15.1 (9.4)	–30.1	3	–28.9 (15.9)	–86.7
PRK 283	12	–3.2 (2.7)	–38.4	3	–16.6 (10.1)	–49.9	2	–5.8 (1.9)	–11.5
PRK 300	14	–3.9 (4.4)	–54.6	3	–12.8 (5.9)	–38.5	2	–6.4 (1.8)	–12.8
PRK 343	12	–4.6 (3.7)	–55.5	3	–20.1 (10.1)	–60.3	2	–9.4 (2.0)	–19.8
AQN 283	8	–2.8 (3.0)	–22.4	2	–7.6 (7.0)	–15.1	2	–11.6 (7.0)	–23.1
AQN 300	9	–3.4 (3.5)	–30.6	2	–7.5 (7.3)	–14.9	2	–13.5 (7.2)	–27.0
AQN 343	10	–4.2 (3.4)	–42.0	2	–7.7 (7.3)	–15.4	2	–17.6 (9.9)	–35.1

[a]Protein and simulation temperature.

[b]Number of salt bridges, salt-bridge networks, or calcium ions.

[c]Average value of electrostatic free energy contribution (or electrostatic strength). SD is in parentheses.

[d]Total value (or cumulative sum) of electrostatic free energy contribution.

Table S7 Comparison of flexibility of structural regions of interest among VPR, PRK and AQN.

Structural region	C_α RMSF of VPR[a] (Å)			C_α RMSF of PRK[a] (Å)			C_α RMSF of AQN[a] (Å)		
	283 K	300 K	343 K	283 K	300 K	343 K	283 K	300 K	343 K
All[b]	0.70 (0.70)	0.77 (0.82)	0.97 (1.04)	0.52 (0.29)	0.61 (0.37)	0.68 (0.36)	0.51 (0.29)	0.56 (0.36)	0.64 (0.42)
Catalytic triad[c]	0.50 (0.03)	0.53 (0.10)	0.65 (0.13)	0.42 (0.03)	0.51 (0.12)	0.55 (0.11)	0.49 (0.05)	0.43 (0.11)	0.52 (0.11)
S1 site[d]	0.72 (0.43)	0.74 (0.41)	0.84 (0.39)	0.60 (0.34)	0.67 (0.36)	0.76 (0.40)	0.55 (0.18)	0.64 (0.31)	0.68 (0.39)
S2 site[e]	0.88 (0.63)	1.10 (0.98)	1.26 (1.02)	0.80 (0.79)	0.85 (0.46)	0.78 (0.39)	0.73 (0.47)	0.92 (0.83)	1.11 (0.98)
S2-loop[f]	1.15 (0.57)	1.58 (0.92)	1.68 (0.96)	1.11 (0.71)	1.17 (0.49)	0.98 (0.37)	1.13 (0.49)	1.31 (0.76)	1.54 (0.91)
PSL[g]	0.68 (0.12)	0.81 (0.14)	0.95 (0.14)	0.58 (0.09)	1.02 (0.28)	0.89 (0.16)	0.83 (0.30)	0.82 (0.17)	0.98 (0.18)

[a]C_α root mean square fluctuation (RMSF) value of each residue was calculated over the concatenated equilibrium MD trajectories of the three proteases at the three simulation temperatures, and then was averaged over the residues in the structural region of interest (SD is in parentheses).

[b]All residues.

[c]Residues D39, H69 and S224.

[d]Residues 132-135, 158-161, 162, 169, and 222-225.

[e]Residues 39, 40, 67, 69, 96, 100.

[f]Residues 95-101.

[g]Surface polar loop (PSL) is composed of residues 58-68.