

**Doctoral Thesis**

**Protein thermostabilization strategy by  
ion–ion interactions at temperatures of  
over 100 °C**

Yoshinori Matsuura

2020

# Contents

Abbreviations .....	III
Chapter 1: General introduction .....	4
1-1. Protein stability.....	4
1-2. CutA1 proteins. ....	5
1-3. Relationship of the CutA1 proteins and the CvP-bias. ....	7
1-4. Organization of this thesis. ....	8
1-5. References. ....	9
Chapter 2: Role of charged residues in stabilization of <i>Pyrococcus horikoshii</i> CutA1	14
2-1. Introduction. ....	14
2-2. Experimental methods. ....	15
2-3. Results. ....	17
2-4. Discussion. ....	19
2-5. Conclusions. ....	30
2-6. Supporting information. ....	38
2-7. References. ....	49
Chapter 3: Evaluation of salt bridges in CutA1 using molecular dynamic simulations.	52
3-1. Introduction. ....	52
3-2. Experimental methods. ....	53
3-3. Results and discussion. ....	55
3-4. Concluding remarks. ....	71
3-5. Supporting information. ....	83
3-6. References. ....	109
Chapter 4: Stabilization of <i>Escherichia coli</i> CutA1 by rational protein design .....	113
4-1. Introduction. ....	113
4-2. Experimental methods. ....	114
4-3. Results and discussion. ....	119
4-4. Conclusions. ....	128
4-5. References. ....	138
Chapter 5: Stabilization of <i>Escherichia coli</i> CutA1 by introduction of charged residues	140
5-1. Introduction. ....	140
5-2. Experimental methods. ....	141
5-3. Results. ....	142

5-4. Discussion. ....	146
5-5. Supporting information. ....	157
5-6. References. ....	163
Chapter 6: Stabilization of <i>Escherichia coli</i> CutA1 by ion–ion interactions .....	167
6-1. Introduction. ....	167
6-2. Experimental methods. ....	168
6-3. Results. ....	171
6-4. Discussion. ....	173
6-5. Supporting information. ....	183
6-6. References. ....	197
Chapter 7: Conclusions.....	201
List of publications .....	203
Acknowledgements .....	204

## Abbreviations

<i>E. coli</i>	<i>Escherichia coli</i>
<i>P. horikoshii</i>	<i>Pyrococcus horikoshii</i>
EcCutA1	CutA1 protein from an mesophile, <i>Escherichia coli</i>
PhCutA1	CutA1 from hyperthermophile, <i>Pyrococcus horikoshii</i>
TtCutA1	CutA1 from thermophile, <i>Thermus thermophilus</i>
OsCutA1	CutA1 from mesophile, <i>Oryza sativa</i>
HsCutA1	CutA1 from <i>Homo sapiens</i>
$T_d$	denaturation temperature
DSC	differential scanning calorimetry
PDB	protein data bank
EDTA	ethylene diamine tetra acetic acid
SDS-PAGE	sodium dodecyl sulfate-poly-acrylamide gel electrophoresis
$pI$	isoelectric point
RMSD	root mean square deviation
RMSF	root mean square fluctuations
$R_g$	radius of gyration
MD	molecular dynamics
GROMACS	groningen machine for chemical simulations
Amber	assisted model building and energy refinement
Charmm	chemistry at harvard macromolecular mechanics
Gromos	groningen molecular simulation
spc/e	simple point charge/extended
tip3p	transferable intermolecular potential 3p
PCR	polymerase chain reaction
SPMP	stability profile of mutant protein

## **Chapter 1: General introduction.**

### ***1-1. Protein stability.***

The stability of globular proteins is called “marginal stability” because of the low conformational stability, which is delicately balanced by a combination of many stabilizing and destabilizing interactions, such as hydrophobic interactions, hydrogen bonds, electrostatic interactions, and entropic effects<sup>1, 2</sup>. The denaturation Gibbs energy ( $\Delta G$ ) of globular proteins is approximately 50 kJ/mol, which equals energy of only a few hydrogen bonds. Hence, only single mutations can drastically change the stability of the protein. Two types of mutagenesis are used for protein engineering to improve the stability of the protein. One is random mutagenesis, known as “directed evolution of enzymes”, research that was awarded the Nobel Prize in Chemistry in 2018. The other is site specific mutagenesis. Recently, much of the tertiary structure of the protein has been determined by various methods, such as X-ray crystallography, nuclear magnetic resonance, and electron microscopy. Therefore, these data can be applied to various types of research, including thermostabilization of the protein.

In nature, there are organisms that live in various temperature ranges. The proteins isolated from hyperthermophiles, which grow preferentially at temperatures near the boiling point of water, are extremely stable as compared with those isolated from mesophiles. The structures of many proteins from hyperthermophiles have been determined, and several factors responsible for their extreme thermostability have been proposed, including increases in the number of ion pairs<sup>3-10</sup> and hydrogen bonds<sup>11, 12</sup>, core

hydrophobicity<sup>13</sup>, and packing density<sup>14</sup>, as well as the oligomerization of several subunits<sup>15-17</sup>, and an entropic effect due to relatively shorter surface loops and peptide chains<sup>18</sup>. However, there are no established methods to rationally design enhanced conformational stability. However, it is important to develop these techniques because stability-enhanced proteins may be highly useful for industrial and biotechnological processes<sup>19</sup>.

### ***1-2. CutA1 proteins.***

In 2006, it was reported that the CutA1 protein (*PhCutA1*) from the hyperthermophile, *Pyrococcus horikoshii*, had an unusually high stability at pH 7.0 with a denaturation temperature ( $T_d$ ) of nearly 150 °C, which is approximately 30 °C greater than the highest record determined by differential scanning calorimetry (DSC)<sup>9</sup>. The CutA1 protein was originally identified as the product of the *cutA* gene locus of *Escherichia coli*, and it is involved in divalent metal tolerance<sup>20</sup>.

The X-ray structures and stabilities of CutA1 proteins from species with various growth temperatures have been also examined, including those of *Pyrococcus horikoshii* (*PhCutA1*)<sup>21</sup>, *Thermus thermophilus* (*TtCutA1*)<sup>21</sup>, *Oryza sativa* (*OsCutA1*), *Homo sapiens* (brain) (*HsCutA1*)<sup>22</sup>, and *Escherichia coli* (*EcCutA1*)<sup>23</sup>. Structure analysis of the CutA1 family demonstrated that each CutA1 protein is quite similar.

A monomer structure of about 12 kDa consists of three  $\alpha$ -helices and five  $\beta$ -strands. Three monomers are assembled into a trimer through interactions between the edges of three  $\beta$ -strands. This tightly intertwined interaction contributes to the stabilization of the trimer structures for the CutA1 proteins<sup>9</sup>. Their  $T_d$  values are also unusually high relative

to the growth temperatures of each species: 113.9 °C for *Tt*CutA1<sup>24</sup>, 98.9 °C for *Os*CutA1<sup>24</sup>, 96.2 °C for *Hs*CutA1<sup>25</sup>, and 89.9 °C for *Ec*CutA1<sup>26</sup>.

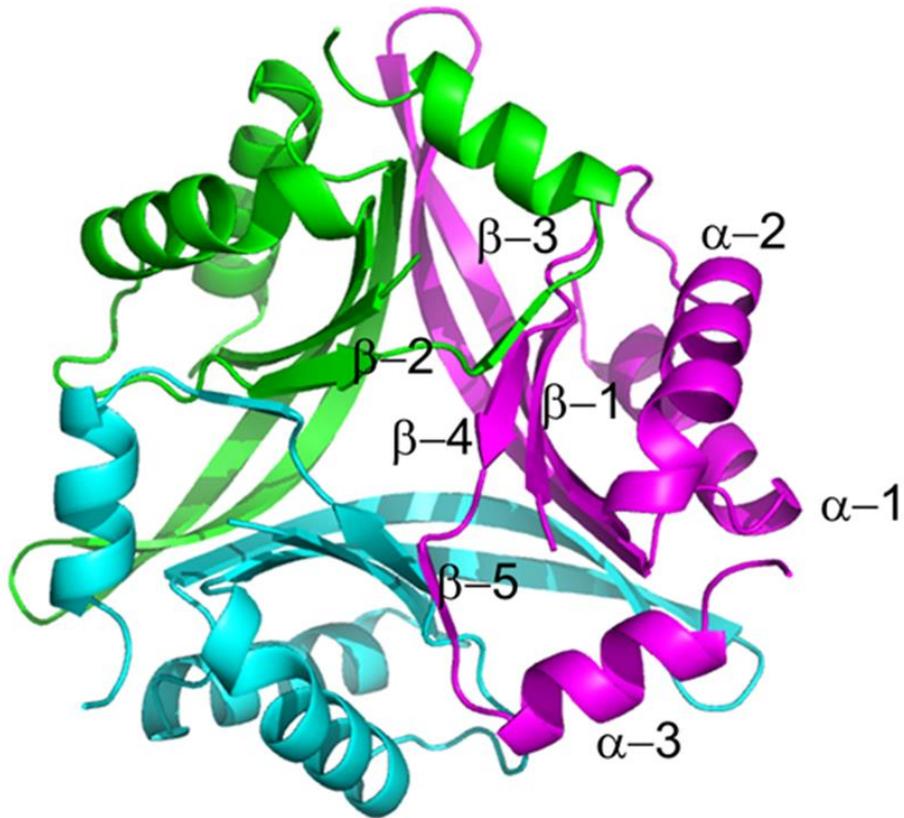


Fig. 1. The trimer structure of *Ph*CutA1 (A, B, and C subunits of PDB ID 4noy). Different colors represent different chains.  $\alpha$  and  $\beta$  represent  $\alpha$ -helix and  $\beta$ -strand, respectively.

### ***1-3. Relationship of the CutA1 proteins and the CvP-bias.***

Genomic analysis performed by Suhre and Claverie showed that a large difference between the proportions of the numbers of charged residues (Asp, Glu, Lys, Arg) and polar residues (Asn, Gln, Ser, Thr), which is called *CvP*-bias, is the dominant proteome characteristic of microorganisms adapted to hyperthermophilic growth<sup>27</sup>. Many reports have shown that an abundance of ion pairs and ionic networks formed by charged residues contributes to the stabilization of proteins from hyperthermophiles<sup>28-38</sup>. Negative contributions of salt bridges to protein stability have also been reported<sup>39-40</sup>. However, how the abundance of ion pairs contributes to the conformational stability of proteins from hyperthermophiles remains unclear.

Although the X-ray crystal structure of *PhCutA1* clearly resembles that of *EcCutA1*, the amino acid compositions of the two proteins are quite different. The difference is especially apparent in the proportions of charged and polar residues for the *CvP*-bias: 40.2% (8 Asp, 16 Glu, 11 Lys and 6 Arg in 102 residues) and 7.8% (1 Ser, 5 Thr, 2 Asn, and 0 Gln) in *PhCutA1*, respectively; 18.8% and 23.2% in *EcCutA1*, respectively. Thus, the *CvP*-bias values of *PhCutA1* and *EcCutA1* are 32.4% and -4.4%, respectively. The *CvP*-bias of *PhCutA1* is more than double that of the average *CvP*-bias (14.7%) of proteins from *P. horikoshii*<sup>27</sup>, suggesting that the content of charged residues in *PhCutA1* is unusually large even among proteins from hyperthermophiles. Therefore, *PhCutA1* may represent a good model for investigating the role of charged residues in the stability of proteins at temperatures greater than 100 °C.

#### ***1-4. Organization of this thesis.***

This thesis consists of five chapters as described below.

In Chapter 2, to elucidate the role of ion-ion interactions in protein stability, the mutant proteins of *PhCutA1*, in which charged residues were substituted by noncharged residues, were comprehensively examined. In Chapter 3, we compared the strengths of salt bridges among force fields by performing MD simulations using *PhCutA1*. In Chapter 4, to enhance the heat stability of the *EcCutA1* protein to that of the *PhCutA1* with a denaturation temperature ( $T_d$ ) of 150 °C, we first examined the structure-sequence (3D-1D) compatibility between the conformation of *EcCutA1* and its native sequence using the stability profile of mutant protein (SPMP). In Chapter 5, to experimentally obtain thermodynamic parameters of protein denaturation at temperatures over 100 °C, we designed certain hydrophobic mutant proteins of *EcCutA1*, which have denaturation temperatures ( $T_d$ ) of 101-113 °C and show a reversible heat denaturation. Furthermore, hyperthermostable mutant proteins ( $T_d = 137$  °C) were also evaluated by substituting six residues with charged residues. In Chapter 6, to elucidate the contribution of charged residues to protein stabilization at temperatures of over 100 °C, we constructed more than 100 *EcCutA1* mutants. The goal was to determine if one can achieve the same stability as that of *CutA1* from the hyperthermophile *Pyrococcus horikoshii*, which has a denaturation temperature near 150 °C.

### ***1-5. References.***

1. Pace, C. N., Shirley, B. A., McNutt, M., and Gajiwala, K. (1996) Forces contributing to the conformational stability of proteins. *FASEB J.* **10**, 75-832.
2. Privalov, P. L. and Gill, S. J. (1988) Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem.* **39**, 191-234.
3. Perutz, M. F. and Raidt, H. (1975) Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2, *Nature* **255**, 256-259.
4. Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M., and Pearl, L. H. (1997) Crystal structure of the  $\beta$ -glycosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*: resilience as a key factor in thermostability, *J. Mol. Biol.* **271**, 789-802.
5. Tahirov, T. H., Oki, H., Tsukihara, T., Ogasahara, K., Yutani, K., Ogata, K., Izu, Y., Tsunasawa, S., and Kato, I. (1998) Crystal structure of methionine aminopeptidase from hyperthermophile, *Pyrococcus furiosus*. *J. Mol. Biol.* **284**, 101-124.
6. Elcock, A. H. (1998) The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins, *J. Mol. Biol.* **284**, 489-502.
7. Yamagata, Y., Ogasahara, K., Hioki, Y., Lee, S. J., Nakagawa, A., Nakamura, M., Ishida, M., Kuramitsu, S., and Yutani, K. (2001) Entropic stabilization of the tryptophan synthase  $\alpha$ -subunit from a hyperthermophile, *Pyrococcus furiosus*. X-ray analysis and calorimetry, *J. Biol. Chem.* **276**, 11062-11071.
8. Vieille, C. and Zeikus, G. J. (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* **65**, 1-43.
9. Tanaka, T., Sawano, M., Ogasahara, K., Sakaguchi, Y., Bagautdinov, B., Katoh, E., Kuroishi, C., Shinkai, A., Yokoyama, S., and Yutani, K. (2006) Hyperthermostability of CutA1 protein, with a denaturation temperature of nearly 150 °C. *FEBS Lett.* **580**, 4224-4230.
10. Matsuura, Y., Takehira, M., Sawano, M., Ogasahara, K., Tanaka, T., Yamamoto, H., Kunishima, N., Katoh, E., and Yutani, K. (2012) Role of charged residues in stabilization of *Pyrococcus horikoshii* CutA1, which has a denaturation temperature of nearly 150 °C, *FEBS J.* **279**, 78-90.
11. Tanner, J. J., Hecht, R. M., and Krause, K. L. (1996). Determinants of enzyme

- thermostability observed in the molecular structure of *Thermus aquaticus* D-glyceraldehyde-3-phosphate dehydrogenase at 2.5 Å resolution, *Biochemistry* **35**, 2597-2609.
12. Bush, J. and Makhatadze, G. I. (2011) Statistical analysis of protein structures suggests that buried ionizable residues in proteins are hydrogen bonded or form salt bridges. *Proteins* **79**, 2027-2032.
  13. Schumann, J., Bohm, G. Schumacher, G., Rudolph, R., and Jaenicke, R. (1993) Stabilization of creatinase from *Pseudomonas putida* by random mutagenesis. *Protein Sci.* **2**, 1612-1620.
  14. Russell, R. J. M., Hough, D. W., Danson, M. J., and Taylor, G. L. (1994) The crystal structure of citrate synthase from the thermophilic archaeon *Thermoplasma acidophilum*. *Structure* **2**, 1157-1167.
  15. Jaenicke, R., Schurig, H., Beaucamp, N., and Ostendorp, R. (1996) Structure and stability of hyperstable proteins: Glycolytic enzymes from hyperthermophilic bacterium *Thermotoga maritima*. *Adv. Protein Chem.* **48**, 181-269.
  16. Sterner, R., Kleemann, G. R., Szadkowski, H., Lustig, A., Hennig, M., and Kirschner, K. (1996) Phosphoribosyl anthranilate isomerase from *Thermotoga maritima* is an extremely stable and active homodimer. *Protein Sci.* **5**, 2000-2008.
  17. Dams, T. and Jaenicke, R. (1999) Stability and folding of dihydrofolate reductase from the hyperthermophilic bacterium *Thermotoga maritima*. *Biochemistry* **38**, 9169-9178.
  18. Yamagata, Y., Ogasahara, K., Hioki, Y., Lee, S. J., Nakagawa, A., Nakamura, H., Ishida, M., Kuramitsu, S., and Yutani, K. (2001) Entropic stabilization of the tryptophan synthase  $\alpha$ -subunit from a hyperthermophile, *Pyrococcus furiosus*. X-ray analysis and calorimetry. *J. Biol. Chem.* **276**, 11062-11071.
  19. Kirk., O., Borchert, T. V., and Fuglsang, C. C. (2002) Industrial enzyme applications. *Curr. Opin. Biotechnol.* **13**, 345-351.
  20. Fong, S.T., Camakaris, J., and Lee, B.T. (1995) Molecular genetics of a chromosomal locus involved in copper tolerance in *Escherichia coli* K-12. *Mol. Microbiol.* **15**, 1127-1137.

21. Bagautdinov, B. (2014) The structures of the CutA1 proteins from *Thermus thermophilus* and *Pyrococcus horikoshii*: characterization of metal-binding sites and metal-induced assembly. *Acta Crystallogr. F Struct. Biol. Commun.* **70**, 404-413.
22. Bagautdinov, B., Matsuura Y, Bagautdinova S, Kunishima N, and Yutani K. (2008) Structure of putative CutA1 from Homo sapiens determined at 2.05 Å resolution. *Acta Crystallogr. F Struct. Biol. Commun.* **64**, 351-357.
23. Arnesano, F., Banci, L., Benvenuti, M., Bertini, I., Calderone, V., Mangani, S., and Viezzoli, M. S. (2003) The evolutionarily conserved trimeric structure of CutA1 proteins suggests a role in signal transduction. *J. Biol. Chem.* **278**, 45999-46006.
24. Sawano, M., Yamamoto, H., Ogasahara, K., Kidokoro, S., Katoh, S., Ohnuma, T., Katoh, E., Yokoyama, S., and Yutani, K. (2008) Thermodynamic basis for the stabilities of three CutA1s from *Pyrococcus horikoshii*, *Thermus thermophilus*, and *Oryza sativa*, with unusually high denaturation temperatures. *Biochemistry* **47**, 721-730.
25. Bagautdinov, B., Matsuura, Y., Yamamoto, H., Sawano, M., Ogasahara, K., Takehira, M., Kunishima N., Katoh, E., and Yutani, K. (2015) Thermodynamic analysis of unusually thermostable CutA1 protein from human brain and its protease susceptibility. *J. Biochem.* **157**, 169-176.
26. Matsuura, Y., Ota, M., Tanaka, T., Takehira, M., Ogasahara, K., Bagautdinov, B., Kunishima, N., and Yutani, K. (2010) Remarkable improvement in the heat stability of CutA1 from *Escherichia coli* by rational protein design. *J. Biochem.* **148**, 449-458.
27. Suhre K and Claverie J.-M. (2003) Genomic correlates of hyperthermostability, an update. *J. Biol. Chem.* **278**, 17198-171202.
28. Korndörfer, I., Steipe, B., Huber, R., Tomschy, A., and Jaenicke, R. (1995) The Crystal Structure of Holo-glyceraldehyde-3-phosphate Dehydrogenase from the Hyperthermophilic Bacterium *Thermotoga maritime* at 2.5 Å Resolution. *J. Mol. Biol.* **246**, 511-521.
29. Hennig, M., Darimont, B., Sterner, R., Kirschner, K., and Jansonius, J. N. (1995) 2.0 Å structure of indole-3-glycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus*: possible determinants of protein stability. *Structure* **3**, 1295-1306.

30. Ermler, U., Merckel, M. C., Thauer, R. K., and Shima, S. (1997) Formylmethanofuran: tetrahydromethanopterin formyltransferase from *Methanopyrus kandleri* - new insights into salt-dependence and thermostability. *Structure* **5**, 635-646.
31. Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M., and Pearl, L. H. (1997) Crystal structure of the  $\beta$ -glycosidase from the hyperthermophilic archeon *Sulfolobus solfataricus*: resilience as a key factor in thermostability. *J. Mol. Biol.* **271**, 789-802.
32. Ogasahara, K., Lapshina, Sakai, M., Izu, Y., Tsunasawa, S., Kato, I., and Yutani, K., (1998) Electrostatic stabilization in methionine aminopeptidase from hyperthermophile *Pyrococcus furiosus*. *Biochemistry* **37**, 5939-5946.
33. Vetriani, C., Maeder, D. L., Tolliday, N., Yip, K. S.-P., Stillman, T. J., Britton, K. L., Rice, D. W., Klump, H. H., and Robb, F. T. (1998) Protein thermostability above 100 °C: a key role for ionic interactions. *Proc. Natl. Acad. Sci. USA* **95**, 12300-12305.
34. Tahirov, T. H., Oki, H., Tsukihara, T., Ogasahara, K., Yutani, K., Ogata, K., Izu, Y., Tsunasawa, S., and Kato, I. (1998) Crystal structure of methionine aminopeptidase from hyperthermophile, *Pyrococcus furiosus*. *J. Mol. Biol.* **284**, 101-124.
35. de Bakker, P. I. W., Hünenberger, P. H., and McCammon, J. A. (1999) Molecular dynamics simulations of the hyperthermophilic protein Sac7d from *Sulfolobus acidocaldarius*: contribution of salt bridges to thermostability. *J. Mol. Biol.* **285**, 1811-1830.
36. Karshikoff, A., and Ladenstein, R. (2001) Ion pairs and the thermotolerance of proteins from hyperthermophiles: a "traffic rule" for hot roads. *Trends Biochem. Sci.* **26**, 550-556.
37. Yamagata, Y., Ogasahara, K., Hioki, Y., Lee, S. J., Nakagawa, A., Nakamura, H., Ishida, M., Kuramitsu, S., and Yutani, K. (2001) Entropic stabilization of the tryptophan Synthase  $\alpha$ -Subunit from a hyperthermophile, *Pyrococcus furiosus*: x-ray analysis and calorimetry. *J. Biol. Chem.* **276**, 11062-11071.
38. Robinson-Rechavi, M., Alibés, A., and Godzik, A. (2006) Contribution of electrostatic interactions, compactness and quaternary structure to protein thermostability: Lessons from structural genomics of *Thermotoga maritima*. *J. Mol. Biol.* **356** 547-557.

39. Honig, B., and Yang, A.-S. (1995) Free energy balance in protein folding. *Adv. Protein Chem.* **46**, 27-58.
40. Waldburger, C. D., Schildbach, J. F., and Sauer, R. T. (1995) Are buried salt bridges important for protein stability and conformational specificity? *Nat. Struct. Biol.* **2**, 122-128.

## Chapter 2: Role of charged residues in stabilization of *Pyrococcus horikoshii* CutA1

### 2-1. Introduction.

In this study, using *Ph*CutA1, we performed systematic and exhaustive analysis of the stability of mutant proteins by DSC to elucidate the role of ion–ion interactions in protein stability. The mutant proteins were constructed with the following in mind. (a) Previous analysis has shown that the increased number of ion pairs in the monomeric structure of *Ph*CutA1 contributes to the stabilization of the trimeric structure<sup>1</sup>. Therefore, Glu and Asp residues forming ion pairs within 4 Å in the same monomer were substituted to a single amino acid (Glu to Ala or Gln, Asp to Ala or Asn). (b) To confirm the contribution of charged residues to protein stability, other charged residues, which have relatively higher orders of favorable (or un-favorable) electrostatic energy (Table 2), were substituted to a single non-charged residue. The electrostatic energy of ion–ion interactions was evaluated by FoldX (<http://foldx.crg.es/>)<sup>2</sup>. (c) To confirm the contribution of the charged residues examined in single mutant proteins, double or multiple mutant proteins were constructed on the basis of the results obtained with single mutant proteins. From the changes in stability of more than 50 *Ph*CutA1 mutants, we discuss the role of ion–ion interactions in hyperthermophile proteins in conformational stability at temperatures greater than 100 °C.

## **2-2. Experimental methods.**

### **Construction, Expression, and Purification of Mutant Proteins.**

Mutant versions of *PhCutA1* were constructed by site-directed mutagenesis<sup>3</sup>, expressed, and purified as described previously<sup>1</sup>. All purified mutant proteins ran as single bands in SDS-PAGE. Protein concentrations were determined using an absorption coefficient of  $E^{1\%}_{1\text{cm}} = 27.1$  at 280 nm, which is based on number of aromatic amino acids<sup>4</sup>. All chemical reagents were of analytical grade.

### **Measurement of thermal stability.**

Thermal denaturation of proteins was measured using VP-DSC/ETR microcalorimeter (Microcal LLC, Northampton, MA, USA) for temperatures up to 150 °C or Nano-DSC 6300Y microcalorimeter (TA Instruments, USA) for temperatures up to 160 °C. Protein samples were dialyzed for more than 20 h against 50 mM potassium phosphate (pH 7.0) including 2 mM EDTA. Samples were filtered through 0.22- $\mu\text{m}$  pore membranes following dialysis and were degassed before measurements. Protein concentrations were typically around 1.0 mg/ml. The heating rate (scan rate) of DSC measurements was 60 °C/h.

**Estimation of unfolding Gibbs energy and electrostatic energy due to ion–ion interactions from the tertiary structure of a protein using FoldX.**

The computer algorithm, FoldX<sup>2</sup>, can quantitatively estimate the factors contributing to protein stability and protein interactions. FoldX is available via a web-interface at <http://foldx.crg.es/>. The FoldX energy function at pH 7.0 includes stabilization factors that are important for protein stability. The difference in unfolding Gibbs energy ( $\Delta G$ ) between wild-type and mutant proteins ( $\Delta\Delta G_{\text{FoldX}}$ ) is evaluated on the basis of the following factors by FoldX.

$$\begin{aligned} \Delta\Delta G_{\text{FoldX}} = & \Delta\Delta G_{\text{BH}} + \Delta\Delta G_{\text{SH}} + \Delta\Delta G_{\text{VW}} + \Delta\Delta G_{\text{El}} + \Delta\Delta G_{\text{Sol,P}} + \Delta\Delta G_{\text{Sol,H}} + \Delta\Delta G_{\text{VW,cla}} \\ & + \Delta\Delta G_{\text{Ent,s}} + \Delta\Delta G_{\text{Ent,m}} + \Delta\Delta G_{\text{Tor,cla}} + \Delta\Delta G_{\text{Bac,cla}} + \Delta\Delta G_{\text{Hd}} + \Delta\Delta G_{\text{El,sub}} \end{aligned} \quad (1),$$

where  $\Delta\Delta G_{\text{BH}}$ ,  $\Delta\Delta G_{\text{SH}}$ ,  $\Delta\Delta G_{\text{VW}}$ ,  $\Delta\Delta G_{\text{El}}$ ,  $\Delta\Delta G_{\text{Sol,P}}$ ,  $\Delta\Delta G_{\text{Sol,H}}$ ,  $\Delta\Delta G_{\text{VW,cla}}$ ,  $\Delta\Delta G_{\text{Ent,s}}$ ,  $\Delta\Delta G_{\text{Ent,m}}$ ,  $\Delta\Delta G_{\text{Tor,cla}}$ ,  $\Delta\Delta G_{\text{Bac,cla}}$ ,  $\Delta\Delta G_{\text{Hd}}$ , and  $\Delta\Delta G_{\text{El,sub}}$  represent the energies of the backbone hydrogen bonds (BH), side-chain hydrogen bonds (SH), Van der Waals interactions (VW), electrostatic interactions (El), solvation of polar atoms (Sol,P), solvation of hydrophobic atoms (Sol,H), Van der Waals clashes (VW,cla), entropy of the side chains (Ent,s), entropy of the main chain (Ent,m), torsional clashes (Tor,cla), backbone clashes (Bac,cla), helix dipoles (Hd), and electrostatic interactions between the subunits (El,sub), respectively. FoldX recommends that the RepairPDB command should be run prior to energy calculation, because the crystal structures are models based on electron density and have errors produced during refinement that will result in nonstandard angles or distances.

In the reported structure of wild-type *PhCutA1*, Cys29 was replaced by cysteine-s-dioxide (PDB entry 4nyo). Therefore, we generated a standard structure of *PhCutA1* by modeling a replacement of cysteine-s-dioxide with cysteine for energy calculations of the mutant proteins. Models of this standard structure (wild-type protein) and the necessary

mutant structures were built using FoldX.  $\Delta\Delta G_{\text{FoldX}}$  values for the mutant proteins were generated using FoldX. The electrostatic energy of each residue between charged residues was sorted out from “AllAtoms\_Electro” file in FoldX. The electrostatic energies of the wild-type *PhCutA1* are shown in Table S1.

### **2-3. Results.**

#### **Design of *PhCutA1* mutants.**

All of the mutant variants of *PhCutA1* in this work are listed in Table 1. The variants were constructed with the following intentions. (a) Previous analysis has shown that the increased number of ion pairs in the monomeric structure of *PhCutA1* contributes to the stabilization of the trimeric structure<sup>1</sup>. Therefore, Glu and Asp residues (Glu15, Glu24, Glu47, Asp60, Glu63, Glu64, Glu67, Glu71, Asp84, and Glu99) having ion pairs within 4 Å on the same subunit of *PhCutA1* were substituted by Gln or Ala and Asn or Ala, respectively, in order to confirm the role of intra-subunit ion pairs on the same subunit in the stability of *PhCutA1*. (b) The electrostatic energies between the ion groups formed by each charged residue, evaluated by FoldX, are shown in Fig. S1 and are listed in order of favorable and unfavorable interactions (15 residues each) in Table 2. To elucidate how much the ion–ion interactions of charged residues contribute to protein stability, the charged residues listed in Table 2 were single-substituted by non-charged residues. (c) Double mutants of ion-paired residues and multiple mutants were also constructed (Table 1) in order to confirm the contribution of the charged residues as a pair to stability.

### The thermal stability of *PhCutA1* mutants.

To examine changes in stability due to mutations, the heat stability of mutant variants was measured using DSC. Typical DSC curves at pH 7.0 are shown in Fig. 1. As shown in the figure, the DSC curves seem to be slightly aggregated after heat denaturation; therefore, the apparent peak temperatures of the DSC curves were taken as the denaturation temperature ( $T_d$ ). The  $T_d$  values listed in Table 1 are averages of more than two experiments. The  $T_d$  value of E99Q was 154.9 °C, the highest among examined mutant proteins, and 6.4 °C greater than that of the wild-type protein. In fact, the E99Q mutant set a new world record for the protein with the highest heat stability<sup>1</sup>. The  $T_d$  values of 13 variants among the 45 single mutant proteins were higher than that of the wild-type protein, as shown in Table 1. These results suggest that even a protein with a  $T_d$  of 150 °C has many sites that could be altered to further increase its stability. The greatest decrease in  $T_d$  due to a single amino acid substitution was 12.4 °C for R25A. The *Ph14S* mutant is the *PhCutA1* variant E12Q/K16A/K44A/E46Q/D48N/K49A/R58A/D60N/E63Q/E67Q/E71Q/D84N/E90Q/K94A, which was substituted at all 14 sites of intra-subunit ion interactions within 4 Å, a pair of which is closely located in the sequence (Fig. 2). The  $\Delta T_d$  of *Ph14S* was  $-7.2$  °C.

The changes in  $T_d$  of Glu and Asp mutant variants (at Glu15, Glu24, Glu47, Asp60, Glu63, Glu64, Glu67, Glu71, Asp84, and Glu99) forming intra-subunit ion pairs on the same subunit of *PhCutA1* ranged from  $-3.5$  (E64Q and D60N) to 6.4 °C (E99Q) (Table 1). The average value of these 20 samples forming intra-subunit ion pairs on the same subunit was  $-1.0 \pm 1.5$  °C. However, the average changes in  $T_d$  of all 21 Glu mutants in Table 1 was  $0.03 \pm 2.05$  °C, while the average changes of 11 Asp and 13 positively

charged (Lys and Arg) mutants were  $-2.17 \pm 3.27$  °C and  $-5.14 \pm 3.38$ °C, respectively. These results suggest that the Glu residues of *PhCutA1* contribute less to protein stability than other ionic residues.

#### **2-4. Discussion.**

The total energies of ion–ion interactions for each charged residue evaluated by FoldX are shown in Fig. S1, and Table S1 lists all values of each pair of interactions. It might be important to elucidate how electrostatically favorable (or unfavorable) residues contribute to protein stability. For example, *PhCutA1* has 30 intra-subunit ion pairs within 5 Å whereas *EcCutA1* has only one. By contrast, *PhCutA1* has 16 inter-subunit ion pairs within 5 Å, whereas *EcCutA1* has 17<sup>1</sup> even though the number of inter-subunit ion pairs for oligomeric proteins seem to increase with thermostability<sup>6-8</sup>. Fig. 2 shows intra/inter-subunit ion pairs within 5 Å of *PhCutA1*<sup>1</sup>. Solid and broken lines represent intra- and inter-subunit ion pairs, respectively. In this study, the stability of *PhCutA1* was extensively examined using mutant proteins in which charged residues were changed to noncharged residues in order to elucidate the contribution of electrostatic interactions to protein stability at temperatures over 100 °C. The changes in stability can be elucidated based on relationships with their local structures as follows.

#### **The role of intra-subunit ion pairs in protein stability.**

Eight *PhCutA1* Glu residues at positions 15, 24, 47, 63, 64, 67, 71 and 99, which form intra-subunit ion interactions within 4 Å with positively charged residues, were

substituted by Gln and Ala. Two Asp residues at positions 60 and 84 were substituted by Asn and Ala. The  $T_d$  values of the Glu24, Glu64, and Asp60 variants decreased by  $\sim 3$  °C. Changes in stability of the other mutations, except for Glu99, were slight ( $< 3$  °C). These results seem to indicate that intra-subunit interactions at the examined positions do not play a very important role in stability, contrary to the expectations in the previous report<sup>1</sup>, because changes in  $T_d$  due to the deletion of charged residues were small and changes in  $T_d$  due to Ala substitutions were similar to those of Gln or Asn substitutions.

The energy of ion–ion interactions estimated by FoldX suggests that charged residues at positions 24, 60, 71, and 99 have favorable interaction (Fig. S1 and Table S1). The decrease in stability of mutant proteins at positions 24 and 60 corresponded to the FoldX estimates, but the  $T_d$  of E99Q was 154.9 °C ( $\Delta T_d = 6.4$  °C), showing that Glu99 is unfavorable for stability even though the electrostatic energy estimated by FoldX is favorable at this position. Glu99 is located at the C-terminus of the  $\alpha 3$  helix and forms a salt bridge with Arg25 in loop 2 (Fig. 3A). The changes in  $T_d$  of R25A and R25A/E99Q were  $-12.4$  and  $-2.4$  °C, respectively, indicating that the decrease in stability of the double mutant is largely compensated by the E99Q substitution. The other charged residues near Arg25 and Glu99 provide only limited effect on stability compared with the electrostatic energy between Arg 25 and Glu99 (by the analysis of FoldX) (Table S1). Arg25 forms favorable interactions with Glu99 ( $-12.0$  kJ/mol) and Glu98 ( $-1.3$  kJ/mol), and an unfavorable interaction with Arg58 (1.4 kJ/mol). Glu99 forms favorable interactions with Arg25 ( $-12.0$  kJ/mol) and Arg58 ( $-1.1$  kJ/mol), and unfavorable interactions with Glu42 (1.3 kJ/mol) and Glu98 (2.9 kJ/mol). Therefore, the increase in stability of E99Q is not primarily caused by changes in electrostatic interactions. Protein  $\alpha$  helices have been reported to have dipole moments due to the alignment of peptide bond dipoles; a

negatively charged group at the C-terminus of an  $\alpha$  helix can therefore destabilize the native structure of a protein<sup>9-11</sup>. The mutant E99Q should be stabilized by deletion of the negatively charged group at the C-terminus of the  $\alpha_3$  helix in addition to elimination of the repulsive interactions with Glu42 and Glu98. Furthermore, FoldX suggests that the “side-chain hydrogen bonds” and “solvation of polar atoms” energies in equation 1 in addition to “helix dipole” are favorable for  $\Delta\Delta G$  of E99Q. This may be because the carboxyl group of Glu99 is partially buried. These positive factors for E99Q might largely surpass the strong electrostatic energy of the interaction with Arg25.

#### **The role of inter-subunit ion pairs in protein stability.**

The Glu residues at positions 15 and 47 form inter-subunit ion pairs in addition to intra-subunit pairs, forming an ion network with Lys19 and Arg36 (Fig. 2 and 3B). Single mutants of paired residues and a quadruple mutant of all of the residues in the network were constructed. The difference in  $T_d$  ( $\Delta T_d$ ) of E15Q, K19A, R36A, E47Q, and E15Q/K19A/R36A/E47Q were  $-0.1$ ,  $-2.6$ ,  $-1.9$ ,  $-1.1$ , and  $-2.8$  °C, respectively (Table 1). Fig. 3B shows the location of these four charged residues, which seem to strengthen the interactions between segments of different subunits, the  $\alpha_1$  helix of the A chain and the  $\beta_2$  (or  $\beta_3$ ) strand of the C chain. The change in stability of the quadruple mutant is close to the highest value of the single mutants, indicating that this network contributes only slightly to stability. The ionic energies of Lys19 and Arg36 ( $-3.6$  and  $-13.5$  kJ/mol, respectively) contribute favorably to the stability, but the ionic energies of Glu15 and Glu47 ( $3.8$  and  $3.1$  kJ/mol) contribute negatively, as estimated by FoldX (Fig. S1), resulting in a slight contribution regardless of the ion network between different subunits.

Table S1 indicates that Glu15 forms favorable interactions with Arg36 (−5.4 kJ/mol) and Lys19 (−1.7 kJ/mol) and unfavorable interactions with Glu47 (4.9 kJ/mol), Asp10 (1.5 kJ/mol), Glu12 (3.9 kJ/mol), and Glu34 (1.4 kJ/mol). Glu47 forms favorable interactions with Arg36 (−3.8 kJ/mol) and Lys19 (−3.1 kJ/mol) and unfavorable interactions with Glu46 (2.1 kJ/mol), Asp48 (2.0 kJ/mol), and Glu15 (4.9 kJ/mol) (Table S1). These multiple negative and positive interactions might originate from the unusually high number of charged residues in a small protein (102 residues).

However, substitutions for Asp residues at positions 87 and 91 dramatically affected the  $T_d$ , even though changes in  $T_d$  for other Glu and Asp substitutions were less than about 3 °C, as described (Table 1). Asp87 in the  $\beta$ 5 strand of the A chain forms inter-subunit ion pairs with Arg82 in the  $\beta$ 4 strand and Lys66 in the  $\alpha$ 2 helix of the C chain, and Asp91 at the N-terminus of the  $\alpha$ 3 helix of the A chain interacts with Lys70 in the  $\alpha$ 2 helix of the C chain (Fig. 3C). The  $\Delta T_d$  values of D87N, D91N, K66A, K70A, and R82A were −7.0, −8.4, −6.5, −3.4, and −10.0 °C, respectively.  $\Delta T_d$  values of the K66A/D87N and K70A/D91N double mutants were −9.9 and −5.1 °C, respectively (Table 1). The dramatic decrease in  $T_d$  for the D91N mutant compared to the other Glu and Asp mutants can be explained by decreases in both electrostatic energy and helix dipole moment at the N-terminus of the  $\alpha$ 3 helix. The change in  $T_d$  for D87N might be caused by the close connection of Asp87 with two counter-pairs (Arg82 and Lys66) as described below.

**Changes in denaturation temperature of mutant variants at the top 15 electrostatically favorable and unfavorable sites.**

It is important to elucidate how much the ion–ion interactions of charged residues contribute to protein stability. Table 2 lists the energies of the ion–ion interactions of charged residues in the order of favored to unfavored (and unfavored to favored), evaluated by FoldX. The most favored energy is  $-35.7$  kJ/mol for Arg82 and the most unfavored is  $19.6$  kJ/mol for Asp86. The favored and unfavored residues listed in Table 2 were substituted by noncharged residues. Some charged residues provided favorable electrostatic energies and contributed to stability, such as Arg82. However, the stability of the obtained mutant proteins did not change as expected. For example, the change in  $T_d$  of R68A was small even though the electrostatic energy of Arg68 was quite favorable, and the decrease in  $T_d$  of R33A was large even though the energetically unfavorable charged group of Arg33 was eliminated. Next, changes in stability of these mutant proteins will be discussed on the basis of structural features near the mutation sites.

Arg82 is located in the  $\beta 4$  strand (Fig. 3C), and has the most favorable energy of all the charged residues (Table 2). Arg82 forms strong salt bridges with Asp84 (electrostatic energy of  $-15.6$  kJ/mol), Glu59 ( $-2.7$  kJ/mol), and Glu63 ( $-4.0$  kJ/mol) in the same chain, and with Asp86 ( $-6.2$  kJ/mol) and Asp87 ( $-4.9$  kJ/mol) in the other chain (Table S1). Because these ion–ion interactions are eliminated when Arg82 is replaced by Ala, the  $T_d$  of R82A decreased by  $10.0$  °C. Lys66 in the  $\alpha 2$  helix close to Arg82 in the  $\beta 4$  strand also has high electrostatic energy (Table 2). Lys66 forms strong salt bridges with Glu63 ( $-1.3$  kJ/mol) and Glu67 ( $-3.0$  kJ/mol) in the same  $\alpha 2$  helix, and with Asp87 ( $-7.2$  kJ/mol), Glu90 ( $-9.0$  kJ/mol), and Asp91 ( $-1.5$  kJ/mol) in another subunit (Fig. 3C). Replacement of Lys66 by Ala resulted in a significant decrease in  $T_d$  ( $6.5$  °C). Arg82 links Lys66 via Asp87 in a centrally located ion network that contributes to the unusually high stability of *PhCutA1*.

By contrast, Asp86, Asp84, Glu67, and Glu63, which are involved in the ion network of Arg82 and Lys66, have unfavorable electrostatic energies (Table 2). Substitutions at positions D84N, E67Q, and E63Q showed only slight decreases in  $T_d$ , suggesting that the repulsive energy of these charged residues hardly affected protein stability. In the case of Asp86, the repulsive energy was great, because the Asp86 residues in the A and C chains are close to each other as shown in Fig. 3C. Therefore, D86N had a slightly higher  $T_d$  (Table 2) due to elimination of a charged group. The residues in the least favored 13 of electrostatic energy in Table 2 are negatively charged (Glu and Asp), except for Arg33. These results indicate that the number of negatively charged residues, Glu and Asp, surpasses that of positively charged residues, Lys and Arg, so the ion-ion interactions of many Glu and Asp residues should be forced to be partially repulsive to each other.

The electrostatic energies of Arg68, Glu24, and Glu71 are  $-28.0$ ,  $-15.6$ , and  $-5.7$  kJ/mol, respectively, higher than those of the other residues (Table 2). However, the decreases in  $T_d$  for R68A, E24Q, and E71Q are small,  $-2.1$ ,  $-1.7$ , and  $-1.5$  °C, respectively. As shown in Fig. 3D, Arg68 and Glu24 form a salt bridge between the  $\alpha 1$  and  $\alpha 2$  helices, and Glu71 forms a salt bridge with Arg68 in the same helix. The favorable electrostatic energies of Arg68 with Glu24, Glu71, and Glu64 were evaluated by FoldX to be  $-16.6$ ,  $-11.4$ , and  $-2.2$  kJ/mol, respectively. Arg68 has unfavorable interactions with Lys16 (1.2 kJ/mol) and Lys101 (1.0 kJ/mol). Regardless of the high favorable electrostatic energy of Arg68, the decrease in  $T_d$  of R68A was only 2.1 °C. According to  $\Delta\Delta G$  analysis by FoldX, this compensation is mainly caused by “solvation of polar atoms” and “entropy effect of side-chain” energies in equation 1. Glu24 and Glu71, paired with Arg68, also have favorable electrostatic energies of  $-15.6$  and  $-5.7$  kJ/mol, respectively

(Table 2). The E24Q and E71Q mutant variants showed slight decreases in  $T_d$  of 1.7 and 1.5 °C, respectively. Glu24 and Glu71 are located at the C-terminus of the  $\alpha 1$  helix and in the second position from the C-terminus of the  $\alpha 2$  helix, respectively (Fig. 3D). Because the substitution of Glu with noncharged Gln at the C-terminus of a helix enhances the energy of the helix dipole, the decrease in stability due to the elimination of ion pairs might be compensated by the increase in the helix dipole in the stability of both mutant proteins.

Arg58 and Asp60 have high favorable electrostatic energies of  $-21.8$  and  $-8.8$  kJ/mol, respectively (Table 2). Arg58, which is at the turn between the  $\beta 3$  strand and the  $\alpha 2$  helix, strongly interacts with Asp60 at the N-terminus of the  $\alpha 2$  helix (Fig. 3E). The electrostatic energy was  $-15.4$  kJ/mol between them (Table S1). Furthermore, Arg58 interacts weakly with Glu59 ( $-2.5$  kJ/mol) and the carboxyl group of the C-terminal Lys102 residue ( $-3.2$  kJ/mol). The decreases in  $T_d$  for the R58A, R58A/D60N, D60N, and D60A variants were 6.8, 6.8, 3.5, and 3.3 °C, respectively. Comparably large changes in stability due to substitution of Asp60 with nonionizable residues might result from the sum of the elimination of ion-ion interactions and the decrease in helix dipole moment. The  $\Delta T_d$  of the R58A/D60N double mutant hardly changed, compared with R58A, because of the elimination of repulsive interactions of Asp60 with Glu59, Glu63, Glu64, and C-terminus of Lys102 (Table S1). Glu59 also showed high electrostatic energy ( $-7.3$  kJ/mol) (Table 2), and given its position in the N-terminal region of the  $\alpha 2$  helix, should also stabilize the wild-type protein by the helix dipole moment. However, the  $T_d$  of E59Q increased by 3.7 °C contrary to expectation. This might be explained by the following. (a) The electrostatic energy of Glu59 was primarily evaluated from the interaction with the amino group of the N-terminal Met ( $-12.8$  kJ/mol) at pH 7.0. This value might be

overestimated because the amino group may not be fully ionized at pH 7.0<sup>12</sup>, the value at which the stability of the mutant protein was measured by DSC. (b) Because there are many repulsive residues (Asp60, Asp84, Asp86, Asp87, and Glu63) around Glu59 (Table S1), the elimination of these unfavorable interactions may result in the increase in stability of the mutant protein.

Arg33 is in the  $\beta$ 2 strand, and is almost completely buried in the interior of the molecule (accessible surface area of 5.3%). This residue is close to and is ionically repulsive to the Arg33 residues of the other two chains (Fig. 3F) with a repulsive energy of 12.7 kJ/mol. Arg33 also favorably interacts with Glu34 (−2.9 kJ/mol) and Glu50 (−2.5 kJ/mol) (Table S1). The stabilities of the R33A and R33M variants decreased by 9.2 and 7.7 °C, respectively, due to substitution of the repulsive residue (Table 1). If burial polar residues and repulsive ionic interactions are eliminated, mutant proteins should generally increase in stability. However, the stability of both mutant proteins decreased. This discrepancy may be explained from the crystal structure of *PhCutA1* (4nyo), which indicates that NH1 and NH2 of Arg33 contact ND1 of His35 at a distance of 2.86 Å and OH of Tyr5 at a distance of 3.20 Å, respectively (Fig. 3G). These interactions should stabilize the native structure of the wild-type protein, whereas the mutant proteins (R33A and R33M) were destabilized by the elimination of these interactions. It has been reported that completely buried, nonion-paired glutamic acid contributes favorably to the conformational stability of pyrrolidone carboxyl peptidases from hyperthermophiles because of the formation of a hydrogen bond<sup>13</sup>. It is possible that Arg33 is not ionized in the interior of the molecule<sup>14</sup>.

The locations of Glu42, Glu46, and Asp48, high in the order of unfavorable electrostatic interactions, are shown in Fig. 3H. The three mutants at these sites, E42Q, E46Q, and D48N, increased  $T_d$  values by 4.9, 3.6, and 2.7 °C, respectively. Glu42 mainly interacts with Lys44 (−1.6 kJ/mol), Glu46 (1.8 kJ/mol), and Glu99 (1.3 kJ/mol) (Table S1), resulting in an unfavorable electrostatic energy of 2.5 kJ/mol (Table 2). Glu46 has strong unfavorable interactions with Asp48 (5.1 kJ/mol), Glu47 (2.1 kJ/mol), Glu42 (1.8 kJ/mol), and Glu50 (1.0 kJ/mol), and a favorable interaction with Lys44 (−3.9 kJ/mol). Asp48 has strong unfavorable interactions with Glu46 (5.1 kJ/mol), Glu50 (3.1 kJ/mol), Glu47 (2.0 kJ/mol), and Asp10 (1.2 kJ/mol) and a slightly favorable interaction with Arg36 (−1.1 kJ/mol) (Table S1). The increase in stability of these mutants can be explained by decreased repulsive ionic interactions.

### **Relationships between $\Delta T_d$ and $\Delta\Delta G$ of unfolding and between $\Delta T_d$ and $\Delta\Delta G$ of electrostatic energy.**

Fig. 4 shows the relationship between changes in  $\Delta T_d$  due to mutations of *PhCutA1* and  $\Delta\Delta G$  of unfolding estimated by FoldX (equation 1). The correlation coefficient was 0.737 except for the R33M, R58A, and R58A/D60N variants. In the cases of R58A and R58A/D60N, FoldX estimates unusually high values of  $\Delta\Delta G$  due to decreases in the “side-chain hydrogen bonds” category to 58.8 and 57.1 kJ/mol, respectively. The FoldX values for the “Van der Waals” and “Solvation of polar atoms” categories for R33M were more favorable than those for R33A. These results indicate that the estimated changes in stability for mutant proteins modeled by FoldX are reasonably reliable, although there are a few outliers. On the other hand, the difference between the electrostatic energies of ion-

ion interactions for charged residues of the wild-type and each mutant protein ( $\Delta\Delta G$  of electrostatic energy) was estimated by FoldX. The correlation between  $\Delta T_d$  and  $\Delta\Delta G$  of electrostatic energy was poor, having a correlation coefficient of 0.437 (not shown).

As described above, these studies show that introduction of charged residues into proteins contributes to stability in various ways, including attractive or repulsive ionic interactions, as well as positive or negative effects on other important factors for protein stability, and the effects due to these other factors sometimes surpass the energy of ion–ion interactions. Therefore, the introduction of ion–ion interactions does not directly contribute to protein stability, but  $\Delta\Delta G$  of unfolding, estimated by considering almost all interactions as shown in equation 1, correlates with the experimentally measured stability (Fig. 4). The FoldX webpage suggests that FoldX is accurate for relative energies, although absolute energies are not precise (<http://foldx.crg.es/examples.jsp>). As described, this work indicates that FoldX is useful for evaluating changes in stability of mutant proteins<sup>15, 16</sup>.

### **The role of charged residues of *PhCutA1* in protein stability.**

The large number of charged residues having intra-subunit ion pairs in the same subunit (30 pairs in *PhCutA1* against 1 pair in *EcCutA1*) was anticipated to play an important role in the thermostability of *PhCutA1* at high temperatures around 150 °C<sup>1</sup>. To confirm this hypothesis, single mutant variants replacing Glu or Asp residues with noncharged residues were constructed and their stabilities were measured using DSC. The changes in stability of the mutant proteins were not great: the largest change in  $\Delta T_d$  was

$-3.5\text{ }^{\circ}\text{C}$  and the average  $\Delta T_d$  was  $-1.0 \pm 1.5\text{ }^{\circ}\text{C}$ , suggesting that intra-subunit interactions at the examined positions do not play a very important role in stability. The  $\Delta T_d$  of the *Ph14S* mutant, a variant with replacement of all residues forming intra-subunit ion interactions within  $4\text{ \AA}$ , was  $-7.2\text{ }^{\circ}\text{C}$ . This value is similar to the  $\Delta T_d$  ( $-6.8\text{ }^{\circ}\text{C}$ ) of the R58A mutant, which is included in the *Ph14S* mutant. The great decrease in  $T_d$  of R58A might originate from loss of the interaction with Asp60 and many other neighboring negatively charged ones (Fig. 3E and Table S1). The  $\Delta T_d$  values of R58A and R58A/D60N were not different (Table 1). These results suggest that all ion pairs, which are deleted in *Ph14S*, hardly contribute to protein stability except for the interaction between Arg58 and Asp60. On the other hand, the average changes in  $T_d$  of all Glu mutants examined (Table 1) was  $0.03 \pm 2.05\text{ }^{\circ}\text{C}$ , suggesting that the Glu residues of *PhCutA1* are weak contributors to the protein stability of *PhCutA1*. However, the electrostatic energy ( $-159.3\text{ kJ/mol}$ ) of *PhCutA1* is quite high, compared to that ( $-9.7\text{ kJ/mol}$ ) of *EcCutA1*. The *pI* for *PhCutA1* and *EcCutA1* is 4.91 and 4.85, respectively. In the case of *PhCutA1*, unusually large numbers of negatively charged residues (especially Glu) might be essential in order to make highly efficient interactions with positively charged residues and to generate high electrostatic energy, resulting in a situation in which negatively charged residues surrounding positively charged residues are forced to be partially repulsive to each other.

### **Design of stable proteins by the introduction of ion pairs.**

Ion pairs are effective in stabilizing proteins from mesophiles and thermophiles, when they are exposed in solvent. However, when fully buried, each charged residue represents a desolvation penalty during folding<sup>17</sup>. Recently, the Makhatadze group has

developed a computational method for the rational design of stable proteins by optimization of surface ion–ion interactions<sup>18, 19</sup>. In this method, the energies of ion–ion interactions on the protein surface are calculated using the Tanford–Kirkwood model corrected for solvent accessibility<sup>19, 20</sup>. There are several successful examples including  $\alpha$ -lactalbumin<sup>21</sup>, ubiquitin<sup>18</sup>, and acylphosphatase<sup>22</sup>. However, our study suggests that the structural features of mutation sites should be carefully considered, and particularly residues that are involved in helix dipole moment, even if the residues are on the surface. Therefore, the present findings might be useful for designing stable proteins by the introduction of ion pairs.

## **2-5. Conclusions.**

To examine the role of ion–ion interactions in the stability of *PhCutA1*, a protein with a large number of charged residues and a high denaturation temperature, mutant proteins in which charged residues were substituted by noncharged residues were comprehensively examined. The following features were found.

(1) The average decrease in  $T_d$  of Lys or Arg mutants was greater than that of Glu or Asp mutants, suggesting that several negatively charged residues do not directly contribute to stability. However, many negatively charged residues should be essential in order to make efficient interactions with positively charged residues and to generate high electrostatic energy, because the energy of ion–ion interactions ( $-159.3$  kJ/mol) of *PhCutA1* with  $pI$  of 4.91 was much larger than that ( $-9.7$  kJ/mol) of *EcCutA1* with  $pI$  of 4.85. Therefore, negatively charged residues were forced to be partially repulsive to each other. (2) Changes in stability of mutant proteins with  $T_d$  of around 140–150 °C could be

explained by considering factors important for protein stability and the structural features of mutant sites, suggesting that stabilizing factors already reported are valid on mutant proteins with  $T_d$  around 140–150 °C. These findings are useful for the design of proteins that are stable at temperatures over 100 °C. (3) The  $T_d$  of the E99Q mutant of *PhCutA1* was higher by 6.4 °C than that of the wild-type protein (148.5 °C). The world record for heat stability was updated. (4) FoldX was useful for evaluating changes in stability of mutant proteins.

Table 1.

Denaturation temperatures ( $T_d$  in °C) of *Ph*CutA1 mutants and the difference in  $T_d$  ( $\Delta T_d$ ) between the wild-type and its mutant at pH 7.0.  $T_d$  is the average value of more than two data. Positive values of  $\Delta T_d$  indicate the increase in stability due to mutations. The  $T_d$  of the wild-type is 148.5 °C.

mutants	$T_d$	$\Delta T_d$	mutants	$T_d$	$\Delta T_d$	mutants	$T_d$	$\Delta T_d$
E12Q	149.2	0.7	D10N	149.4	0.9	K19A	145.9	-2.6
E15A	147.7	-0.8	D48N	151.2	2.7	K44A	148.1	-0.4
E15Q	148.4	-0.1	D60A	145.2	-3.3	K49A	149.5	0.8
E24A	145.5	-3.0	D60N	145.0	-3.5	K66A	142.0	-6.5
E24Q	146.8	-1.7	D76N	149.4	0.9	K70A	145.1	-3.4
E34Q	147.1	-1.4	D84A	147.0	-1.5	K101A	144.1	-4.4
E42Q	153.4	4.9	D84N	148.0	-0.5	R25A	136.1	-12.4
E46Q	152.1	3.6	D86N	151.0	2.5	R33A	139.3	-9.2
E47A	146.0	-2.5	D87N	141.5	-7.0	R33M	140.8	-7.7
E47Q	147.4	-1.1	D91A	141.9	-6.6	R36A	146.6	-1.9
E59Q	152.2	3.7	D91N	140.1	-8.4	R58A	141.7	-6.8
E63A	149.0	0.5				R68A	146.4	-2.1
E63Q	147.0	-1.5				R82A	138.5	-10.0
E64A	145.6	-2.9				R25A/E99Q	146.1	-2.4
E64Q	145.0	-3.5				R58A/D60N	141.7	-6.8
E67A	149.0	0.5				R68A/E24Q	146.2	-2.3
E67Q	148.0	-0.5				K70A/D91N	143.4	-5.1
E71A	148.0	-0.5				K66A/D87N	138.6	-9.9
E71Q	147.0	-1.5				K101A/E64Q	145.3	-3.2
E99A	150.0	1.5				E15Q/K19A/R36A/E47Q	145.7	-2.8
E99Q	154.9	6.4				<i>Ph</i> 14S*	141.3	-7.2

\**Ph*14S represents E12Q/K16A/K44A/E46Q/D48N/K49A/R58A/D60N/E63Q/E67Q/E71Q/D84N/E90Q/K94A.

Table 2. The energy of ion-ion interactions and changes in  $T_d$  due to mutations at each charged residue of *PhCutA1*. Positive values of  $\Delta T_d$  indicate the increase in stability due to mutations. The energies of ion-ion interactions come from Table S1.

in the order of favored to un-favored				in the order of un-favored to favored			
positions	Energy of ion interaction (kJ/mol)	mutants	$\Delta T_d$ (°C)	positions	Energy of ion interaction (kJ/mol)	mutants	$\Delta T_d$ (°C)
ARG82	-35.7	R82A	-10.0	ASP86	19.6	D86N	2.5
ARG68	-28.0	R68A	-2.1	GLU12	11.6	E12Q	0.7
ARG58	-21.8	R58A	-6.8	ARG33	10.9	R33A	-9.2
LYS66	-21.5	K66A	-6.5	ASP10	10.6	D10N	0.9
LYS70	-16.2	K70A	-3.4	ASP48	10.5	D48N	2.7
GLU24	-15.6	E24Q	-1.7	ASP84	7.2	D84N	-0.5
ARG36	-13.5	R36A	-1.9	GLU46	5.5	E46Q	3.6
ARG25	-12.0	R25A	-12.4	GLU67	5.2	E67Q	-0.5
GLU34	-11.0	E34Q	-1.4	GLU63	5.1	E63Q	-1.5
LYS49	-10.2	K49A	0.8	GLU15	3.8	E15Q	-0.1
GLU99	-9.1	E99Q	6.4	ASP76	3.3	D76N	0.9
ASP60	-8.8	D60N	-3.5	GLU47	3.1	E47Q	-1.1
GLU59	-7.3	E59Q	3.7	GLU42	2.5	E42Q	4.9
LYS101	-6.7	K101A	-4.4	LYS23	1.7	-	
GLU71	-5.7	E71Q	-1.5	ASP87	1.5	D87N	-7.0

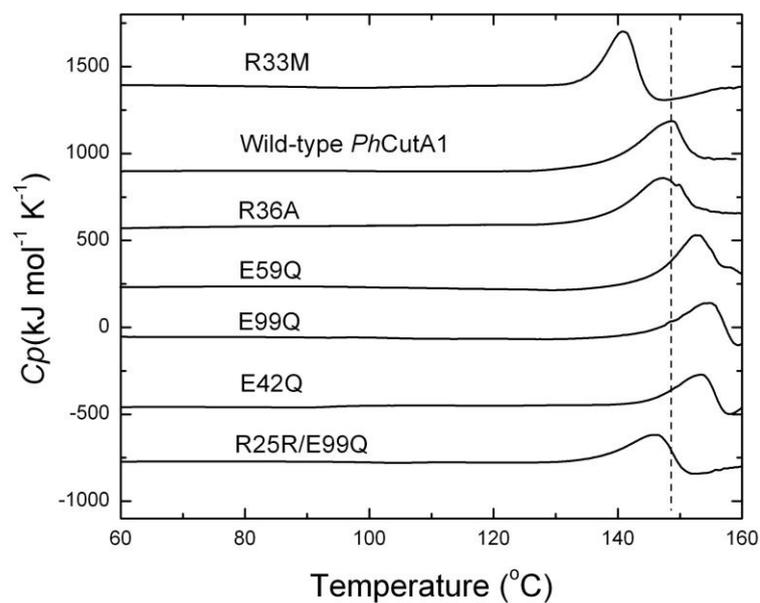


Fig. 1. Typical DSC curves of *PhCutA1* mutant proteins at pH 7.0. Heating rates of DSC were  $60^{\circ}\text{C}/\text{h}$ . The perpendicular line shows the  $T_d$  of the wild-type protein.

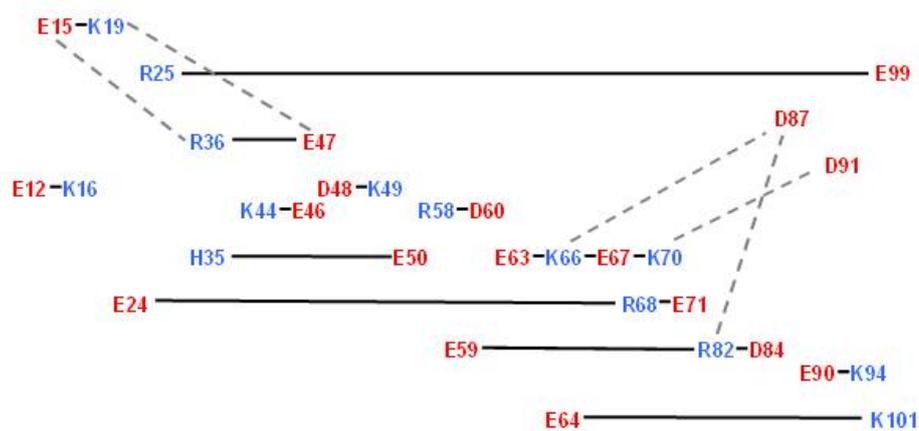
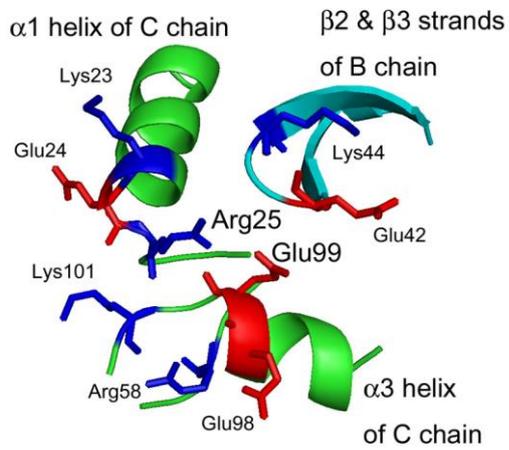
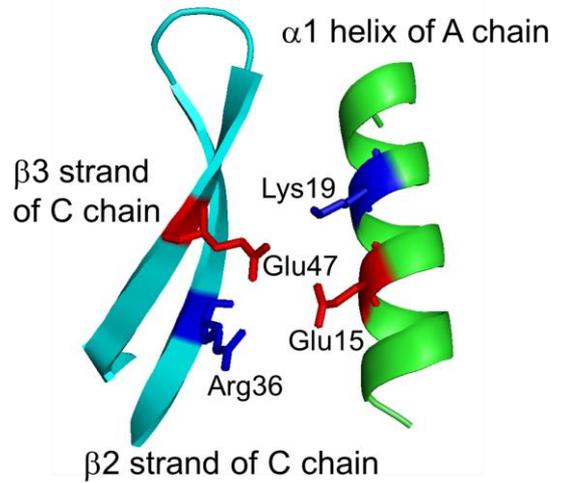


Fig. 2. Intra/inter-subunit ion pairs within  $5 \text{ \AA}$ . Solid and broken lines represent intra- and inter-subunit ion pairs, respectively.

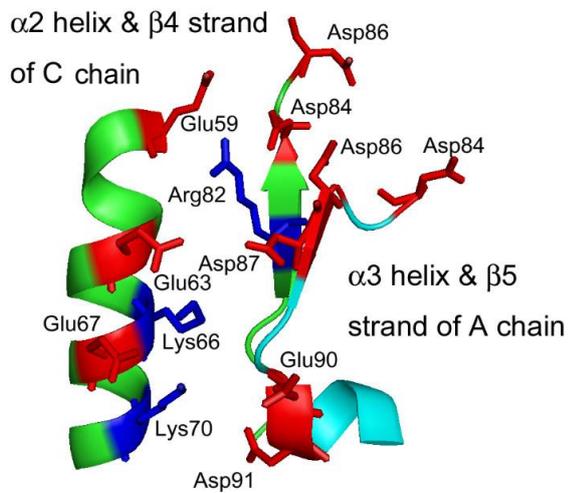
(A)



(B)



(C)



(D)

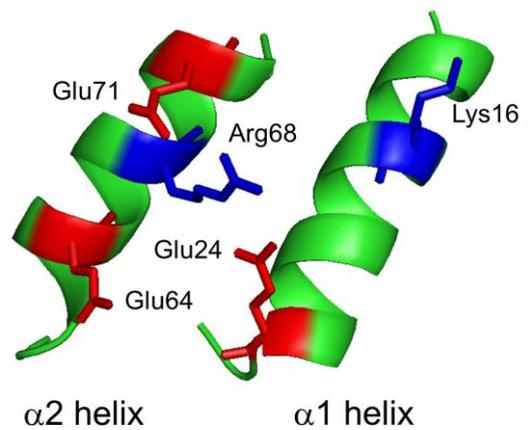
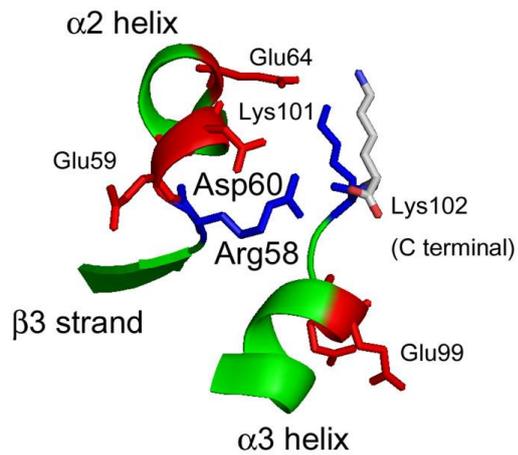


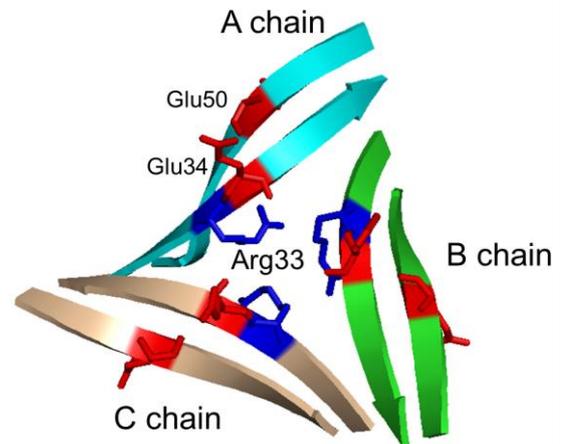
Fig. 3. The structure of *PhCutA1* at specific sites. (A) Near the Arg25-Glu99 ion pair. (B) The Glu-47-Lys19-Glu15-Arg36 ion network. (C) Ion networks of centrally-located Arg82 and Lys66. (D) The Arg68-Glu24 ion pair.

Fig. 3. continued.

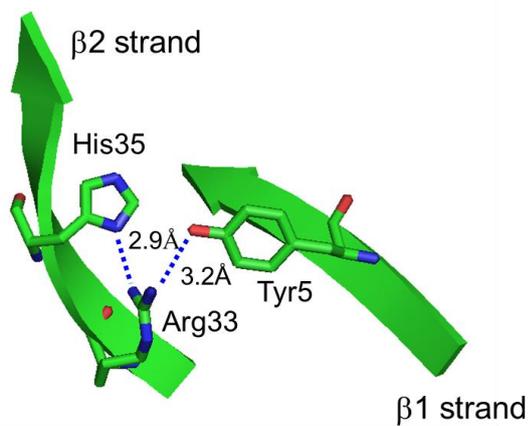
(E)



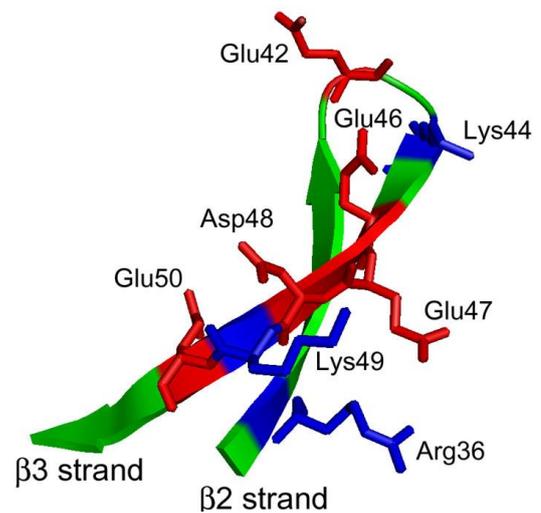
(F)



(G)



(H)



(E) Ion pairs of Arg58. (F) The ion group of Arg33 seems to be repulsive to those of Arg33 in the other chains in the interior of the molecule. (G) Buried Arg33 favorably interacts with other residues. (H) Apparently unfavorable charged residues, Glu42, Glu46, and Asp48.

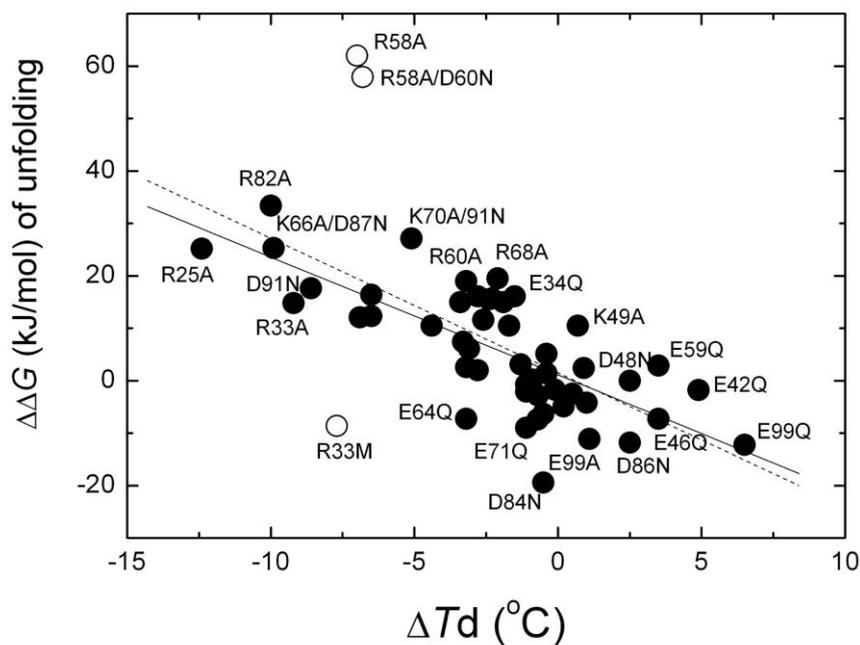


Fig. 4. Relationship between  $\Delta T_d$  of *PhCutA1* variants and  $\Delta\Delta G$  of unfolding obtained from FoldX.  $\Delta T_d$  values were obtained as in Table 1.  $\Delta\Delta G$  of unfolding for each mutant protein was obtained from equation (1). The structures of mutant proteins were constructed by FoldX. Broken and solid lines represent linear regression of all data points with and without the three open circle mutant proteins, respectively. Correlation coefficients of the broken and solid lines were  $-0.639$  and  $-0.737$ , respectively.

2-6. Supporting information.

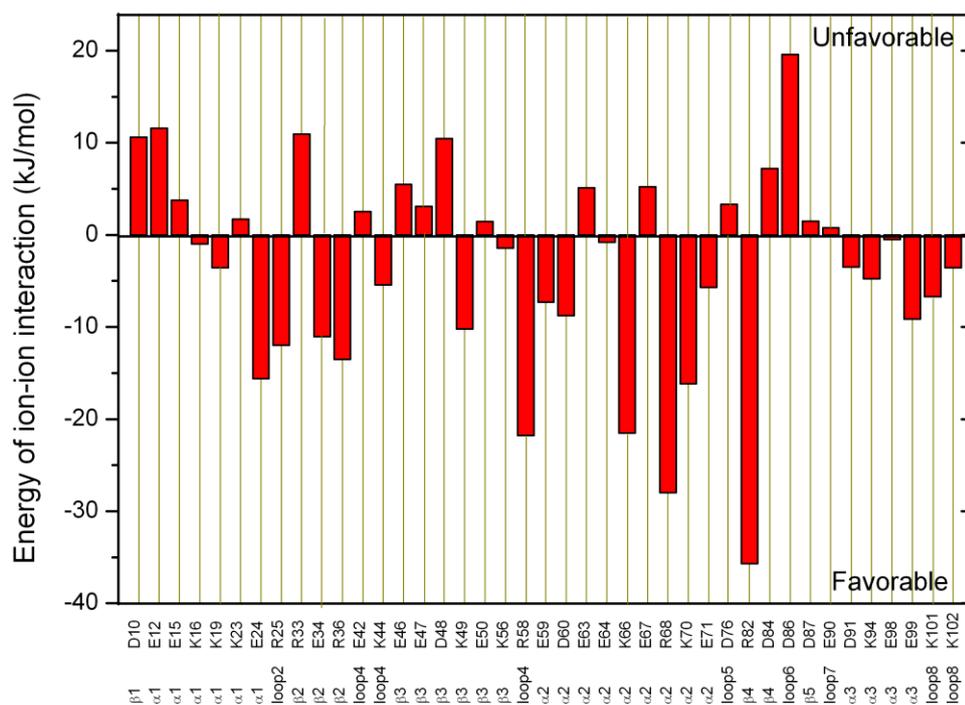


Fig. S1. The electrostatic energies of ion–ion interactions for all charged residues included in the *PhCutA1* trimer. The electrostatic energies between ion groups of the *PhCutA1* trimer were calculated by FoldX at pH 7.0 as listed in Table S1. Negative and positive values represent favorable and un-favorable (repulsive interaction), respectively. The bottom line represents the secondary structure of each charged residue.

Table S1. Electrostatic energies of ion-ion interactions between charged residues of *PhCutA1* including the values of each pair interaction estimated by FoldX. **Red numbers** parenthesized mean negative values that show favorable interaction.

Charged residues	A chain			B chain			C chain			(A+B+C)/2 kJ/mol	per each residue kJ/mol
	Pair residues		kJ/mol	Pair residues		kJ/mol	Pair residues		kJ/mol		
N-terminal	META1	LYSA56	0.13	METB1	LYSB56	0.13	METC1	LYSC56	0.14	0.20	
	META1	ARGA58	0.59	METB1	ARGB58	0.60	METC1	ARGC58	0.62	0.90	
	META1	GLUA59	(8.46)	METB1	GLUB59	(8.57)	METC1	GLUC59	(8.61)	(12.82)	
	META1	ASPA60	(0.54)	METB1	ASPB60	(0.55)	METC1	ASPC60	(0.58)	(0.83)	
	META1	GLUA63	(0.22)	METB1	GLUB63	(0.23)	METC1	GLUC63	(0.25)	(0.35)	
	META1	ARGA82	0.86	METB1	ARGB82	0.75	METC1	ARGC82	0.87	1.24	
	META1	ASPA84	(1.19)	METB1	ASPB84	(1.18)	METC1	ASPC84	(1.23)	(1.80)	
	META1	ASPA86	(0.49)	METB1	ASPB86	(0.49)	METC1	ASPC86	(0.47)	(0.72)	
	META1	ASPA87	(0.62)	METB1	ASPB87	(0.60)	METC1	ASPC87	(0.57)	(0.89)	
	META1	GLUA90	(0.33)	METB1	GLUB90	(0.28)	METC1	GLUC90	(0.22)	(0.42)	
	META1	ASPA91	(0.19)	METB1	ASPB91	(0.20)	METC1	ASPC91	(0.11)	(0.25)	
	META1	LYSA94	0.18	METB1	LYSB94	0.17				0.17	
	META1	GLUA98	(0.08)	METB1	GLUB98	(0.08)	METC1	GLUC98	(0.08)	(0.12)	
	META1	LYSA102	(0.21)	METB1	LYSB102	0.09	METC1	LYSC102	0.11	(0.00)	
	META1	ASPB86	(0.38)	METB1	ASPC86	(0.40)	METC1	ASPA86	(0.40)	(0.59)	
	META1	ASPB87	(0.08)	METB1	ASPC87	(0.09)	METC1	ASPA87	(0.09)	(0.13)	
	META1	LYSC66	0.17	METB1	LYSA66	0.19	METC1	LYSB66	0.08	0.22	
	META1	ARGC82	0.31	METB1	ARGA82	0.30	METC1	ARGB82	0.32	0.47	
	META1	ASPC84	(0.24)	METB1	ASPA84	(0.23)	METC1	ASPB84	(0.23)	(0.35)	(16.08)
	Asp 10	ASPA10	ARGC36	(0.54)	ASPB10	ARGA36	(0.48)	ASPC10	ARGB36	(0.55)	(0.79)
ASPA10		GLUC47	0.08	ASPB10	GLUA47	0.09				0.09	
ASPA10		GLUA12	4.52	ASPB10	GLUB12	5.67	ASPC10	GLUC12	5.45	7.82	
ASPA10		GLUA15	0.97	ASPB10	GLUB15	0.97	ASPC10	GLUC15	0.98	1.46	
ASPA10		LYSA16	(0.37)	ASPB10	LYSB16	(0.48)	ASPC10	LYSC16	(0.21)	(0.53)	
ASPA10		LYSA19	(0.09)	ASPB10	LYSB19	(0.09)	ASPC10	LYSC19	(0.08)	(0.13)	
ASPA10		ARGA33	(0.15)	ASPB10	ARGB33	(0.15)	ASPC10	ARGC33	(0.15)	(0.22)	
ASPA10		GLUA34	0.60	ASPB10	GLUB34	0.60	ASPC10	GLUC34	0.65	0.92	
ASPA10		GLUA46	0.33							0.16	
ASPA10		ASPA48	1.23	ASPB10	ASPB48	0.62	ASPC10	ASPC48	0.56	1.20	
ASPA10		LYSA49	(0.31)	ASPB10	LYSB49	(0.47)	ASPC10	LYSC49	(0.26)	(0.52)	
ASPA10		GLUA50	0.75	ASPB10	GLUB50	0.71	ASPC10	GLUC50	0.69	1.07	
							ASPC10	ASPC76	0.18	0.09	10.63
Glu 12		GLUA12	ARGC36	(1.00)	GLUB12	ARGA36	(0.82)	GLUC12	ARGB36	(0.98)	(1.40)
	GLUA12	GLUC47	0.53	GLUB12	GLUA47	0.46	GLUC12	GLUB47	0.46	0.73	
	GLUA12	ASPA10	4.52	GLUB12	ASPB10	5.67	GLUC12	ASPC10	5.45	7.82	
	GLUA12	GLUA15	2.93	GLUB12	GLUB15	2.44	GLUC12	GLUC15	2.49	3.93	
	GLUA12	LYSA16	(0.26)	GLUB12	LYSB16	(0.27)	GLUC12	LYSC16	(0.20)	(0.37)	
	GLUA12	LYSA19	(0.26)	GLUB12	LYSB19	(0.23)	GLUC12	LYSC19	(0.21)	(0.35)	
	GLUA12	GLUA34	0.77	GLUB12	GLUB34	0.75	GLUC12	GLUC34	0.81	1.16	
	GLUA12	ASPA48	0.48	GLUB12	ASPB48	0.17	GLUC12	ASPC48	0.09	0.37	
	GLUA12	LYSA49	(0.32)	GLUB12	LYSB49	(0.59)	GLUC12	LYSC49	(0.27)	(0.59)	
	GLUA12	GLUA50	0.19	GLUB12	GLUB50	0.19	GLUC12	GLUC50	0.19	0.28	11.59
Glu 15	GLUA15	GLUC34	0.09	GLUB15	GLUA34	0.18	GLUC15	ARGB33	(0.21)	(0.10)	
	GLUA15	ARGC36	(3.81)	GLUB15	ARGA36	(2.96)	GLUC15	GLUB34	0.18	0.23	
	GLUA15	LYSC44	(0.16)				GLUC15	ARGB36	(4.06)	(5.42)	
	GLUA15	GLUC46	0.45	GLUB15	GLUA46	0.08	GLUC15	GLUB46	0.16	0.34	
	GLUA15	GLUC47	3.45	GLUB15	GLUA47	3.79	GLUC15	GLUB47	2.57	4.90	
	GLUA15	ASPC48	0.37	GLUB15	ASPA48	0.29	GLUC15	ASPB48	0.38	0.52	
	GLUA15	LYSC49	(0.25)	GLUB15	LYSA49	(0.25)	GLUC15	LYSB49	(0.10)	(0.30)	
	GLUA15	GLUC50	0.08	GLUB15	GLUA50	0.16	GLUC15	GLUB50	0.10	0.17	
	GLUA15	ASPA10	0.97	GLUB15	ASPB10	0.97	GLUC15	ASPC10	0.98	1.46	
	GLUA15	GLUA12	2.93	GLUB15	GLUB12	2.44	GLUC15	GLUC12	2.49	3.93	
	GLUA15	LYSA16	(0.28)	GLUB15	LYSB16	(0.25)	GLUC15	LYSC16	(0.30)	(0.42)	
	GLUA15	LYSA19	(1.12)	GLUB15	LYSB19	(1.34)	GLUC15	LYSC19	(0.91)	(1.69)	
	GLUA15	LYSA23	(0.22)	GLUB15	LYSB23	(0.11)	GLUC15	LYSC23	(0.09)	(0.21)	
	GLUA15	ARGA33	(0.22)	GLUB15	ARGB33	(0.16)	GLUC15	ARGC33	(0.32)	(0.35)	
	GLUA15	GLUA34	0.69	GLUB15	GLUB34	0.75	GLUC15	GLUC34	0.80	1.12	
	GLUA15	LYSA49	(0.20)	GLUB15	LYSB49	(0.32)	GLUC15	LYSC49	(0.19)	(0.36)	3.75

Lys 16	LYSA16	ASPA10	(0.37)	LYSB16	ASPB10	(0.48)	LYSC16	GLUB47	(0.10)	(0.05)		
	LYSA16	GLUA12	(0.26)	LYSB16	GLUB12	(0.27)	LYSC16	ASPC10	(0.21)	(0.53)		
	LYSA16	GLUA15	(0.28)	LYSB16	GLUB15	(0.25)	LYSC16	GLUC12	(0.20)	(0.37)		
	LYSA16	LYSA19	0.17	LYSB16	LYSB19	0.13	LYSC16	GLUC15	(0.30)	(0.42)		
	LYSA16	LYSA23	0.16	LYSB16	LYSB23	0.08	LYSC16	LYSC19	0.36	0.33		
	LYSA16	GLUA24	(0.44)	LYSB16	GLUB24	(0.40)	LYSC16	LYSC23	0.28	0.26		
	LYSA16	ARGA68	0.79	LYSB16	ARGB68	0.84	LYSC16	GLUC24	(0.77)	(0.81)		
				LYSB16	LYSB70	0.08		ARGC68	0.84	1.24		
				LYSB16	GLUB71	(0.56)	LYSC16	GLUC71	(0.33)	(0.60)		
				LYSB16	ASPB76	(0.20)						
							LYSC16	LYSC101	0.08	0.04	(0.97)	
	Lys 19	LYSA19	ARGC36	0.35	LYSB19	ARGA36	0.39	LYSC19	ARGB36	0.34	0.54	
LYSA19		LYSC44	0.54	LYSB19	LYSA44	0.13	LYSC19	LYSB44	0.13	0.40		
LYSA19		GLUC46	(0.57)	LYSB19	GLUA46	(0.25)	LYSC19	GLUB46	(0.41)	(0.61)		
LYSA19		GLUC47	(2.32)	LYSB19	GLUA47	(1.84)	LYSC19	GLUB47	(1.95)	(3.05)		
LYSA19		ASPC48	(0.22)	LYSB19	ASPA48	(0.20)	LYSC19	ASPB48	(0.31)	(0.36)		
				LYSB19	GLUA50	(0.08)	LYSC19	GLUB50	(0.08)	(0.08)		
LYSA19		ASPA10	(0.09)	LYSB19	ASPB10	(0.09)	LYSC19	ASPC10	(0.08)	(0.13)		
LYSA19		GLUA12	(0.26)	LYSB19	GLUB12	(0.23)	LYSC19	GLUC12	(0.21)	(0.35)		
LYSA19		GLUA15	(1.12)	LYSB19	GLUB15	(1.34)	LYSC19	GLUC15	(0.91)	(1.69)		
LYSA19		LYSA16	0.17	LYSB19	LYSB16	0.13	LYSC19	LYSC16	0.36	0.33		
LYSA19		LYSA23	1.16	LYSB19	LYSB23	0.75	LYSC19	LYSC23	0.80	1.36		
LYSA19		GLUA24	(0.19)	LYSB19	GLUB24	(0.19)	LYSC19	GLUC24	(0.19)	(0.29)		
LYSA19		ARGA25	0.24	LYSB19	ARGB25	0.25	LYSC19	ARGC25	0.28	0.38		
LYSA19		ARGA68	0.08	LYSB19	ARGB68	0.08	LYSC19	ARGC68	0.08	0.12		
LYSA19		GLUA99	(0.08)	LYSB19	GLUB99	(0.08)	LYSC19	GLUC99	(0.09)	(0.13)	(3.56)	
Lys 23		LYSA23	LYSC44	0.67	LYSB23	LYSA44	0.53	LYSC23	LYSB44	0.15	0.67	
		LYSA23	GLUC46	(0.27)	LYSB23	GLUA46	(0.19)	LYSC23	GLUB46	(0.29)	(0.38)	
	LYSA23	GLUC47	(0.32)	LYSB23	GLUA47	(0.24)	LYSC23	GLUB47	(0.24)	(0.40)		
							LYSC23	ASPB48	(0.08)	(0.04)		
	LYSA23	GLUA15	(0.22)	LYSB23	GLUB15	(0.11)	LYSC23	GLUC15	(0.09)	(0.21)		
	LYSA23	LYSA16	0.16	LYSB23	LYSB16	0.08	LYSC23	LYSC16	0.28	0.26		
	LYSA23	LYSA19	1.16	LYSB23	LYSB19	0.75	LYSC23	LYSC19	0.80	1.36		
	LYSA23	GLUA24	(0.44)	LYSB23	GLUB24	(0.32)	LYSC23	GLUC24	(0.32)	(0.54)		
	LYSA23	ARGA25	0.47	LYSB23	ARGB25	0.59	LYSC23	ARGC25	0.70	0.89		
	LYSA23	ARGA68	0.35	LYSB23	ARGB68	0.27	LYSC23	ARGC68	0.19	0.40		
	LYSA23	GLUA99	(0.22)	LYSB23	GLUB99	(0.27)	LYSC23	GLUC99	(0.28)	(0.39)		
	LYSA23	LYSA101	0.11				LYSC23	LYSC101	0.08	0.10	1.71	
Glu 24	GLUA24	LYSA16	(0.44)	GLUB24	LYSB16	(0.40)	GLUC24	LYSC16	(0.77)	(0.81)		
	GLUA24	LYSA19	(0.19)	GLUB24	LYSB19	(0.19)	GLUC24	LYSC19	(0.19)	(0.29)		
	GLUA24	LYSA23	(0.44)	GLUB24	LYSB23	(0.32)	GLUC24	LYSC23	(0.32)	(0.54)		
	GLUA24	ARGA25	(0.60)	GLUB24	ARGB25	(0.60)	GLUC24	ARGC25	(0.61)	(0.90)		
	GLUA24	ARGA58	(0.31)	GLUB24	ARGB58	(0.39)	GLUC24	ARGC58	(0.24)	(0.47)		
	GLUA24	ASPA60	0.08	GLUB24	ASPB60	0.17	GLUC24	ASPC60	0.09	0.17		
				GLUB24	GLUB63	0.08	GLUC24	GLUC63	0.08	0.08		
	GLUA24	GLUA64	1.45	GLUB24	GLUB64	1.48	GLUC24	GLUC64	1.54	2.23		
				GLUB24	LYSB66	(0.08)					(0.04)	
	GLUA24	GLUA67	0.47	GLUB24	GLUB67	0.46	GLUC24	GLUC67	0.39	0.66		
	GLUA24	ARGA68	(11.27)	GLUB24	ARGB68	(10.70)	GLUC24	ARGC68	(11.30)	(16.64)		
	GLUA24	GLUA71	1.72	GLUB24	GLUB71	1.83	GLUC24	GLUC71	1.90	2.72		
	GLUA24	GLUA99	0.08	GLUB24	GLUB99	0.08	GLUC24	GLUC99	0.08	0.12		
GLUA24	LYSA101	(1.66)	GLUB24	LYSB101	(0.76)	GLUC24	LYSC101	(1.62)	(2.02)			
GLUA24	LYSA102	(0.30)	GLUB24	LYSB102	0.27	GLUC24	LYSC102	0.28	0.12	(15.60)		
Arg 25	ARGA25	GLUC42	(0.50)	ARGB25	GLUA42	(0.61)	ARGC25	GLUB42	(0.62)	(0.86)		
	ARGA25	LYSC44	0.27	ARGB25	LYSA44	0.59	ARGC25	LYSB44	0.42	0.64		
	ARGA25	GLUC46	(0.36)	ARGB25	GLUA46	(0.43)	ARGC25	GLUB46	(0.55)	(0.67)		
	ARGA25	LYSA19	0.24	ARGB25	LYSB19	0.25	ARGC25	LYSC19	0.28	0.38		
	ARGA25	LYSA23	0.47	ARGB25	LYSB23	0.59	ARGC25	LYSC23	0.70	0.89		
	ARGA25	GLUA24	(0.60)	ARGB25	GLUB24	(0.60)	ARGC25	GLUC24	(0.61)	(0.90)		
				ARGB25	LYSB56	0.07					0.04	
	ARGA25	ARGA58	0.89	ARGB25	ARGB58	0.90	ARGC25	ARGC58	0.92	1.35		
	ARGA25	ASPA60	(0.31)	ARGB25	ASPB60	(0.32)	ARGC25	ASPC60	(0.32)	(0.47)		
	ARGA25	GLUA64	(0.16)	ARGB25	GLUB64	(0.15)	ARGC25	GLUC64	(0.15)	(0.23)		
	ARGA25	ARGA68	0.13	ARGB25	ARGB68	0.12	ARGC25	ARGC68	0.12	0.18		
							ARGC25	LYSC94	0.32	0.16		
	ARGA25	GLUA98	(0.87)	ARGB25	GLUB98	(0.88)	ARGC25	GLUC98	(0.81)	(1.28)		
	ARGA25	GLUA99	(7.43)	ARGB25	GLUB99	(8.36)	ARGC25	GLUC99	(8.27)	(12.03)		
	ARGA25	LYSA101	0.63	ARGB25	LYSB101	0.42	ARGC25	LYSC101	0.46	0.76		
	ARGA25	LYSA102	(0.62)	ARGB25	LYSB102	0.49	ARGC25	LYSC102	0.29	0.09	(11.97)	

Arg 33	ARGA33	ARGC33	3.79	ARGB33	ARGC33	5.14	ARGC33	ARGA33	3.79	3.79		
				ARGB33	ARGA33	3.75	ARGC33	ARGB33	5.14	5.14		
	ARGA33	ARGB33	3.75	ARGB33	ARGA33	3.75						
	ARGA33	GLUC34	(0.51)	ARGB33	GLUA34	(0.45)	ARGC33	GLUA34	(0.50)	(0.73)		
	ARGA33	ARGC36	0.59	ARGB33	ARGA36	0.46				0.53		
	ARGA33	LYSC49	0.19	ARGB33	LYSA49	0.17				0.18		
	ARGA33	GLUC50	(0.37)	ARGB33	GLUA50	(0.34)	ARGC33	GLUA50	(0.14)	(0.43)		
	ARGA33	LYSC56	0.18	ARGB33	LYSA56	0.27	ARGC33	LYSA56	0.82	0.64		
				ARGB33	GLUC15	(0.21)				(0.10)		
	ARGA33	GLUB34	(0.55)	ARGB33	GLUC34	(0.57)	ARGC33	GLUB34	(0.54)	(0.83)		
	ARGA33	ARGB36	0.06	ARGB33	ARGC36	0.06	ARGC33	ARGB36	0.60	0.36		
				ARGB33	LYSC49	0.16	ARGC33	LYSB49	0.08	0.12		
	ARGA33	GLUB50	(0.15)	ARGB33	GLUC50	(0.16)	ARGC33	GLUB50	(0.38)	(0.35)		
	ARGA33	LYSB56	0.83	ARGB33	LYSC56	0.83	ARGC33	LYSB56	0.18	0.92		
	ARGA33	ASPA10	(0.15)	ARGB33	ASPB10	(0.15)	ARGC33	ASPC10	(0.15)	(0.22)		
	ARGA33	GLUA15	(0.22)	ARGB33	GLUB15	(0.16)	ARGC33	GLUC15	(0.32)	(0.35)		
	ARGA33	GLUA34	(0.86)	ARGB33	GLUB34	(0.93)	ARGC33	GLUC34	(0.84)	(1.32)		
	ARGA33	ARGA36	0.46	ARGB33	ARGB36	0.66	ARGC33	ARGC36	0.53	0.82		
				ARGB33	GLUB47	(0.07)				(0.03)		
	ARGA33	ASPA48	(0.31)	ARGB33	ASPB48	(0.14)	ARGC33	ASPC48	(0.14)	(0.30)		
	ARGA33	LYSA49	0.35	ARGB33	LYSB49	0.33	ARGC33	LYSC49	0.37	0.52		
	ARGA33	GLUA50	(1.15)	ARGB33	GLUB50	(1.20)	ARGC33	GLUC50	(1.17)	(1.76)		
	ARGA33	LYSA56	0.43	ARGB33	LYSB56	0.34	ARGC33	LYSC56	0.39	0.58	10.94	
	Glu 34	GLUA34	GLUB15	0.18	GLUB34	GLUC15	0.18	GLUC34	GLUA15	0.09	0.23	
		GLUA34	ARGB33	(0.45)	GLUB34	ARGC33	(0.54)	GLUC34	ARGA33	(0.51)	(0.75)	
		GLUA34	GLUB34	0.74	GLUB34	GLUC34	0.70	GLUC34	GLUA34	0.60	1.02	
		GLUA34	ARGB36	(0.07)	GLUB34	ARGC36	(0.16)				(0.11)	
		GLUA34	LYSB49	(0.19)	GLUB34	LYSC49	(0.25)	GLUC34	LYSA49	(0.19)	(0.31)	
		GLUA34	ARGC33	(0.50)	GLUB34	ARGA33	(0.55)	GLUC34	ARGB33	(0.57)	(0.81)	
		GLUA34	GLUC34	0.60	GLUB34	GLUA34	0.74	GLUC34	GLUB34	0.70	1.02	
		GLUA34	ARGC36	(2.25)	GLUB34	ARGA36	(1.73)	GLUC34	ARGB36	(2.36)	(3.17)	
		GLUA34	GLUC47	0.27	GLUB34	GLUA47	0.36	GLUC34	GLUB47	0.19	0.41	
		GLUA34	LYSC49	(0.46)	GLUB34	LYSA49	(0.41)	GLUC34	LYSB49	(0.25)	(0.56)	
GLUA34		ASPA10	0.60	GLUB34	ASPB10	0.60	GLUC34	ASPC10	0.65	0.92		
GLUA34		GLUA12	0.77	GLUB34	GLUB12	0.75	GLUC34	GLUC12	0.81	1.16		
GLUA34		GLUA15	0.69	GLUB34	GLUB15	0.75	GLUC34	GLUC15	0.80	1.12		
GLUA34		ARGA33	(0.86)	GLUB34	ARGB33	(0.93)	GLUC34	ARGC33	(0.84)	(1.32)		
GLUA34		ARGA36	(1.30)	GLUB34	ARGB36	(1.20)	GLUC34	ARGC36	(1.09)	(1.79)		
GLUA34		GLUA47	0.32	GLUB34	GLUB47	0.20	GLUC34	GLUC47	0.10	0.31		
GLUA34		ASPA48	0.66	GLUB34	ASPB48	0.28	GLUC34	ASPC48	0.43	0.69		
GLUA34		LYSA49	(6.83)	GLUB34	LYSB49	(6.64)	GLUC34	LYSC49	(6.32)	(9.89)		
GLUA34		GLUA50	0.54	GLUB34	GLUB50	0.50	GLUC34	GLUC50	0.56	0.80	(11.04)	
Arg 36		ARGA36	ASPB10	(0.48)	ARGB36	ASPC10	(0.55)	ARGC36	ASPA10	(0.54)	(0.79)	
		ARGA36	GLUB12	(0.82)	ARGB36	GLUC12	(0.98)	ARGC36	GLUA12	(1.00)	(1.40)	
	ARGA36	GLUB15	(2.96)	ARGB36	GLUC15	(4.06)	ARGC36	GLUA15	(3.81)	(5.42)		
	ARGA36	LYSB19	0.39	ARGB36	LYSC19	0.34	ARGC36	LYSA19	0.35	0.54		
	ARGA36	ARGB33	0.46	ARGB36	ARGC33	0.60	ARGC36	ARGA33	0.59	0.83		
	ARGA36	GLUB34	(1.73)	ARGB36	GLUC34	(2.36)	ARGC36	GLUA34	(2.25)	(3.17)		
	ARGA36	LYSB49	0.52	ARGB36	LYSC49	0.44	ARGC36	LYSA49	0.46	0.71		
				ARGC36	LYSA56	0.08	ARGC36	LYSA56	0.08	0.04		
				ARGB36	ARGA33	0.06	ARGC36	ARGB33	0.06	0.06		
				ARGB36	GLUA34	(0.07)	ARGC36	GLUB34	(0.16)	(0.11)		
	ARGA36	ARGA33	0.46	ARGB36	ARGB33	0.66	ARGC36	ARGC33	0.53	0.82		
	ARGA36	GLUA34	(1.30)	ARGB36	GLUB34	(1.20)	ARGC36	GLUC34	(1.09)	(1.79)		
	ARGA36	GLUA46	(0.16)	ARGB36	GLUB46	(0.14)	ARGC36	GLUC46	(0.46)	(0.37)		
	ARGA36	GLUA47	(3.11)	ARGB36	GLUB47	(2.23)	ARGC36	GLUC47	(2.29)	(3.82)		
	ARGA36	ASPA48	(0.83)	ARGB36	ASPB48	(0.60)	ARGC36	ASPC48	(0.67)	(1.06)		
	ARGA36	LYSA49	1.83	ARGB36	LYSB49	0.78	ARGC36	LYSC49	1.82	2.21		
	ARGA36	GLUA50	(0.54)	ARGB36	GLUB50	(0.54)	ARGC36	GLUC50	(0.53)	(0.81)	(13.52)	
	Glu 42	GLUA42	LYSA44	(0.38)	GLUB42	LYSB44	(2.60)	GLUC42	LYSC44	(0.29)	(1.64)	
GLUA42		GLUA46	1.38	GLUB42	GLUB46	1.27	GLUC42	GLUC46	0.98	1.81		
GLUA42		ASPA48	0.17	GLUB42	ASPB48	0.43	GLUC42	ASPC48	0.47	0.54		
GLUA42		GLUA50	0.07							0.04		
GLUA42		ASPA76	0.91	GLUB42	ASPB76	0.65	GLUC42	ASPC76	0.09	0.82		
GLUA42		ARGB25	(0.61)	GLUB42	ARGC25	(0.62)	GLUC42	ARGA25	(0.50)	(0.86)		
GLUA42		ASPB91	0.09	GLUB42	ASPC91	0.17				0.13		
				GLUB42	LYSC94	(0.54)				(0.27)		
GLUA42		GLUB98	0.67	GLUB42	GLUC98	0.48	GLUC42	GLUA98	0.09	0.62		
GLUA42		GLUB99	1.06	GLUB42	GLUC99	1.05	GLUC42	GLUA99	0.53	1.32	2.52	
Lys 44		LYSA44	LYSB19	0.13	LYSB44	LYSC19	0.13	LYSC44	GLUA15	(0.16)	(0.08)	
	LYSA44	LYSB23	0.53	LYSB44	LYSC23	0.15	LYSC44	LYSA19	0.54	0.40		
	LYSA44	ARGB25	0.59	LYSB44	ARGC25	0.42	LYSC44	LYSA23	0.67	0.67		
				LYSB44	LYSC94	0.10	LYSC44	ARGA25	0.27	0.64		
				LYSB44	GLUC98	(0.08)				0.05		
	LYSA44	GLUB99	(0.37)	LYSB44	GLUC99	(0.52)	LYSC44	GLUA99	(0.08)	(0.49)		
	LYSA44	GLUA42	(0.38)	LYSB44	GLUB42	(2.60)	LYSC44	GLUC42	(0.29)	(1.64)		
	LYSA44	GLUA46	(0.34)	LYSB44	GLUB46	(6.76)	LYSC44	GLUC46	(0.78)	(3.94)		
				LYSB44	GLUB47	(0.23)	LYSC44	GLUC47	(0.42)	(0.32)		
				LYSB44	ASPB48	(0.76)	LYSC44	ASPC48	(0.17)	(0.47)		
				LYSB44	GLUB50	(0.22)				(0.11)		
				LYSB44	ASPB76	(0.19)				(0.09)	(5.42)	



Glu 50	GLUA50	GLUB15	0.16	GLUB50	GLUC15	0.10	GLUC50	GLUA15	0.08	0.17			
	GLUA50	LYSB19	(0.08)	GLUB50	LYSC19	(0.08)				(0.08)			
	GLUA50	ARGB33	(0.34)	GLUB50	ARGC33	(0.38)	GLUC50	ARGA33	(0.37)	(0.55)			
	GLUA50	LYSB56	(2.07)	GLUB50	LYSC56	(1.69)	GLUC50	LYSA56	(1.50)	(2.63)			
	GLUA50	ASPB91	0.26	GLUB50	ASPC91	0.26	GLUC50	ASPA91	0.26	0.39			
	GLUA50	GLUB99	0.16				GLUC50	GLUA99	0.34	0.25			
	GLUA50	ARGC33	(0.14)	GLUB50	ARGA33	(0.15)	GLUC50	ARGB33	(0.16)	(0.23)			
	GLUA50	ASPA10	0.75	GLUB50	ASPB10	0.71	GLUC50	ASPC10	0.69	1.07			
	GLUA50	GLUA12	0.19	GLUB50	GLUB12	0.19	GLUC50	GLUC12	0.19	0.28			
	GLUA50	ARGA33	(1.15)	GLUB50	ARGB33	(1.20)	GLUC50	ARGC33	(1.17)	(1.76)			
	GLUA50	GLUA34	0.54	GLUB50	GLUB34	0.50	GLUC50	GLUC34	0.56	0.80			
	GLUA50	ARGA36	(0.54)	GLUB50	ARGB36	(0.54)	GLUC50	ARGC36	(0.53)	(0.81)			
	GLUA50	GLUA42	0.07							0.04			
					GLUB50	LYSB44	(0.22)				(0.11)		
	GLUA50	GLUA46	1.02	GLUB50	GLUB46	0.59	GLUC50	GLUC46	0.43	1.02			
	GLUA50	GLUA47	0.51	GLUB50	GLUB47	0.51	GLUC50	GLUC47	0.40	0.71			
	GLUA50	ASPA48	2.32	GLUB50	ASPB48	1.99	GLUC50	ASPC48	1.80	3.05			
	GLUA50	LYSA49	(0.31)	GLUB50	LYSB49	(0.29)	GLUC50	LYSC49	(0.31)	(0.46)			
	GLUA50	ASPA76	0.19	GLUB50	ASPB76	0.18	GLUC50	ASPC76	0.20	0.29	1.44		
	Lys 56	LYSA56	ARGB33	0.27	LYSB56	ARGC33	0.18	LYSC56	ARGA33	0.18	0.32		
		LYSA56	ARGC33	0.82	LYSB56	ARGA33	0.83	LYSC56	ARGB33	0.83	1.24		
		LYSA56	ARGC36	0.08							0.04		
				0.00	LYSB56	GLUA46	(0.21)					(0.10)	
		LYSA56	ASPC48	(0.24)	LYSB56	ASPA48	(0.24)	LYSC56	ASPB48	(0.23)	(0.36)		
		LYSA56	GLUC50	(1.50)	LYSB56	GLUA50	(2.07)	LYSC56	GLUB50	(1.69)	(2.63)		
							LYSC56	LYSB70	0.07	0.04			
LYSA56		META1	0.13	LYSB56	METB1	0.13	LYSC56	METC1	0.14	0.20			
			0.00	LYSB56	ARB25	0.07					0.04		
LYSA56		ARGA33	0.43	LYSB56	ARGB33	0.34	LYSC56	ARGC33	0.39	0.58			
LYSA56		GLUA59	(0.09)	LYSB56	GLUB59	(0.09)	LYSC56	GLUC59	(0.09)	(0.13)			
							LYSC56	ASPC86	(0.08)	(0.04)			
							LYSC56	ASPC87	(0.08)	(0.04)			
LYSA56		ASPA91	(0.17)	LYSB56	ASPB91	(0.21)	LYSC56	ASPC91	(0.23)	(0.31)			
LYSA56		GLUA99	(0.22)	LYSB56	GLUB99	(0.18)	LYSC56	GLUC99	(0.16)	(0.28)	(1.44)		
Arg 58	ARGA58	META1	0.59	ARGB58	METB1	0.60	ARGC58	METC1	0.62	0.90			
	ARGA58	GLUA24	(0.31)	ARGB58	GLUB24	(0.39)	ARGC58	GLUC24	(0.24)	(0.47)			
	ARGA58	ARGA25	0.89	ARGB58	ARB25	0.90	ARGC58	ARGC25	0.92	1.35			
	ARGA58	GLUA59	(1.61)	ARGB58	GLUB59	(1.65)	ARGC58	GLUC59	(1.79)	(2.52)			
	ARGA58	ASPA60	(10.05)	ARGB58	ASPB60	(10.33)	ARGC58	ASPC60	(10.39)	(15.38)			
	ARGA58	GLUA63	(0.66)	ARGB58	GLUB63	(0.73)	ARGC58	GLUC63	(0.85)	(1.12)			
	ARGA58	GLUA64	(0.87)	ARGB58	GLUB64	(0.76)	ARGC58	GLUC64	(0.71)	(1.17)			
	ARGA58	GLUA67	(0.07)							(0.04)			
	ARGA58	ARGA68	0.05							0.03			
	ARGA58	ARGA82	0.67	ARGB58	ARB82	0.64	ARGC58	ARGC82	0.66	0.99			
	ARGA58	ASPA84	(0.46)	ARGB58	ASPB84	(0.46)	ARGC58	ASPC84	(0.46)	(0.69)			
	ARGA58	LYSA94	0.43	ARGB58	LYSB94	0.41	ARGC58	LYSC94	0.17	0.50			
	ARGA58	GLUA98	(0.66)	ARGB58	GLUB98	(0.66)	ARGC58	GLUC98	(0.82)	(1.07)			
	ARGA58	GLUA99	(0.70)	ARGB58	GLUB99	(0.68)	ARGC58	GLUC99	(0.74)	(1.06)			
	ARGA58	LYSA101	0.65	ARGB58	LYSB101	1.14	ARGC58	LYSC101	0.65	1.22			
	ARGA58	LYSA102	(6.82)	ARGB58	LYSB102	(0.19)	ARGC58	LYSC102	0.56	(3.23)	(21.76)		
	Glu 59	GLUA59	ASPB86	0.63	GLUB59	ASPC86	0.64	GLUC59	ASPA86	0.65	0.96		
		GLUA59	ASPB87	0.19	GLUB59	ASPC87	0.29	GLUC59	ASPA87	0.28	0.38		
		GLUA59	LYSC66	(0.20)	GLUB59	LYSA66	(0.22)				(0.21)		
		GLUA59	ARGC82	(0.29)	GLUB59	ARGA82	(0.28)	GLUC59	ARGB82	(0.29)	(0.43)		
GLUA59		ASPC84	0.25	GLUB59	ASPA84	0.25	GLUC59	ASPB84	0.24	0.37			
GLUA59		META1	(8.46)	GLUB59	METB1	(8.57)	GLUC59	METC1	(8.61)	(12.82)			
GLUA59		LYSA56	(0.09)	GLUB59	LYSB56	(0.09)	GLUC59	LYSC56	(0.09)	(0.13)			
GLUA59		ARGA58	(1.61)	GLUB59	ARGB58	(1.65)	GLUC59	ARGC58	(1.79)	(2.52)			
GLUA59		ASPA60	2.11	GLUB59	ASPB60	2.13	GLUC59	ASPC60	2.39	3.31			
GLUA59		GLUA63	0.62	GLUB59	GLUB63	0.64	GLUC59	GLUC63	0.72	0.99			
GLUA59		GLUA64	0.08							0.04			
GLUA59		ARGA82	(1.95)	GLUB59	ARB82	(1.61)	GLUC59	ARGC82	(1.90)	(2.73)			
GLUA59		ASPA84	2.12	GLUB59	ASPB84	2.03	GLUC59	ASPC84	2.02	3.08			
GLUA59		ASPA86	0.61	GLUB59	ASPB86	0.60	GLUC59	ASPC86	0.58	0.89			
GLUA59		ASPA87	0.63	GLUB59	ASPB87	0.62	GLUC59	ASPC87	0.59	0.92			
GLUA59		GLUA90	0.44	GLUB59	GLUB90	0.39	GLUC59	GLUC90	0.17	0.50			
GLUA59		ASPA91	0.08	GLUB59	ASPB91	0.07	GLUC59	ASPC91	0.08	0.12			
GLUA59		LYSA94	(0.34)	GLUB59	LYSB94	(0.33)				(0.33)			
GLUA59		GLUA98	0.08	GLUB59	GLUB98	0.09	GLUC59	GLUC98	0.17	0.17			
				GLUB59	LYSB101	(0.08)				(0.04)			
GLUA59		LYSA102	0.44	GLUB59	LYSB102	(0.05)	GLUC59	LYSC102	(0.01)	0.19	(7.30)		

Asp 60	ASPA60	META1	(0.54)	ASPB60	METB1	(0.55)	ASPC60	METC1	(0.58)	(0.83)		
	ASPA60	GLUA24	0.08	ASPB60	GLUB24	0.17	ASPC60	GLUC24	0.09	0.17		
	ASPA60	ARGA25	(0.31)	ASPB60	ARGB25	(0.32)	ASPC60	ARGC25	(0.32)	(0.47)		
	ASPA60	ARGA58	(10.05)	ASPB60	ARGB58	(10.33)	ASPC60	ARGC58	(10.39)	(15.38)		
	ASPA60	GLUA59	2.11	ASPB60	GLUB59	2.13	ASPC60	GLUC59	2.39	3.31		
	ASPA60	GLUA63	1.26	ASPB60	GLUB63	1.48	ASPC60	GLUC63	2.00	2.37		
	ASPA60	GLUA64	0.96	ASPB60	GLUB64	0.79	ASPC60	GLUC64	0.73	1.24		
				ASPB60	LYSB66	(0.17)					(0.08)	
	ASPA60	GLUA67	0.31	ASPB60	GLUB67	0.08	ASPC60	GLUC67	0.08	0.24		
	ASPA60	ARGA82	(0.88)	ASPB60	ARGB82	(0.83)	ASPC60	ARGC82	(0.88)	(1.29)		
	ASPA60	ASP84	0.51	ASPB60	ASP84	0.52	ASPC60	ASPC84	0.51	0.77		
	ASPA60	LYSA94	(0.24)	ASPB60	LYSB94	(0.22)					(0.23)	
	ASPA60	GLUA98	0.16	ASPB60	GLUB98	0.16	ASPC60	GLUC98	0.30	0.31		
	ASPA60	GLUA99	0.08	ASPB60	GLUB99	0.08	ASPC60	GLUC99	0.23	0.20		
	ASPA60	LYSA101	(0.37)	ASPB60	LYSB101	(0.82)	ASPC60	LYSC101	(0.44)	(0.81)		
	ASPA60	LYSA102	1.32	ASPB60	LYSB102	0.78	ASPC60	LYSC102	0.80	1.45		
	ASPA60	ASP87	0.18	ASPB60	ASPC87	0.20	ASPC60	ASP87	0.19	0.28		
											(8.75)	
	Glu 63	GLUA63	ASPB86	0.48	GLUB63	ASPC86	0.47	GLUC63	ASP86	0.41	0.68	
		GLUA63	ASP87	2.27	GLUB63	ASPC87	2.29	GLUC63	ASP87	1.27	2.92	
		GLUA63	GLUB90	0.74	GLUB63	GLUC90	0.61	GLUC63	GLUA90	0.57	0.96	
		GLUA63	ASP91	0.17	GLUB63	ASPC91	0.16					0.16
		GLUA63	META1	(0.22)	GLUB63	METB1	(0.23)	GLUC63	METC1	(0.25)	(0.35)	
					GLUB63	GLUB24	0.08	GLUC63	GLUC24	0.08	0.08	
		GLUA63	ARGA58	(0.66)	GLUB63	ARGB58	(0.73)	GLUC63	ARGC58	(0.85)	(1.12)	
		GLUA63	GLUA59	0.62	GLUB63	GLUB59	0.64	GLUC63	GLUC59	0.72	0.99	
		GLUA63	ASP60	1.26	GLUB63	ASP60	1.48	GLUC63	ASPC60	2.00	2.37	
		GLUA63	GLUA64	0.87	GLUB63	GLUB64	0.90	GLUC63	GLUC64	1.04	1.40	
GLUA63		LYSA66	(0.92)	GLUB63	LYSB66	(1.17)	GLUC63	LYSC66	(0.51)	(1.30)		
GLUA63		GLUA67	1.51	GLUB63	GLUB67	0.80	GLUC63	GLUC67	0.82	1.57		
GLUA63		ARGA68	(0.07)	GLUB63	ARGB68	(0.07)	GLUC63	ARGC68	(0.07)	(0.10)		
GLUA63		LYSA70	(0.19)	GLUB63	LYSB70	(0.16)					(0.18)	
GLUA63		GLUA71	0.17								0.08	
GLUA63		ARGA82	(2.88)	GLUB63	ARGB82	(2.88)	GLUC63	ARGC82	(2.31)	(4.03)		
GLUA63		ASP84	0.77	GLUB63	ASP84	0.78	GLUC63	ASPC84	0.69	1.12		
GLUA63		LYSA101	(0.10)	GLUB63	LYSB101	(0.29)	GLUC63	LYSC101	(0.27)	(0.33)		
GLUA63		LYSA102	(0.22)	GLUB63	LYSB102	0.17	GLUC63	LYSC102	0.40	0.17		
											5.10	
Glu 64		GLUA64	ASP87	0.17	GLUB64	ASPC87	0.10	GLUC64	ASP87	0.10	0.18	
					GLUB64	GLUC90	0.19	GLUC64	GLUA90	0.17	0.18	
		GLUA64	GLUA24	1.45	GLUB64	GLUB24	1.48	GLUC64	GLUC24	1.54	2.23	
		GLUA64	ARGA25	(0.16)	GLUB64	ARGB25	(0.15)	GLUC64	ARGC25	(0.15)	(0.23)	
		GLUA64	ARGA58	(0.87)	GLUB64	ARGB58	(0.76)	GLUC64	ARGC58	(0.71)	(1.17)	
		GLUA64	GLUA59	0.08							0.04	
		GLUA64	ASP60	0.96	GLUB64	ASP60	0.79	GLUC64	ASPC60	0.73	1.24	
		GLUA64	GLUA63	0.87	GLUB64	GLUB63	0.90	GLUC64	GLUC63	1.04	1.40	
	GLUA64	LYSA66	(0.20)	GLUB64	LYSB66	(0.40)	GLUC64	LYSC66	(0.22)	(0.41)		
	GLUA64	GLUA67	2.12	GLUB64	GLUB67	1.66	GLUC64	GLUC67	1.37	2.57		
	GLUA64	ARGA68	(1.37)	GLUB64	ARGB68	(1.38)	GLUC64	ARGC68	(1.58)	(2.17)		
	GLUA64	LYSA70	(0.09)	GLUB64	LYSB70	(0.10)	GLUC64	LYSC70	(0.09)	(0.14)		
	GLUA64	GLUA71	1.30	GLUB64	GLUB71	0.93	GLUC64	GLUC71	1.03	1.63		
	GLUA64	ARGA82	(0.29)	GLUB64	ARGB82	(0.22)	GLUC64	ARGC82	(0.22)	(0.37)		
	GLUA64	LYSA101	(1.97)	GLUB64	LYSB101	(3.79)	GLUC64	LYSC101	(4.31)	(5.04)		
	GLUA64	LYSA102	(2.57)	GLUB64	LYSB102	0.52	GLUC64	LYSC102	0.56	(0.75)		
											(0.79)	
	Lys 66	LYSA66	METB1	0.19				LYSC66	META1	0.17	0.18	
				0.00	LYSB66	GLUB24	(0.08)				(0.04)	
		LYSA66	GLUB59	(0.22)				LYSC66	GLUA59	(0.20)	(0.21)	
					LYSB66	ASP60	(0.17)					(0.08)
		LYSA66	ASPB86	(0.29)	LYSB66	ASPC86	(0.16)	LYSC66	ASP86	(0.25)	(0.35)	
		LYSA66	ASP87	(8.28)	LYSB66	ASPC87	(1.55)	LYSC66	ASP87	(4.58)	(7.20)	
		LYSA66	GLUB90	(5.83)	LYSB66	GLUC90	(4.29)	LYSC66	GLUA90	(7.91)	(9.01)	
		LYSA66	ASP91	(1.02)	LYSB66	ASPC91	(0.80)	LYSC66	ASP91	(1.21)	(1.52)	
		LYSA66	LYSB94	0.13				LYSC66	LYSA94	0.14	0.14	
		LYSA66	GLUA63	(0.92)	LYSB66	GLUB63	(1.17)	LYSC66	GLUC63	(0.51)	(1.30)	
		LYSA66	GLUA64	(0.20)	LYSB66	GLUB64	(0.40)	LYSC66	GLUC64	(0.22)	(0.41)	
LYSA66		GLUA67	(0.71)	LYSB66	GLUB67	(3.80)	LYSC66	GLUC67	(1.41)	(2.96)		
				LYSB66	ARGB68	0.17					0.08	
LYSA66		LYSA70	0.60	LYSB66	LYSB70	0.61	LYSC66	LYSC70	0.41	0.81		
LYSA66		GLUA71	(0.20)	LYSB66	GLUB71	(0.21)	LYSC66	GLUC71	(0.08)	(0.24)		
				LYSB66	ASP76	(0.08)					(0.04)	
LYSA66		ARGA82	0.73	LYSB66	ARGB82	0.51	LYSC66	ARGC82	0.62	0.93		
LYSA66		ASP84	(0.29)	LYSB66	ASP84	(0.19)	LYSC66	ASPC84	(0.26)	(0.37)		
				LYSB66	LYSB101	0.07					0.04	
				LYSB66	METC1	0.08					0.04	
											(21.51)	

Glu 67	GLUA67	ASPB87	0.66	GLUB67	ASPC87	0.57	GLUC67	ASPA87	0.72	0.98	
	GLUA67	GLUB90	0.97	GLUB67	GLUC90	1.90	GLUC67	GLUA90	1.73	2.30	
	GLUA67	ASPB91	0.56	GLUB67	ASPC91	0.86	GLUC67	ASPA91	0.98	1.20	
	GLUA67	GLUA24	0.47	GLUB67	GLUB24	0.46	GLUC67	GLUC24	0.39	0.66	
	GLUA67	ARGA58	(0.07)								(0.04)
	GLUA67	ASPA60	0.31	GLUB67	ASPB60	0.08	GLUC67	ASPC60	0.08	0.24	
	GLUA67	GLUA63	1.51	GLUB67	GLUB63	0.80	GLUC67	GLUC63	0.82	1.57	
	GLUA67	GLUA64	2.12	GLUB67	GLUB64	1.66	GLUC67	GLUC64	1.37	2.57	
	GLUA67	LYSA66	(0.71)	GLUB67	LYSB66	(3.80)	GLUC67	LYSC66	(1.41)	(2.96)	
	GLUA67	ARGA68	(0.76)	GLUB67	ARGB68	(0.80)	GLUC67	ARGC68	(0.70)	(1.13)	
	GLUA67	LYSA70	(0.66)	GLUB67	LYSB70	(1.12)	GLUC67	LYSC70	(1.12)	(1.45)	
	GLUA67	GLUA71	2.04	GLUB67	GLUB71	1.25	GLUC67	GLUC71	1.01	2.15	
				GLUB67	ASPB76	0.20	GLUC67	ASPC76	0.08	0.14	
				GLUB67	ARGB82	(0.24)	GLUC67	ARGC82	(0.34)	(0.50)	
	GLUA67	ARGA82	(0.43)	GLUB67	LYSB101	(0.22)	GLUC67	LYSC101	(0.21)	(0.33)	
	GLUA67	LYSA101	(0.23)							(0.15)	
	GLUA67	LYSA102	(0.30)								5.23
Arg 68	ARGA68	LYSA16	0.79	ARGB68	LYSB16	0.84	ARGC68	LYSC16	0.84	1.24	
	ARGA68	LYSA19	0.08	ARGB68	LYSB19	0.08	ARGC68	LYSC19	0.08	0.12	
	ARGA68	LYSA23	0.35	ARGB68	LYSB23	0.27	ARGC68	LYSC23	0.19	0.40	
	ARGA68	GLUA24	(11.27)	ARGB68	GLUB24	(10.70)	ARGC68	GLUC24	(11.30)	(16.64)	
	ARGA68	ARGA25	0.13	ARGB68	ARGB25	0.12	ARGC68	ARGC25	0.12	0.18	
	ARGA68	ARGA58	0.05							0.03	
	ARGA68	GLUA63	(0.07)	ARGB68	GLUB63	(0.07)	ARGC68	GLUC63	(0.07)	(0.10)	
	ARGA68	GLUA64	(1.37)	ARGB68	GLUB64	(1.38)	ARGC68	GLUC64	(1.58)	(2.17)	
				ARGB68	LYSB66	0.17				0.08	
	ARGA68	GLUA67	(0.76)	ARGB68	GLUB67	(0.80)	ARGC68	GLUC67	(0.70)	(1.13)	
	ARGA68	LYSA70	0.19	ARGB68	LYSB70	0.19	ARGC68	LYSC70	0.19	0.28	
	ARGA68	GLUA71	(6.47)	ARGB68	GLUB71	(8.07)	ARGC68	GLUC71	(8.28)	(11.41)	
				ARGB68	ASPB76	(0.07)				(0.03)	
	ARGA68	LYSA101	0.73	ARGB68	LYSB101	0.48	ARGC68	LYSC101	0.85	1.03	
ARGA68	LYSA102	0.29							0.14	(27.97)	
Lys 70				LYSB70	LYSB16	0.08				0.04	
	LYSA70	GLUA63	(0.19)	LYSB70	GLUB63	(0.16)				(0.18)	
	LYSA70	GLUA64	(0.09)	LYSB70	GLUB64	(0.10)	LYSC70	GLUC64	(0.09)	(0.14)	
	LYSA70	LYSA66	0.60	LYSB70	LYSB66	0.61	LYSC70	LYSC66	0.41	0.81	
	LYSA70	GLUA67	(0.66)	LYSB70	GLUB67	(1.12)	LYSC70	GLUC67	(1.12)	(1.45)	
	LYSA70	ARGA68	0.19	LYSB70	ARGB68	0.19	LYSC70	ARGC68	0.19	0.28	
	LYSA70	GLUA71	(0.45)	LYSB70	GLUB71	(0.30)	LYSC70	GLUC71	(0.32)	(0.54)	
	LYSA70	ASPA76	(0.57)	LYSB70	ASPB76	(0.86)	LYSC70	ASPC76	(0.80)	(1.11)	
				LYSB70	LYSC56	0.07				0.04	
	LYSA70	ASPB87	(0.36)	LYSB70	ASPC87	(0.27)	LYSC70	ASPA87	(0.23)	(0.43)	
	LYSA70	GLUB90	(1.65)	LYSB70	GLUC90	(1.89)	LYSC70	GLUA90	(0.77)	(2.15)	
	LYSA70	ASPB91	(7.15)	LYSB70	ASPC91	(8.61)	LYSC70	ASPA91	(6.90)	(11.33)	
	LYSA70	LYSB94	0.14	LYSB70	LYSC94	0.19	LYSC70	LYSA94	0.10	0.22	
	LYSA70	GLUB98	(0.19)	LYSB70	GLUC98	(0.08)	LYSC70	GLUA98	(0.20)	(0.23)	(16.18)
Glu 71	GLUA71	LYSA16	(0.31)	GLUB71	LYSB16	(0.56)	GLUC71	LYSC16	(0.33)	(0.60)	
	GLUA71	GLUA24	1.72	GLUB71	GLUB24	1.83	GLUC71	GLUC24	1.90	2.72	
	GLUA71	GLUA63	0.17							0.08	
	GLUA71	GLUA64	1.30	GLUB71	GLUB64	0.93	GLUC71	GLUC64	1.03	1.63	
	GLUA71	LYSA66	(0.20)	GLUB71	LYSB66	(0.21)	GLUC71	LYSC66	(0.08)	(0.24)	
	GLUA71	GLUA67	2.04	GLUB71	GLUB67	1.25	GLUC71	GLUC67	1.01	2.15	
	GLUA71	ARGA68	(6.47)	GLUB71	ARGB68	(8.07)	GLUC71	ARGC68	(8.28)	(11.41)	
	GLUA71	LYSA70	(0.45)	GLUB71	LYSB70	(0.30)	GLUC71	LYSC70	(0.32)	(0.54)	
	GLUA71	ASPA76	0.17	GLUB71	ASPB76	0.36	GLUC71	ASPC76	0.18	0.36	
	GLUA71	LYSA101	(0.30)	GLUB71	LYSB101	(0.22)	GLUC71	LYSC101	(0.32)	(0.42)	
	GLUA71	LYSA102	(0.19)							(0.09)	
	GLUA71	GLUB90	0.26	GLUB71	GLUC90	0.08				0.17	
	GLUA71	ASPB91	0.44	GLUB71	ASPC91	0.26	GLUC71	ASPA91	0.26	0.48	(5.71)
	Asp 76	ASPA76	GLUB90	0.24	ASPB76	GLUC90	0.29				0.27
ASPA76		ASPB91	1.55	ASPB76	ASPC91	1.33	ASPC76	ASPA91	1.08	1.98	
ASPA76		LYSB94	(0.08)	ASPB76	LYSC94	(0.34)				(0.21)	
ASPA76		GLUB98	0.53				ASPC76	GLUA98	0.28	0.41	
ASPA76		GLUB99	0.30				ASPC76	GLUA99	0.08	0.19	
							ASPC76	ASPC10	0.18	0.09	
				ASPB76	LYSB16	(0.20)					(0.10)
ASPA76		GLUA42	0.91	ASPB76	GLUB42	0.65	ASPC76	GLUC42	0.09	0.82	
				ASPB76	LYSB44	(0.19)				(0.09)	
ASPA76		GLUA46	0.43	ASPB76	GLUB46	0.17				0.30	
ASPA76		ASPA48	0.08	ASPB76	ASPB48	0.08				0.08	
ASPA76		GLUA50	0.19	ASPB76	GLUB50	0.18	ASPC76	GLUC50	0.20	0.29	
				ASPB76	LYSB66	(0.08)				(0.04)	
				ASPB76	GLUB67	0.20	ASPC76	GLUC67	0.08	0.14	
			ASPB76	ARGB68	(0.07)				(0.03)		
ASPA76	LYSA70	(0.57)	ASPB76	LYSB70	(0.86)	ASPC76	LYSC70	(0.80)	(1.11)		
ASPA76	GLUA71	0.17	ASPB76	GLUB71	0.36	ASPC76	GLUC71	0.18	0.36	3.32	

Arg 82	ARGA82	META1	0.86	ARGB82	METB1	0.75	ARGC82	METC1	0.87	1.24		
	ARGA82	ARGA58	0.67	ARGB82	ARGB58	0.64	ARGC82	ARGC58	0.66	0.99		
	ARGA82	GLUA59	(1.95)	ARGB82	GLUB59	(1.61)	ARGC82	GLUC59	(1.90)	(2.73)		
	ARGA82	ASPA60	(0.88)	ARGB82	ASPB60	(0.83)	ARGC82	ASPC60	(0.88)	(1.29)		
	ARGA82	GLUA63	(2.88)	ARGB82	GLUB63	(2.88)	ARGC82	GLUC63	(2.31)	(4.03)		
	ARGA82	GLUA64	(0.29)	ARGB82	GLUB64	(0.22)	ARGC82	GLUC64	(0.22)	(0.37)		
	ARGA82	LYSA66	0.73	ARGB82	LYSB66	0.51	ARGC82	LYSC66	0.62	0.93		
	ARGA82	GLUA67	(0.43)	ARGB82	GLUB67	(0.24)	ARGC82	GLUC67	(0.34)	(0.50)		
	ARGA82	ASPA84	(10.49)	ARGB82	ASPB84	(10.37)	ARGC82	ASPC84	(10.38)	(15.62)		
	ARGA82	ASPA86	(0.74)	ARGB82	ASPB86	(0.73)	ARGC82	ASPC86	(0.72)	(1.10)		
	ARGA82	ASPA87	(0.07)				ARGC82	ASPC87	(0.07)	(0.07)		
	ARGA82	METB1	0.30	ARGB82	METC1	0.32	ARGC82	META1	0.31	0.47		
	ARGA82	GLUB59	(0.28)	ARGB82	GLUC59	(0.29)	ARGC82	GLUA59	(0.29)	(0.43)		
	ARGA82	ASPB84	(0.49)	ARGB82	ASPC84	(0.51)	ARGC82	ASPA84	(0.51)	(0.76)		
	ARGA82	ASPB86	(3.44)	ARGB82	ASPC86	(3.70)	ARGC82	ASPA86	(3.65)	(5.40)		
	ARGA82	ASPB87	(2.75)	ARGB82	ASPC87	(4.04)	ARGC82	ASPA87	(3.03)	(4.91)		
	ARGA82	GLUB90	(0.57)	ARGB82	GLUC90	(0.46)	ARGC82	GLUA90	(0.66)	(0.84)		
				ARGB82	ASPC91	(0.14)					(0.07)	
		ARGA82	ASPC84	(0.29)	ARGB82	ASPA84	(0.21)	ARGC82	ASPB84	(0.28)	(0.39)	
		ARGA82	ASPC86	(0.54)	ARGB82	ASPA86	(0.53)	ARGC82	ASPB86	(0.52)	(0.79)	(35.68)
	Asp 84	ASPA84	METB1	(0.23)	ASPB84	METC1	(0.23)	ASPC84	META1	(0.24)	(0.35)	
		ASPA84	GLUB59	0.25	ASPB84	GLUC59	0.24	ASPC84	GLUA59	0.25	0.37	
		ASPA84	ARGB82	(0.21)	ASPB84	ARGC82	(0.28)	ASPC84	ARGA82	(0.29)	(0.39)	
ASPA84		ASPB84	0.54	ASPB84	ASPC84	0.55	ASPC84	ASPA84	0.57	0.83		
ASPA84		ASPB86	9.00	ASPB84	ASPC86	9.47	ASPC84	ASPA86	10.02	14.25		
ASPA84		ASPB87	1.03	ASPB84	ASPC87	1.23	ASPC84	ASPA87	1.08	1.67		
ASPA84		GLUB90	0.09	ASPB84	GLUC90	0.08	ASPC84	GLUA90	0.26	0.21		
ASPA84		ARGC82	(0.51)	ASPB84	ARGA82	(0.49)	ASPC84	ARGB82	(0.51)	(0.76)		
ASPA84		ASPC84	0.57	ASPB84	ASPA84	0.54	ASPC84	ASPB84	0.55	0.83		
ASPA84		ASPC86	0.71	ASPB84	ASPA86	0.67	ASPC84	ASPB86	0.70	1.04		
ASPA84		META1	(1.19)	ASPB84	METB1	(1.18)	ASPC84	METC1	(1.23)	(1.80)		
ASPA84		ARGA58	(0.46)	ASPB84	ARGB58	(0.46)	ASPC84	ARGC58	(0.46)	(0.69)		
ASPA84		GLUA59	2.12	ASPB84	GLUB59	2.03	ASPC84	GLUC59	2.02	3.08		
ASPA84		ASPA60	0.51	ASPB84	ASPB60	0.52	ASPC84	ASPC60	0.51	0.77		
ASPA84		GLUA63	0.77	ASPB84	GLUB63	0.78	ASPC84	GLUC63	0.69	1.12		
ASPA84		LYSA66	(0.29)	ASPB84	LYSB66	(0.19)	ASPC84	LYSC66	(0.26)	(0.37)		
ASPA84		ARGA82	(10.49)	ASPB84	ARGB82	(10.37)	ASPC84	ARGC82	(10.38)	(15.62)		
ASPA84		ASPA86	1.62	ASPB84	ASPB86	1.61	ASPC84	ASPC86	1.57	2.40		
ASPA84		ASPA87	0.41	ASPB84	ASPB87	0.39	ASPC84	ASPC87	0.41	0.61	7.20	
Asp 86		ASPA86	METC1	(0.40)	ASPB86	META1	(0.38)	ASPC86	METB1	(0.40)	(0.59)	
	ASPA86	GLUC59	0.65	ASPB86	GLUA59	0.63	ASPC86	GLUB59	0.64	0.96		
	ASPA86	GLUC63	0.41	ASPB86	GLUA63	0.48	ASPC86	GLUB63	0.47	0.68		
	ASPA86	LYSC66	(0.25)	ASPB86	LYSA66	(0.29)	ASPC86	LYSB66	(0.16)	(0.35)		
	ASPA86	ARGC82	(3.65)	ASPB86	ARGA82	(3.44)	ASPC86	ARGB82	(3.70)	(5.40)		
	ASPA86	ASPC84	10.02	ASPB86	ASPA84	9.00	ASPC86	ASPB84	9.47	14.25		
	ASPA86	ASPC86	1.96	ASPB86	ASPA86	1.95	ASPC86	ASPB86	2.18	3.05		
	ASPA86	ASPC87	0.24	ASPB86	ASPA87	0.16	ASPC86	ASPB87	0.24	0.32		
	ASPA86	META1	(0.49)	ASPB86	METB1	(0.49)	ASPC86	METC1	(0.47)	(0.72)		
							ASPC86	LYSC56	(0.08)	(0.04)		
	ASPA86	GLUA59	0.61	ASPB86	GLUB59	0.60	ASPC86	GLUC59	0.58	0.89		
	ASPA86	ARGA82	(0.74)	ASPB86	ARGB82	(0.73)	ASPC86	ARGC82	(0.72)	(1.10)		
	ASPA86	ASPA84	1.62	ASPB86	ASPB84	1.61	ASPC86	ASPC84	1.57	2.40		
	ASPA86	ASPA87	1.12	ASPB86	ASPB87	1.06	ASPC86	ASPC87	1.17	1.68		
	ASPA86	GLUA90	0.34	ASPB86	GLUB90	0.17				0.25		
	ASPA86	ARGB82	(0.53)	ASPB86	ARGC82	(0.52)	ASPC86	ARGA82	(0.54)	(0.79)		
	ASPA86	ASPB84	0.67	ASPB86	ASPC84	0.70	ASPC86	ASPA84	0.71	1.04		
	ASPA86	ASPB86	1.95	ASPB86	ASPC86	2.18	ASPC86	ASPA86	1.96	3.05	19.58	
	Asp 87	ASPA87	METC1	(0.09)	ASPB87	META1	(0.08)	ASPC87	METB1	(0.09)	(0.13)	
ASPA87		GLUC59	0.28	ASPB87	GLUA59	0.19	ASPC87	GLUB59	0.29	0.38		
ASPA87		ASPC60	0.19	ASPB87	ASPA60	0.18	ASPC87	ASPB60	0.20	0.28		
ASPA87		GLUC63	1.27	ASPB87	GLUA63	2.27	ASPC87	GLUB63	2.29	2.92		
ASPA87		GLUC64	0.10	ASPB87	GLUA64	0.17	ASPC87	GLUB64	0.10	0.18		
ASPA87		LYSC66	(4.58)	ASPB87	LYSA66	(8.28)	ASPC87	LYSB66	(1.55)	(7.20)		
ASPA87		GLUC67	0.72	ASPB87	GLUA67	0.66	ASPC87	GLUB67	0.57	0.98		
ASPA87		LYSC70	(0.23)	ASPB87	LYSA70	(0.36)	ASPC87	LYSB70	(0.27)	(0.43)		
ASPA87		ARGC82	(3.03)	ASPB87	ARGA82	(2.75)	ASPC87	ARGB82	(4.04)	(4.91)		
ASPA87		ASPC84	1.08	ASPB87	ASPA84	1.03	ASPC87	ASPB84	1.23	1.67		
ASPA87		ASPB86	0.16	ASPB87	ASPC86	0.24	ASPC87	ASPA86	0.24	0.32		
ASPA87		META1	(0.62)	ASPB87	METB1	(0.60)	ASPC87	METC1	(0.57)	(0.89)		
							ASPC87	LYSC56	(0.08)	(0.04)		
ASPA87		GLUA59	0.63	ASPB87	GLUB59	0.62	ASPC87	GLUC59	0.59	0.92		
ASPA87		ARGA82	(0.07)				ASPC87	ARGC82	(0.07)	(0.07)		
ASPA87		ASPA84	0.41	ASPB87	ASPB84	0.39	ASPC87	ASPC84	0.41	0.61		
ASPA87		ASPA86	1.12	ASPB87	ASPB86	1.06	ASPC87	ASPC86	1.17	1.68		
ASPA87		GLUA90	4.17	ASPB87	GLUB90	3.21	ASPC87	GLUC90	1.68	4.53		
ASPA87		ASPA91	0.62	ASPB87	ASPB91	0.70	ASPC87	ASPC91	0.60	0.96		
ASPA87	LYSA94	(0.24)	ASPB87	LYSB94	(0.23)				(0.23)	1.50		

Glu 90	GLUA90	GLUC63	0.57	GLUB90	GLUA63	0.74	GLUC90	GLUB63	0.61	0.96		
	GLUA90	GLUC64	0.17				GLUC90	GLUB64	0.19	0.18		
	GLUA90	LYSC66	(7.91)	GLUB90	LYSA66	(5.83)	GLUC90	LYSB66	(4.29)	(9.01)		
	GLUA90	GLUC67	1.73	GLUB90	GLUA67	0.97	GLUC90	GLUB67	1.90	2.30		
	GLUA90	LYSC70	(0.77)	GLUB90	LYSA70	(1.65)	GLUC90	LYSB70	(1.89)	(2.15)		
				GLUB90	GLUA71	0.26	GLUC90	GLUB71	0.08	0.17		
				GLUB90	ASPA76	0.24	GLUC90	ASPB76	0.29	0.27		
	GLUA90	ARGC82	(0.66)	GLUB90	ARGA82	(0.57)	GLUC90	ARGB82	(0.46)	(0.84)		
	GLUA90	ASPC84	0.26	GLUB90	ASPA84	0.09	GLUC90	ASPB84	0.08	0.21		
	GLUA90	META1	(0.33)	GLUB90	METB1	(0.28)	GLUC90	METC1	(0.22)	(0.42)		
	GLUA90	GLUA59	0.44	GLUB90	GLUB59	0.39	GLUC90	GLUC59	0.17	0.50		
	GLUA90	ASPA86	0.34	GLUB90	ASPB86	0.17				0.25		
	GLUA90	ASPA87	4.17	GLUB90	ASPB87	3.21	GLUC90	ASPC87	1.68	4.53		
	GLUA90	ASPA91	2.33	GLUB90	ASPB91	2.60	GLUC90	ASPC91	3.45	4.19		
	GLUA90	LYSA94	(0.54)	GLUB90	LYSB94	(0.56)	GLUC90	LYSC94	(0.28)	(0.69)		
	GLUA90	GLUA98	0.20	GLUB90	GLUB98	0.20	GLUC90	GLUC98	0.24	0.32	0.76	
	Asp 91				ASPB91	GLUA42	0.09	ASPC91	GLUB42	0.17	0.13	
		ASPA91	GLUC50	0.26	ASPB91	GLUA50	0.26	ASPC91	GLUB50	0.26	0.39	
					ASPB91	GLUA63	0.17	ASPC91	GLUB63	0.16	0.16	
		ASPA91	LYSC66	(1.21)	ASPB91	LYSA66	(1.02)	ASPC91	LYSB66	(0.80)	(1.52)	
ASPA91		GLUC67	0.98	ASPB91	GLUA67	0.56	ASPC91	GLUB67	0.86	1.20		
ASPA91		LYSC70	(6.90)	ASPB91	LYSA70	(7.15)	ASPC91	LYSB70	(8.61)	(11.33)		
ASPA91		GLUC71	0.26	ASPB91	GLUA71	0.44	ASPC91	GLUB71	0.26	0.48		
ASPA91		ASPC76	1.08	ASPB91	ASPA76	1.55	ASPC91	ASPB76	1.33	1.98		
							ASPC91	ARGB82	(0.14)	(0.07)		
ASPA91		META1	(0.19)	ASPB91	METB1	(0.20)	ASPC91	METC1	(0.11)	(0.25)		
ASPA91		LYSA56	(0.17)	ASPB91	LYSB56	(0.21)	ASPC91	LYSC56	(0.23)	(0.31)		
ASPA91		GLUA59	0.08	ASPB91	GLUB59	0.07	ASPC91	GLUC59	0.08	0.12		
ASPA91		ASPA87	0.62	ASPB91	ASPB87	0.70	ASPC91	ASPC87	0.60	0.96		
ASPA91		GLUA90	2.33	ASPB91	GLUB90	2.60	ASPC91	GLUC90	3.45	4.19		
ASPA91		LYSA94	(0.36)	ASPB91	LYSB94	(0.39)	ASPC91	LYSC94	(0.63)	(0.69)		
ASPA91		GLUA98	0.62	ASPB91	GLUB98	0.60	ASPC91	GLUC98	0.42	0.82		
ASPA91		GLUA99	0.17	ASPB91	GLUB99	0.17	ASPC91	GLUC99	0.18	0.26	(3.48)	
Lys 94								LYSC94	GLUB42	(0.54)	(0.27)	
								LYSC94	LYSB44	0.10	0.05	
		LYSA94	LYSC66	0.14	LYSB94	LYSA66	0.13				0.14	
	LYSA94	LYSC70	0.10	LYSB94	LYSA70	0.14	LYSC94	LYSB70	0.19	0.22		
				LYSB94	ASPA76	(0.08)	LYSC94	ASPB76	(0.34)	(0.21)		
	LYSA94	META1	0.18	LYSB94	METB1	0.17				0.17		
							LYSC94	ARGC25	0.32	0.16		
	LYSA94	ARGA58	0.43	LYSB94	ARGB58	0.41	LYSC94	ARGC58	0.17	0.50		
	LYSA94	GLUA59	(0.34)	LYSB94	GLUB59	(0.33)				(0.33)		
	LYSA94	ASPA60	(0.24)	LYSB94	ASPB60	(0.22)				(0.23)		
	LYSA94	ASPA87	(0.24)	LYSB94	ASPB87	(0.23)				(0.23)		
	LYSA94	GLUA90	(0.54)	LYSB94	GLUB90	(0.56)	LYSC94	GLUC90	(0.28)	(0.69)		
	LYSA94	ASPA91	(0.36)	LYSB94	ASPB91	(0.39)	LYSC94	ASPC91	(0.63)	(0.69)		
	LYSA94	GLUA98	(1.04)	LYSB94	GLUB98	(1.35)	LYSC94	GLUC98	(3.14)	(2.77)		
LYSA94	GLUA99	(0.18)	LYSB94	GLUB99	(0.21)	LYSC94	GLUC99	(0.76)	(0.57)			
LYSA94	LYSA102	(0.36)	LYSB94	LYSB102	0.13	LYSC94	LYSC102	0.19	(0.02)	(4.77)		
Glu 98	GLUA98	GLUC42	0.09	GLUB98	GLUA42	0.67	GLUC98	GLUB42	0.48	0.62		
							GLUC98	LYSB44	(0.08)	(0.04)		
				GLUB98	GLUA46	0.08				0.04		
	GLUA98	LYSC70	(0.20)	GLUB98	LYSA70	(0.19)	GLUC98	LYSB70	(0.08)	(0.23)		
	GLUA98	ASPC76	0.28	GLUB98	ASPA76	0.53				0.41		
	GLUA98	META1	(0.08)	GLUB98	METB1	(0.08)	GLUC98	METC1	(0.08)	(0.12)		
	GLUA98	ARGA25	(0.87)	GLUB98	ARGB25	(0.88)	GLUC98	ARGC25	(0.81)	(1.28)		
	GLUA98	ARGA58	(0.66)	GLUB98	ARGB58	(0.66)	GLUC98	ARGC58	(0.82)	(1.07)		
	GLUA98	GLUA59	0.08	GLUB98	GLUB59	0.09	GLUC98	GLUC59	0.17	0.17		
	GLUA98	ASPA60	0.16	GLUB98	ASPB60	0.16	GLUC98	ASPC60	0.30	0.31		
	GLUA98	GLUA90	0.20	GLUB98	GLUB90	0.20	GLUC98	GLUC90	0.24	0.32		
	GLUA98	ASPA91	0.62	GLUB98	ASPB91	0.60	GLUC98	ASPC91	0.42	0.82		
	GLUA98	LYSA94	(1.04)	GLUB98	LYSB94	(1.35)	GLUC98	LYSC94	(3.14)	(2.77)		
	GLUA98	GLUA99	1.82	GLUB98	GLUB99	2.32	GLUC98	GLUC99	1.69	2.91		
	GLUA98	LYSA102	0.69	GLUB98	LYSB102	(0.66)	GLUC98	LYSC102	(1.20)	(0.58)	(0.49)	

Charged residues	A chain		B chain		C chain		(A+B+C)/2 kJ/mol	per each residue kJ/mol			
	Pair residues	kJ/mol	Pair residues	kJ/mol	Pair residues	kJ/mol					
Glu 99	GLUA99	GLUC42	0.53	GLUB99	GLUA42	1.06	GLUC99	GLUB42	1.05	1.32	
	GLUA99	LYSC44	(0.08)	GLUB99	LYSA44	(0.37)	GLUC99	LYSB44	(0.52)	(0.49)	
	GLUA99	GLUC46	0.41	GLUB99	GLUA46	0.60	GLUC99	GLUB46	0.57	0.79	
	GLUA99	ASPC48	0.29				GLUC99	ASPB48	0.18	0.24	
	GLUA99	GLUC50	0.34	GLUB99	GLUA50	0.16				0.25	
	GLUA99	ASPC76	0.08	GLUB99	ASPA76	0.30				0.19	
	GLUA99	LYSA19	(0.08)	GLUB99	LYSB19	(0.08)	GLUC99	LYSC19	(0.09)	(0.13)	
	GLUA99	LYSA23	(0.22)	GLUB99	LYSB23	(0.27)	GLUC99	LYSC23	(0.28)	(0.39)	
	GLUA99	GLUA24	0.08	GLUB99	GLUB24	0.08	GLUC99	GLUC24	0.08	0.12	
	GLUA99	ARGA25	(7.43)	GLUB99	ARGB25	(8.36)	GLUC99	ARGC25	(8.27)	(12.03)	
	GLUA99	LYSA56	(0.22)	GLUB99	LYSB56	(0.18)	GLUC99	LYSC56	(0.16)	(0.28)	
	GLUA99	ARGA58	(0.70)	GLUB99	ARGB58	(0.68)	GLUC99	ARGC58	(0.74)	(1.06)	
	GLUA99	ASPA60	0.08	GLUB99	ASPB60	0.08	GLUC99	ASPC60	0.23	0.20	
	GLUA99	ASPA91	0.17	GLUB99	ASPB91	0.17	GLUC99	ASPC91	0.18	0.26	
	GLUA99	LYSA94	(0.18)	GLUB99	LYSB94	(0.21)	GLUC99	LYSC94	(0.76)	(0.57)	
	GLUA99	GLUA98	1.82	GLUB99	GLUB98	2.32	GLUC99	GLUC98	1.69	2.91	
	GLUA99	LYSA101	(0.22)	GLUB99	LYSB101	(0.10)	GLUC99	LYSC101	(0.10)	(0.21)	
	GLUA99	LYSA102	0.58	GLUB99	LYSB102	(0.56)	GLUC99	LYSC102	(0.50)	(0.24)	
Lys 101							LYSC101	LYSC16	0.08	0.04	
	LYSA101	LYSA23	0.11				LYSC101	LYSC23	0.08	0.10	
	LYSA101	GLUA24	(1.66)	LYSB101	GLUB24	(0.76)	LYSC101	GLUC24	(1.62)	(2.02)	
	LYSA101	ARGA25	0.63	LYSB101	ARGB25	0.42	LYSC101	ARGC25	0.46	0.76	
	LYSA101	ARGA58	0.65	LYSB101	ARGB58	1.14	LYSC101	ARGC58	0.65	1.22	
				LYSB101	GLUB59	(0.08)				(0.04)	
	LYSA101	ASPA60	(0.37)	LYSB101	ASPB60	(0.82)	LYSC101	ASPC60	(0.44)	(0.81)	
	LYSA101	GLUA63	(0.10)	LYSB101	GLUB63	(0.29)	LYSC101	GLUC63	(0.27)	(0.33)	
	LYSA101	GLUA64	(1.97)	LYSB101	GLUB64	(3.79)	LYSC101	GLUC64	(4.31)	(5.04)	
				LYSB101	LYSB66	0.07				0.04	
	LYSA101	GLUA67	(0.23)	LYSB101	GLUB67	(0.22)	LYSC101	GLUC67	(0.21)	(0.33)	
	LYSA101	ARGA68	0.73	LYSB101	ARGB68	0.48	LYSC101	ARGC68	0.85	1.03	
	LYSA101	GLUA71	(0.30)	LYSB101	GLUB71	(0.22)	LYSC101	GLUC71	(0.32)	(0.42)	
	LYSA101	GLUA99	(0.22)	LYSB101	GLUB99	(0.10)	LYSC101	GLUC99	(0.10)	(0.21)	
	LYSA101	LYSA102	0.74	LYSB101	LYSB102	(1.33)	LYSC101	LYSC102	(0.76)	(0.67)	
	Lys 102	LYSA102	META1	(0.21)	LYSB102	METB1	0.09	LYSC102	METC1	0.11	(0.00)
		LYSA102	GLUA24	(0.30)	LYSB102	GLUB24	0.27	LYSC102	GLUC24	0.28	0.12
		LYSA102	ARGA25	(0.62)	LYSB102	ARGB25	0.49	LYSC102	ARGC25	0.29	0.09
LYSA102		ARGA58	(6.82)	LYSB102	ARGB58	(0.19)	LYSC102	ARGC58	0.56	(3.23)	
LYSA102		GLUA59	0.44	LYSB102	GLUB59	(0.05)	LYSC102	GLUC59	(0.01)	0.19	
LYSA102		ASPA60	1.32	LYSB102	ASPB60	0.78	LYSC102	ASPC60	0.80	1.45	
LYSA102		GLUA63	(0.22)	LYSB102	GLUB63	0.17	LYSC102	GLUC63	0.40	0.17	
LYSA102		GLUA64	(2.57)	LYSB102	GLUB64	0.52	LYSC102	GLUC64	0.56	(0.75)	
LYSA102		GLUA67	(0.30)							(0.15)	
LYSA102		ARGA68	0.29							0.14	
LYSA102		GLUA71	(0.19)							(0.09)	
LYSA102		LYSA94	(0.36)	LYSB102	LYSB94	0.13	LYSC102	LYSC94	0.19	(0.02)	
LYSA102		GLUA98	0.69	LYSB102	GLUB98	(0.66)	LYSC102	GLUC98	(1.20)	(0.58)	
LYSA102		GLUA99	0.58	LYSB102	GLUB99	(0.56)	LYSC102	GLUC99	(0.50)	(0.24)	
LYSA102		LYSA101	0.74	LYSB102	LYSB101	(1.33)	LYSC102	LYSC101	(0.76)	(0.67)	
Total value									(159.27)	(159.27)	

## 2-7. References.

1. Tanaka T., Sawano M., Ogasahara K., Sakaguchi Y., Bagautdinov B., Katoh E., Kuroishi C., Shinkai A., Yokoyama S., and Yutani K. (2006) Hyper-thermostability of CutA1 protein, with a denaturation temperature of nearly 150 °C. *FEBS Lett.* **580**, 4224-4230.
2. Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. (2005) The FoldX web server: an online force field. *Nucleic Acids Res.* **33**, W382-W388.
3. Matsuura, Y., Ota M., Tanaka T., Takehira M., Ogasahara K., Bagautdinov B., Kunishima N., and Yutani K. (2010) Remarkable improvement in the heat stability of CutA1 from *Escherichia coli* by rational protein design. *J. Biochem.* **148**, 449-458.
4. Pace, C.N., Vajdos, F., Fee, L., Grimsley G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411-2423.
5. Sawano, M., Yamamoto, H., Ogasahara, K., Kidokoro, S., Katoh, S., Ohnuma, T., Katoh E, Yokoyama. S., and Yutani K. (2008) Thermodynamic basis for the stabilities of three CutA1s from *Pyrococcus horikoshii*, *Thermus thermophilus*, and *Oryza sativa*, with unusually high denaturation temperatures. *Biochemistry* **47**, 721-730.
6. Korndörfer, I., Steipe, B., Huber, R., Tomschy, A., and Jaenicke, R. (1995) The crystal structure of holo-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima* at 2.5 Å resolution. *J. Mol. Biol.* **246**, 511-521.
7. Hennig, M., Darimont, B., Sterner, R., Kirschner, K., and Jansonius, J.N. (1995) 2.0 Å structure of indole-3-glycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus*: possible determinants of protein stability. *Structure* **3**, 1295-1306.
8. Karshikoff, A. and Ladenstein, R. (2001) Ion pairs and the thermotolerance of proteins from hyperthermophiles: a "traffic rule" for hot roads. *Trends Biochem. Sci.* **26**, 550-556.
9. Sheridan, R.P., Levy, R. M., and Salemme, F. R. (1982)  $\alpha$ -Helix dipole model and electrostatic stabilization of 4- $\alpha$ -helical proteins. *Proc. Natl. Acad. Sci. USA* **79**, 4545-4549.

10. Sali, D., Bycroft, M. and Fersht, A. R. (1988) Stabilization of protein structure by interaction of alpha-helix dipole with a charged side chain. *Nature* **335**, 740-743.
11. Aqvist, J., Luecke, H., Quioco, F. A. and Warshel, A. (1991) Dipoles localized at helix termini of proteins stabilize charges. *Proc. Natl. Acad. Sci. USA* **88**, 2026-2030.
12. Tanford C. (1962) The interpretation of hydrogen ion titration curves of proteins. *Adv. Protein Chem.* **17**, 69-165.
13. Kaushik, J. K., Iimura, S., Ogasahara, K., Yamagata, Y., Segawa, S., and Yutan, K. (2006) Completely buried, non-ion-paired glutamic acid contributes favorably to the conformational stability of pyrrolidone carboxyl peptidases from hyperthermophiles. *Biochemistry* **45**, 7100-7112.
14. Bush, J. and Makhatadze, G. I. (2011) Statistical analysis of protein structures suggests that buried ionizable residues in proteins are hydrogen bonded or form salt bridges. *Proteins* **79**, 2027-2032.
15. Guerois, R., Nielsen, J. E. and Serrano, L. (2002) Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J. Mol. Biol.* **320**, 369-387.
16. Schymkowitz, J. W., Rousseau, F., Martins, I. C., Ferkinghoff-Borg, J., Stricher, F., and Serrano, L. (2005) Prediction of water and metal binding sites and their affinities by using the Fold-X force field. *Proc. Natl. Acad. Sci. USA* **102**, 10147-10152.
17. Elcock, A. H. (1998) The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. *J. Mol. Biol.* **284**, 489-502.
18. Strickler, S. S., Gribenko, A. V., Keiffer, T. R., Tomlinson, J., Reihle, T., Loladze, V. V., and Makhatadze, G. I. (2006) Protein stability and surface electrostatics: a charged relationship. *Biochemistry* **45**, 2761-2766.
19. Schweiker, K. L. and Makhatadze, G. I. (2009) A computational approach for the rational design of stable proteins and enzymes: optimization of surface charge-charge interactions. *Methods Enzymol.* **454**, 175-211.
20. Tanford, C. and Kirkwood, J. G. (1957) Theory of protein titration curves. I. General equations for impenetrable spheres. *J. Am. Chem. Soc.* **79**, 5333-5339.
21. Permyakov, S. E., Makhatadze, G. I., Owenius, R., Uversky, V. N., Brooks, C. L., Permyakov, E. A., and Berliner LJ (2005) How to improve nature: study of the

electrostatic properties of the surface of  $\alpha$ -lactalbumin. *Protein Eng. Des. Sel.* **18**, 425-433.

22. Gribenko, A. V., Patel, M. M., Liu, J., McCallum, S. A., Wang, C., and Makhatadze, G. I, (2009) Rational stabilization of enzymes by computational redesign of surface charge-charge interactions. *Proc. Natl. Acad. Sci. USA* **106**, 2601-2606.

## Chapter 3: Evaluation of salt bridges in CutA1 using molecular dynamic simulations

### 3-1. Introduction.

Many charged residues are located on the surface of protein molecules; however, the solution structures of surface residues are often poorly reflected by crystallographic data, largely due to artifacts of crystallization and cooling to 100 K. Furthermore, in X-ray snapshots residues visualized are fixed, even though in solution they might fluctuate; this is particularly true of charged residues on the protein surface. Molecular dynamics (MD) simulations can be used to elucidate the structural features of these residues<sup>1-7</sup>. However, because the quality of MD results depends strongly on the force fields used<sup>8</sup>, it is important to understand how the strengths of ion-ion interactions (salt bridges) are influenced by different force fields.

Charged residues in *PhCutA1* engage in many intra- and inter-subunit interactions, whose stabilities have been extensively examined in mutation studies<sup>9</sup>. Close examination of these residues might provide insight into how the strengths of salt bridges change in MD simulations using different force fields.

In this study, we used the trimer structure of *PhCutA1* (102×3 residues) to compare the strengths of salt bridges in simulations using various force fields (Table 1). Gromos43a1<sup>10</sup> and Gromos53a6<sup>11</sup> are united-atom representations for aliphatic CH<sub>n</sub> groups, whereas Charmm27<sup>12</sup>, Amber99sb<sup>13</sup>, and Amber14sb<sup>14</sup> are all-atom

representations. For water models, we used tip3p (transferable intermolecular potential 3P)<sup>15</sup> and spc/e (extended simple point charge)<sup>16</sup>. Charmm27, Amber99sb, and Amber14sb were used with tip3p, and Amber99sb, Gromos43a1, and Gromos53a6 were used with spc/e, yielding a total of six combinations (hereafter, referred to simply as ‘the six force fields’). We performed 400-ns MD simulations of *PhCutA1* (trimer) at 300 K using these six force fields (Table 1), and analyzed the influence of the various force fields on the strengths of salt bridges on the basis of the locations of charged residues in the *PhCutA1* X-ray structure.

### **3-2. Experimental methods.**

For MD simulations, we used the trimer structure of *PhCutA1*, which contains three identical subunits (A, B, and C subunits; PDB ID 4nyo). Buried ratio and pKa of negatively and positively charged residues in the crystal structure (4nyo) of *PhCutA1* are listed in Table S9.

MD simulations were performed using the GROMACS software (ver. 4.5.5)<sup>17, 18</sup> at 300 K, using running conditions as described<sup>19, 20</sup>. In this study, we compared the structures of *PhCutA1* in MD simulations using six force fields (listed in Table 1). Hydrogen atoms were added to each protein. The models were solvated in water boxes with a minimum distance of 1.2 nm between the protein and the box. Counter-ions were added to the model to neutralize any net charge. Salt concentration was set to 150 mM. The number of Na<sup>+</sup>, Cl<sup>-</sup>, and H<sub>2</sub>O in the simulation box and size of the box during MD simulations are listed in Table S10. The periodic boundary condition was adopted and the

long-range electrostatic interactions were computed using the Particle-Mesh-Eward (PME) method<sup>21</sup>. The system was weakly coupled to a heat bath by velocity rescaling<sup>22</sup> with a relaxation time of 0.1 ps. A Parrinello-Rahman barostat<sup>23</sup> was used to maintain a pressure constant at 1 atm with a relaxation time of 0.5 ps. Hydrogen atoms were constrained using LINCS<sup>24</sup>, and MD simulations at 300 K were conducted with an integration time step of 1 fs. Energy minimizations were done to remove bad van der Waals contacts. Next, the temperature was raised from 50 to 300 K in increments of 50 K, with 10,000 integration steps at each temperature and a harmonic constraint of C-alpha atoms. Thereafter, the ensemble was equilibrated through four 100-ps cycles with gradually released harmonic constraints: 1000, 100, 10, and 1 kJ mol<sup>-1</sup> nm<sup>-2</sup>. The subsequent MD stages for the *PhCutA1* were carried out without any restraint at 300 K. The resultant MD trajectories were analyzed using the GROMACS software, as described previously<sup>19, 20</sup>. For estimation of salt bridges, distance was calculated (using the command ‘gmx distance’) between the C $\epsilon$  atom of Lys (or C $\zeta$  of Arg) and the C $\gamma$  atom of Asp (or C $\delta$  of Glu).

We confirmed how robust the results for the formation of salt bridges are (a) within the single force field during 400-ns MD simulations and (b) when different initial geometry is introduced. In the case of (a), percent occupancies of intra-subunit salt bridges in *PhCutA1* were examined at each 100-ns during 400-ns MD. As shown in Table S5, the robustness in the single force field was confirmed except for the early stage (< 100-ns) of Amber99sb\_spc/e. In the case of (b), percent occupancies of intra-subunit salt bridges for each of three identical subunits in *PhCutA1* were examined during 400-ns MD at 300 K using the six force fields. As shown in Table S6, the standard deviation of

average values among three subunits was < 2.1 % in the six force fields, suggesting the robustness of our results among the different initial structures.

We also examined the validation for the different population of the rotamers in employed force fields for residues involved in salt bridges presented on Fig. S4 - S14 using a software MolProbity (<http://molprobity.biochem.duke.edu>). As shown in Table S7, the most residues examined show rotamers of the favored region, indicating validated conformation preferences of residues involved in the bonds for the studied X-ray and MDs models.

To evaluate the energy of ion-ion interactions in *PhCutA1*, we used the algorithm FoldX<sup>25</sup>, which can quantitatively estimate the factors that are important for protein stability. FoldX is available via a web interface at <http://foldxsuite.crg.eu>. Electrostatic energies due to ion-ion interactions between charged residues were calculated using the “AllAtoms\_Electro” file in FoldX. The electrostatic energy in FoldX is calculated from a simple implementation of Coulomb’s law, in which the dielectric constant scales with the burial of the bond under consideration<sup>21</sup>. For calculations of ion-ion interactions, structural snapshots from MD simulations of *PhCutA1* were taken every 20- for 400-ns after an initial 100-ns run (total, 16 structures).

### ***3-3. Results and discussion.***

#### **Differences in general characteristics of *PhCutA1* among MD simulations using six force fields.**

Fig. 1A shows the trajectories of the RMSDs of all  $C_{\alpha}$  atoms of *PhCutA1* in MD simulations performed at 300 K using six different force fields. As shown in the figure, RMSDs were similar after 50 ns, except in the case of *Charmm27\_tip3p*, for which the value was smaller RMSD (Table 1). On the other hand, the radius of gyration ( $R_g$ ) of *PhCutA1* differed significantly among the six force fields (Fig. 1B): lowest for *Gromos43a1\_spc/e* (a compact structure) and highest for *Amber99sb\_spc/e* and *Amber99sb\_tip3p* (a loose structure) (Table 1). The largest difference in  $R_g$  was 0.096 nm, about 5% of the highest value. This large difference between force fields was also observed in a mutant *EcCutA1* in an MD simulation at 300 K<sup>26</sup>. The difference in  $R_g$  among force fields might affect the strengths of salt bridges between the subunits of *PhCutA1*.

The average residue numbers of secondary structures in *PhCutA1* in MD simulations at 300 K using each of the six force fields are listed in Table S2. The MD-simulated *PhCutA1* proteins exhibited similar secondary structures, including  $\beta$ -sheet,  $\alpha$ -helix,  $\beta$ -bridge, and turn. The numbers of residues in  $\beta$ -sheet and  $\alpha$ -helix structures were similar among the six force fields, although there were some slight differences. The difference in helicity at each residue of *PhCutA1* was examined in 50-400-ns MD simulations at 300 K using the six force fields (Fig. S1). In the regions of the  $\alpha$ -1,  $\alpha$ -2, and  $\alpha$ -3 helices, *Charmm27\_tip3p* yielded relatively high helicity, as shown in the figure, suggesting that the  $\alpha$ -helical structure of *Charmm27\_tip3p* is more stable in a long region of  $\alpha$ -helix than those of the other force fields, although the average number of residues in  $\alpha$ -helix structure was not the highest for *Charmm27\_tip3p* (Table S2). On the other hand, the  $\alpha$ -helical structure of *Amber99sb\_tip3p* seemed to be unstable in all three  $\alpha$ -helix

regions. Yoda *et al.*<sup>27</sup> reported that for small peptides, the Amber force field favors  $\alpha$ -helix, whereas Gromos favors  $\beta$ -hairpin.

The average root-mean-square fluctuations (RMSF) of the  $C_{\alpha}$  atoms of *PhCutA1* from 50 to 400-ns MD simulations at 300 K using the six force fields are shown in Fig. S2A. The differences in average RMSF of each  $C_{\alpha}$  atom were obtained by subtracting the average values of  $C_{\alpha}$  atoms from the value for each  $C_{\alpha}$  atom (Fig. S2B). In the N-terminal region, the fluctuations of Amber99sb\_spc/e and Amber99sb\_tip3p were larger than those in the other force fields, whereas in the C-terminal region, the fluctuations of Gromos43a1\_spc/e and Gromos53a6\_spc/e were larger. The loop region between positions 41 and 45 fluctuated considerably in all cases, but the fluctuations were larger in the three used with the spc/e water model, and smaller in the cases of Amber99sb\_tip3p and Amber14sb\_tip3p. In the loop regions near positions 75 and 90, fluctuations were larger in the case of Amber99sb\_tip3p than the other force fields. The average values of the fluctuations in Charmm27\_tip3p and Amber14sb\_tip3p were smaller than those of the other force fields (Table 1).

### **Evaluation of the strengths of salt bridges in *PhCutA1* during MD simulations at 300 K.**

In MD simulations, the strengths of salt bridges in proteins differ depending on the force fields used<sup>28-30</sup>. Hence, we compared the formation of salt bridges in *PhCutA1* in 300 K MD simulations using six different force fields. Specifically, we examined favorable intra-subunit interactions of 186 residues and favorable inter-subunit

interactions of 60 residues with Arg or Lys in the trimer of *PhCutA1* (Table S1A, B). Interactions were selected when the distance between favorable pairs of charged residues was  $< 0.6$  nm at least once among 31 structural snapshots obtained from MD simulations of *PhCutA1* using Gromos43a1\_spc/e; the snapshots were acquired every 10 ns for 400 ns after an initial 100-ns run.

Fig. 2A, B shows typical trajectories of the distances of intra- and inter-subunits, respectively, during 400-ns MD simulations at 300 K. Fig. 2A represents the distance between the  $C_{\epsilon}$  atom of Lys101 and the  $C_{\delta}$  atom of Glu64 in the A-subunit of *PhCutA1* during MD simulations using all six force fields. In the cases of Charmm27\_tip3p and Amber99sb\_spc/e, large fluctuations were not detected, and the lengths of salt bridges with Lys101 and Glu64 were  $0.47 \pm 0.10$  nm and  $0.46 \pm 0.10$  nm, respectively (Table S3A). The trajectories of Gromos43a1\_spc/e and Gromos53a6\_spc/e fluctuated considerably; the lengths of ion-ion interactions were  $1.20 \pm 0.34$  nm and  $1.27 \pm 0.32$  nm, respectively. On the other hand, Fig. 2B shows the trajectory of the distance between the  $C_{\epsilon}$  atom of Lys70 in the C-subunit of *PhCutA1* and the  $C_{\gamma}$  atom of Asp91 in the A-subunit during MD simulations using the six force fields. The inter-subunit interactions of Amber99sb\_spc/e fluctuated the most, and the length of the ion-ion interaction was  $0.78 \pm 0.22$  nm. By contrast, that of Gromos43a1\_spc/e was  $0.53 \pm 0.14$  (Table S3B), indicating formation of a stable salt bridge between subunits.

A salt bridge is considered to form when the distance between favorable pairs of charged residues is  $< 0.6$  nm<sup>3</sup>. We examined the percent occupancies of ion pairs with lengths below 0.6 nm for targeted intra- and inter interactions of 246 [186 residues for intra- (Table S1A) and 60 ones for inter-subunit interactions (Table S1B)] of *PhCutA1*,

as shown in Fig. S3. In this analysis, 100% occupancy indicates that all lengths of a favorable ion pair were  $< 0.6$  nm during a 400-ns MD simulation at 300 K. For example, as shown in Fig. S3A, Arg68 was expected to interact with five favorable residues (Glu24, Glu64, Glu67, Glu71, and C-terminal) within the same subunit; however, in simulations using Charmm27\_tip3p, Amber99sb\_spec, and Amber 99sb\_tip3p, Arg68 interacted with Glu24 with almost 100% occupancy, but barely interacted with Glu64, Glu67, or the C-terminus. The three bars in the figure depicting the interaction of Arg68 and Glu24 represent data from the A, B, and C subunits of *PhCutA1*. Fig. S3 also shows the distance of salt bridges obtained from the crystal structure of *PhCutA1*.

#### **Difference in the strengths of salt bridges at specific sites in *PhCutA1* during MD simulations using the six force fields.**

Table 2 shows percent occupancies of intra-subunit salt bridges of *PhCutA1* during 400-ns MD simulations at 300 K. The average percent occupancy of intra-subunit salt bridges for each positively charged residue ranged from 74.0% to 83.6% among the six force fields (Table 2A), and the average value was  $79.4 \pm 2.5\%$ . For inter-subunit salt bridges, the value ranged from 23.7% to 38.0% (Table 2B) with a mean of  $30.3 \pm 4.3\%$ . These results indicate that the average strength of salt bridges for each positively charged residue of *PhCutA1* was similar among the six force fields examined, within experimental error, and that the difference in force fields barely affected the average strength of salt bridges. However, the strengths of salt bridges at specific sites in the structure were significantly affected by the force field used, as described below.

As described in the previous section, Arg68 formed a strong salt bridge with Glu24 and Glu71 in simulation using Charmm27\_tip3p, Amber99sb\_spc/e, and Amber99sb\_tip3p (Fig. S3A). On the other hand, when using Amber14sb\_tip3p, Gromos43a1\_spc/e, and Gromos53a6\_spc/e, Arg68 formed salt bridges with Glu64 and Glu67 as well as Glu24 and Glu71, although their percent occupancies were reduced (Table 2A). These results can be observed in snapshots of *PhCutA1* acquired in 100-ns MD simulations at 300 K. As shown in Fig. S4, in Amber99sb\_tip3p, Arg68 interacted strongly with Glu24 and Glu71, and these interactions remained stable over 400 ns. In the case of Amber14sb\_tip3p, Arg68 interacted strongly with Glu24 and Glu64 at 100 ns, but the rotational isomer of Arg68 located in the middle of  $\alpha$ -2 helix could easily interact with Glu71 (Table S7). These tendencies were also observed in simulations using Gromos43a1\_spc/e and Gromos53a6\_spc/e (Table 2A).

Arg82 of *PhCutA1* formed strong salt bridges with many favorable ion pairs due to intra- and inter-subunit interactions. In the crystal structure of the trimer (A, B, and C subunits of 4nyo), Arg82 was close to Glu59, Glu63, and Asp84 in the same subunit and to Asp86 and Asp87 in another subunit (Fig. S5A). Many negatively charged residues are located near Arg82, and these ion-ion interactions fluctuated strongly during 400-ns MD simulations. In simulations using Charmm27\_tip3p, Amber14sb\_tip3p, and Gromos53a6\_spc/e, Arg82 remained engaged in a strong interaction with Asp84, but when using the other force fields, the occupancies were quite small (Table 2A). Because it is difficult to alter the movement of Arg82, the difference in this interaction might be due to fluctuation of Asp84. Occupancies of salt bridges between Arg82 and Glu59 decreased to 28-59%, in contrast to the strong interactions in the crystal structures, indicating that this effect might be caused by fluctuation of Glu59, which is located in a

loop region. On the other hand, the inter-subunit interactions of Arg82 with Asp87 in the crystal structure exchanged partners (Asp86 or Asp87) depending on the force fields used in the MD simulations. The total occupancy of the inter-subunit interaction was lowest when using Charmm27\_tip3p (44%) (Table 2B), but highest when using Amber99sb\_spc/e or Amber99sb\_tip3p (170%). On the other hand, Gromos43a1\_spc/e and Gromos53a6\_spc/e differed in terms of the favorable pairs (Asp86 and Asp87) engaged in interactions with Arg82 (Table 2B). This difference among force fields was also apparent in the intra-subunit interaction between Arg82 and Asp84: the occupancies were 7% and 87% in the case of Gromos43a1\_spc/e and Gromos53a6\_spc/e, respectively. Fig. S5B shows a snapshot of the configuration around Arg82 at 200 ns in a simulation using Gromos53a6\_spc/e. The difference within the Gromos group might be correlated with the difference in  $R_g$  (Fig. 1B and Table 1). Overall, these results indicate that ion-ion interactions in the crystal structure, in which negatively charged residues crowd around Arg82, are stronger than those in solution.

The amino group of the N-terminal residue Met1 in the crystal structure formed a salt bridge with Glu59, but Asp84 and Asp86 in the loop region also formed salt bridges with the N terminus during 400-ns MD simulations. As shown in Fig. S6, in the case of Amber99sb\_tip3p, the N terminus formed tight salt bridges with Asp84. The occupancies of the interaction between the N terminus and Glu59 were lower for Amber99sb\_spc/e and Amber99sb\_tip3p than for the other four force fields, whereas those of Asp84 were higher (Table 2A). The occupancy of Asp84 was only 2.7% in simulations using Amber14sb\_tip3p. In the case of Gromos43a1\_spc/e and Gromos53a6, the N terminus also forms salt bridges with Asp84 and Asp86, despite the almost 100% occupancy of

Glu59. Because these residues interacting with the N terminus are not involved in  $\alpha$ -helix or  $\beta$ -sheet, dependence on force fields might be strengthened.

Lys44 formed salt bridges with Glu42 and Glu46, which are located in a loop region between the  $\beta$ -2 and  $\beta$ -3 sheets. As shown in the snapshots in Fig. S7A, B, in the case of Charmm27\_tip3p, Lys44 formed strong salt bridges with both residues, whereas in the case of Gromos53a6\_spc/e, Glu42 was far from Lys44, and the occupancy of salt bridges was only 1.8%. The sum of occupancy of Lys44 was lowest for Amber14sp\_tip3p (41%) and highest for Charmm27\_tip3p (85%) (Table 2A). On the other hand, in simulations using Gromos43a1\_spc/e and Gromos53a6\_spc/e, Lys44 formed salt bridges with Glu47 with occupancies of 1.8% and 0.5%, respectively, although these values are not large. The expansion of flexibility in the loop region in the cases of Gromos43a1\_spc/e and Gromos53a6\_spc/e was also observed in the case of the interaction with Asp84 or Asp86, described above in the discussion of the N terminus.

Lys70 formed salt bridges with Glu67 in the same subunit and with Glu90 and Glu91, which are located in the N terminus of the  $\alpha$ -3 helix in another subunit. In regard to intra-subunit interactions, the occupancies of salt bridges between Lys70 and Asp76 were 18% and 16%, respectively, in the cases of Amber99sb\_spc/e and Amber99sb\_tip3p, but < 1% in simulations using the other force fields. In terms of inter-subunit interactions, in the cases of Amber99sb\_spc/e and Amber99sb\_tip3p, the sum of occupancy of Lys70 was 12% and 24%, respectively, whereas for other force fields the value ranged from 60% to 69%. The distances between Lys70 and Glu90 or Glu91 were quite close in the case of Gromos43a1\_spc/e (Fig. S8A), but in Amber99sb\_tip3p they were far (Fig. S8B). That is, the N terminus of the  $\alpha$ -3 helix was far from Lys70 in the adjacent subunit in the case

of Amber99sb\_tip3p, suggesting a difference in inter-subunit interactions that can be attributed to the use of different force fields. Consequently, Lys70 had the opportunity to form a salt bridge with Asp76 in the same subunit when using Amber99sb\_spc/e and Amber99sb\_tip3p. This observation is correlated with the difference in  $R_g$ , which was larger for Amber99sb\_spc/e and Amber99sb\_tip3p than for the other force fields (Fig. 1B).

In the crystal structures, Arg36 in the middle of the  $\beta$ -2 strand formed a tight salt bridge with Glu15, located in the middle of the  $\alpha$ -1 helix in another subunit. Arg36 also formed a salt bridge with Glu47 in the middle of the  $\beta$ -3 strand in the same subunit. Fig. S9 shows snapshots of Arg36 in the cases of Amber99sb\_tip3p and Gromos53a6\_spc/e at 200 ns. The occupancies of the Arg36-Glu15 interaction were higher with Amber99sb\_spc/e and Amber99sb\_tip3p than the other force field; in particular, the occupancy of inter-subunit salt bridges in between Arg36 and Glu15 was lowest for Gromos53a6\_spc/e among the six force fields. On the other hand, the sum of occupancy of Arg36 in the same subunit was the highest (99%) for Charmm27\_tip3p and the lowest (59%) for Amber14sb\_tip3p. In the case of Arg36, it is quite difficult to explain the difference in the strength of salt bridges among the six force fields, although one possibility is difference in the combination of rotamers (Table S7).

Lys101 and Lys102 are located in the C terminus of *PhCutA1*. In the cases of Charmm27\_tip3p, Amber99sb\_spc/e, and Amber99sb\_tip3p, Lys101 formed salt bridges with Glu64 with percent occupancies above 80%, whereas in the case of Gromos43a1\_spc/e and Gromos53a6\_spc/e, the corresponding values were only 7-8 % (Table 2A). On the other hand, in the cases of Gromos43a1\_spc/e and Gromos53a6\_spc/e,

Lys102 formed a salt bridge with Glu64 (occupancies of 54-60 %), but the occupancies were barely detectable when using the other force fields. Fig. S10A, B shows typical snapshots at 100 ns in simulations using Charmm27\_spc/e and Groms53a6\_spc/e, respectively. The two amino groups of Lys102 were located in opposite directions in simulations using these two force fields, as shown in the figure. In the cases of Gromos43a1\_spc/e and Gromos53a6\_spc/e, Lys102 forms salt bridges with several favorable residues in addition to Glu64 and Glu98, whereas in the cases of the other force fields, Lys102 only forms a salt bridge with Glu98 (Table 2A). These results indicate that the two C-terminal residues fluctuate much more intensely in the two Gromos force fields than in the other cases (Fig. S2B).

Arg58, located in a small loop between the  $\beta$ -3 sheet and  $3_{10}$ -helix, formed salt bridges with Glu60 with occupancies of 80-100%, and with the C-terminal carboxyl group (Lys102) with occupancies of 17-47%. For both salt bridges, the occupancy was lowest for Gromos43a1\_spc/e, which might be related to the flexibility of loop region. The snapshots of the C terminus and Arg58 at 200 ns in simulations using Amber14sb\_tip3p and Gromos43a1\_spc/e are shown in Fig. S11A, B, respectively.

Lys56 formed a salt bridge with Glu50 in another subunit in simulations using Gromos43a1\_spc/e (Fig. S12A), but not when using Charmm27\_tip3p, Amber99sb\_tip3p, or Amber14sb\_tip3p (Fig. S3B and Table 2B). A snapshot of the inter-subunit interaction between Lys56 and Glu50 (Fig. S12B) in a simulation using Amber99sb\_spc/e shows that Glu50 was stably located in the middle of the  $\beta$ -2 sheet, although the salt bridge was longer than in a simulation using Gromos43a1 (0.69 vs.0.33 nm). On the other hand, in the case of Gromos43a1\_spc/e (Fig. S12A), the  $\beta$ -2 sheet around Glu50 was absent. The

fluctuation (instability) in the middle of  $\beta$ -2 sheet might create the opportunity to form a salt bridge  $< 0.6$  nm in length.

### **Reevaluation of the electrostatic energies of charged residues for *PhCutA1*.**

The electrostatic energies of charged residues for *PhCutA1* have been estimated based on the crystal structures in order to evaluate the contribution of charged residues to conformational stability. In terms of unfavorable interactions, the worst three residues are Asp86, Glu12, and Arg33, with ion-ion interaction energies of 19.6, 11.6, and 10.9 kJ/mol, respectively<sup>9</sup>. As shown in Fig. S13A, in the crystal structure Asp86 interacts repulsively with Asp84 in another subunit, and the distance between Asp86 and Asp84 ranges from 0.34 to 0.41 nm. However, this repulsion energy was weakened during MD simulations at 300 K: the average electrostatic energy for Asp86 became 6.5 kJ/mol in the case of Gromos43a1\_spc/e, as compared to 19.6 kJ/mol in the crystal structure (Table S4). A snapshot of the configuration around Asp86 confirmed that the distances between Asp86 and Asp84 in the adjacent subunit increased from 0.45 to 0.74 nm (Fig. S13B). In the cases of Charmm27\_tip3p and Amber99sb\_spc/e, the electrostatic energies for Asp86 were 16.5 and 4.5 kJ/mol, respectively (Table S4).

Glu12 engages in repulsive interactions with Asp10 and Glu15. During 400-ns MD simulations, the electrostatic energies of Glu12 became 8.2, 7.9, and 6.5 kJ/mol for Gromos43a1, Charmm27\_tip3p, and Amber99sb\_spc/e, respectively, indicating slight relaxations.

In X-ray crystal structure of *PhCutA1*, Arg33 is completely buried in the interior of the molecule. However, the side chain of Arg33 appeared on the surface of the trimer in MD simulations using Gromacs43a1\_spc/e (Fig. S14B), whereas in the case of Amber99sb\_tip3p, this residue was still located in the interior of the trimer. The superimposed structure shown in Fig. S14B suggests that both structures are rotational isomers of Arg33 (Table S7). On the surface, the ionic group of Arg33 can form ion-ion interactions with Glu34 residues in the same or adjacent subunits (Fig. S14A). Except for Gromos43a1\_spc/e and Gromos53a6\_spc/e, occupancies of salt bridges were not detected in intra- or inter-subunit interactions (Tables 2A, B), indicating that the side chains of Arg33 remain buried in the interior of the trimer during 400-ns MD simulations. When the ionic groups of charged residues are buried in the interior of protein molecules, the difference in strengths of salt bridges among force fields might stand out between Gromos groups with the united-atom representation and other groups with the all-atom representation. During 400-ns MD simulations, the electrostatic energy of Arg33 significantly improved to  $-4.9$  kJ/mol in the case of Gromos43a1\_spc/e and to  $5.8$  and  $2.3$  kJ/mol in the cases of Charmm27\_tip3p and Amber99sb\_spc/e, respectively (Table S4). The adjustment of unfavorable energy for Arg33 was the highest for Gromos43a1\_spc/e.

Fig. 3 shows the correlation between electrostatic energy at mutation sites estimated by FoldX using various structures and the difference in denaturation temperatures ( $\Delta T_d$ )<sup>9</sup> due to mutations of charged residues to noncharged residues. As mentioned above, the figure shows that the electrostatic energies at Asp86 and Arg33 are significantly reduced in structures obtained from MD simulations. Detailed data are provided in Table S4. On the other hand, in the cases of Gromos43a1\_spc/e and Gromos53a6\_spc/e, the

electrostatic energies at Arg58, Lys66, and Arg68 are higher than those in the crystal structures (Fig. 3 and Table S4). These data indicate that the electrostatic energies of structures from MD simulations approach the line generated by regressing electrostatic energies at mutation sites vs.  $\Delta T_d$  due to mutations (Fig. 3). Through molecular adaptation, changes in conformational stability due to mutations should optimize factors such as hydrophobic interactions, electrostatic interactions, and entropic effects. Consequently,  $\Delta T_d$  and electrostatic energies at mutation sites should not exhibit a strict linear correlation. However, our data show a weak linear correlation between them (Fig. 3 and Table 3); because the energies of ion-ion interaction with MD simulation of Gromos group approach nearer the linear regression lines than those of the crystal structures, we can conclude that the configurations of charged residues in structures obtained by MD simulations using the Gromos group are better than those obtained with other force fields. Table 3 shows a comparison of linear regression error coefficients between electrostatic energy at a targeted residue and the difference in denaturation temperature among eight different structures. The error coefficients of Gromos group were lower than those of the others, suggesting that the configurations of charged residues in MD simulations performed using the Gromos group are superior.

### **The interaction of charged residues with counter-ions of salts.**

In the present MD system, salt concentration (NaCl) was set to 150 mM. The percent occupancies of salt bridges (< 0.6 nm) of positively charged residues in *PhCutA1* with chloride (Cl<sup>-</sup>) ion were examined (Table 4). In the case of Charmm27\_tip3p and Amber14sb\_tip3p, Arg33 which was completely buried in the interior of a molecule in

the initial state formed a salt bridge with one Cl<sup>-</sup> ion by almost 100 %, and salt bridges of its Arg33 with negatively charged residues were not observed (Table 4). As mentioned above, in the case of Gromos43a1\_spc/e, Arg33 which was located on the surface formed a salt bridge with a negatively charged residue by percent occupancy of 121.3 %, but percent occupancy (summation) of salt bridges of its Arg33 with several Cl<sup>-</sup> ions was 13.2 % (Table 4). Trajectories of distance between Arg33 in *PhCutA1* and Cl<sup>-</sup> ion during 400-ns MD simulations at 300 K for three force fields are shown in Fig. S15. In the case of Charmm27\_tip3p and Amber99sb\_tip3p, percent occupancies of salt bridges of Arg33 with one Cl<sup>-</sup> ion were 100.0% and 86.0 % during 400-ns MD simulations, respectively. In the case of Gromos43a1\_spc/e, that was only 1.7 % (Fig. S15): the summation of occupancies of Arg33 with several Cl<sup>-</sup> ions was 13.2 % as mentioned above. These results suggest that charged residues which are completely buried in a molecule are neutralized by counter-ions.

Except for Arg33, the average percent occupancy of intra- and inter-subunit salt bridges for each positively charged residues ranged from 93.5 % to 112.1 % among the six force fields (average 2 in Table 4), and the average value was  $105.1 \pm 6.6$  %. On the other hand, the value of interactions with Cl<sup>-</sup> ions ranged from 3.0 % to 5.4 % (average 2 in Table 4) with mean of  $3.8 \pm 0.8$  %. The average values of the interaction between charged residues of amino-acids were greater by about 30 times than others, suggesting that the ion interaction between amino acids is considerably stronger than that of charged residues with counter-ions of salts. The big difference in the interaction of charged residues with counter-ions on the surface of a molecule was not detected among six force fields.

Furthermore, the average percent occupancies of intra- and inter-subunit salt bridges for each negatively charged residues were also examined: they were from 60.2 to 69.7 % among the six force fields (Table S8), and the average value was  $65.0 \pm 4.0$  %. The big difference between the average percent occupancies between negatively and positively charged residues (Table 4 and Table S8) is caused by the difference in residue numbers between them. On the other hand, the value of interactions of negatively charged residues with  $\text{Na}^+$  ions ranged from 19.4 % to 59.6 % (Table S8) with mean of  $31.9 \pm 16.0$  %. This bigger value compared with that of Cl ion suggests the neutralization of excess negatively charged residues.

### **Re-evaluation of salt bridges obtained from X-ray crystal structures of *PhCutA1*.**

Fig. S3 also shows the distance of salt bridges  $< 0.6$  nm, obtained from X-ray crystal structures of *PhCutA1*. The 4nyo.pdb structure of *PhCutA1* was obtained in normal buffer solutions, whereas the 1umj.pdb structure of the same protein was obtained in the presence of 3 M guanidine hydrochloride, a denaturant<sup>31</sup>. *PhCutA1* is not denatured in 3 M guanidine hydrochloride<sup>31, 32</sup>, but the local structures on the surface of the protein, including salt bridges, might be perturbed by such severe conditions. Focusing on the main pair residues (strong salt bridges) with positively charged residues in Fig S3, the pair residues of salt bridges observed from six kind MD simulations coincided with those from three kind crystal structures except for three pairs, which are Lys49 and Lys94 in ABC subunits of 4nyo, and Lys102 in three crystals. Ion pairs of Lys49 with Glu34 and Lys94 with Glu91 in crystal structures coincided with those of MD simulations except for ABC subunits of 4nyo, in which B factors (suggesting fluctuation of crystal atoms)

were also high (Fig. S3A), indicating that the pairs involving Lys49 and Lys94 in ABC of 4noy are questionable. Lys102, a C-terminal residue with a high B factor, is not detectable in 1umj, which was crystallized in the presence of 3 M guanidine hydrochloride<sup>31</sup>. These results indicate that if the crystal data are probed strictly, the main pairs of salt bridges obtained from MD simulations at 300 K completely agree with X-ray crystal data measured at 100 K. However, charged residues in solution constantly fluctuate, as shown in Fig. 2, and would have many chances to form salt bridges with other favorable ionic pairs.

#### **Difference in the strengths of salt bridges of *PhCutA1* using six force fields.**

Next, we counted the numbers of ion pairs forming salt bridges < 0.6 nm in length during 400-ns MD simulations for 186 targeted intra-subunit interactions and 60 targeted inter-subunit interactions. Table 5 lists the numbers and percentages of salt bridges vs. percent occupancy. The number of ion-ion interactions with percent occupancies > 0.1% was 106 of 186 for Charmm27\_tip3p, and 161 of 186 for Gromos43a1\_spc/e (57% and 87%, respectively). These results indicate that in the case of Gromos43a1\_spc/e, *PhCutA1* forms much more favorable salt bridges than in the other cases. On the other hand, occupancies > 90% were observed for 21 salt bridges in the case of Charmm27\_tip3p and only six in the case of Gromos43a1\_spc/e (Table 5A), indicating that the salt bridges of Charmm27\_tip3p at each site are more stable than those of other force fields, although the average values of occupancy for each positively charged residue were similar (Table 2).

In regard to inter-subunit interactions (Table 5B), Amber14\_tip3p had the lowest percent occupancy ( $>0.1$ ) and Gromos43a1\_spc/e the highest. In the case of higher occupancies, the data for the inter-subunit interactions were more complicated due to differences in the strengths of subunit-subunit interactions among different force fields. However, in the case of the Gromos group, charged residues of *PhCutA1* might be able to interact with many more favorable charged residues, both within and between subunits, than in the case of the other force fields.

### **3-4. Conclusions.**

1. We investigated the influence of six different force fields on the formation of salt bridges involving positively charged residues (Arg or Lys) and favorable partners in MD simulations of the *PhCutA1* trimer. The force fields used were Charmm27\_tip3p, Amber99sb\_spc/e, Amber99sb\_tip3p, Amber14sb\_tip3p, Gromos43a1\_spc/e, and Gromos53a6\_spc/e.
2. We examined the effects of force fields on RMSD, RMSF,  $R_g$ , and secondary structures. The average RMSD was lowest for Charmm27\_tip3p, and similar for the other force fields.  $R_g$  was lowest for Gromos43a1\_spc/e and highest for Amber99sb\_spc/e and Amber99sb\_tip3p. The average number of residues in each type of secondary structures was similar among the six force fields, but percent helicity was highest for Charmm27\_tip3p.
3. Percent occupancies of salt bridges  $< 0.6$  nm for all the targeted pairs of 246 residues (186 residues for intra- and 60 ones for inter-subunit interactions) (Fig. S3) and for all positively charged residues (Table 2) were used to estimate the strengths of salt

- bridges. Furthermore, the average length with its error bar of salt bridges suggests that its average value considerably fluctuates during 400-ns MD simulations (Table S3).
4. The average percent occupancy of intra-subunit salt bridges for each positively charged residue ranged from 83.6% for Amber99sb\_spc/e to 74.0% for Gromos53a6\_spc/e; the average was  $79.4 \pm 2.5\%$  (Table 2A). For inter-subunit interactions, the value ranged from 38.0% for Gromos43a1\_spc/e to 23.7% for Charmm27\_spc/e; the average was  $30.3 \pm 4.3\%$  (Table 2B). Thus, the average strength of salt bridges for positively charged residues did not differ significantly among the six force fields. However, in the case of the Gromos group, positively charged residues of *PhCutA1* were able to interact with many more favorable residues (Glu or Asp) in the other force fields, and this was true for both intra- and inter-subunit interactions (Table 5). On the other hand, the strengths of salt bridges at specific sites within structures were significantly affected by the force field used.
  5. The orientation of the ionizable group of a charged residue is determined by the rotational isomer. For example, in the  $\alpha$ -2 helix, a rotational isomer of Arg68 forms a salt bridge with Glu64 in the cases of Charmm27\_tip3p, Amber99sb\_spc/e, and Amber99sb\_tip3p or with Glu71 in the other cases (Table S7). The side chain of Arg33 in the  $\beta$ -2 sheet is buried in the interior of the molecule in the initial structure, but during MD simulations using Gromos43a1\_spc/e and Gromos53a6\_spc/e, the side chain (which is located on the protein surface) rotates to form salt bridges with Glu34 in the same and neighboring subunits. By contrast, in the other cases, Arg33 remains buried in the interior of the molecule. Arg82 forms a salt bridge with Asp84 in the cases of Charmm27\_tip3p, Amber14sb\_tip3p, and Gromos53a6\_spc/e, but not otherwise. For Arg82, this difference seems to be caused by fluctuation

(isomerization) for Asp84 of the pair residue. It is possible that the charged residues fluctuate over all possible rotational isomers in solution, depending on the differences in the energy levels of rotamer among the six force fields.

6. The occupancies of a salt bridge between Lys70 and Glu90 in the inter-subunit interaction were smaller in the cases of Amber99sb\_spc/e and Amber99sb\_tip3p than the other force fields. This might be correlated with the larger  $R_g$  values (i.e., loose structures) obtained with Amber99sb\_spc/e and Amber99sb\_tip3p.
7. Lys102 forms salt bridges with five different residues in the cases of Gromos43a1\_spc/e and Gromos53a6\_spc/e, but with only two residues in the other force fields. The same effects were observed for many other interactions, such as N terminus-Asp86 and Lys101-Glu98, which are located in loop regions. The greater flexibility of side chains of charged residues in the case of the Gromos group might be correlated with differences in the atomic representations of aliphatic  $CH_n$  (i.e., united vs. all-atom representation).
8. The occupancies of salt bridges for Amber99sb\_spc/e and Amber99sb\_tip3p were similar, indicating that the choice of water model barely affects the strength of salt bridges in the case of the Amber99sb force field.
9. The electrostatic energies of charged residues at mutation sites were re-evaluated for two structures obtained by X-ray crystal analysis and six structures obtained by MD simulations. From the correlation between electrostatic energies at mutation sites and the difference in denaturation temperatures of mutants in which charged residues were replaced with noncharged residues, the configurations of charged residues in structures in simulations using Gromos43a1\_spc/e and Gromos53a6\_spc/e were superior to those in the six other structures.

10. If the crystal data were probed strictly, the main pairs of salt bridges obtained from MD simulations at 300 K completely agreed with the X-ray crystal data measured at 100 K, although their strengths differed among the six force fields. However, many unfavorable interactions with high repulsive energies were observed in the crystal structures. These unfavorable interactions were diminished during MD simulations.
11. Finally, in the case of the Gromos group, positively charged residues could interact with many more favorable residues (Glu or Asp) than in the other force fields, especially in loop regions, causing the apparent strength at each site to be weaker. The strength at each site was highest for Charmm27\_tip3p. That is, the Gromos group has the advantage that a charged residue has more opportunities to seek for favorable interaction partners in solution during the limited period of the MD simulation, although this notion requires confirmation.

Table 1. List of six force fields used in MD simulations, and comparison of RMSD,  $R_g$ , and RMSF for each force field (50-400ns).

	force fields	water model	RMSD (nm)	$R_g$ (nm)	RMSF (nm)
1	Charmm27	tip3p	0.09 ± 0.01	1.93 ± 0.01	0.06 ± 0.02
2	Amber99sb	spc/e	0.17 ± 0.01	1.99 ± 0.01	0.08 ± 0.03
3	Amber99sb	tip3p	0.17 ± 0.02	1.98 ± 0.01	0.08 ± 0.03
4	Amber14sb	tip3p	0.17 ± 0.02	1.94 ± 0.01	0.06 ± 0.02
5	Gromos43a1	spc/e	0.17 ± 0.01	1.90 ± 0.01	0.08 ± 0.04
6	Gromos53a6	spc/e	0.17 ± 0.02	1.95 ± 0.01	0.08 ± 0.04

Table 2.

(A) Percent occupancies of intra-subunit salt bridges in *PhCutA1* during 400-ns MD simulations at 300 K using the indicated force fields. (B) Percent occupancies of inter-subunit salt bridges in *PhCutA1* during 400-ns MD simulations at 300 K using the indicated force fields. Data show average values of three subunits.

(A)

Targeted residues		Force fields					
Positively charged Residues	pair residues	Charmm27 _tip3p	Amber99sb _spc/e	Amber99sb _tip3p	Amber14sb _tip3p	Gromos_43a1 _spc/e	Gromos53a6 _spc/e
N-terminal	Glu59	97.7	84.3	61.0	95.0	98.4	98.9
	Asp84	13.5	44.9	41.7	2.7	37.6	15.1
	Asp86	0.7	0.0	0.0	0.0	1.3	13.4
	sum	111.8	129.2	102.7	97.7	137.3	127.3
Lys16	Asp10	0.0	0.1	0.0	0.0	0.0	0.0
	Glu12	28.6	29.4	32.9	40.3	54.2	47.8
	Glu15	2.0	1.2	2.1	1.3	0.9	0.8
	sum	30.6	30.7	35.0	41.6	55.1	48.6
Lys19	Glu12	0.0	0.0	0.0	0.0	0.0	0.0
	Glu15	90.4	94.0	95.4	91.8	90.1	90.1
	sum	90.4	94.0	95.4	91.8	90.1	90.1
Lys23	Glu24	4.3	1.2	1.4	7.5	6.4	8.3
Arg25	Glu24	0.0	0.0	0.0	0.0	0.3	0.0
	Glu98	0.0	0.1	0.0	0.1	0.4	0.0
	Glu99	97.2	96.1	91.3	99.3	86.5	94.9
	C-terminal	0.0	0.0	0.0	0.0	2.9	1.1
	sum	97.3	96.2	91.3	99.4	90.1	95.9
Arg33	Glu34	0.0	0.9	0.2	0.0	54.8	35.3
Arg36	Glu34	0.0	0.3	0.1	0.1	0.3	0.1
	Glu46	0.0	0.0	0.0	0.0	0.0	0.0
	Glu47	89.9	74.9	79.5	56.7	65.8	71.1
	Asp48	8.8	5.6	2.8	2.3	1.2	0.2
	sum	98.7	80.8	82.4	59.0	67.4	71.3
Lys44	Glu42	19.6	12.2	14.4	8.3	5.6	1.8
	Glu46	64.9	49.8	40.4	32.2	41.0	60.4
	Glu47	0.0	0.0	0.0	0.1	1.8	0.5
	sum	84.5	62.0	54.7	40.6	48.4	62.7
Lys49	Asp10	5.0	4.3	6.0	0.4	7.8	0.8
	Glu12	0.0	0.0	0.0	0.0	0.0	0.0
	Glu34	82.6	92.2	84.3	56.9	50.5	63.4
	Glu47	0.6	0.1	0.1	0.5	8.9	2.8
	Asp48	1.9	13.4	5.3	2.5	8.0	2.4
	sum	90.1	110.0	95.8	60.2	75.2	69.3
Arg58	Glu59	0.2	0.0	0.3	0.2	5.4	1.4
	Asp60	99.9	98.5	99.8	99.8	79.5	93.0
	Glu98	0.0	0.0	0.0	0.0	0.8	0.0
	C-terminal	32.4	35.9	30.6	47.0	16.7	17.1
	sum	132.5	134.4	130.6	147.0	102.4	111.5
Lys66	Glu63	52.3	60.0	72.2	89.0	71.2	44.3
	Glu67	7.5	24.1	9.6	8.2	19.3	11.0
	sum	59.8	84.1	81.8	97.2	90.5	55.2

Table 2A. continued.

Arg68	Glu24	99.5	99.5	95.6	53.7	58.2	63.2
	Glu64	9.0	3.0	3.5	37.4	24.4	19.6
	Glu67	1.8	0.7	3.5	45.3	14.5	6.4
	Glu71	83.6	92.2	93.0	72.1	46.4	27.8
	C-terminal	0.0	0.0	0.0	0.0	0.2	0.0
	sum	193.9	195.5	195.6	208.5	143.7	117.0
Lys70	Glu63	0.0	0.0	0.0	0.0	0.0	0.0
	Glu67	29.8	54.5	46.3	40.1	50.0	26.5
	Glu71	2.7	1.0	1.9	3.0	4.3	2.2
	Asp76	0.3	18.1	15.9	0.5	0.9	0.4
	sum	32.8	73.6	64.1	43.6	55.2	29.1
Arg82	Glu59	27.9	46.0	50.0	45.7	51.7	58.9
	Asp60	0.0	0.0	0.0	0.0	0.0	0.0
	Glu63	0.0	0.0	0.1	0.0	0.1	0.0
	Asp84	99.6	2.3	1.1	80.8	6.7	86.8
	sum	127.5	48.4	51.1	126.5	58.6	145.8
Lys94	Glu90	27.2	43.2	43.5	20.5	31.7	16.7
	Asp91	26.0	28.9	41.7	25.5	35.1	21.1
	Glu98	7.4	10.5	4.1	7.4	10.8	15.8
	Glu99	0.0	0.0	0.0	0.0	0.3	0.0
	sum	60.7	82.7	89.3	53.4	78.0	53.6
Lys101	Glu24	10.8	10.4	12.6	15.6	3.7	4.2
	Asp60	0.0	0.0	0.0	0.0	4.3	1.2
	Glu64	86.8	88.0	80.2	14.5	7.7	6.5
	Glu98	0.0	0.0	0.0	0.0	32.9	10.1
	Glu99	0.0	0.0	0.0	0.0	4.0	6.9
	C-terminal	38.7	52.0	46.0	58.6	9.7	7.2
	sum	136.3	150.4	138.8	88.6	62.1	36.0
Lys102	Glu24	0.0	0.0	0.0	0.0	4.4	4.5
	Glu59	0.0	0.0	0.0	0.0	0.0	0.0
	Asp60	0.5	0.3	0.4	0.2	37.3	29.6
	Glu63	0.0	0.0	0.0	0.0	1.5	0.6
	Glu64	0.0	0.1	0.0	0.1	59.1	54.0
	Glu98	51.2	47.2	44.9	58.7	19.4	6.9
	Glu99	0.0	0.0	0.0	0.0	6.8	4.6
	sum	51.7	47.5	45.3	58.9	128.5	100.2
	Average <sup>a</sup>	82.5	83.6	79.7	77.7	79.0	74.0

<sup>a</sup> Average values per positively charged residue.

(B)

Targeted residues		Force fields					
Positively charged Residues	pair residues	Charmm27 _tip3p	Amber99sb _spc/e	Amber99sb _tip3p	Amber14sb _tip3p	Gromos_43a1 _spc/e	Gromos53a6 _spc/e
Lys19	Glu46	0.0	0.1	0.0	0.0	1.1	0.1
	Glu47	72.3	69.1	91.6	57.4	35.2	32.2
	sum	72.3	69.2	91.6	57.4	36.3	32.3
Lys23	Glu42	0.0	0.0	0.0	0.0	0.2	0.0
Arg25	Glu42	0.0	0.0	0.0	0.0	2.3	0.8
Arg33	Glu34	0.0	0.7	3.7	0.0	66.5	38.3
	Glu47	0.0	0.0	0.0	0.0	0.0	0.0
	sum	0.0	0.7	3.7	0.0	66.5	38.3
Arg36	Glu15	33.2	53.1	64.9	34.3	34.1	13.1
	Glu34	0.1	0.0	0.1	0.4	3.4	3.2
	sum	33.3	53.1	64.9	34.7	37.5	16.3
Lys44	Glu15	0.0	0.0	0.0	0.0	0.1	0.0
Lys56	Glu50	0.2	12.6	0.3	0.0	59.7	13.9
Lys66	Asp87	30.3	46.0	51.8	63.5	25.8	11.1
	Glu90	44.9	9.6	12.4	17.1	28.6	36.5
	Asp91	0.0	0.0	0.0	0.0	0.0	0.1
	sum	75.2	55.5	64.2	80.6	54.4	47.7
Lys70	Glu90	2.2	1.2	1.1	18.6	10.1	9.5
	Asp91	57.1	11.0	22.6	44.1	58.5	56.4
	sum	59.3	12.1	23.7	62.7	68.6	65.9
Arg82	Asp86	24.8	98.0	99.8	57.9	90.3	8.5
	Asp87	19.3	71.7	70.1	45.9	39.5	88.2
	sum	44.1	169.8	169.9	103.8	129.8	96.8
Lys101	Glu42	0.0	0.0	0.0	0.0	0.3	0.3
Lys102	Glu42	0.0	0.0	0.0	0.0	0.4	0.1
average <sup>a</sup>		23.7	31.1	34.9	28.3	38.0	26.0

<sup>a</sup> Average values per positively charged residue.

Table 3. Linear regression error coefficients between electrostatic energy at targeted residue and difference in denaturation temperature of mutant *PhCutA1* among different structures. *R*, *SD*, and *P* represent correlation coefficient, standard deviation, and *P* value, respectively, in linear regression using data of Table S4. Three lines in Fig. 3 are linear regressions between  $\Delta T_d$  and electrostatic energy for three structures described.

		<i>R</i>	<i>SD</i>	<i>P</i>
crystal structures	ABC of 4nyo	0.465	11.2	0.008
	DEF of 4nyo	0.483	10.5	0.006
force fields in MD simulation	Charmm27_tip3p	0.464	10.0	0.009
	Amber99sb_spc/e	0.442	9.6	0.013
	Amber99sb_tip3p	0.445	9.8	0.012
	Amber14sb_tip3p	0.477	10.1	0.007
	Gromos43a1_spc/e	0.498	8.6	0.004
	Gromos53a6_spc/e	0.522	8.0	0.003

Table 4. Percent occupancies of salt bridges (< 0.6 nm) of positively charged residues in *PhCutA1* with Cl<sup>-</sup> ions during 400-ns MD simulations at 300 K using the indicated force fields. Highlight shows the interaction with Arg33.

Positively charged Residues	Force fields											
	Charmm27_tip3p		Amber99sb_spc/e		Amber99sb_tip3p		Amber14sb_tip3p		Gromos43a1_spc/e		Gromos53a6_spc/e	
	N. C. Residue*	Cl <sup>-</sup> ion	N. C. Residue*	Cl <sup>-</sup> ion	N. C. Residue*	Cl <sup>-</sup> ion	N. C. Residue*	Cl <sup>-</sup> ion	N. C. Residue*	Cl <sup>-</sup> ion	N. C. Residue*	Cl <sup>-</sup> ion
N-terminal	111.8	0.4	129.2	1.7	102.7	2.1	97.7	0.8	137.3	0.6	127.3	0.2
Lys16	30.6	10.1	30.7	7.8	35.0	5.9	41.6	6.1	55.1	4.0	48.6	4.7
Lys19	90.4	7.4	94.0	4.2	95.4	2.6	91.8	4.2	90.1	4.1	90.1	5.2
Lys23	4.3	14.7	1.2	13.1	1.4	8.4	7.5	9.8	6.4	8.3	8.3	10.2
Arg25	97.3	2.3	96.2	1.7	91.3	1.3	99.4	1.5	90.1	0.9	95.9	0.4
<b>Arg33</b>	<b>0.0</b>	<b>99.9</b>	<b>0.9</b>	<b>58.8</b>	<b>0.2</b>	<b>62.1</b>	<b>0.0</b>	<b>99.8</b>	<b>54.8</b>	<b>13.2</b>	<b>35.3</b>	<b>18.0</b>
Arg36	98.7	5.8	80.8	2.0	82.4	1.4	59.0	3.6	67.4	6.3	71.3	8.5
Lys44	84.5	6.7	62.0	6.9	54.7	5.6	40.6	5.5	48.4	6.4	62.7	6.2
Lys49	90.1	10.4	110.0	4.6	95.8	3.8	60.2	7.0	75.2	7.2	69.3	10.4
Arg58	132.5	1.1	134.4	0.4	130.6	0.3	147.0	0.5	102.4	2.6	111.5	1.8
Lys66	59.8	2.0	84.1	1.6	81.8	1.4	97.2	0.9	90.5	0.6	55.2	0.9
Arg68	193.9	4.9	195.5	1.9	195.6	2.2	208.5	2.3	143.7	1.0	117.0	1.6
Lys70	32.8	3.0	73.6	1.7	64.1	1.9	43.6	1.9	55.2	0.9	29.1	1.5
Arg82	127.5	1.6	48.4	0.6	51.1	1.0	126.5	1.2	58.6	0.4	145.8	0.4
Lys94	60.7	6.2	82.7	3.0	89.3	3.1	53.4	4.3	78.0	3.6	53.6	3.6
Lys101	136.3	3.0	150.4	1.7	138.8	1.7	88.6	3.0	62.1	3.0	36.0	3.5
Lys102	51.7	6.3	47.5	4.7	45.3	4.5	58.9	4.3	128.5	2.0	100.2	2.0
average1 <sup>a</sup>	82.5	10.9	83.6	6.8	79.7	6.4	77.7	9.2	79.0	3.8	74.0	4.6
average2 <sup>b</sup>	87.7	5.4	88.8	3.6	84.7	3.0	82.6	3.6	80.6	3.2	76.4	3.8

<sup>a</sup> N. C. Residue represents percent occupancies with negatively charged residues. Data come from Table 2A, B.

<sup>b</sup> average1 and average2 represent the average of 17 positively charged residues and that of 16 residues except for data of Arg33, respectively.

Table 5. (A) Number of targeted ion-ion interactions forming intra-subunit salt bridges in simulations using the indicated force fields. (B) The number of targeted ion-ion interactions forming inter-subunit salt bridges in simulations using the indicated force fields. Data represent number and percentage of targeted ion-ion interactions forming salt bridges (< 0.6 nm).

(A)

% <sup>a</sup>	Force fields											
	Charmm27_tip3p		Amber99_spc/e		Amber99_tip3p		Amber14_tip3p		Gromos43a1_spc/e		Gromos53a6_spc/e	
	number	% <sup>b</sup>	number	% <sup>b</sup>	number	% <sup>b</sup>	number	% <sup>b</sup>	number	% <sup>b</sup>	number	% <sup>b</sup>
>0.1	106	57.0	115	61.8	114	61.3	118	63.4	161	86.6	142	76.3
>1	94	50.5	96	51.6	98	52.7	94	50.5	135	72.6	121	65.1
>10	67	36.0	74	39.8	70	37.6	71	38.2	81	43.5	77	41.4
>40	40	21.5	51	27.4	46	24.7	48	25.8	45	24.2	40	21.5
>70	29	15.6	26	14.0	25	13.4	20	10.8	15	8.1	18	9.7
>80	28	15.1	24	12.9	21	11.3	17	9.1	10	5.4	15	8.1
>90	21	11.3	17	9.1	14	7.5	15	8.1	6	3.2	10	5.4

(B)

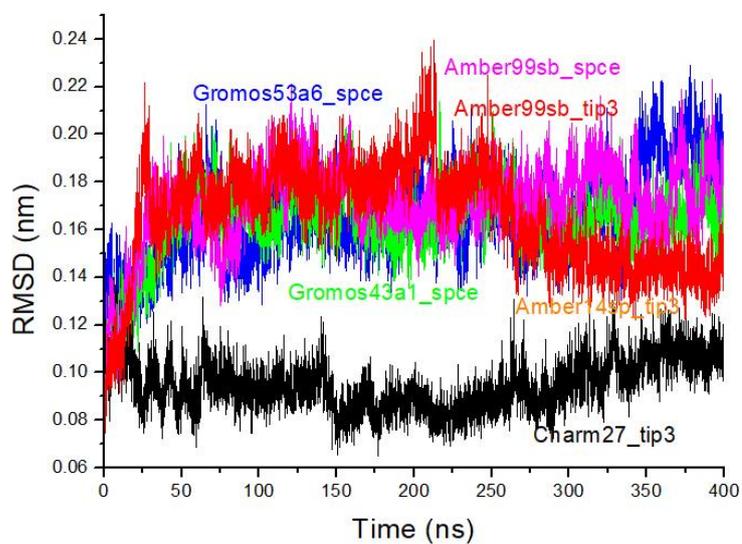
% <sup>a</sup>	Force fields											
	Charmm27_tip3p		Amber99_spc/e		Amber99_tip3p		Amber14_tip3p		Gromos43a1_spc/e		Gromos53a6_spc/e	
	number	% <sup>c</sup>	number	% <sup>c</sup>	number	% <sup>c</sup>	number	% <sup>c</sup>	number	% <sup>c</sup>	number	% <sup>c</sup>
>0.1	28	46.7	30	50.0	28	46.7	26	43.3	48	80.0	46	76.7
>1	24	40.0	27	45.0	22	36.7	24	40.0	39	65.0	34	56.7
>10	21	35.0	21	35.0	18	30.0	14	23.3	28	46.7	26	43.3
>40	8	13.3	15	25.0	16	26.7	2	3.3	13	21.7	8	13.3
>70	2	3.3	7	11.7	10	16.7	0	0	6	10.0	5	5.0
>80	1	1.7	3	5.0	8	13.3	0	0	5	8.3	5	5.0
>90	0	0	3	5.0	5	8.3	0	0	3	5.0	0	0

<sup>a</sup> Occupancy of each ion-ion interaction.

<sup>b</sup> Percentage of 186 targeted ion-ion interactions.

<sup>c</sup> Percentage of 60 targeted ion-ion interactions.

(A)



(B)

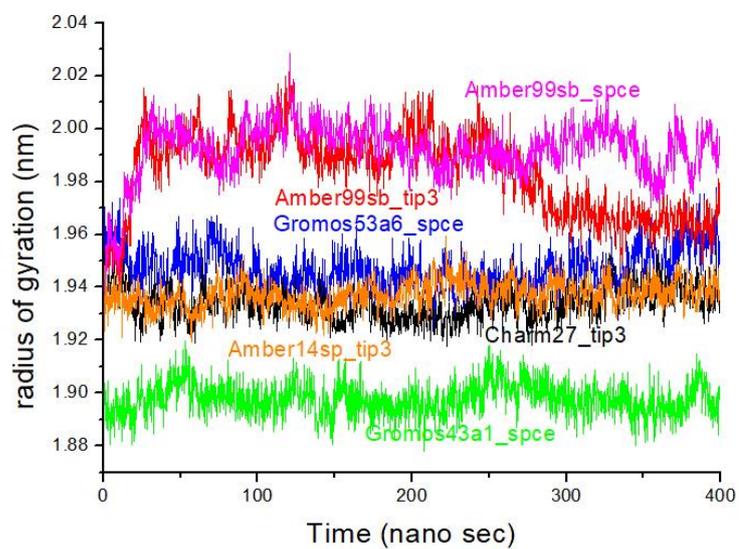
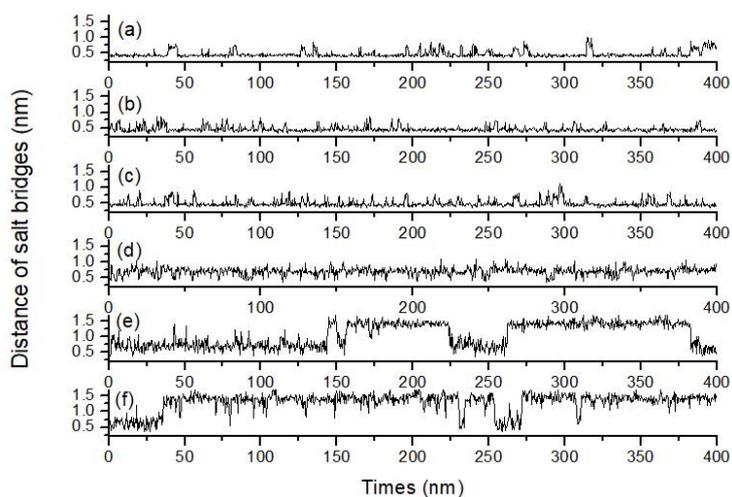


Fig. 1. Trajectories of RMSD (A) of  $C\alpha$  atoms and  $R_g$  (B) for *PhCutA1* at six different force fields over 400 ns of MD simulations. Black, magenta, red, orange, green, and blue represent Charmm27\_tip3p, Amber99sb\_spce/e, Amber99sb\_tip3p, Amber14sb\_tip3p, Gromos43a1\_spce/e, and Gromo53a6\_spce/e, respectively.

(A)



(B)

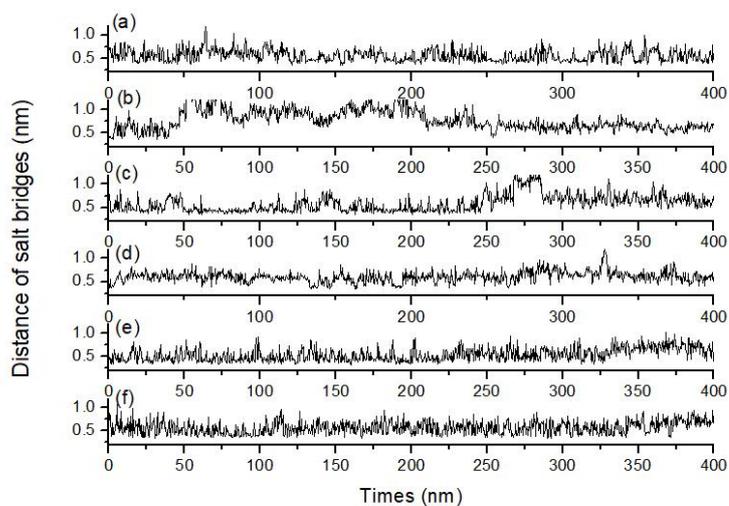


Fig. 2. Trajectory of the length of typical salt bridges in *PhCutA1* at different force fields. (a), (b), (c), (d), (e), and (f) represent Charmm27\_tip3p, Amber99sb\_spc/e, Amber99sb\_tip3p, Amber14sb\_tip3p, Gromos43a1\_spc/e, and Gromo53a6\_spc/e, respectively. (A) The distances between the  $C_{\epsilon}$  atom of Lys101 and the  $C_{\delta}$  atom of Glu64 in the A subunit. (B) The distances between the  $C_{\epsilon}$  atom of Lys70 in the C subunit and the  $C_{\gamma}$  atom of Asp91 in the A-subunit.

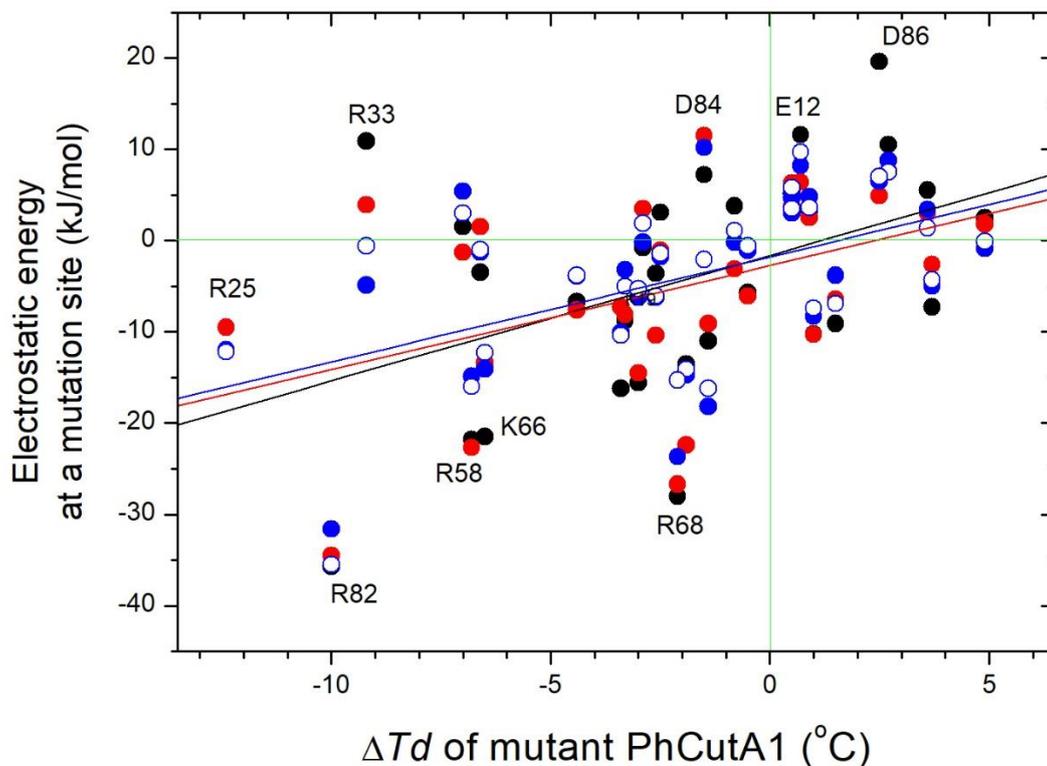


Fig. 3. Correlation between electrostatic energy at mutation sites estimated from FoldX using different structures and difference in denaturation temperatures due to mutation of *PhCutA1*. Black represents the structure obtained from ABC subunits from 4nyo (crystal structures). Red, blue, and open circles represent from structures in MD simulations using Amber99sb\_tip3p, Gromos43a1\_spc/e, and Gromos53a6\_spc/e, respectively. All data are listed in Table S4. Linear lines of black, red, and blue represent linear regressions of the data for ABC subunits from 4nyo, Amber99sb\_tip3p, and Gromos53a6\_spc/e, respectively.

### 3-5. Supporting information.

Table S1.

(A) 186 intra-subunit interactions between favorable ion pairs in *PhCutA1*.

1 Met1AN_Glu59ACD	61 Lys44ACE_Glu47ACD	121 Lys70ACE_Asp76ACG
2 Met1BN_Glu59BCD	62 Lys44BCE_Glu47BCD	122 Lys70BCE_Asp76BCG
3 Met1CN_Glu59CCD	63 Lys44CCE_Glu47CCD	123 Lys70CCE_Asp76CCG
4 Met1AN_Asp84ACG	64 Lys49ACE_Asp10ACG	124 Arg82ACZ_Glu59ACD
5 Met1BN_Asp84BCG	65 Lys49BCE_Asp10BCG	125 Arg82BCZ_Glu59BCD
6 Met1CN_Asp84CCG	66 Lys49CCE_Asp10CCG	126 Arg82CCZ_Glu59CCD
7 Met1AN_Asp86ACG	67 Lys49ACE_Glu12ACD	127 Arg82ACZ_Asp60ACG
8 Met1BN_Asp86BCG	68 Lys49BCE_Glu12BCD	128 Arg82BCZ_Asp60BCG
9 Met1CN_Asp86CCG	69 Lys49CCE_Glu12CCD	129 Arg82CCZ_Asp60CCG
10 Lys16ACE_Asp10ACG	70 Lys49ACE_Glu34ACD	130 Arg82ACZ_Glu63ACD
11 Lys16BCE_Asp10BCG	71 Lys49BCE_Glu34BCD	131 Arg82BCZ_Glu63BCD
12 Lys16CCE_Asp10CCG	72 Lys49CCE_Glu34CCD	132 Arg82CCZ_Glu63CCD
13 Lys16ACE_Glu12ACD	73 Lys49ACE_Glu47ACD	133 Arg82ACZ_Asp84ACG
14 Lys16BCE_Glu12BCD	74 Lys49BCE_Glu47BCD	134 Arg82BCZ_Asp84BCG
15 Lys16CCE_Glu12CCD	75 Lys49CCE_Glu47CCD	135 Arg82CCZ_Asp84CCG
16 Lys16ACE_Glu15ACD	76 Lys49ACE_Asp48ACG	136 Lys94ACE_Glu90ACD
17 Lys16BCE_Glu15BCD	77 Lys49BCE_Asp48BCG	137 Lys94BCE_Glu90BCD
18 Lys16CCE_Glu15CCD	78 Lys49CCE_Asp48CCG	138 Lys94CCE_Glu90CCD
19 Lys19ACE_Glu12ACD	79 Arg58ACZ_Glu59ACD	139 Lys94ACE_Asp91ACG
20 Lys19BCE_Glu12BCD	80 Arg58BCZ_Glu59BCD	140 Lys94BCE_Asp91BCG
21 Lys19CCE_Glu12CCD	81 Arg58CCZ_Glu59CCD	141 Lys94CCE_Asp91CCG
22 Lys19ACE_Glu15ACD	82 Arg58ACZ_Asp60ACG	142 Lys94ACE_Glu98ACD
23 Lys19BCE_Glu15BCD	83 Arg58BCZ_Asp60BCG	143 Lys94BCE_Glu98BCD
24 Lys19CCE_Glu15CCD	84 Arg58CCZ_Asp60CCG	144 Lys94CCE_Glu98CCD
25 Lys23ACE_Glu24ACD	85 Arg58ACZ_Glu98ACD	145 Lys94ACE_Glu99ACD
26 Lys23BCE_Glu24BCD	86 Arg58BCZ_Glu98BCD	146 Lys94BCE_Glu99BCD
27 Lys23CCE_Glu24CCD	87 Arg58CCZ_Glu98CCD	147 Lys94CCE_Glu99CCD
28 Arg25ACZ_Glu24ACD	88 Arg58ACZ_Lys102AC	148 Lys101ACE_Glu24ACD
29 Arg25BCZ_Glu24BCD	89 Arg58BCZ_Lys102BC	149 Lys101BCE_Glu24BCD
30 Arg25CCZ_Glu24CCD	90 Arg58CCZ_Lys102CC	150 Lys101CCE_Glu24CCD
31 Arg25ACZ_Glu98ACD	91 Lys66ACE_Glu63ACD	151 Lys101ACE_Asp60ACG
32 Arg25BCZ_Glu98BCD	92 Lys66BCE_Glu63BCD	152 Lys101BCE_Asp60BCG
33 Arg25CCZ_Glu98CCD	93 Lys66CCE_Glu63CCD	153 Lys101CCE_Asp60CCG
34 Arg25ACZ_Glu99ACD	94 Lys66ACE_Glu67ACD	154 Lys101ACE_Glu64ACD
35 Arg25BCZ_Glu99BCD	95 Lys66BCE_Glu67BCD	155 Lys101BCE_Glu64BCD
36 Arg25CCZ_Glu99CCD	96 Lys66CCE_Glu67CCD	156 Lys101CCE_Glu64CCD
37 Arg25ACZ_Lys102AC	97 Arg68ACZ_Glu24ACD	157 Lys101ACE_Glu98ACD
38 Arg25BCZ_Lys102BC	98 Arg68BCZ_Glu24BCD	158 Lys101BCE_Glu98BCD
39 Arg25CCZ_Lys102CC	99 Arg68CCZ_Glu24CCD	159 Lys101CCE_Glu98CCD
40 Arg33ACZ_Glu34ACD	100 Arg68ACZ_Glu64ACD	160 Lys101ACE_Glu99ACD
41 Arg33BCZ_Glu34BCD	101 Arg68BCZ_Glu64BCD	161 Lys101BCE_Glu99BCD
42 Arg33CCZ_Glu34CCD	102 Arg68CCZ_Glu64CCD	162 Lys101CCE_Glu99CCD
43 Arg36ACZ_Glu34ACD	103 Arg68ACZ_Glu67ACD	163 Lys101ACE_Lys102AC
44 Arg36BCZ_Glu34BCD	104 Arg68BCZ_Glu67BCD	164 Lys101BCE_Lys102BC
45 Arg36CCZ_Glu34CCD	105 Arg68CCZ_Glu67CCD	165 Lys101CCE_Lys102CC
46 Arg36ACZ_Glu46ACD	106 Arg68ACZ_Glu71ACD	166 Lys102ACE_Glu24ACD
47 Arg36BCZ_Glu46BCD	107 Arg68BCZ_Glu71BCD	167 Lys102BCE_Glu24BCD
48 Arg36CCZ_Glu46CCD	108 Arg68CCZ_Glu71CCD	168 Lys102CCE_Glu24CCD
49 Arg36ACZ_Glu47ACD	109 Arg68ACZ_Lys102AC	169 Lys102ACE_Glu59ACD
50 Arg36BCZ_Glu47BCD	110 Arg68BCZ_Lys102BC	170 Lys102BCE_Glu59BCD
51 Arg36CCZ_Glu47CCD	111 Arg68CCZ_Lys102CC	171 Lys102CCE_Glu59CCD
52 Arg36ACZ_Asp48ACG	112 Lys70ACE_Glu63ACD	172 Lys102ACE_Asp60ACG
53 Arg36BCZ_Asp48BCG	113 Lys70BCE_Glu63BCD	173 Lys102BCE_Asp60BCG
54 Arg36CCZ_Asp48CCG	114 Lys70CCE_Glu63CCD	174 Lys102CCE_Asp60CCG
55 Lys44ACE_Glu42ACD	115 Lys70ACE_Glu67ACD	175 Lys102ACE_Glu63ACD
56 Lys44BCE_Glu42BCD	116 Lys70BCE_Glu67BCD	176 Lys102BCE_Glu63BCD
57 Lys44CCE_Glu42CCD	117 Lys70CCE_Glu67CCD	177 Lys102CCE_Glu63CCD
58 Lys44ACE_Glu46ACD	118 Lys70ACE_Glu71ACD	178 Lys102ACE_Glu64ACD
59 Lys44BCE_Glu46BCD	119 Lys70BCE_Glu71BCD	179 Lys102BCE_Glu64BCD
60 Lys44CCE_Glu46CCD	120 Lys70CCE_Glu71CCD	180 Lys102CCE_Glu64CCD
		181 Lys102ACE_Glu98ACD
		182 Lys102BCE_Glu98BCD
		183 Lys102CCE_Glu98CCD
		184 Lys102ACE_Glu99ACD
		185 Lys102BCE_Glu99BCD
		186 Lys102CCE_Glu99CCD

(B) 60 inter-subunit interactions between favorable ion pairs in *PhCutA1*.

1	Lys19ACE_Glu46CCD	21	Arg33CCZ_Glu47BCD	41	Lys66BCE_Asp91CCG
2	Lys19BCE_Glu46ACD	22	Arg36ACZ_Glu15BCD	42	Lys66CCE_Asp91ACG
3	Lys19CCE_Glu46BCD	23	Arg36BCZ_Glu15CCD	43	Lys70ACE_Glu90BCD
4	Lys19ACE_Glu47CCD	24	Arg36CCZ_Glu15ACD	44	Lys70BCE_Glu90CCD
5	Lys19BCE_Glu47ACD	25	Arg36ACZ_Glu34BCD	45	Lys70CCE_Glu90ACD
6	Lys19CCE_Glu47BCD	26	Arg36BCZ_Glu34CCD	46	Lys70ACE_Asp91BCG
7	Lys23ACE_Glu42CCD	27	Arg36CCZ_Glu34ACD	47	Lys70BCE_Asp91CCG
8	Lys23BCE_Glu42ACD	28	Lys44ACE_Glu15BCD	48	Lys70CCE_Asp91ACG
9	Lys23CCE_Glu42BCD	29	Lys44BCE_Glu15CCD	49	Arg82ACZ_Asp86BCG
10	Arg25ACZ_Glu42CCD	30	Lys44CCE_Glu15ACD	50	Arg82BCZ_Asp86CCG
11	Arg25BCZ_Glu42ACD	31	Lys56ACE_Glu50CCD	51	Arg82CCZ_Asp86ACG
12	Arg25CCZ_Glu42BCD	32	Lys56BCE_Glu50ACD	52	Arg82ACZ_Asp87BCG
13	Arg33ACZ_Glu34BCD	33	Lys56CCE_Glu50BCD	53	Arg82BCZ_Asp87CCG
14	Arg33ACZ_Glu34CCD	34	Lys66ACE_Asp87BCG	54	Arg82CCZ_Asp87ACG
15	Arg33BCZ_Glu34ACD	35	Lys66BCE_Asp87CCG	55	Lys101BCE_Glu42ACD
16	Arg33BCZ_Glu34CCD	36	Lys66CCE_Asp87ACG	56	Lys101CCE_Glu42BCD
17	Arg33CCZ_Glu34ACD	37	Lys66ACE_Glu90BCD	57	Lys101ACE_Glu42CCD
18	Arg33CCZ_Glu34BCD	38	Lys66BCE_Glu90CCD	58	Lys102ACE_Glu42CCD
19	Arg33ACZ_Glu47CCD	39	Lys66CCE_Glu90ACD	59	Lys102BCE_Glu42ACD
20	Arg33BCZ_Glu47ACD	40	Lys66ACE_Asp91BCG	60	Lys102CCE_Glu42BCD

Table S2. Number of residues of *PhCutA1* in each type of secondary structure in MD simulations (50-400 ns). Values represent the average number of residues in each type of secondary structures among 3 subunits.

force fields	structure*	$\beta$ -sheet	$\alpha$ -helix
Charmm27_tip3p	77.0 $\pm$ 1.2	39.4 $\pm$ 0.7	33.3 $\pm$ 0.6
Amber99sb_spc/e	77.5 $\pm$ 1.5	39.5 $\pm$ 0.8	31.7 $\pm$ 1.4
Amber99sb_tip3p	76.6 $\pm$ 1.4	38.7 $\pm$ 0.9	30.9 $\pm$ 1.5
Amber14sp_tip3p	78.1 $\pm$ 0.9	39.9 $\pm$ 0.5	33.4 $\pm$ 0.7
Gromos43a1_spc/e	76.6 $\pm$ 2.0	38.0 $\pm$ 1.7	33.6 $\pm$ 0.5
Gromos53a6_spc/e	75.6 $\pm$ 1.8	37.3 $\pm$ 1.4	33.6 $\pm$ 0.4

\*structure =  $\beta$ -sheet +  $\alpha$ -helix +  $\beta$ -bridge + turn.

Table S3.

(A) Average distance between favorable intra-subunit salt bridges in *PhCutA1* during 400-ns MD simulations using the indicated force fields. (B) Average distance between favorable inter-subunit salt bridges in *PhCutA1* during 400-ns MD simulations using the indicated force fields. These values represent the average of three subunits. The unit of the distance is nm. These data are shown when ion-pairs (less than 0.7 nm) were detected at least once among six force fields. Yellow and orange represent the lowest and highest values of the distance among six force fields, respectively.

(A)

targeted residues	pair residues	Force fields					
		Charmm27_tip3p	Amber99sb_spc/e	Amber99sb_tip3p	Amber14sb_tip3p	Gromos43a1_spc/e	Gromos53a6_spc/e
N-term	Glu59	0.37 ± 0.08	0.44 ± 0.13	0.54 ± 0.17	0.37 ± 0.09	0.39 ± 0.07	0.38 ± 0.04
	Asp84	0.66 ± 0.05	0.62 ± 0.08	0.72 ± 0.17	0.72 ± 0.07	0.62 ± 0.06	0.67 ± 0.07
	Asp86	0.96 ± 0.05	0.98 ± 0.07	1.08 ± 0.14	0.99 ± 0.05	0.90 ± 0.09	0.70 ± 0.12
Lys16	Glu12	0.74 ± 0.19	0.75 ± 0.22	0.72 ± 0.21	0.70 ± 0.23	0.62 ± 0.18	0.64 ± 0.18
Lys19	Glu15	0.45 ± 0.10	0.45 ± 0.09	0.43 ± 0.07	0.45 ± 0.10	0.49 ± 0.08	0.49 ± 0.08
Arg25	Glu99	0.46 ± 0.06	0.48 ± 0.05	0.49 ± 0.07	0.49 ± 0.03	0.50 ± 0.11	0.48 ± 0.07
Arg33	Glu34	0.93 ± 0.04	1.09 ± 0.11	1.04 ± 0.09	0.91 ± 0.07	0.68 ± 0.20	0.83 ± 0.16
Arg36	Glu47	0.47 ± 0.09	0.53 ± 0.15	0.53 ± 0.14	0.64 ± 0.22	0.56 ± 0.12	0.56 ± 0.11
Lys44	Glu46	0.56 ± 0.15	0.63 ± 0.19	0.68 ± 0.20	0.72 ± 0.21	0.69 ± 0.23	0.59 ± 0.19
Lys49	Glu34	0.49 ± 0.15	0.47 ± 0.10	0.50 ± 0.14	0.61 ± 0.21	0.65 ± 0.19	0.58 ± 0.16
Arg58	Asp60	0.40 ± 0.02	0.42 ± 0.03	0.41 ± 0.02	0.41 ± 0.02	0.52 ± 0.11	0.48 ± 0.07
Lys66	C-term	0.62 ± 0.09	0.61 ± 0.07	0.62 ± 0.08	0.59 ± 0.08	0.79 ± 0.19	0.74 ± 0.16
	Glu63	0.58 ± 0.15	0.57 ± 0.18	0.51 ± 0.16	0.43 ± 0.11	0.55 ± 0.11	0.62 ± 0.11
	Glu67	0.78 ± 0.11	0.73 ± 0.16	0.82 ± 0.14	0.83 ± 0.14	0.70 ± 0.13	0.77 ± 0.13
Arg68	Glu24	0.41 ± 0.03	0.41 ± 0.03	0.43 ± 0.08	0.59 ± 0.18	0.58 ± 0.13	0.57 ± 0.12
Lys70	Glu71	0.52 ± 0.10	0.49 ± 0.08	0.49 ± 0.08	0.54 ± 0.14	0.65 ± 0.17	0.71 ± 0.16
	Glu67	0.71 ± 0.19	0.60 ± 0.20	0.64 ± 0.20	0.66 ± 0.19	0.63 ± 0.18	0.71 ± 0.17
	Glu59	0.64 ± 0.10	0.63 ± 0.13	0.61 ± 0.15	0.64 ± 0.11	0.61 ± 0.08	0.60 ± 0.07
Arg82	Asp84	0.40 ± 0.02	0.74 ± 0.07	0.77 ± 0.08	0.48 ± 0.07	0.67 ± 0.06	0.47 ± 0.08
Lys94	Glu90	0.77 ± 0.23	0.70 ± 0.24	0.70 ± 0.23	0.81 ± 0.23	0.69 ± 0.20	0.77 ± 0.20
Lys101	Asp91	0.68 ± 0.13	0.67 ± 0.13	0.63 ± 0.15	0.69 ± 0.15	0.67 ± 0.15	0.72 ± 0.14
	Glu64	0.47 ± 0.10	0.46 ± 0.10	0.49 ± 0.12	0.70 ± 0.12	1.20 ± 0.34	1.27 ± 0.32
	C-term	0.62 ± 0.07	0.60 ± 0.07	0.61 ± 0.08	0.59 ± 0.07	0.75 ± 0.10	0.76 ± 0.09
Lys102	Glu98	0.66 ± 0.27	0.68 ± 0.26	0.70 ± 0.28	0.61 ± 0.23	1.27 ± 0.44	1.43 ± 0.37

(B)

targeted residues	pair residues	Force fields					
		Charmm27_tip3p	Amber99sb_spc/e	Amber99sb_tip3p	Amber14sb_tip3p	Gromos43a1_spc/e	Gromos63a6_spc/e
Lys19 A	Glu47 C	0.62 ± 0.24	0.58 ± 0.20	0.47 ± 0.09	0.61 ± 0.17	0.70 ± 0.19	0.69 ± 0.16
	B	0.50 ± 0.13	0.51 ± 0.12	0.46 ± 0.07	0.57 ± 0.15	0.61 ± 0.18	0.65 ± 0.17
	C	0.46 ± 0.10	0.50 ± 0.12	0.46 ± 0.07	0.57 ± 0.16	0.69 ± 0.16	0.67 ± 0.17
Arg33 A	Glu34 C	1.24 ± 0.07	1.12 ± 0.18	1.22 ± 0.12	1.20 ± 0.14	0.65 ± 0.14	0.98 ± 0.16
	B	1.22 ± 0.07	1.26 ± 0.17	1.26 ± 0.10	1.21 ± 0.11	0.52 ± 0.09	0.71 ± 0.25
	C	1.15 ± 0.11	1.05 ± 0.20	0.96 ± 0.23	1.19 ± 0.14	0.58 ± 0.16	0.59 ± 0.11
Arg36 A	Glu15 B	0.66 ± 0.13	0.60 ± 0.18	0.59 ± 0.17	0.62 ± 0.16	0.66 ± 0.15	0.75 ± 0.14
	B	0.63 ± 0.14	0.59 ± 0.18	0.59 ± 0.18	0.70 ± 0.14	0.66 ± 0.15	0.75 ± 0.15
	C	0.63 ± 0.15	0.60 ± 0.18	0.47 ± 0.11	0.63 ± 0.16	0.68 ± 0.15	0.74 ± 0.14
Lys56 A	Glu50 C	0.76 ± 0.07	0.69 ± 0.11	0.90 ± 0.12	0.85 ± 0.09	0.63 ± 0.08	0.66 ± 0.05
	B	0.76 ± 0.06	0.67 ± 0.08	0.81 ± 0.12	0.83 ± 0.10	0.53 ± 0.05	0.66 ± 0.07
	C	0.73 ± 0.05	0.68 ± 0.09	0.85 ± 0.11	0.85 ± 0.10	0.60 ± 0.07	0.67 ± 0.07
Lys66 A	Asp87 B	0.61 ± 0.08	0.59 ± 0.11	0.60 ± 0.11	0.55 ± 0.07	0.61 ± 0.07	0.71 ± 0.10
	B	0.62 ± 0.06	0.60 ± 0.10	0.54 ± 0.09	0.57 ± 0.07	0.64 ± 0.06	0.70 ± 0.10
	C	0.62 ± 0.07	0.59 ± 0.09	0.59 ± 0.09	0.58 ± 0.07	0.63 ± 0.07	0.69 ± 0.10
	A	0.77 ± 0.24	0.86 ± 0.20	1.13 ± 0.16	0.72 ± 0.16	0.69 ± 0.18	0.68 ± 0.18
	B	0.60 ± 0.22	0.94 ± 0.22	1.02 ± 0.29	0.71 ± 0.15	0.72 ± 0.18	0.66 ± 0.18
	C	0.66 ± 0.25	0.79 ± 0.20	0.73 ± 0.23	0.69 ± 0.16	0.69 ± 0.18	0.64 ± 0.17
Lys70 A	Asp91 B	0.61 ± 0.14	0.84 ± 0.21	1.04 ± 0.18	0.65 ± 0.13	0.63 ± 0.17	0.58 ± 0.13
	B	0.54 ± 0.12	0.98 ± 0.21	0.89 ± 0.25	0.59 ± 0.11	0.55 ± 0.13	0.58 ± 0.13
	C	0.56 ± 0.13	0.78 ± 0.22	0.59 ± 0.18	0.61 ± 0.12	0.53 ± 0.14	0.57 ± 0.14
Arg82 A	Asp86 B	0.64 ± 0.09	0.45 ± 0.05	0.44 ± 0.04	0.59 ± 0.05	0.50 ± 0.08	0.84 ± 0.13
	B	0.63 ± 0.05	0.45 ± 0.05	0.44 ± 0.03	0.61 ± 0.05	0.51 ± 0.08	0.83 ± 0.13
	C	0.63 ± 0.05	0.44 ± 0.03	0.44 ± 0.03	0.50 ± 0.10	0.52 ± 0.11	0.82 ± 0.14
	A	0.62 ± 0.05	0.55 ± 0.08	0.60 ± 0.11	0.60 ± 0.06	0.62 ± 0.07	0.50 ± 0.07
	B	0.62 ± 0.05	0.54 ± 0.08	0.54 ± 0.09	0.59 ± 0.06	0.60 ± 0.07	0.50 ± 0.07
	C	0.62 ± 0.05	0.52 ± 0.08	0.51 ± 0.07	0.57 ± 0.08	0.60 ± 0.08	0.51 ± 0.07

Table S4. Electrostatic energy of targeted residues for two structures from crystal analysis and six structures from MD simulation of *PhCutA1*.

Targeted sites	Mutations	$\Delta T_d$ (°C)*	Electrostatic energy of targeted residues (kJ/mol)								
			Crystal structures of 4nyo		Structures of MD simulations						
			ABC	DEF	Charmm27 _tip3p	Amber99sb _spc/e	Amber99sb _tip3p	Amber14sb _tip3p	Gromos43a1 _spc/e	Gromos53a6 _spc/e	
D48	D48N	2.7	10.5	10.8	8.2	6.1	7.6	10.4	8.8	7.5	
D60	D60A	-3.3	-8.8	-7.3	-3.3	-6.9	-8.1	-6.1	-3.2	-5.0	
D76	D76N	0.9	3.3	3.1	3.3	1.8	2.5	4.1	4.8	3.6	
D84	D84A	-1.5	7.2	6.8	4.1	11.1	11.5	9.0	10.2	-2.1	
D86	D86N	2.5	19.6	19.0	16.5	4.5	4.9	15.2	6.5	7.0	
D87	D87N	-7.0	1.5	2.2	2.3	-1.8	-1.3	1.1	5.4	3.0	
D91	D91A	-6.6	-3.5	0.5	-0.4	3.6	1.5	1.9	-1.3	-1.0	
E12	E12Q	0.7	11.6	10.8	7.9	6.5	6.4	7.7	8.2	9.7	
E15	E15A	-0.8	3.8	1.0	0.0	-2.8	-3.1	-1.4	-0.2	1.1	
E24	E24A	-3.0	-15.6	-15.9	-15.1	-15.3	-14.5	-9.0	-6.2	-5.3	
E34	E34Q	-1.4	-11.0	-14.4	-10.3	-8.7	-9.1	-10.9	-18.2	-16.2	
E42	E42Q	4.9	2.5	3.7	1.8	1.0	1.8	2.4	-0.9	-0.1	
E46	E46Q	3.6	5.5	6.2	3.3	1.9	3.0	5.8	3.4	1.4	
E47	E47A	-2.5	3.1	0.8	0.3	-2.5	-1.1	1.6	-1.8	-1.4	
E59	E59Q	3.7	-7.3	-6.6	-4.7	-2.5	-2.6	-4.5	-5.0	-4.3	
E63	E63A	0.5	5.1	5.5	7.6	5.5	6.3	5.8	4.7	5.8	
E64	E64A	-2.9	-0.8	3.2	4.7	2.1	3.5	3.3	-0.2	1.9	
E67	E67A	0.5	5.2	5.4	5.5	3.5	3.7	0.3	3.0	3.5	
E71	E71A	-0.5	-5.7	-4.4	-4.6	-5.5	-6.1	-5.0	-1.1	-0.6	
E99	E99A	1.5	-9.1	-5.1	-7.3	-7.4	-6.4	-7.6	-3.8	-6.9	
K101	K101A	-4.4	-6.7	-4.5	-7.1	-8.6	-7.6	-5.9	-3.9	-3.8	
K19	K19A	-2.6	-3.6	-6.5	-7.2	-8.5	-10.4	-5.7	-6.2	-6.1	
K49	K49A	1.0	-10.2	-8.6	-8.9	-10.3	-10.3	-6.0	-8.3	-7.4	
K66	K66A	-6.5	-21.5	-18.7	-16.1	-15.0	-13.4	-14.2	-14.1	-12.3	
K70	K70A	-3.4	-16.2	-10.8	-10.1	-7.7	-7.3	-8.9	-10.1	-10.4	
R25	R25A	-12.4	-12.0	-6.9	-10.1	-10.4	-9.5	-9.6	-12.0	-12.2	
R33	R33A	-9.2	10.9	3.5	5.8	2.3	3.9	1.6	-4.9	-0.6	
R36	R36A	-1.9	-13.5	-11.9	-18.8	-22.9	-22.4	-17.3	-14.7	-14.1	
R58	R58A	-6.8	-21.8	-25.2	-21.7	-22.5	-22.7	-22.9	-14.9	-16.0	
R68	R68A	-2.1	-28.0	-26.0	-25.2	-27.4	-26.7	-31.3	-23.7	-15.3	
R82	R82A	-10.0	-35.7	-34.8	-34.2	-32.6	-34.5	-34.6	-31.6	-35.5	

\*Difference in denaturation temperatures of *PhCutA1* mutants<sup>9</sup>.

Table S5. Comparison of percent occupancy of intra-subunit salt bridges in *PhCutA1* at each 100 ns during 400 ns MD simulation at 300 K using indicated force fields. Data show average values of percent occupancies of 17 positively charged residues indicated in Table 2A. Average (1) and (2) represent average values from 0 to 400-ns and 100 to 400-ns, respectively.

	Structures of MD simulations					
	Charmm27 _tip3p	Amber99sb _spc/e	Amber99sb _tip3p	Amber14sb _tip3p	Gromos43a1 _spc/e	Gromos53a6 _spc/e
0-100 ns	81.5	102.6	76.0	77.2	84.1	73.2
100-200 ns	83.5	83.5	78.3	77.6	74.2	74.2
200-300 ns	83.8	74.9	81.6	74.4	77.6	74.9
300-400 ns	81.3	73.6	83.1	81.6	80.3	73.6
Average(1)	82.5 ± 1.3	83.6 ± 13.4	79.7 ± 3.2	77.7 ± 3.0	79.0 ± 4.2	74.0 ± 0.7
Average(2)	82.9 ± 1.4	77.3 ± 5.4	81.0 ± 2.4	77.9 ± 3.6	77.4 ± 3.0	74.2 ± 0.6

Table S6. Comparison of percent occupancy of intra-subunit salt bridges in each subunit of *PhCutA1* during 400 ns MD simulation at 300 K using indicated force fields. Data show average values of percent occupancies of 17 positively charged residues indicated in Table 2A. STDEV represent the standard deviation of average values for 3 subunits.

	Force fields					
	Charmm27 _tip3p	Amber99sb _spc/e	Amber99sb _tip3p	Amber14sb _tip3p	Gromos43a1 _spc/e	Gromos53a6 _spc/e
A-subunit	82.9	84.3	78.4	79.0	79.9	71.8
B-subunit	82.4	82.6	79.2	78.9	78.7	74.8
C-subunit	82.3	84.0	81.6	75.3	78.6	75.4
Average	82.5	83.6	79.7	77.7	79.0	74.0
STDEV	0.3	0.9	1.6	2.1	0.7	1.9

Table S7. Side-chain rotamer criteria of charged residues in *PhCutA1*. The side chains of the charged residues in *PhCutA1* shown in Fig. S4 to Fig. S14 were examined by MolProbity (<http://molprobity.biochem.duke.edu>). (p (plus, centered near +60°), t (trans, centered near 180°), and m (minus, centered near -60°)).

Figure	Force field	Residue	Rotamer's configuration	Validation		Dihedral angles, chi1, chi2, chi3, chi4
				(Favored (>2.0%), Allowed (0.3 - 2.0%), Outlier (≠0.3%))		
S4	Amber99sb_tip3p (at 100 ns)	Arg68B	tpt170	25.9	178.9, 56.9, 186.7, 149.9	
		Glu64B	tt0	26.9	189.6, 183.6, 117.2	
		Glu24B	tt0	1.1	153.9, 163.7, 249.5	
	Amber14sb_tip3p (at 100 ns)	Glu71B	mt-10	8.5	314.0, 193.3, 298.6	
		Arg68B	mmm-85	34	293.2, 291.5, 282.5, 287.1	
		Glu64B	tp30	30.6	185.6, 66.4, 177.1	
		Glu24B	mm-30	33.8	289.8, 282.7, 320.3	
Glu71B	mm-30	77.9	293.9, 307.4, 136.7			
S5	4nyo (crystal structure)	Arg82A	ttt180	35.5	164.0, 181.5, 169.0, 164.9	
		Asp86B	m-30	36	297.3, 287	
		Asp87B	t0	30.6	198.4, 196.4	
		Glu63A	mm-30	77.4	293.8, 302.9, 327.9	
		Asp84A	m-30	18.8	293.2, 187.8	
	Gromos53a6_spc/e (at 200ns)	Glu59A	mt-10	36.1	289.5, 166.8, 43.7	
		Arg82A	ttt180	11	183, 189.5, 160.4, 145.8	
		Asp86B	m-30	5.8	275.3, 128.7	
		Asp87B	m-30	2.5	290.6, 29.3	
		Glu63A	mt-10	50.3	294.7, 177.6, 228.1	
		Asp84A	t70	33.5	188.6, 57.8	
Glu59A	pm20	20.1	63.7, 282.7, 8.5			
S6	Amber99sb_tip3p (at 100ns)	Met1	ntt	20.8	282.8, 199.3, 182.4	
		Glu59	tt0	38	186.2, 171.1, 243.5	
		Asp84	tt0	9	171.7, 55.7	
		ASP86	T70	32.7	180.9, 239.1	
S7	Charmm27_tip3p (at 100ns)	Lys44A	pttp	3.4	62.6, 189.4, 165.7, 35.5	
		Glu42A	tt0	8.3	201, 164.1, 243.1	
	Gromos53a6_spc/e (at 100ns)	Glu46A	mm-30	7.9	270.1, 296.6, 329.8	
		Lys44A	mtmt	45.7	297.3, 183.4, 301.6, 183.9	
Glu42A	mt-10	13.8	298.6, 200.3, 82.8			
Glu46A	mm-30	41.6	295.6, 283.4, 324.5			
S8	Gromos43a1_spc/e (at 100ns)	Lys70B	pttt	44.8	69.4, 176.7, 179.2, 166.7	
		Glu67B	mt-10	30	303.6, 186.3, 233.4	
		Asp76B	p0	1.1	45.1, 106.1	
		Glu90C	mp0	1.7	288.6, 60, 258	
	Amber99sb_tip3p (at 100 ns)	Asp91C	m-30	6.1	300, 263.5	
		Lys70B	mtmt	9.4	283.3, 178, 307.6, 208.5	
		Glu67B	tp30	4.7	200, 59.1, 77.1	
		Asp76B	t70	18	193.6, 250.2	
Glu90C	mt-10	35.1	287.6, 185.4, 99.3			
Asp91C	t70	27.6	191.7, 242.4			
S9	Amber99sb_tip3p (at 200ns)	Arg36A	ntp180	81.3	294, 186, 62.1, 190.5	
		Glu34A	tt0	23.1	172.9, 187.5, 76.9	
		Glu47A	outlier	0.1	58.2, 257.7, 296.3	
		Glu15B	mt-10	38.7	292.4, 183.4, 78.5	
	Gromos53a6_spc/e (at 200ns)	Glu34B	tt0	11.5	190.1, 193.8, 98.5	
		Arg36A	ntt180	49.9	289.6, 191.2, 191.4, 155.1	
		Glu34A	tt0	8.4	197.9, 193.6, 62.2	
		Glu47A	pm20	1.4	51.4, 279.5, 64.3	
Glu15B	tt0	31.5	187.6, 176.6, 266.3			
Glu34B	mp0	1.5	314.9, 82.8, 121.2			

Table S7 continued.

S10	Charmm27_tip3p	Lys101B	pttt	27.9	73.9, 162.8, 190, 177.4
	(at 100ns)	Lys102B	mmtm	18.3	305.7, 306.1, 183.4, 314.5
		Glu64B	tt0	37.4	174.5, 165.2, 25.7
		Glu98B	mm-30	25	298.4, 285.2, 303.8
	Gromos53a6_spc/e	Lys101B	pttt	37.9	50.2, 190.9, 176.9, 185.5
	(at 100ns)	Lys102B	tttt	18.4	205.7, 196.5, 175.5, 171.3
		Glu64B	mt-10	0.3	253.7, 165.3, 268.4
	Glu98B	mm-30	18.2	289.9, 279.2, 131.8	
S11	Amber14sb_tip3p	Arg58A	mtm-85	4.7	310.2, 194.1, 311.7, 232.2
	(at 200 ns)	Asp60A	p0	2.5	75.6, 303.3
		Lys102A	mtp	5.9	308.8, 201.9, 72.7, 69.9
	Gromos43a1_spc/e	Arg58A	ptt180	22.6	57.6, 174.9, 165.3, 169.7
	(at 200ns)	Asp60A	m-30	5.4	299.4, 80.5
	Lys102A	pttp	10	76.3, 194.4, 175.5, 55.1	
S12	Gromos43a1_spc/e	Glu50A	mt-10	34.3	281.9, 187, 134.4
	(at 100ns)	Lys56B	mtp	3.1	311, 191.1, 180.5, 118.8
	Amber99sb_tip3p	Glu50A	tt0	33.5	187.5, 186.6, 230
	(at 100 ns)	Lys56B	mttt	90.6	302.5, 177, 181.7, 176.7
S13	4nyo	Arg82A	ttt180	35.5	164, 181.5, 169, 164.9
	(crystal structure)	Asp84A	m-30	18.8	293.2, 187.8
		Asp86A	m-30	18.7	297.1, 277.6
		Arg82B	ttt180	49.9	166.8, 187.2, 171.3, 177.8
		Asp84B	m-30	6.6	287.1, 21.2
		Asp86B	m-30	36	297.3, 287
		Arg82C	ttt180	56.5	168.6, 182.1, 174.3, 169.4
		Asp84C	m-30	3.4	282.7, 29.3
		Asp86C	m-30	31.8	296.6, 285.2
	Gromos43a1_spc/e	Arg82A	ttt180	11	183, 189.5, 160.4, 145.8
	(at 200ns)	Asp84A	t70	33.5	188.6, 57.8
		Asp86A	t0	22.5	198.2, 34.2
		Arg82B	ttt180	18.1	185.9, 177.3, 153.3, 159.2
		Asp84B	m-30	19	290.6, 103.1
		Asp86B	m-30	5.8	275.3, 128.7
	Arg82C	ttt180	24.1	171.7, 181.1, 166, 151.2	
	Asp84C	m-30	1.8	283.1, 263.9	
	Asp86C	t70	7.6	197.7, 261.3	
S14	Gromos43a1_spc/e	Arg33A	ttt-90	1.6	205.7, 145.5, 189.3, 235.2
	(at 100ns)	Glu34A	pm20	0.3	53.5, 286.5, 81.3
		Arg33B	tpt-90	1.8	207.5, 79.6, 189.9, 243.5
		Glu34B	pm20	0.9	76, 266.6, 69.5
		Arg33C	tpt-90	4.7	195, 180.4, 151.8, 232.2
		Glu34C	pm20	0.9	223.6, 176.6, 107.2
	Amber99sb_tip3p	Arg33A	ptt90	53.5	68.7, 171.7, 176.9, 85.3
	(at 100ns)	Glu34A	tt0	23.4	192.7, 175.3, 296.8
		Arg33B	ptt-90	21.6	71.5, 164, 184.1, 266.5
		Glu34B	tt0	1.1	153.9, 163.7, 249.5
	Arg33C	pmt-80	1.6	83.4, 302.1, 177.8, 293	
	Glu34C	tt0	4.7	209.6, 173.4, 130.3	

Table S8. Percent occupancies of salt bridges (less than 0.6 nm) of negatively charged residues in *PhCutA1* with Na<sup>+</sup> ions during 400-ns MD simulations at 300 K using the indicated force fields.

Negatively Charged Residues	Force fields											
	Charmm27_tip3p		Amber99sb_spc/e		Amber99sb_tip3p		Amber14sb_tip3p		Gromos43a1_spc/e		Gromos53a6_spc/e	
	P.C Residue*	Na <sup>+</sup> ion	P.C Residue*	Na <sup>+</sup> ion	P.C Residue*	Na <sup>+</sup> ion	P.C Residue*	Na <sup>+</sup> ion	P.C Residue*	Na <sup>+</sup> ion	P.C Residue*	Na <sup>+</sup> ion
ASP 10	5.0	32.6	4.4	13.5	6.0	22.0	0.4	40.0	7.8	10.1	0.8	12.8
GLU 12	28.6	50.3	29.4	23.1	32.9	23.1	40.3	67.2	54.2	16.8	47.8	18.8
GLU 15	125.7	67.3	148.4	14.7	162.3	16.9	127.4	78.2	125.2	10.9	104.1	10.2
GLU 24	114.5	5.7	111.2	3.2	109.5	4.1	76.7	16.5	73.0	12.8	80.2	9.3
GLU 34	82.7	11.0	94.2	4.8	88.3	5.0	57.4	12.8	175.5	1.2	140.2	1.6
GLU 42	19.6	37.0	12.2	20.6	14.4	19.4	8.3	48.3	8.8	17.9	3.2	19.4
GLU 46	64.9	61.5	49.9	27.8	40.4	24.2	32.2	107.0	42.1	26.3	60.5	27.8
GLU 47	162.9	68.1	144.1	18.9	171.3	20.2	114.6	102.4	111.7	13.5	106.6	12.5
ASP 48	10.7	87.8	18.9	47.1	8.2	54.0	4.7	114.7	9.2	36.8	2.5	35.5
GLU 50	0.2	2.4	12.6	2.3	0.3	4.3	0.0	14.7	59.7	7.9	13.9	2.5
GLU 59	28.0	23.6	46.1	28.9	50.2	19.4	46.0	19.6	57.2	23.7	60.3	14.2
ASP 60	100.4	21.0	98.8	12.9	100.1	11.4	99.9	24.5	121.1	12.9	123.8	8.7
GLU 63	52.3	36.0	60.0	22.2	72.3	19.1	89.0	53.3	72.8	23.6	44.9	22.8
GLU 64	95.9	30.0	91.1	22.5	83.7	24.5	51.9	72.9	91.2	17.6	80.2	18.7
GLU 67	39.1	50.6	79.4	23.6	59.4	25.2	93.7	81.9	83.8	19.6	43.9	24.9
GLU 71	86.3	47.6	93.3	18.2	95.0	18.7	75.1	65.5	50.7	19.7	30.1	21.8
ASP 76	0.3	50.0	18.1	40.0	15.9	35.4	0.5	59.4	0.9	43.7	0.4	28.2
ASP 84	113.1	87.8	47.2	104.3	42.8	87.4	83.4	125.5	44.3	79.2	101.9	46.1
ASP 86	25.5	159.4	98.0	74.9	99.8	56.7	57.9	153.2	91.6	63.1	21.9	63.6
ASP 87	49.6	17.8	117.7	14.6	121.9	13.1	109.4	49.3	65.3	22.4	99.4	22.0
GLU 90	74.3	28.2	53.9	15.8	56.9	17.3	56.2	63.0	70.4	11.7	62.7	9.7
ASP 91	83.1	12.2	39.9	16.1	64.4	17.2	69.6	32.5	93.6	11.1	77.6	8.1
GLU 98	58.7	27.7	57.8	16.1	49.1	19.4	66.1	38.7	64.3	16.0	32.8	18.9
GLU 99	97.2	9.5	96.1	6.4	91.3	7.5	99.3	16.1	97.5	12.1	106.3	12.4
C-terminal	71.2	45.8	87.9	17.8	76.6	18.1	105.5	33.7	29.3	17.3	25.4	13.8
Average	63.3	42.8	67.6	24.4	68.2	23.4	60.8	59.6	69.7	21.9	60.2	19.4

\*P. C. Residue represents percent occupancies with positively charged residues. Data come from Table 2A and 2B.

Table S9. Buried ratio and pKa of negatively and positively charged residues in the crystal structure of *PhCutA1*\*.

Residues	Buried ratio			pKa		
	A-subunit	B-subunit	C-subunit	A-subunit	B-subunit	C-subunit
ASP 10	0.55	0.25	0.84	3.58	3.13	2.7
ASP 48	0.58	0.51	0.9	5.12	4.53	8.26
ASP 60	0.02	0.07	0.06	2.07	2.43	2.17
ASP 76	0	0	0	4.04	4.01	3.97
ASP 84	0.28	0.31	0.3	3.01	3.06	3.31
ASP 86	0.36	0.39	0.41	4.36	4.59	4.94
ASP 87	0.23	0.24	0.23	3.23	3.21	3.43
ASP 91	0.26	0.36	0.34	3.87	3.27	3.09
GLU 12	0.42	0	0.49	3.8	4.77	4.7
GLU 15	1	0.38	0.93	11.09	5.06	1.42
GLU 24	0.04	0.07	0.06	3.07	3.84	3.25
GLU 34	1	1	1	5.77	7.6	7.27
GLU 42	0	0	0	4.73	4.45	4.45
GLU 46	0	0	0.5	4.67	3.67	3.58
GLU 47	0.33	0.67	1	4.48	5.57	6.37
GLU 50	1	1	1	6.78	6.46	6.23
GLU 59	0.16	0.06	0.2	3.91	5.22	3.9
GLU 63	0.01	0	0	3.89	3.92	4.33
GLU 64	0	0.03	0.04	3.54	3.74	3.93
GLU 67	0	0	0	4.54	4.78	3.36
GLU 71	0	0	0	4.06	4.21	4.28
GLU 90	0	0	0	4.48	4.75	4.89
GLU 98	0	0	0	4.67	3.89	3.43
GLU 99	0.35	0.28	0.26	5.76	4.85	4.72
C- 102	0	0	0	3.21	2.87	2.77
LYS 16	0	0	0.12	10.57	10.41	11.49
LYS 19	0.62	0.06	0.57	11.71	10.64	10.88
LYS 23	0	0	0	10.32	10.2	10.43
LYS 44	0	0	0	10.62	11.31	10.62
LYS 49	0.33	0.74	1	10.62	10.32	12.05
LYS 56	1	1	1	7.3	7.7	7.8
LYS 66	0.14	0.2	0.2	11.74	10.67	12.01
LYS 70	0.07	0.18	0.07	11	11.56	11.06
LYS 94	0	0	0	10.66	11.06	11.49
LYS 101	0	0	0	11.33	11.5	11.48
LYS 102	0	0.01	0	10.4	10	10.48
ARG 25	0.13	0.14	0.14	12.19	12.61	12.53
ARG 33	1	1	1	13.57	14.68	12.39
ARG 36	0.69	1	1	12.57	12.1	14.23
ARG 58	0.15	0.18	0.17	13.81	13.65	13.37
ARG 68	0	0	0	14.78	14.07	14.61
ARG 82	0.27	0.27	0.27	13.65	13.81	13.43

\*These values in the crystal structure (4nyo) were estimated using a software, propka3.0, revision 182 in [http://nbc-222.ucsd.edu/pdb2pqr\\_2.0.0/](http://nbc-222.ucsd.edu/pdb2pqr_2.0.0/).

Table S10. The number of Na<sup>+</sup>, Cl<sup>-</sup>, and H<sub>2</sub>O in the simulation box and size of the box during MD simulations at 300 K using indicated force fields.

	Charmm27 _tip3p	Amber99sb _spc/e	Amber99sb _tip3p	Amber14sb _tip3p	Gromos43a1 _spc/e	Gromos53a6 _spc/e
Na <sup>+</sup>	59	59	58	59	59	59
Cl <sup>-</sup>	38	39	38	38	38	38
H <sub>2</sub> O	12011	12011	12011	12011	12023	12022
Box-X*	8.10715	8.07810	8.10875	8.10591	8.07218	8.09331
Box-Y*	7.64355	7.61617	7.64507	7.64238	7.61058	7.63050
Box-Z*	6.61970	6.59599	6.62102	6.61869	6.59115	6.60840

\*The average values (nm) of simulation box during 400 ns at 300 K.

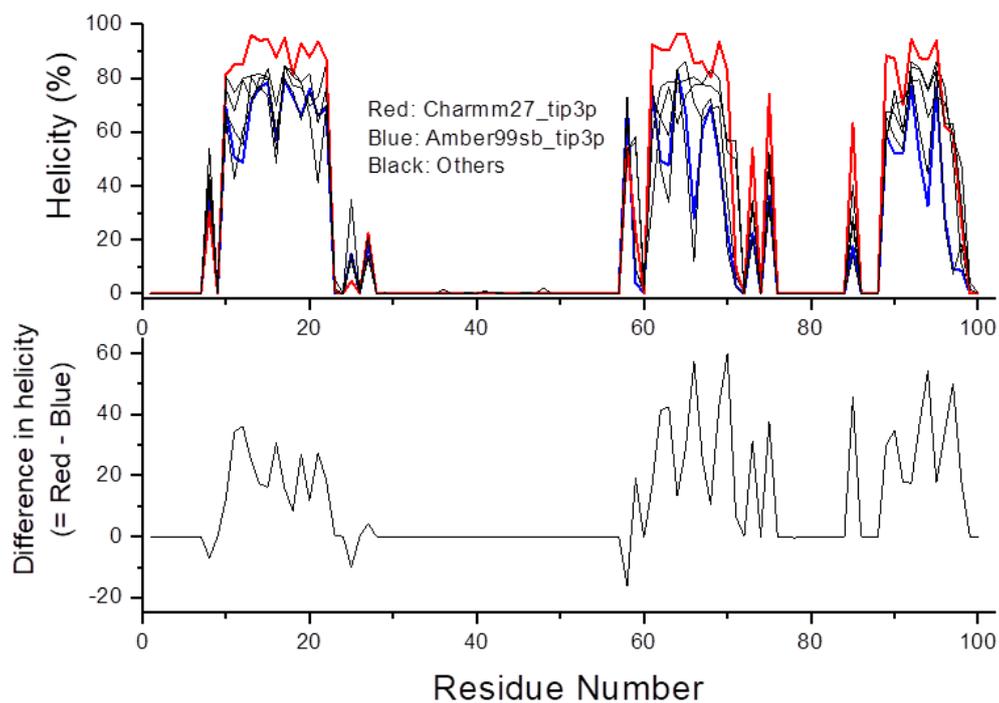
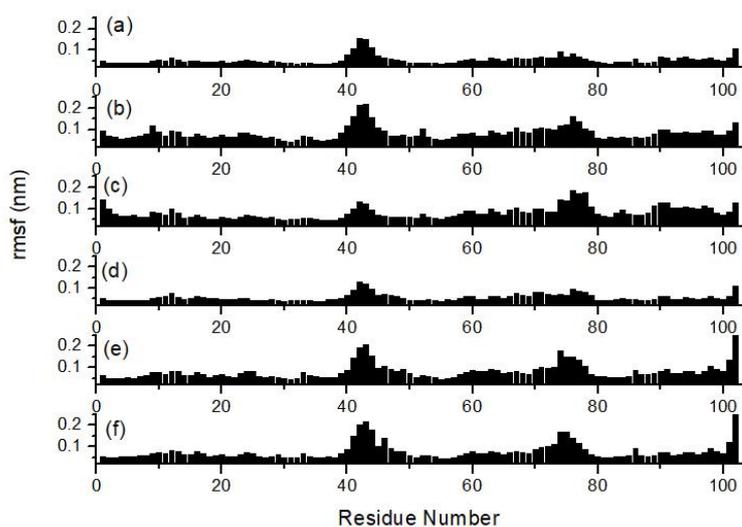


Fig. S1. Comparison of helicity for *PhCutA1* among six force fields in 300 K MD simulations (50-400 ns). (A) Percent helicity shows average values for each residue of *PhCutA1*. Red, Blue, and Black represent Charmm27\_tip3p, Amber99sb\_tip3p, and others, respectively. (B) Difference in helicity (Helicity subtraction of Amber99sb\_tip3p from Charmm27\_tip3p).

(A)



(B)

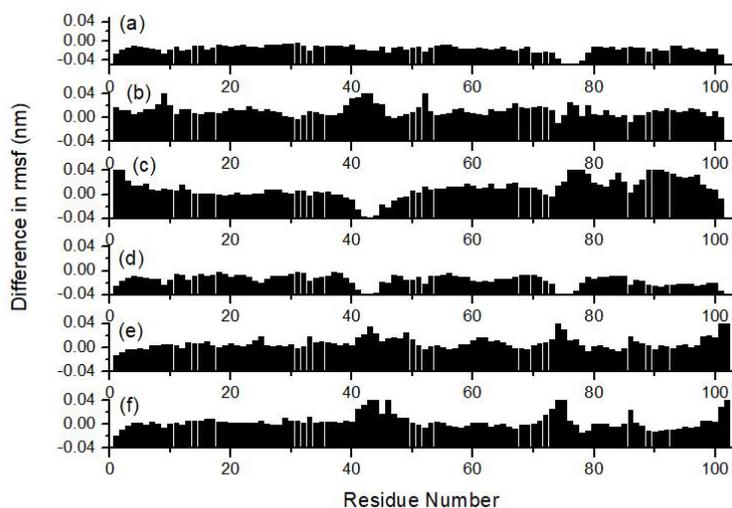


Fig. S2. Comparison of root-mean-square fluctuation (RMSF) for the C $\alpha$  atoms of *PhCutA1* among six force fields in 300 K MD simulations (50-400 ns). (a), (b), (c), (d), (e), and (f) represent Charmm27\_tip3p, Amber99sb\_spc/e, Amber99sb\_tip3p, Amber14sb\_tip3p, Gromos43a1\_spc/e, and Gromo53a6\_spc/e, respectively. (A) Average RMSF values at each residue. (B) The difference values at each residue. The differences in average RMSF at each C $\alpha$  atom was obtained by subtracting the average value of C $\alpha$  atoms in six force fields from a value of each C $\alpha$  atom.

Fig. S3A

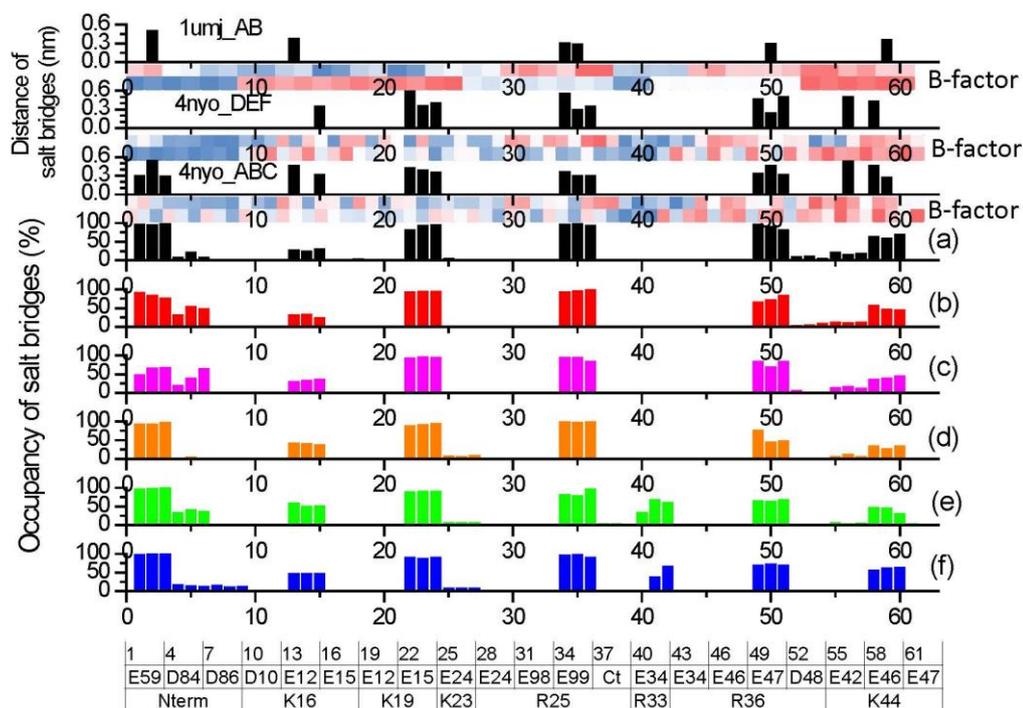


Fig. S3. Percent occupancy of salt bridges for each of targeted ionic pairs at six different force fields and the distance of salt bridges obtained from crystal structures. (A) Intra-subunit interaction, (B) Inter-subunit interaction. Targeted ionic pairs are listed in Tables S1A and S1B for intra- and inter-subunit interaction, respectively. Three bars represent the data for each targeted pair in A, B, and C-subunits. The lengths of salt bridges were calculated by the CCP4 software using the structures of ABC and DEF for PDB ID 4nyo and AB for 1umj. B-factors of charged residues in crystal analysis are indicated by a color gradient: blue indicates the lowest B-factors and red the highest. The upper line, with the color gradient showing B-factors, represents the B-factors of negatively charged residues; the other one is positively charged. (a), (b), (c), (d), (e), and (f) represent Charmm27\_tip3p, Amber99sb\_spc/e, Amber99sb\_tip3p, Amber14sb\_tip3p, Gromos43a1\_spc/e, and Gromos53a6\_spc/e, respectively.

Fig. S3A continued

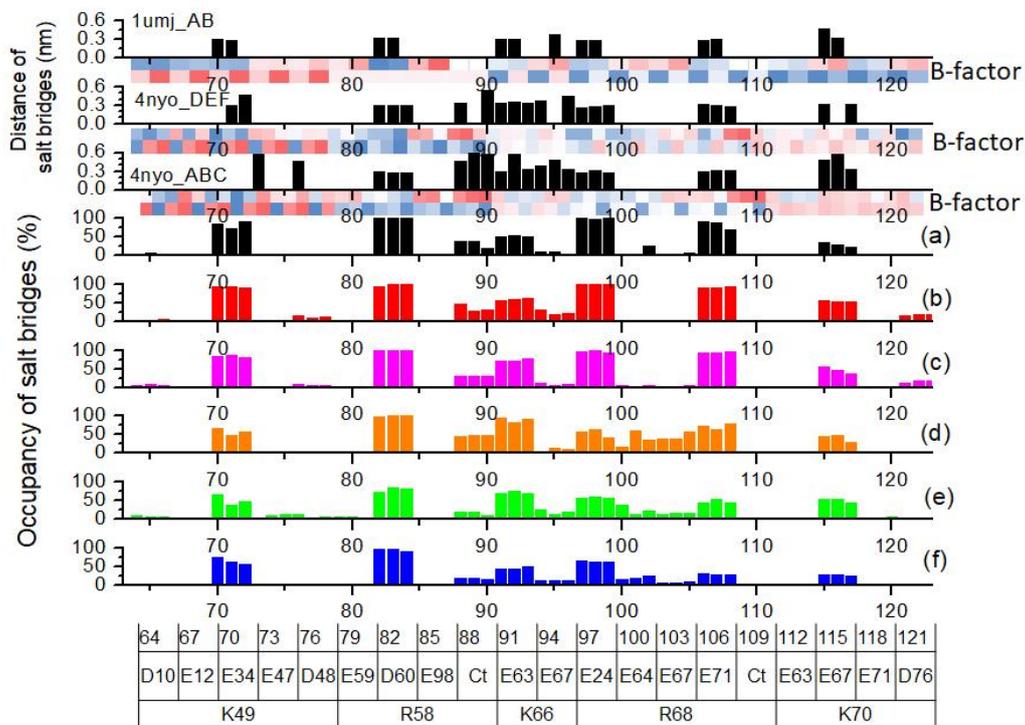


Fig. S3A continued

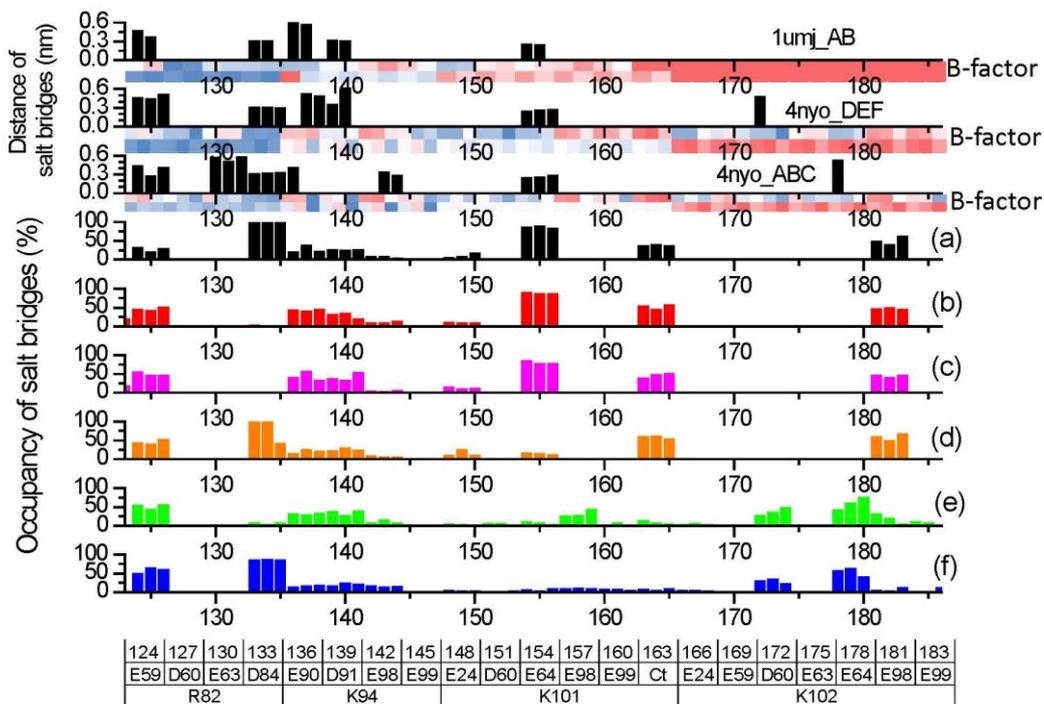


Fig. S3B

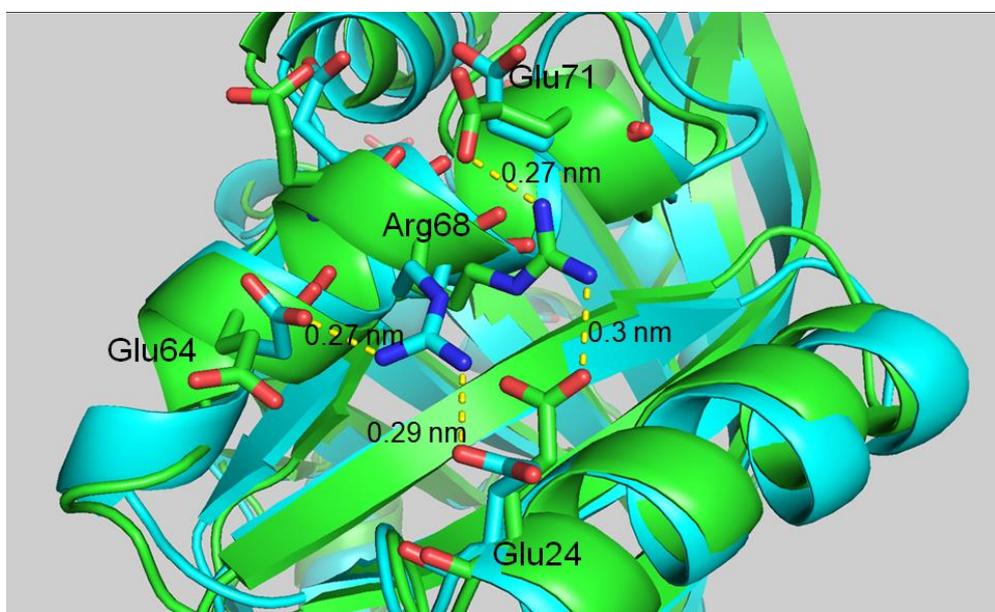
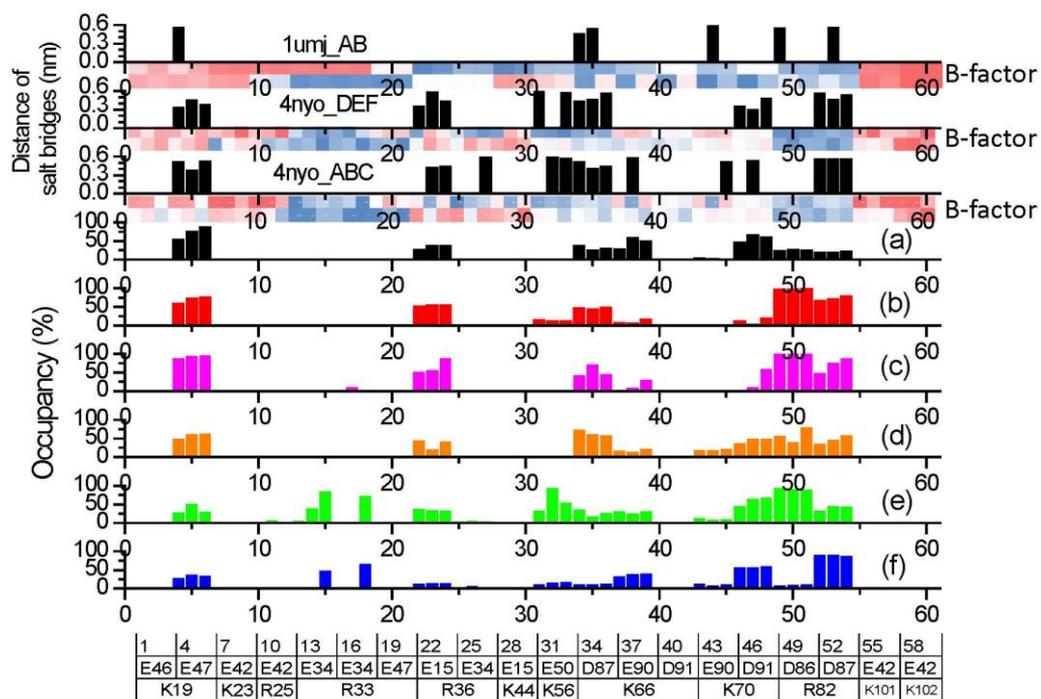
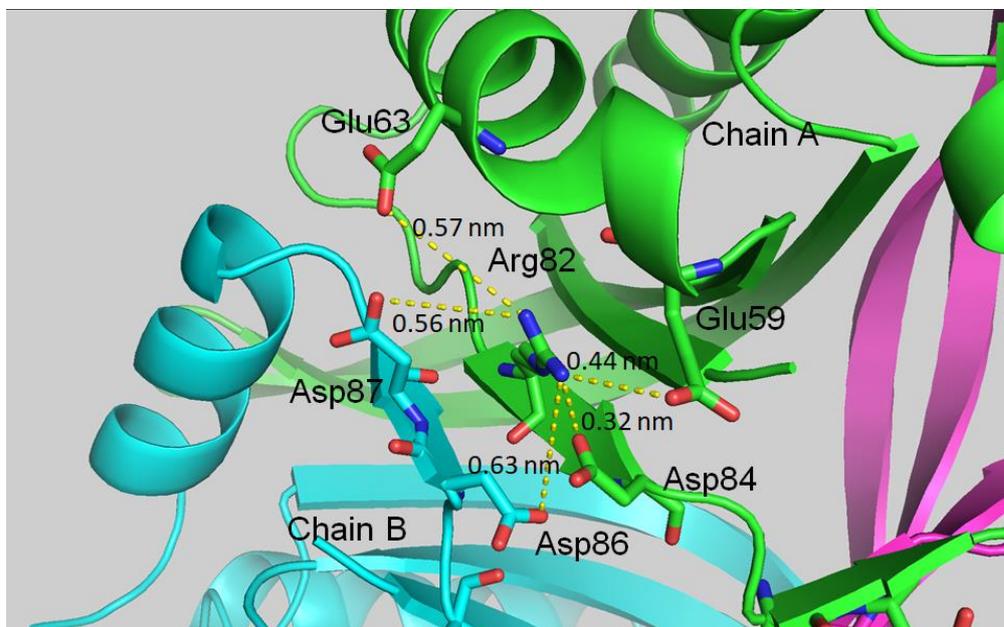


Fig. S4. Snapshot of the configuration around Arg68 of *PhCutA1*. Green and cyan represent snapshots at 100 ns of 300 K MD simulations for Amber99sb\_tip3p and Amber14sb\_tip3p, respectively.

(A)



(B)

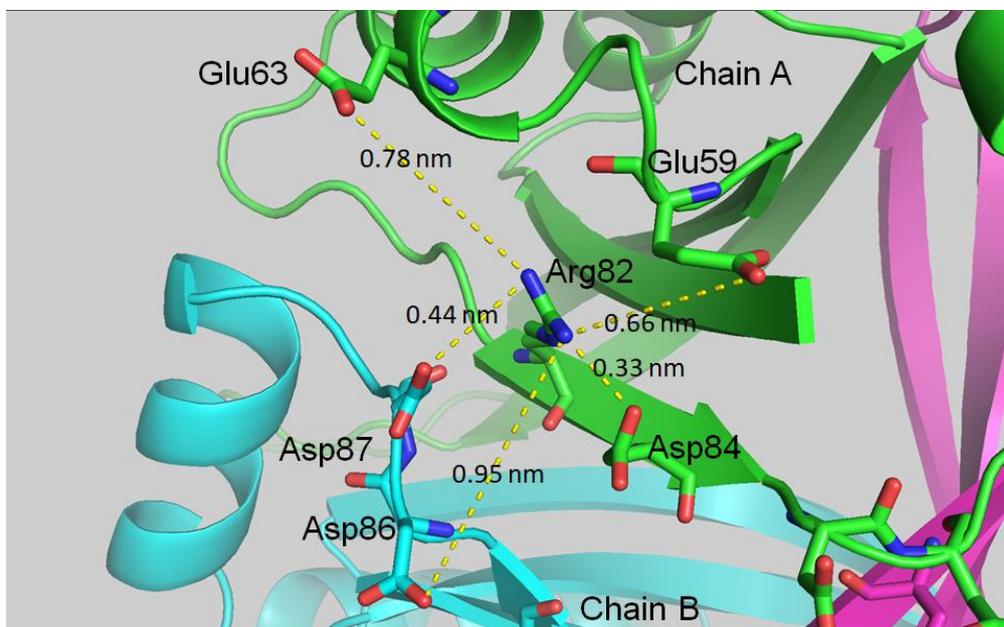


Fig. S5. The configuration around Arg82 of *PhCutA1*. Green, cyan, and magenta represent A, B, and C-subunits of *PhCutA1*, respectively. (A) The crystal structure of *PhCutA1* (A, B, and C-subunits of 4nyo). (B) The snapshot around Arg82 of *PhCutA1* at 200 ns of an MD simulation in the case of Gromos53a6\_spc/e.

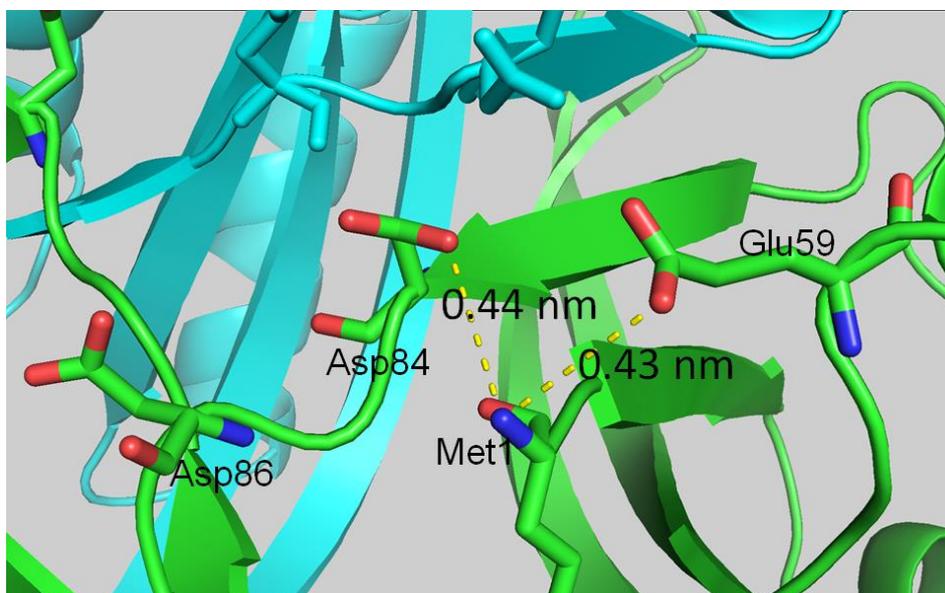
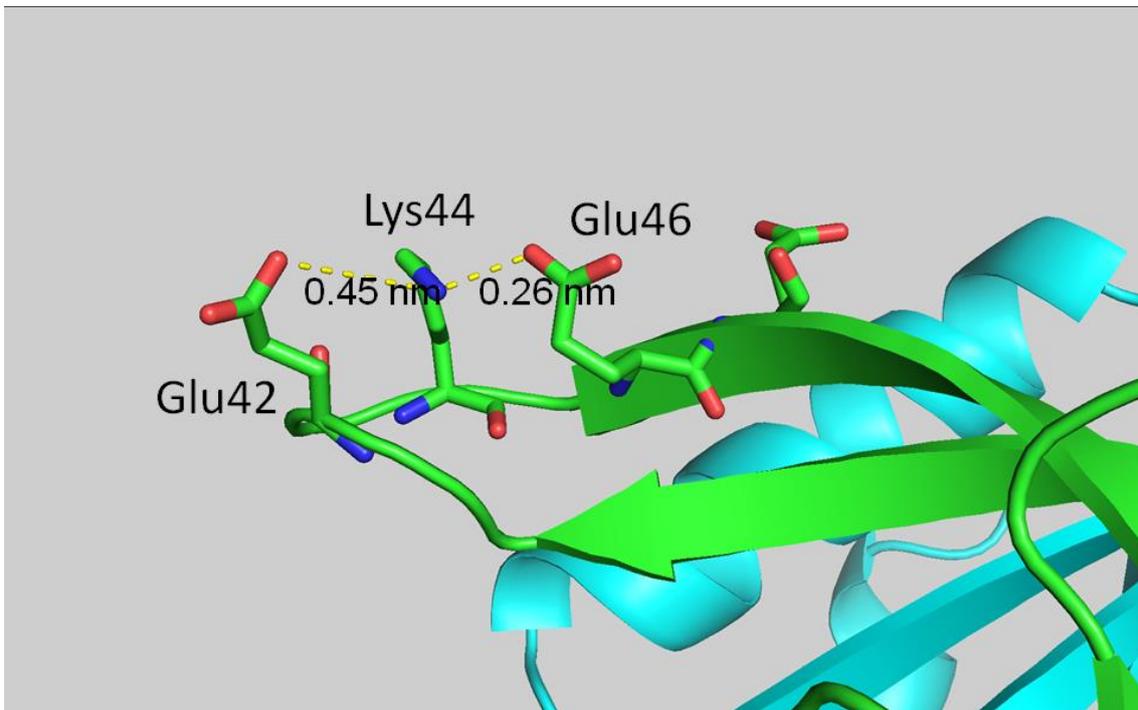


Fig. S6. The snapshot of the configuration around N-terminal (Met1) of *PhCutA1* at 100 ns of an MD simulation using Amber99sb\_tip3p.

(A)



(B)

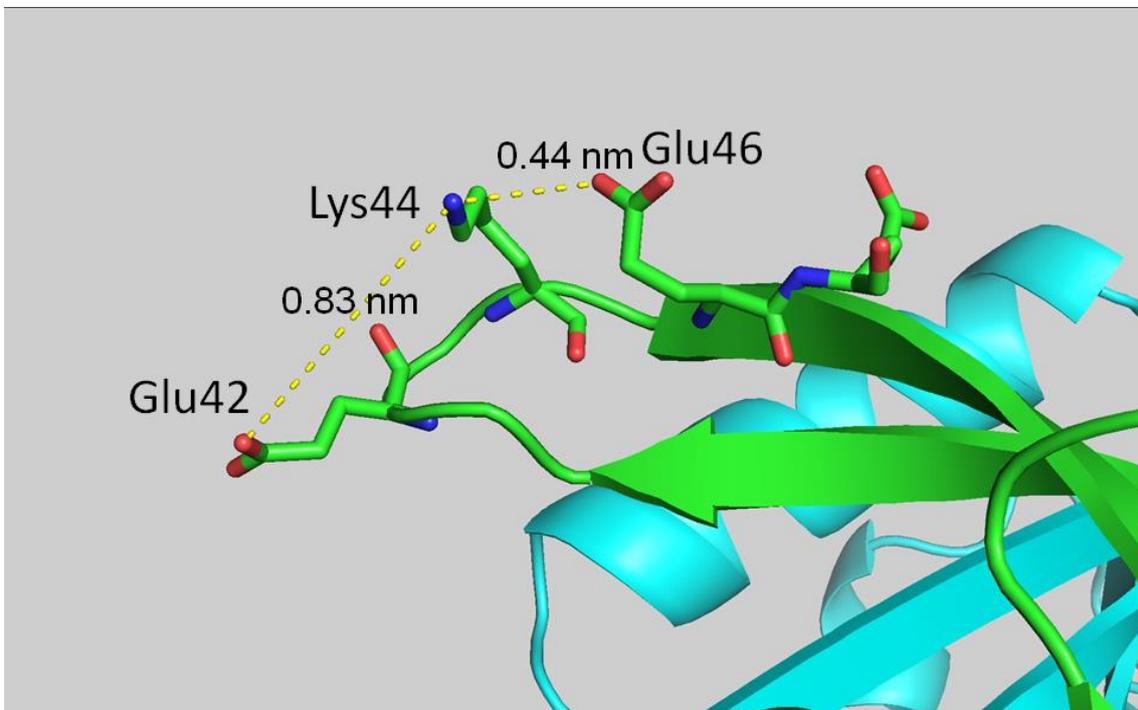
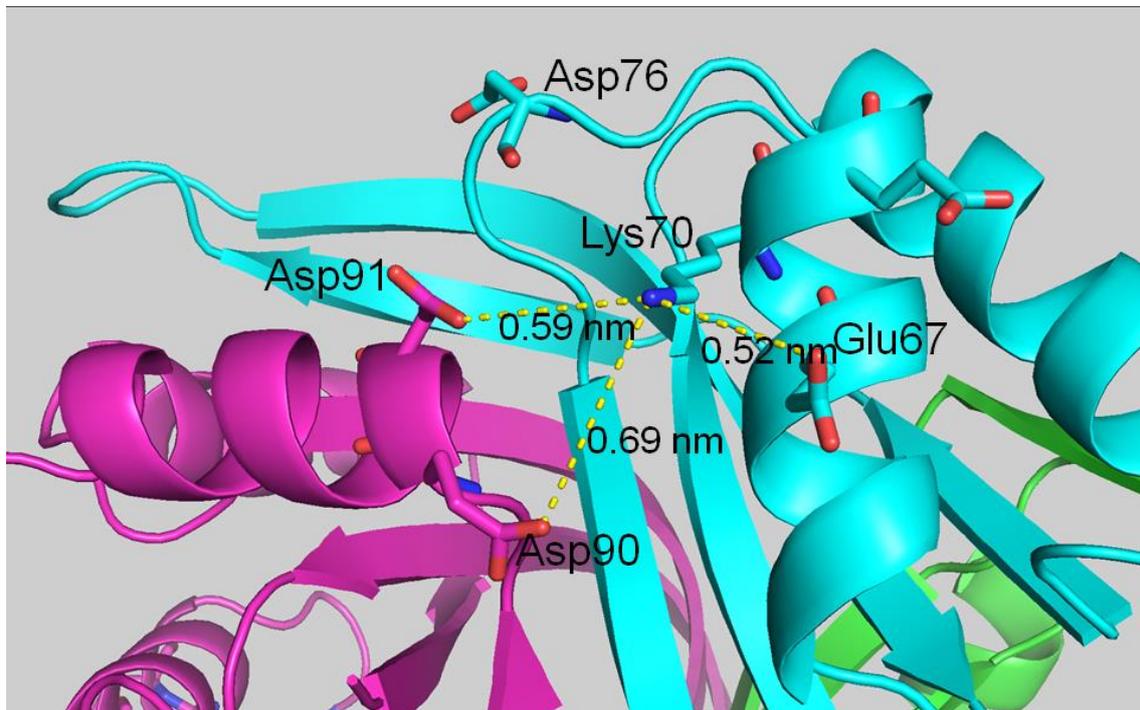


Fig. S7. The snapshots of the configuration around Lys44 of *PhCutA1* at 100 ns in an MD simulation using Charmm27\_tip3p (A) and Gromos53a6\_spc/e (B).

(A)



(B)

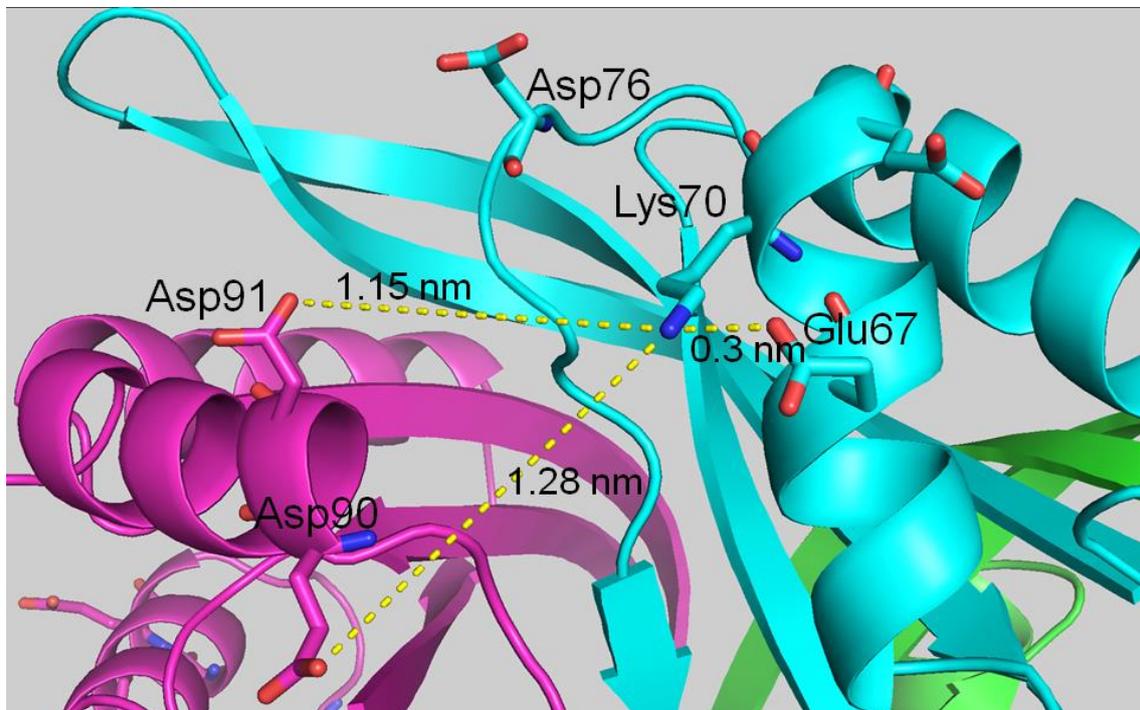
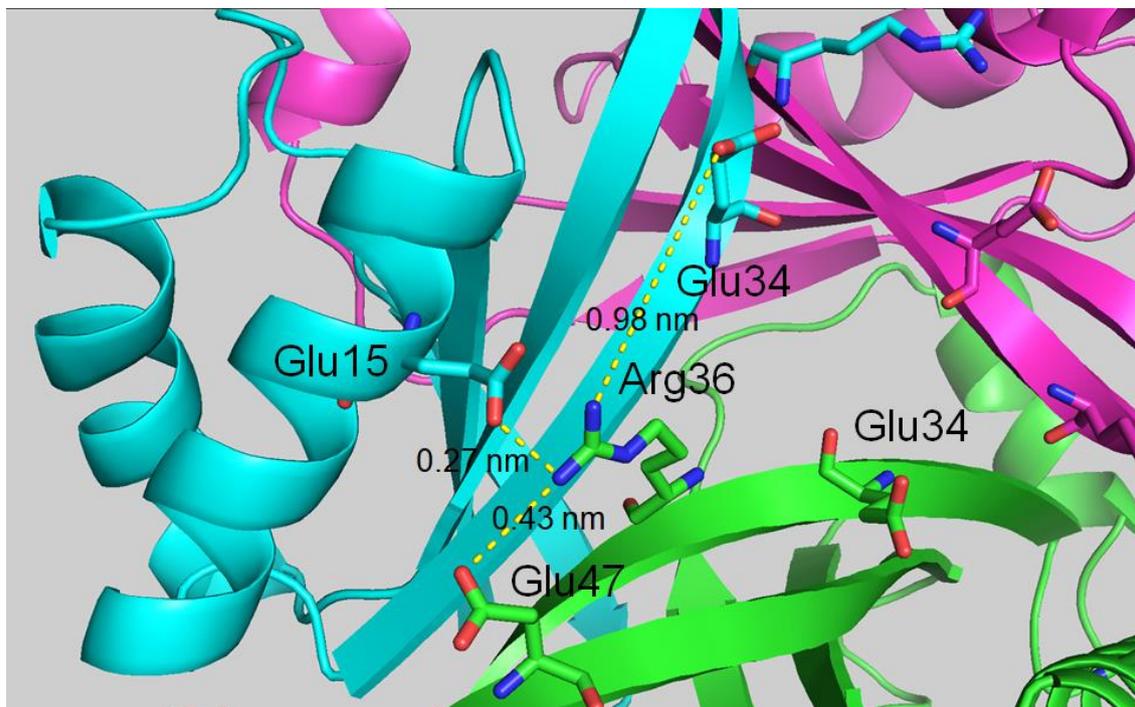


Fig. S8. The snapshots of the configuration around Lys70 of *PhCutA1* at 100 ns in an MD simulation using Gromos43a1\_spc/e (A) and Amber99sb\_tip3p (B).

(A)



(B)

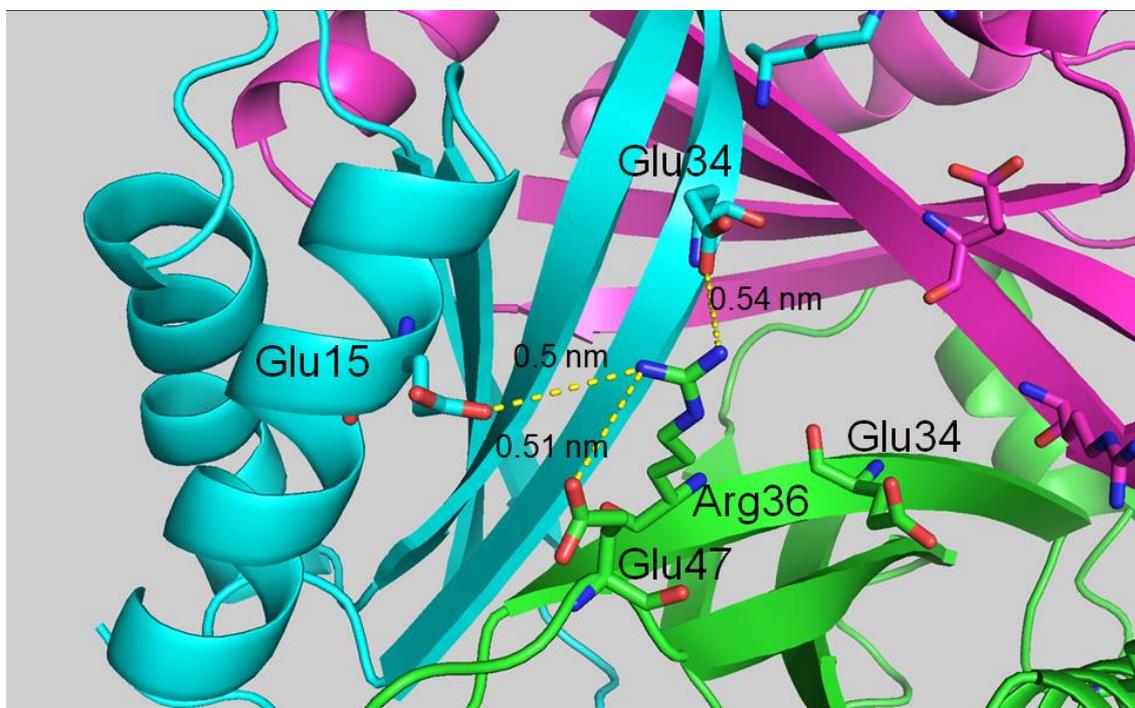
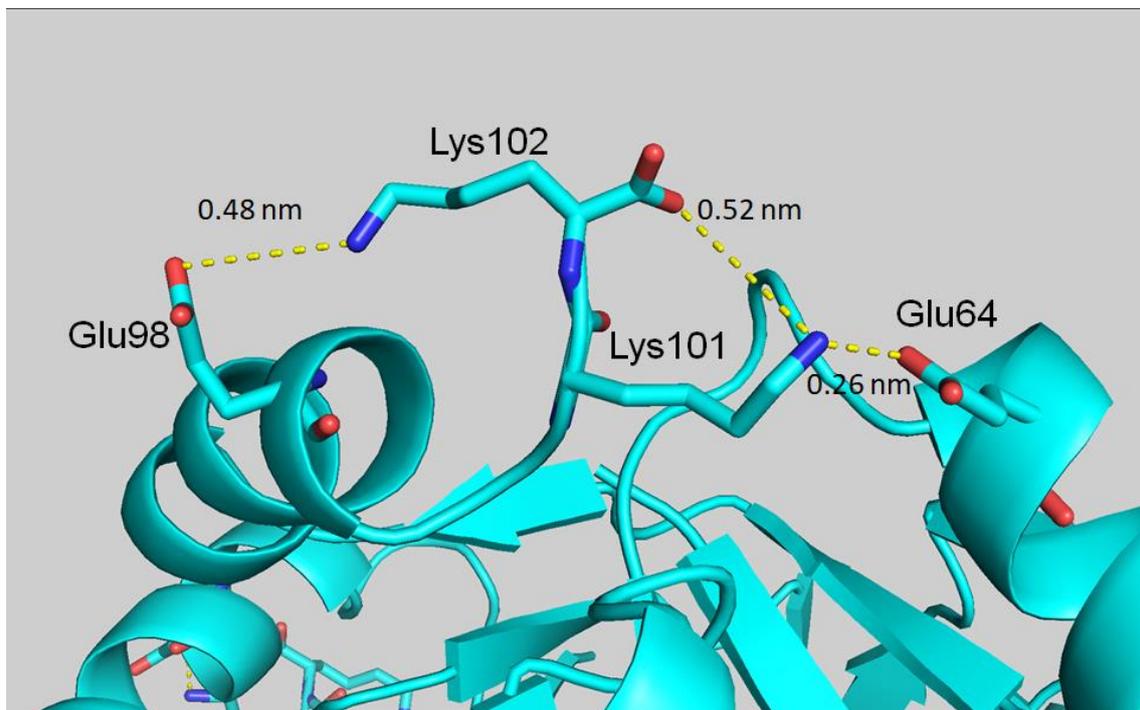


Fig. S9. The snapshots of the configuration around Arg36 of *PhCutA1* at 200 ns in an MD simulation using Amber99sb\_tip3p (A) and Gromos53a6\_spc/e (B).

(A)



(B)

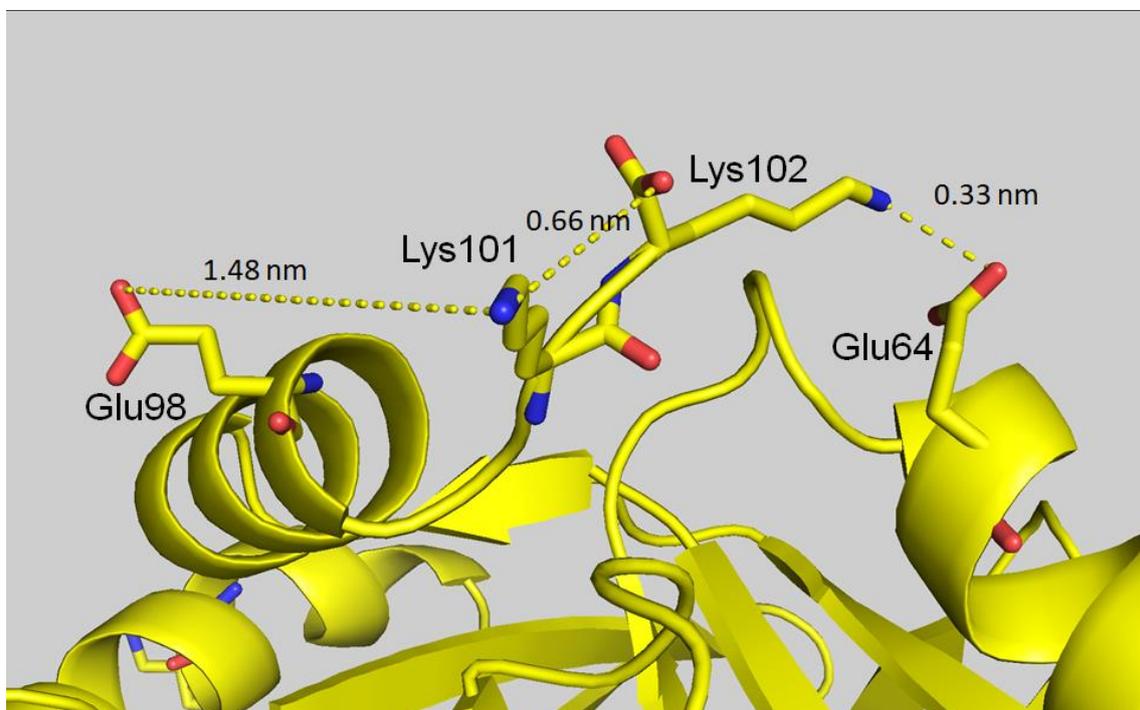
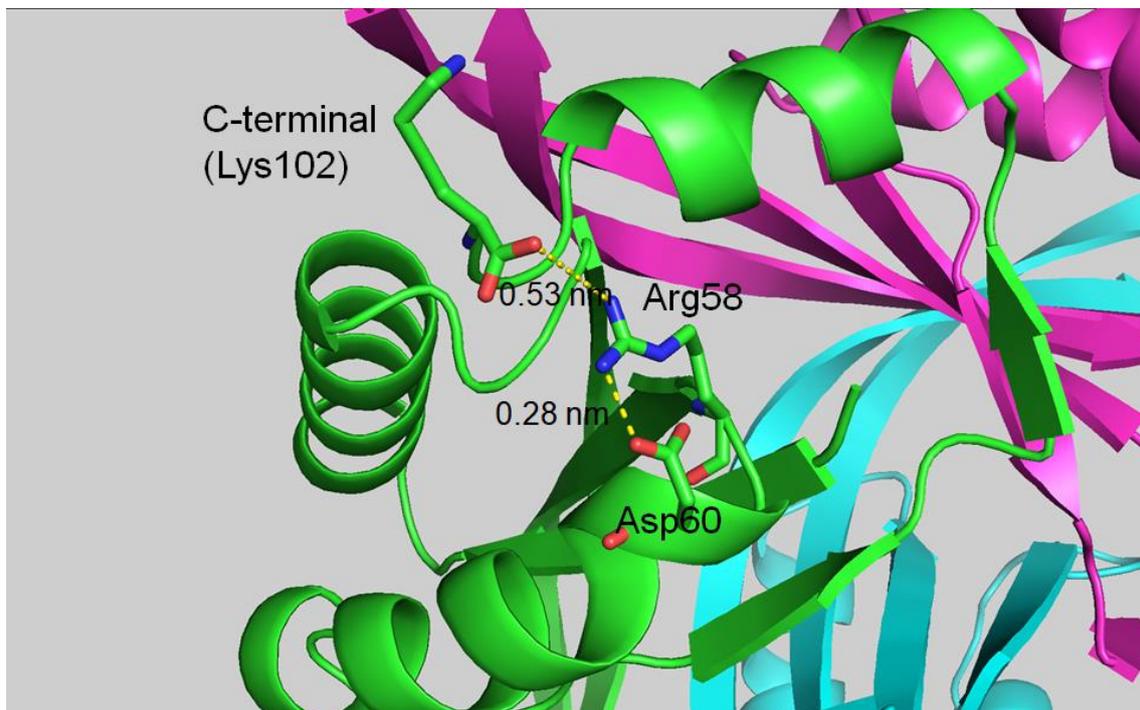


Fig. S10. The snapshots of the configuration around Lys101 and Lys102 of *PhCutA1* at 100 ns in an MD simulation using Charmm27\_tip3p (A) and Gromos53a6\_spc/e (B).

(A)



(B)

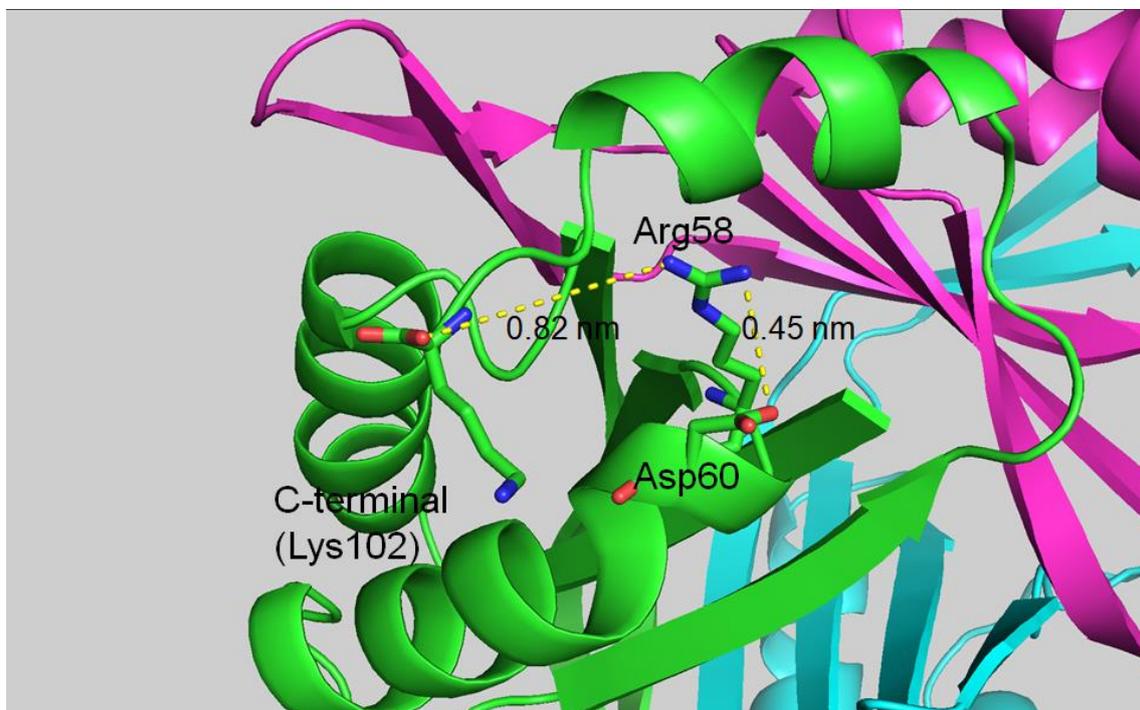
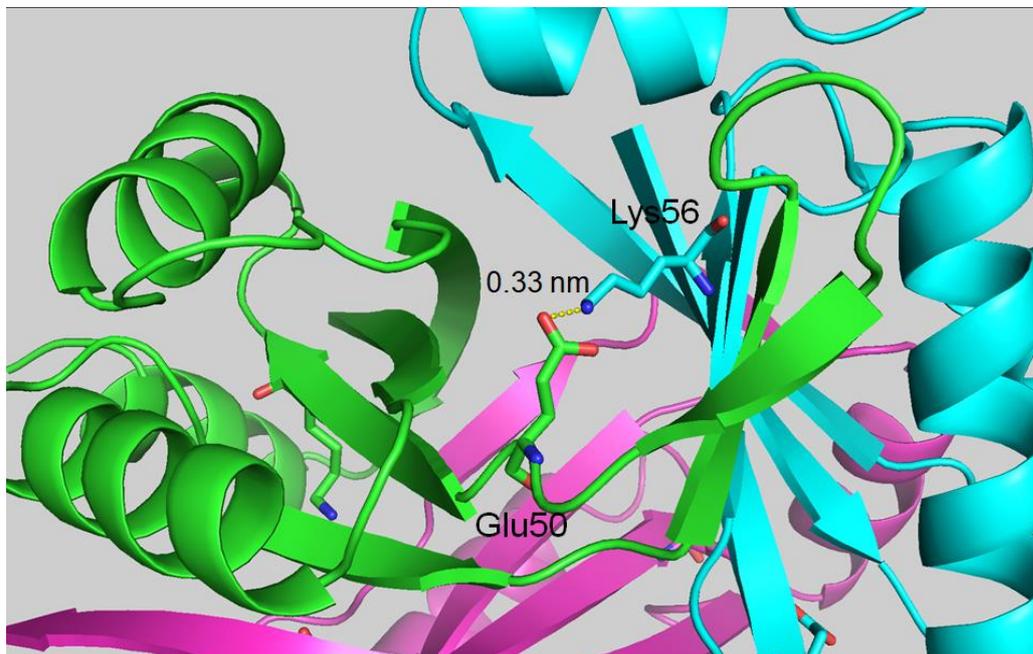


Fig. S11. The snapshots of the configuration around Arg58 of *PhCutA1* at 200 ns in an MD simulation using Amber14\_tip3p (A) and Gromos43a1\_spc/e (B).

(A)



(B)

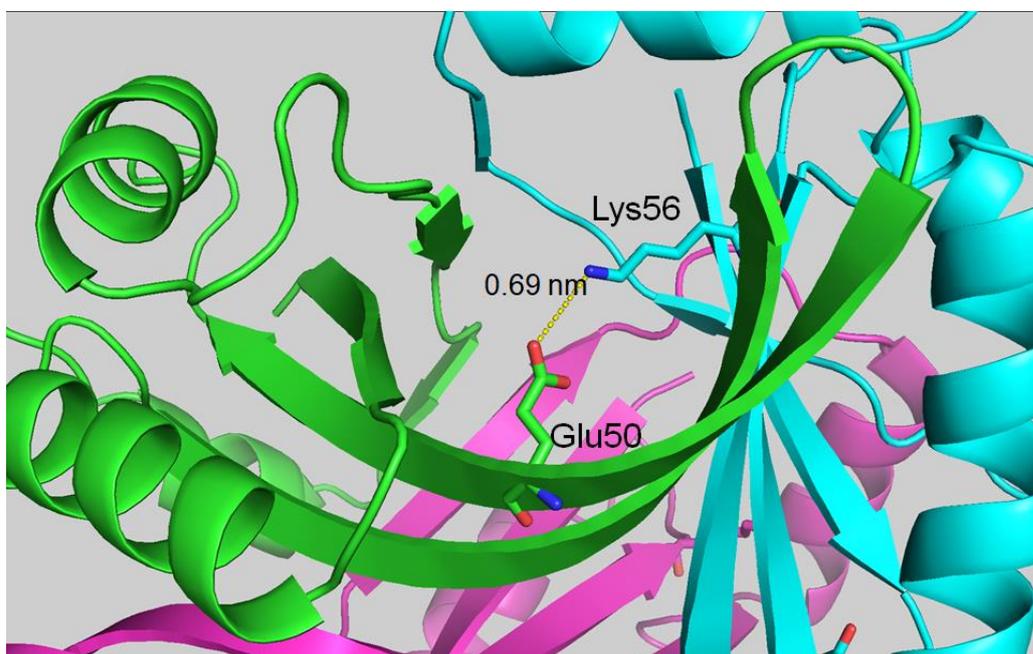


Fig. S12. The snapshots of the configuration around Glu50 interacting with Lys56 of *PhCutA1* at 100 ns in an MD simulation using Gromos43a1\_spc/e (A) and Amber99sb\_tip3p (B). Green, cyan, and magenta represent A, B, and C-subunits of *PhCutA1*, respectively.

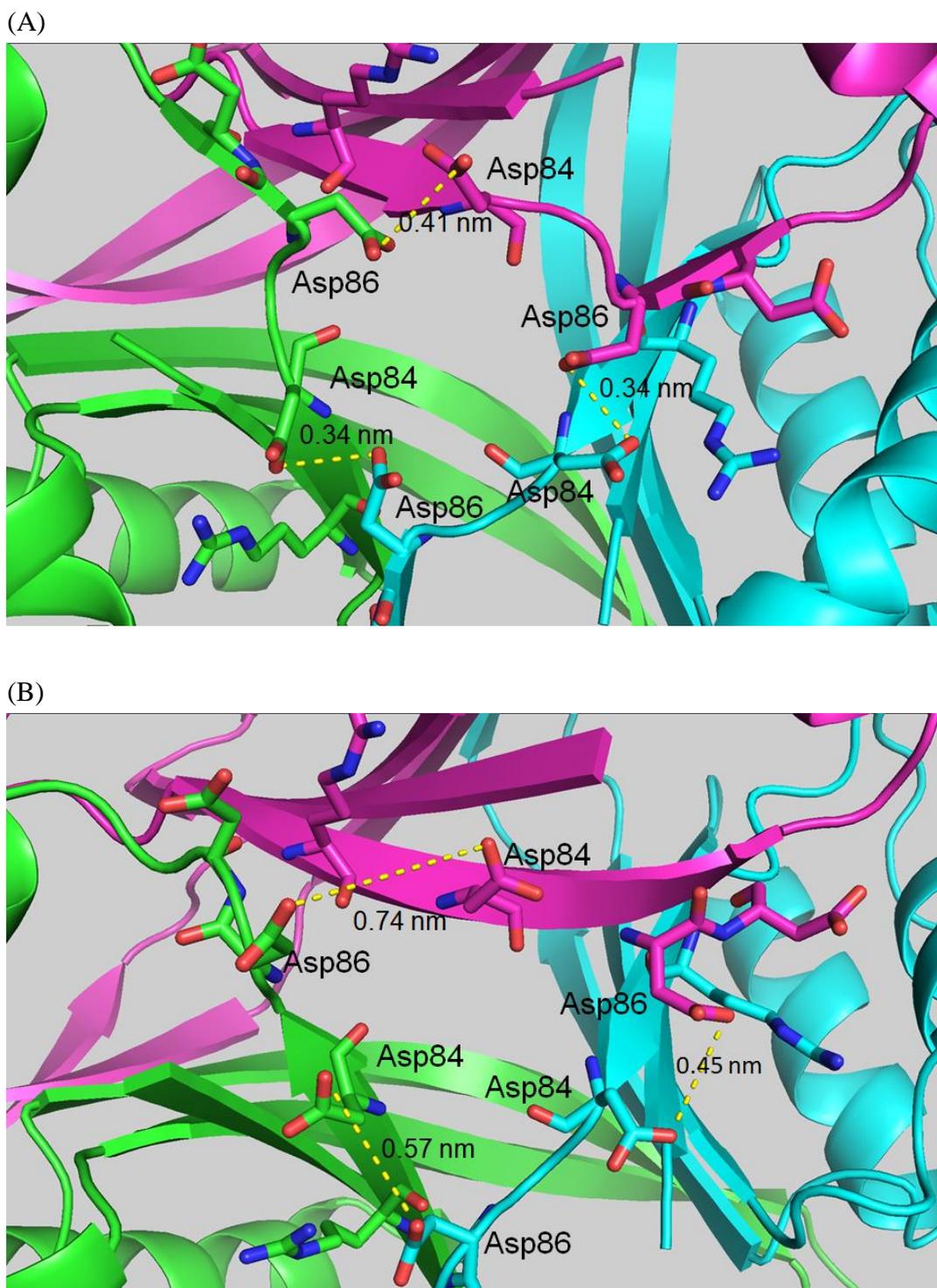
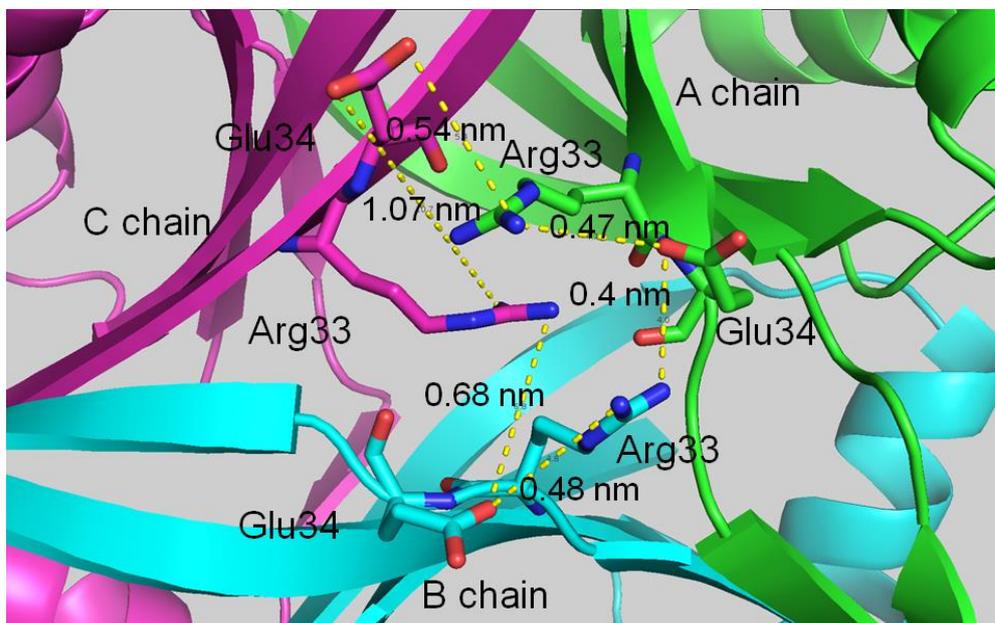


Fig. S13. The configuration around Asp84 and Asp86 of *PhCutA1*. Green, cyan, and magenta represent A, B, and C-subunits of *PhCutA1*, respectively. (A) The crystal structure of *PhCutA1* (A, B, and C-subunits of 4nyo). (B) The snapshot of *PhCutA1* at 200 ns of an MD simulation in the case of Gromos43a1\_spc/e.

(A)



(B)

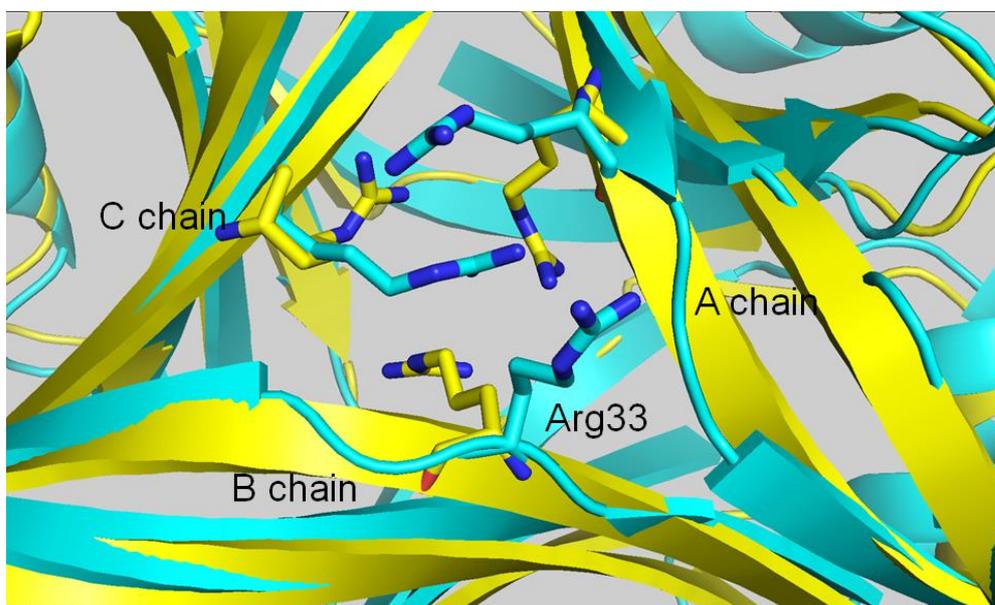


Fig. S14. The configuration around Arg33 of *PhCutA1*. (A) Snapshots around Arg33 and Glu34 of *PhCutA1* at 100 ns of an MD simulation using Gromos43a1\_spc/e. Green, cyan, and magenta represent A, B, and C-subunits of *PhCutA1*, respectively. (B) Snapshots around Arg33 of *PhCutA1* at 100 ns of MD simulation in the cases of Gromos43a1\_spc/e (cyan) and Amber99sb\_tip3p (yellow). Both structures are superimposed.

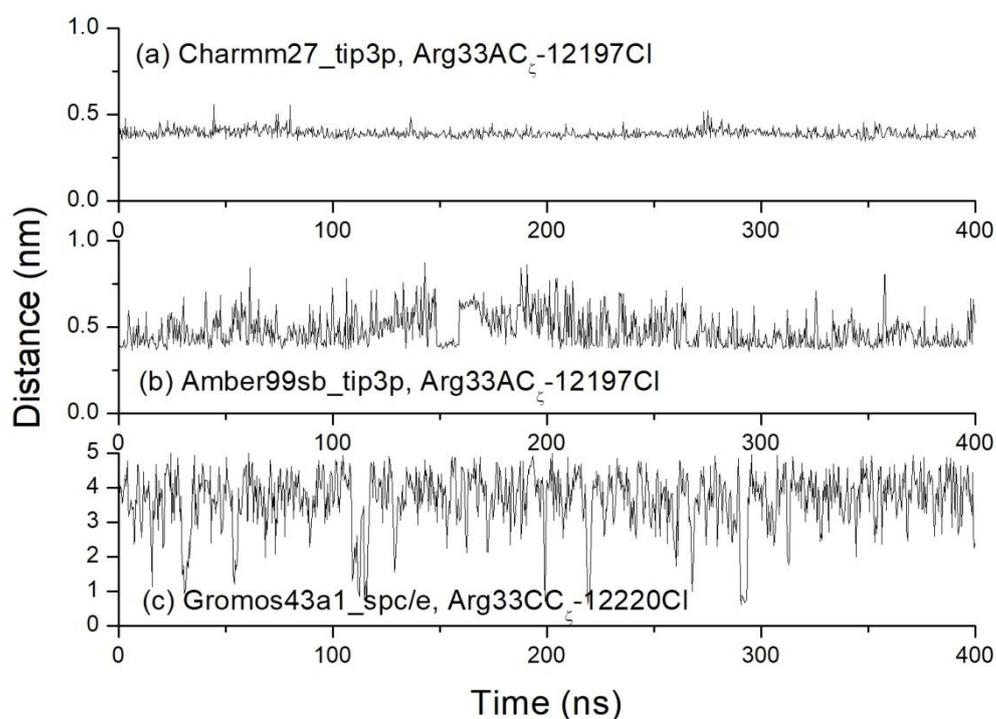


Fig. S15. Trajectories of distance between Arg33 in *PhCutA1* and  $\text{Cl}^-$  ion during 400-ns MD simulations at 300 K using indicated force fields. (a) Charmm27\_tip3p, the distance between  $\text{C}_\zeta$  of Arg33 in A-subunit and  $\text{Cl}^-$  ion of the number 12197. The percent occupancy of distance (less than 0.6 nm) between them was 100.0 %. (b) Amber99sb\_tip3p, the distance between  $\text{C}_\zeta$  of Arg33 in A-subunit and  $\text{Cl}^-$  ion of the number 12197. The percent occupancy of distance (less than 0.6 nm) between them was 86.0 %. (c) Gromos43a1\_spc/e, the distance between  $\text{C}_\zeta$  of Arg33 in C-subunit and  $\text{Cl}^-$  ion of the number 12220. The percent occupancy of distance (less than 0.6 nm) between them was 1.7 %.

### 3-6. References.

1. Mayor, U., Johnson, C. M., Daggett, V., and Fersht, A. R. (2000) Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation. *Proc. Natl. Acad. Sci. USA* **97**, 13518-13522.
2. Mayor, U., Guydosh, N. R., Johnson, C. M., Grossmann, J. G., Sato, S., Jas, G. S., Freund, S. M., Alonso, D. O., Daggett, V., and Fersht, A. R. (2003) The complete folding pathway of a protein from nanoseconds to microseconds. *Nature* **421**, 863-867.
3. Thomas, A. S. and Elcock, A. H. (2004) Molecular simulations suggest protein salt bridges are uniquely suited to life at high temperatures. *J. Am. Chem. Soc.* **126**, 2208-2214.
4. Pang, J. and Allemann R. K. (2007) Molecular dynamics simulation of thermal unfolding of *Thermatoga maritima* DHFR. *Phys. Chem. Chem. Phys.* **9**, 711-718.
5. Voelz, V. A., Singh, V. R., Wedemeyer, W. J., Lapidus, L. J., and Pande, V. S. (2008) Unfolded-state dynamics and structure of protein L characterized by simulation and experiment. *J. Am. Chem. Soc.* **132**, 4702-4709.
6. Lindorff-Larsen, K., Trbovic, N., Maragakis, P., Piana, S., and Shaw, D. E. (2012) Structure and dynamics of an unfolded protein examined by molecular dynamics simulation. *J. Am. Chem. Soc.* **134**, 3787-3791.
7. Sterpone, F., and Melchionna, S. (2012) Thermophilic proteins: insight and perspective from in silico experiments. *Chem. Soc. Rev.* **41**, 1665-1676.
8. Robustelli, P., Piana, S., and Shaw, D. E., (2018) Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci. USA* **115**, E4758-E4766.
9. Matsuura, Y., Takehira, M., Sawano M., Ogasahara, K., Tanaka, T., Yamamoto, H., Kunishima, N., Katoh, E., and Yutani, K. (2012) Role of charged residues in stabilization of *Pyrococcus horikoshii* CutA1, which has a denaturation temperature of nearly 150 °C, *FEBS J.* **279**, 78-90.
10. van Gunsteren, W. F., Billeter, S. R., Eising, A. A., Hunenberger, P. H., Kruger, P., Scott, W. R. P., and Tironi, I. G. (1996) *Biomolecular Simulation: The Gromos96 Manual and Use Guide*; Vdf Hochschulverlag AG an der Zurich Switzerland.

11. Oostenbrink, C., Villa, A., Mark, A. E., and van Gunsteren, W. F. A. (2004) Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-Field Parameter Sets 53A5 and 53A6. *J. Comput. Chem.* **25**, 1656-1676.
12. MacKerell, A. D., Feig, M., and Brooks, C. L. (2004) III Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamic Simulations. *J. Comput. Chem.* **26**, 1400-1415.
13. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbond Parameters. *Proteins* **65**, 712-725.
14. Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, *J. Chem. Theory Comput. Chem.* **11** 3696-3713.
15. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R., W., and Klein, M. L. (1983) Comparison of Somple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **79**, 926-935.
16. Berendsen, H. J. C., Grigera, L. R., and Straatsma, T. P. (1987) The Missing Term in Effective Pair Potentials. *J. Phys. Chem.*, **91**, 6269-6271.
17. Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435-447.
18. van der Spoel, D., Lindahl, E., Hess, B., van Buuren, A. R., Apol, E., Meulenhoff, P. J., Tieleman, D. P., Sijbers, A. L. T. M., Feenstra, K. A., van Drunen, R., and Berendsen, H. J. C. (2010) Gromacs User Manual version 4.5.4, [www.gromacs.org](http://www.gromacs.org).
19. Matsuura, Y., Takehira, M., Makhatadze, G. I., Joti, Y., Naitow, H., Kunishima, N., and Yutani, K. (2018) Strategy for stabilization of CutA1 proteins due to ion-ion interactions at temperatures of over 100 °C. *Biochemistry* **57**, 2649-2656.
20. Yutani, K., Matsuura, Y., Naitow, H., and Joti, Y. (2018) Ion-ion interactions in the denatured state contribute to the stabilization of CutA1 proteins. *Sci. Rep.* **8**, 7613-7622.

21. Darden, T., York, D. and Pedersen, L. (1993) Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089-10092.
22. Bussi, G., Donadio, D. and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101-014107.
23. Parrinello, M. and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182-7190.
24. Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations. *J. Comp. Chem.* **18**, 1463-1472.
25. Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. (2005) The FoldX web server: an online force field. *Nucleic Acids Res.* **33**, W382-W388.
26. Yutani K., Matsuura Y., and Joti Y. (2019) Confirmation of the formation of salt bridges in the denatured state of CutA1 protein using molecular dynamics simulations. *BPPB* **16**, 176-184.
27. Yoda, T., Sugita, Y., and Okamoto, Y. (2004) Comparisons of force fields for proteins by generalized-ensemble simulations. *Chem. Phys. Lett.* **386**, 460-467.
28. Debiec, K. T., Gronenborn, A. M., and Chong, L. T. (2014) Evaluating the strength of salt bridges: a comparison of current biomolecular force fields. *J. Phys. Chem. B* **118**, 6561-6569.
29. Zhou, X., Xi, W., Luo, Y., Cao, S., and Wei, G. (2014) Interactions of a water-soluble fullerene derivative with amyloid- $\beta$  protofibrils: dynamics, binding mechanism, and the resulting salt-bridge disruption. *J. Phys. Chem. B* **118**, 6733-6741.
30. Ahmed, M. C., Papaleo, E., and Lindorff-Larsen, K. (2018) How well do force fields capture the strength of salt bridges in proteins? *Peer J.* **6**, e4967-e4984.
31. Tanaka, Y., Tsumoto, K., Umetsu, M., Nakanishi, T., Yasutake, Y., Sakai, N., Yao, M., Tanaka, I., Arakawa, T., and Kumagai, I. (2004) Structural evidence for guanidine-protein side chain interactions: crystal structure of *Pyrococcus horikoshii* CutA in the presence of 3M guanidine hydrochloride. *Biochem. Biophys. Res. Commun.* **323**, 185-191.
32. Sawano, M., Yamamoto, H., Ogasahara, K., Kidokoro, S., Katoh, S., Ohnuma, T., Katoh, E., Yokoyama, S., and Yutani, K. (2008) Thermodynamic basis for the

stabilities of three CutA1s from *Pyrococcus horikoshii*, *Thermus thermophilus*, and *Oryza sativa*, with unusually high denaturation temperatures. *Biochemistry* **47**, 721-730.

## Chapter 4: Stabilization of *Escherichia coli* CutA1 by rational protein design

### 4-1. Introduction.

In order to determine the key elements that improve the  $T_d$ , we used protein engineering to enhance the heat stability of *Ec*CutA1 so that it has a  $T_d$  that is comparable to the  $T_d$  of 150 °C for *Ph*CutA1. The reverse strategy, in which the unstable mutants are created from *Ph*CutA1, is not promising because it is likely that every mutation will destabilize the protein, making it difficult to identify the essential residues for stabilization. Therefore, creating a protein with a  $T_d = 150$  °C from *Ec*CutA1 is a better approach to confirm the thermo-stabilization mechanism of *Ph*CutA1 and to establish a method to rationally improve the conformational stability of a protein.

As a first step in the mutation strategy to enhance the heat stability of the CutA1 protein from *Ec*CutA1, we examined the structure-sequence (3D-1D) compatibility between the conformation of *Ec*CutA1 and its native sequence using SPMP (stability profile of mutant protein). SPMP estimates changes in the stability of 19 mutant proteins at every position of a protein based on the X-ray crystal structure<sup>1-3</sup>. We chose seven incompatible positions in *Ec*CutA1 which were predicted from SPMP, and then introduced single and multiple point mutations at these locations. The stabilities of the constructed *Ec*CutA1 mutants were evaluated by heat and denaturant denaturation, and their structures were determined by X-ray crystallography. We will discuss the

stabilization mechanism of the mutant proteins based on the X-ray crystalline structure and 3D-1D compatibility.

#### **4-2. Experimental methods.**

##### **Cloning, expression and purification of CutA1 from *E. coli*.**

The *EcCutA1* gene was PCR amplified using *PfuTurbo* DNA polymerase (Stratagene) and *E. coli* K-12 W3110 genomic DNA. The primers used to amplify *EcCutA1* gene were 5'-GATATACATATGCTTGATGAAAAAAGTTTCG-3' and 5'-AAAGGATCCTCAGCGTAAAGATGCGTTGAGC-3'. The full-length PCR products were digested with *NdeI* and *BamHI*, and the fragment was inserted into the pET-11a expression vector (Novagen) that had been linearized with *NdeI* and *BamHI*. All of the prepared mutants using the *EcCutA1* gene were generated using a QuikChange II XL site-directed mutagenesis kit (Stratagene). The sequence was confirmed by DNA sequencing.

For protein production, *E. coli* BL21-CodonPlus (DE3)-RIL competent cells were transformed with the recombinant plasmid and grown at 37 °C in Luria-Bertani medium containing 50 µg/ml ampicillin for 20 h. The cells were harvested by centrifuging at 4,500g for 5 min at 4 °C, suspended in 20 mM Tris-HCl, pH 8.0, containing 0.5 M NaCl and 2 mM EDTA, and finally disrupted by sonicating and heating at 70-90 °C for 15 min. The cell debris and heat-denatured proteins were removed by centrifuging at 20,000g for 30 min. The supernatant was used as the crude extract for purification.

The crude extract was desalted on a HiPrep 26/10 desalting column (GE Healthcare) and applied to a Super Q Toyopearl 650M (Tosoh) column equilibrated with 20 mM Tris-HCl, pH 8.0, containing 2 mM EDTA (buffer A). After eluting with a linear gradient of 0.2-1.4 M NaCl, the fractions containing *EcCutA1* were collected. Ammonium sulfate was added to the sample to a final concentration of 1.5 M, and the soluble fraction then applied to a Resource PHE column (GE Healthcare) equilibrated with 50 mM sodium phosphate buffer, pH 7.0, containing 1.5 M ammonium sulfate and 2 mM EDTA. The *EcCutA1*-containing fractions were eluted with a linear gradient of 0.75-0 M ammonium sulfate. The sample was concentrated by ultrafiltration (VivaSpin, 5 kDa cut-off) and loaded onto a HiLoad 16/60 Superdex 75 prep-grade column (GE Healthcare) equilibrated with buffer A containing 0.2 M NaCl. The homogeneity and identity of the purified sample were assessed by SDS-PAGE. The concentrations of the wild-type and mutant proteins were estimated from the absorbance at 280 nm, assuming  $E^{1\text{cm}}_{1\%} = 14.96$ , which is based on the number of aromatic amino acids<sup>4</sup>.

### **DSC experiments.**

DSC was performed at scan rate of 60 °C/h using a VP-capillary DSC platform. For the measurements, the protein concentrations were 0.25 to 0.68 mg/ml in 50 mM glycine buffer at pH 9.0 containing 2 mM EDTA and 5 mM  $\beta$ -mercaptoethanol. All samples were dialyzed against the buffers overnight at 4 °C and then filtered through a 0.22- $\mu\text{m}$  pore size membrane. The  $T_d$  represents the peak temperature of the DSC curves. The  $T_d$  values in this study represent the average of at least six experiments.

## Guanidine hydrochloride-induced unfolding and refolding.

For the denaturant unfolding experiment, the wild-type and mutant *EcCutA1* were incubated in various guanidine hydrochloride (GuHCl) concentrations at pH 8.0 and 37 °C for various time frames. For the refolding experiment, the proteins were completely unfolded in the presence of 7 M GuHCl for 15 min at 95 °C and then diluted with 20 mM Tris-HCl buffer at pH 8.0 containing 0.2 mM EDTA, 0.1 mM DTE (dithioerythritol), and various concentrations of GuHCl at 37 °C. The unfolding and refolding reactions were monitored by measuring the CD values at 220 nm. The CD measurements were carried out using a Jasco J-725 spectropolarimeter. The unfolded fraction was calculated using equation 1 reported in our previous paper<sup>5</sup>. The refolding curves were analyzed by a linear extrapolation model assuming a two-state transition for the unfolding of a trimer protein, according to the following equations<sup>5,6</sup>.



$$K = 27Ct^2 f_u^3 / (1 - f_u) \quad (2)$$

$$\Delta G^0 = -RT \ln K \quad (3)$$

$$\Delta G^0 = \Delta G_{H_2O}^0 + m[x] \quad (4)$$

$$f_u = \left( \frac{1}{2} \left( \frac{\exp(-(\Delta G_{H_2O}^0 + mx)/(RT))}{27Ct^2} \right) + \frac{1}{18} \left( \frac{\exp(-(\Delta G_{H_2O}^0 + mx)/(RT))}{27Ct^2} \right) + \frac{1}{27} \right)^{1/2} \left( \frac{\exp(-(\Delta G_{H_2O}^0 + mx)/(RT))}{27Ct^2} \right)^{1/3} - \frac{1}{3} \left( \frac{\exp(-(\Delta G_{H_2O}^0 + mx)/(RT))}{27Ct^2} \right) \left( \frac{1}{2} \left( \frac{\exp(-(\Delta G_{H_2O}^0 + mx)/(RT))}{27Ct^2} \right) + \frac{1}{18} \left( \frac{\exp(-(\Delta G_{H_2O}^0 + mx)/(RT))}{27Ct^2} \right) + \frac{1}{27} \right)^{1/2} \left( \frac{\exp(-(\Delta G_{H_2O}^0 + mx)/(RT))}{27Ct^2} \right)^{1/3} \quad (5)$$

where  $K$ ,  $Ct$ , and  $\Delta G^0$  are the equilibrium constant of the unfolding reaction, the molar concentration of the protein expressed in trimer equivalents, and the Gibbs energy

change upon unfolding (standard value), respectively. Factor  $m$  is the slope of the linear correlation between  $\Delta G^0$  and the GuHCl concentration ( $x$ ). Factor  $f_u$  in equation 5 is represented as a function of the GuHCl concentration. We used the data analysis software in MicroCal Origin (Northampton, MA, USA) to produce a least squares fit of the experimental data in the GuHCl unfolding curves to obtain  $\Delta G^0_{H_2O}$ , where  $\Delta G^0_{H_2O}$  is the standard value (in 1 M protein concentration) in the absence of GuHCl (in water).

### **Crystallization and X-ray data collection.**

A 1.5  $\mu$ l aliquot of S11V (23 mg/ml), E61V (21 mg/ml) and S11V/E61V (19 mg/ml) in 20 mM Tris buffer at pH 8.0 including 0.2M NaCl and 2mM EDTA was mixed with an equal volume of reservoirs A, B and C, respectively. Reservoir A contains ~76% of the Crystal Screen 1 (no. 38) from Hampton Research which comprises to 0.1 M HEPES-NaOH (pH 7.5) and 1.4 M tri-sodium citrate. Reservoirs B and C contain ~80% of the SaltRx no. 22 and 21 from Hampton Research, respectively, which comprises 1.2 M tri-sodium citrate containing 0.1 M Tris (pH 8.5) and 0.1 M Bis-Tris propane (pH 7.0), respectively. Crystals suitable for data collection were obtained using the hanging-drop vapor diffusion method over several days at 20 °C. X-ray diffraction data for S11V and S11V/E61V were collected at the RIKEN structural genomics beamline BL26B2 of SPring-8, and data for E61V were collected using an in-house R-axis VII system (Rigaku).

### **Structure solution and refinement.**

The structures of S11V, E61V and S11V/E61V were determined by the molecular replacement method using the wild-type CutA1 from *E. coli* (PDB code, 1NAQ) as the search model. The solution was determined by using automated-MOLREP within the CCP4 program suite<sup>7</sup>, and then refined using CNS<sup>8</sup>. The protein model was built using Coot<sup>9</sup>. The quality of the models was inspected by Procheck<sup>10</sup>. The quality of both structural residues in the most favored regions was over 90%, and there were no residues in the generously allowed or disallowed regions. The statistics for the data collection and refinement are summarized in Table 1.

### **Calculation of changes in unfolding Gibbs energy due to the hydrophobic interaction.**

To estimate the difference ( $\Delta\Delta G_{HP}$ ) in the unfolding Gibbs energy ( $\Delta G$ ) between the wild-type and mutant proteins due to hydrophobic interactions, the following equation is proposed using structural information such as ASA (accessible surface area)<sup>11</sup>.

$$\Delta\Delta G_{HP} = 0.146\Delta\Delta ASA_{non-polar} + 0.021\Delta\Delta ASA_{polar} \quad (6)$$

where  $\Delta\Delta ASA_{non-polar}$  and  $\Delta\Delta ASA_{polar}$  represent the differences in the  $\Delta ASA$  of the non-polar and polar atoms of all residues, respectively, between the wild-type and the mutant proteins upon denaturation. The parameters of equation 6 have been obtained using the stability/structure database upon the denaturation of mutant human lysozymes and T4 lysozyme. To calculate the ASA value, the carbon and sulfur atoms in the residues are assigned to  $ASA_{non-polar}$ , and the nitrogen and oxygen atoms are represented by  $ASA_{polar}$ . The ASA values of the native state were calculated following the procedure proposed by Connolly<sup>12</sup>, using the X-ray crystal structures of the mutant CutA1 proteins. The ASA

values for the denatured forms were estimated using the extended structures of each protein, which were generated from the native structures using the programme Insight II.

#### ***4-3. Results and discussion.***

##### ***EcCutA1 mutants***

Ota *et al.*<sup>3</sup> proposed that changes in the stability of 19 mutant proteins at every position in a protein can be estimated by knowledge-based potential if the X-ray crystal structure data are available. This method is called SPMP analysis, and we used SPMP to examine the structure-sequence (3D-1D) compatibility of *EcCutA1*. A pseudo-energy potential ( $\Delta G_{\text{SPMP}}$ ) derived from a number of PDB structures is used in SPMP. This potential was originally developed to predict protein structures and consists of four elements, *i.e.*, the side-chain packing ( $\Delta G_{\text{SP}}$ ), hydration ( $\Delta G_{\text{Hyd}}$ ), local structure ( $\Delta G_{\text{LC}}$ ) and back-bone-side-chain repulsion ( $\Delta G_{\text{BR}}$ ) as expressed by the following equation.

$$\Delta G_{\text{SPMP}} = \Delta G_{\text{SP}} + \Delta G_{\text{Hyd}} + \Delta G_{\text{LC}} + \Delta G_{\text{BR}} \quad (7)$$

Table 2 shows the SPMP scores of the native amino acids in the chain A of trimer *EcCutA1* (PDB code: 1NAQ) for the four SPMP elements in order of worst (the most incompatible amino acid) to best. The “rank” column in Table 2 shows the rank position of the wild-type amino acid at the specified position out of 20 amino acids, and the “fitness of amino acid” column lists the 20 amino acids from the most compatible to the worst. The most incompatible amino acid estimated by SPMP was Lys67, but Lys67 and Glu90 form a salt bridge in the wild-type structure. Generally, the estimation of electrostatic interaction is known to be difficult for the knowledge-based potentials, as was employed by SPMP. Then, we considered that Lys67 would not necessarily be incompatible. The rank positions of the wild-type amino acids listed in Table 2 were very

low and they were ranked worse than 12<sup>th</sup>, except for Val96, which was ranked 9<sup>th</sup> (see the rank column in Table 2). Therefore, seven positions except for these three residues at positions 67, 90, and 96 were mutated to improve the protein stability. The bold and bold-underlined letters in the fitness column of amino acids in Table 2 show the wild-type and substituted residues, respectively. Single and multiple mutations were constructed at these sites, which are highlighted in the crystal structures of wild-type *EcCutA1* in Fig. 1. The first rank amino acid at position 45, 61, and 73 in Table 2 was not chosen because the side-chains are more bulky than that of the wild-type residues. The SPMP was also applied to evaluate mutations based on *EcCutA1* mutant structures determined in this study (Table 3, described later).

### **DSC experiments on mutant *EcCutA1***

To measure the changes in stability due to the mutations, DSC experiments were conducted with the mutant proteins at pH 9.0. Typical DSC curves for the wild-type and mutant proteins are shown in Fig. 2, and the  $T_d$  of the wild-type and all of the mutant proteins are listed in Table 4. The  $T_d$  values were obtained from the peak temperature of the DSC curves because heat denaturation is not reversible. The stability of the constructed *EcCutA1* mutants was improved at five out of the chosen seven positions. The  $T_d$  values for S11A, E59K, Y60T, E61V, and Q73V were 106.4, 104.2, 91.9, 103.4 and 94.0 °C at pH 9.0, respectively. S11A was improved by 16.5 °C compared to the wild-type (89.9 °C), while G45Q (82.1 °C) and Y86D (76.8 °C) were lower. Multiple mutants at the improved positions were also constructed; the  $T_d$  values of S11V/E61T, S11V/E61V, and S11V/E61V/Q73V were 112.3, 113.5, and 116.5 °C, respectively. The highest one was improved by 26.6 °C. These results indicate that the stability of *EcCutA1*

is remarkably improved by single and multiple substitutions, even though the  $T_d$  of the wild-type protein is considerably high, near 90 °C.

### **Denaturant denaturation of mutant *EcCutA1***

DSC experiments showed only the apparent denaturation temperatures of all of the mutant proteins because the heat denaturation process was irreversible. Therefore, in order to evaluate the thermodynamic parameters of denaturation, the unfolding and refolding curves for the wild-type and mutant proteins were analyzed using the denaturant guanidine hydrochloride (GuHCl). The unfolding curves for all of the proteins versus the various GuHCl concentrations did not reach a constant value even after several days, indicating that the GuHCl-induced denaturation rate is remarkably slow and similar to that for CutA1 proteins from other sources<sup>5</sup>. However, the refolding curves for the wild-type and mutant proteins were highly similar after 1- and 2-day incubations at 37 °C, suggesting that the refolding reaction reached equilibrium after a 1-day incubation at 37 °C. Fig. 3 shows the typical refolding curves for the wild-type and mutant proteins after 1 day at pH 8.0 and 37 °C. These refolding curves were analyzed using the linear extrapolation model assuming a two-state transition for the unfolding of a trimer protein ( $N_3 \leftrightarrow 3U$ ) using equation 5. The solid curves in Fig. 3 represent the fitting lines using equation 5. The unfolding thermodynamic parameters that were obtained from the GuHCl refolding curves for the examined proteins are listed in Table 4. The midpoint of the S11V/E61V/Q73V refolding curve was 4.67 M GuHCl, which is remarkably higher than that of the wild-type protein (2.72 M). The  $\Delta G^0_{H_2O}$  of S11V/E61V/Q73V was 258.9 kJ/mol (Table 4). These thermodynamic parameters were higher than the  $T_d$  (112.8 °C),  $\Delta G^0_{H_2O}$  (207.7 kJ/mol), and midpoint (3.33 M) of CutA1 (*TiCutA1*) from an extreme

thermophile, *Thermus thermophilus*, which grows at approximately 75 °C<sup>5</sup>. The thermodynamic parameters of the double mutant S11V/E61V were also higher than those of *Tt*CutA1. These results indicate that substitutions at only two positions in the mesophile protein greatly improve the protein stability to a level that exceeds the stability of an extreme thermophile protein.

The relationship between the thermodynamic parameters from the GuHCl denaturation and the  $T_d$  from the DSC experiments were examined as shown in Fig. 4. There was a good correlation between  $\Delta G^0_{H_2O}$  and  $T_d$  with a correlation coefficient of 0.86 (Fig. 4). Overall, these results indicate that changes in the heat stability of these mutants roughly parallel those for the denaturant denaturation at 37 °C.

### **Crystal structures of mutant CutA1 and SPMP evaluations**

The stabilities of the proteins with mutations at positions 11 and 61 were remarkably improved as shown in Table 4. Therefore, the tertiary structures of three mutant proteins, S11V, E61V and S11V/E61V, were determined by X-ray crystallography to confirm the structural changes due to these mutations. The overall crystal structures of these three mutant proteins were highly similar to the reported wild-type *Ec*CutA1 structure (1NAQ)<sup>13</sup>. The structures of S11V and S11V/E61V were superimposed with a root mean square deviation (RMSD) of  $0.35 \pm 0.17$  Å between the equivalent C $\alpha$  atoms from positions 10 to 110 of three subunits in both structures, and the RMSD value between E61V and S11V/E61V was  $0.38 \pm 0.17$  Å. On the other hand, the RMSD values of S11V, E61V, and S11V/E61V versus the wild-type structure (1NAQ) were  $0.77 \pm 0.40$  Å,  $0.81 \pm 0.41$  Å, and  $0.80 \pm 0.40$  Å, respectively. The values among the three mutant proteins were approximately half compared to those between the

mutant and wild-type proteins. This difference is due to the high RMSD values of  $\sim 4$  Å around position 86. In the reported wild-type *EcCutA1* structure<sup>13</sup>, each Cys (3 Cys per subunit) binds a mercury atom or a hydroxyl-mercuribenzoic acid. When the structures of the wild-type protein (1NAQ) and the mutant proteins without mercury are compared, the loop around position 86 of the wild-type protein is pushed out into the solvent, to escape from crushing the mercuribenzoic acids bound to Cys16, resulting in great RMSD values.

Table 3 shows the SPMP evaluation of three determined structures of *EcCutA1* proteins with mutations at positions 11 and 61. The SPMP scores of the mutant proteins represent average values of each subunit since the *EcCutA1* structure consists of three identical subunits. The SPMP values ( $\Delta G_{\text{SPMP}}$ ) for S11V at position 11 and E61V at position 61 were remarkably improved from 1.76 to -6.16 kJ/mol and from 0.88 to -2.09 kJ/mol, respectively, and the rank among the 20 amino acid residues increased from 13<sup>th</sup> to 1<sup>st</sup> and 12<sup>th</sup> to 7<sup>th</sup>, respectively. For the double mutant, the scores at positions 11 and 61 were similar to those of the single mutant protein at each position, indicating that the mutations at each position are independent. The SPMP evaluations suggest that substituting Ser with Val at position 11 improves the stability due to changes in the hydration effect ( $\Delta G_{\text{hyd}}$ ), local structure ( $\Delta G_{\text{LC}}$ ), and side-chain packing ( $\Delta G_{\text{SP}}$ ), whereas substituting Glu with Val at position 61 mainly results in changes in  $\Delta G_{\text{hyd}}$  and  $\Delta G_{\text{LC}}$ . These SPMP evaluations are consistent with the experimental DSC and denaturant denaturation results. The  $T_d$  values for S11V and E61V were 105.0 and 103.4 °C, respectively, and the  $T_d$  value for the double mutant was 113.5 °C, suggesting a cumulative effect for each single mutation.

### **Evaluation of stability changes due to mutations based on tertiary structures**

The contribution of some stabilization factors has been quantitatively derived as parameters such as a hydrophobic interaction ( $\Delta\Delta G_{HP}$ ) and hydrogen bond ( $\Delta\Delta G_{HB}$ ) by a unique equation that considers conformational changes due to the mutation(s) using mutant human lysozymes<sup>13</sup>. The enhanced stabilities of S11V and E61V were expected to improve the local structure based on the SPMP evaluation (Table 3), suggesting the conversion to favorable secondary structure. The changes in the propensity to form a secondary structure ( $\Delta\Delta G_{pro}$ ) among the parameters estimated by Funahashi *et al.*<sup>13</sup> are useful to elucidate changes in the stability of the two mutants. The  $\Delta\Delta G_{pro}$  values of S11V and E61V were calculated to be 1.95 (5.84) and 2.73 kJ/mol (8.18 kJ/mol of a trimer) (Table 5), respectively, using the estimated parameter<sup>13</sup>, suggesting that these mutations improve the protein stability. This occurs because both positions are located in the  $\beta$ -sheet and the two substitutions include residues with a higher propensity to form the  $\beta$ -sheet, which indicates enhanced stabilization. These values are compatible with the experimental results and the SPMP evaluation.

Funahashi *et al.*<sup>13</sup> also evaluated the parameters of the hydrophobic interaction. The changes ( $\Delta\Delta G_{HP}$ ) in unfolding Gibbs energy ( $\Delta G$ ) due to mutations, which was caused by hydrophobic interactions, were calculated using equation 6 in the Experimental Methods. Table 5 shows the energy of the hydrophobic interaction ( $\Delta G_{HP}$ ) at the mutated residues in the three mutant proteins and the difference in  $\Delta G_{HP}$  due to these mutations ( $\Delta\Delta G_{HP}$ ). For example, the  $\Delta\Delta G_{HP}$  value of S11V was calculated to be 8.6 kJ/mol, subtracting the average value (7.3) of  $\Delta G_{HP}$  at Ser11 in E61V (as the wild-type value) from the average value (15.9) of  $\Delta G_{HP}$  at Val11 in S11V (Table 5). This method was used because  $\Delta G_{HP}$  at Glu61 of the reported wild-type structure (1NAQ) might be slightly affected by ligands

binding to the Cys residues that were previously described, although the average value of  $\Delta G_{HP}$  at Ser11 in E61V was similar to that of 1NAQ.

These  $\Delta\Delta G_{HP}$  values were compared with  $\Delta\Delta G_{H_2O}^0$  obtained from the GuHCl denaturation experiments (Table 4). The  $\Delta\Delta G_{H_2O}^0$  values for S11V and S11V/E61V, 40.7 and 79.8 kJ/mol, were remarkably larger than the  $\Delta\Delta G_{HP}$  values for these two proteins, 25.8 and 54.0 kJ/mol, respectively, suggesting that other factors such as the propensity to form secondary structures contribute to the stabilization of these proteins. In the case of S11V/E61V, the summation of  $\Delta\Delta G_{HP}$  and  $\Delta\Delta G_{pro}$  was 68.0 kJ/mol and comparable to the  $\Delta\Delta G_{H_2O}^0$  value (79.8 kJ/mol), suggesting that the stability of this mutant protein is mainly stabilized by these two factors. On the other hand, the  $T_d$  and midpoint of the GuHCl denaturation of E61V are close to those of S11V, but the  $\Delta\Delta G_{H_2O}^0$  of E61V was less than half of S11V due to a decrease in the  $m$  (slope) value (Table 4). It may be necessary to evaluate the changes in the  $m$  values in more detail.

### **Characteristics of stabilized *Ec*CutA1 due to mutations**

The tested CutA1 mutants were stabilized at five out of the seven positions that were identified by SPMP. Ser11 in the wild-type protein is almost completely buried in the interior of the molecule and located in the N-terminus of the  $\beta$ -strand. SPMP analysis indicates that the hydration and local structure at Ser11 should be improved (Table 2). The “local structure” score in SPMP evaluates the dihedral angle of each residue in the structure<sup>2</sup>. In the local structure function of SPMP, the propensity of a single residue to form a favorable secondary structure in the backbone conformation is evaluated by classifying the structure into five states (such as  $\alpha$ -helix and  $\beta$ -strand) according to both its position in the  $\phi\psi$  space (*i.e.*, the Ramachandran plot)<sup>14</sup> and the secondary structure

based on the definition by Kabsch and Sander<sup>15</sup>. Therefore, Ser11 was changed to three residues (Val, Ala, or Ile) with a higher hydrophobicity and propensity to form a  $\beta$ -sheet. Compared to the wild-type protein, the  $T_d$  values of the three mutant proteins, S11A, S11V, and S11I, remarkably improved by 16.5, 15.1, and 12.8 °C, respectively, and the stabilities against denaturant also increased. Both Val and Ile have a higher hydrophobicity<sup>16</sup> and propensity to form a  $\beta$ -sheet<sup>17</sup> than Ala. SPMP predicted that S11V has the highest stability among the three mutations (Table 2). For S11V, the summation of  $\Delta\Delta G$  of the hydrophobicity and the propensity to form a secondary structure based on the structure were also comparable with the experimentally obtained  $\Delta\Delta G^0_{H_2O}$  value (Tables 4 and 5), indicating that drastic changes in the stability due to this mutation are mainly caused by changes in the native structure due to the mutation.

Positions 59, 60, and 61 in a sequential residue number were selected by SPMP to improve the protein stability (Fig. 1). These residues are located in the middle of the  $\beta$ -strand and their side-chains are almost exposed to a solvent. The SPMP scores of Glu59 and Glu61 were reduced due to the penalties associated with local structure, while the score for Tyr60 was reduced due to hydration. The Glu residue at position 59 was substituted with a hydrophobic (Leu), a polar-non-charged (Gln), and an opposite-charged residue (Lys). The  $T_d$  values of the three mutants, E59Q, E59L, and E59K, improved by 8.4, 10.3, and 14.3 °C, respectively, compared to the wild-type protein. The propensity to form a  $\beta$ -sheet is in the order of Leu, Gln, Lys, and Glu. The mutant proteins at position 59 might be stabilized by improving the propensity to form a  $\beta$ -sheet, but the  $T_d$  of E59K remarkably increased, suggesting that removing the negative charge contributes to the stability of E59L and E59Q. The mutant Y60T, which does not have a penalty for hydration as estimated by SPMP, was stabilized by only 2.0 °C. At position

61, Glu was substituted by residues that had the three highest evaluations by SPMP, except for Trp. The  $T_d$  values of these three mutants, E61V, E61T, and E61H, improved by 13.5, 8.7, and 4.4 °C, respectively. The propensity to form a  $\beta$ -sheet was in the order of Val, Thr, His, and Glu, which corresponds to the changes in stability. Furthermore, the hydrophobic interaction was estimated to remarkably improve based on the structural changes due to these mutations (Table 5). Therefore, in the case of E61V, both the propensity to form a  $\beta$ -sheet and hydration might predominantly contribute to the increased stability. The  $T_d$  value of Q73V was increased by 4.1 °C. This mutation was expected to reduce the penalty associated with the side-chain packing evaluated by SPMP (Table 2).

The stability of the double and triple mutants roughly shows a cumulative effect of each mutant at positions 11, 61, and 73. The  $T_d$  of the triple mutant, S11V/E61V/Q73V, was 116.5 °C, the highest among the examined mutant proteins, and the  $\Delta T_d$  between this triple mutant and the wild-type protein was 26.6 °C. However, the stabilities of E59K/E61V (104.3 °C) and S11V/E59K/E61T (112.4 °C) were highly similar to E59K (104.2 °C) and S11V/E61T (112.3 °C), respectively. In this case, it seems that the stabilization effect of the mutations at position 61 compensates for removing the electrostatic interaction between Lys59 and Glu61 due to mutations at Glu61, resulting in no cumulative effect on the overall stability. Overall, the drastic increases in stability in the present studies were quantitatively elucidated based on the native structures, although the denatured structures might be affected by the introduced mutations.

Only G45Q and Y86D were remarkably destabilized contrary to the SPMP predictions. Other mutations at these positions were not examined because they are not expected to have greater stability than the wild-type protein.

#### 4-4. Conclusions.

- (1) In order to elucidate the stabilization mechanism of *Ph*CutA1 with a  $T_d$  of 150 °C, we have tried to improve the heat stability of *Ec*CutA1. First, we chose seven positions in *Ec*CutA1 where the native amino acids were incompatible with the structure according to SPMP.
- (2) The stability of the *Ec*CutA1 mutants for five out of the worst seven positions were remarkably improved compared to the wild-type protein. The highest  $T_d$  values for the single (S11A), double (S11V/E61V), and triple (S11V/E61V/Q73V) mutants were 106.4, 113.5, and 116.5 °C respectively, at pH 9.0. The  $T_d$  values of the double and triple mutants exceeded that for CutA1 from *Thermus thermophilus*. The present results indicate that the stability of *Ec*CutA1 is remarkably improved by a few substitutions, even though the stability of the wild-type protein is unusually high with a  $T_d$  of 90 °C.
- (3) The heat stabilities ( $T_d$ ) of the mutant proteins roughly parallel the  $\Delta G^0_{H_2O}$  at 37 °C that was obtained from GuHCl denaturation. The changes in stability ( $\Delta\Delta G^0_{H_2O}$ ) of the three mutant proteins (S11V, E61V, and S11V/E61V) were drastic, but were quantitatively elucidated based on the newly solved native structure.
- (4) As expected, the SPMP scores at positions 11 and 61 of the mutant structures were improved. The present results confirm that SPMP is a powerful tool to explore amino acid substitutions at specific positions in order to improve the stability of a protein.
- (5) The present *Ec*CutA1 mutants with deleted incompatible structures would be good templates to improve the heat stability of *Ec*CutA1 with a  $T_d$  (150 °C) comparable to that of *Ph*CutA1 with many ionic residues.

Table 1.

Data collection and refinement statistics of mutant CutA1 proteins from *E. coli*.

	S11V	E61V	S11V/E61V
<b>Data collection</b>			
Wavelength (Å)	1.000	1.54178	1.000
Space group	$P2_12_12_1$	$P2_12_12_1$	$P6_1$
Cell dimension (Å)	a = 62.239 b = 96.872 c = 106.352	a = 38.187 b = 50.101 c = 147.424	a = b = 52.664 c = 178.361
No. of molecule/asu	6	3	3
Resolution range (Å)	40.0-2.40(2.49-2.40)	40.0-2.30(2.38-2.30)	40.0-2.30(2.38-2.30)
Unique reflections	25891	13267	12446
Redundancy	7.0(7.0)	6.4(6.1)	9.6(6.2)
Completeness (%)	100	99.4	99.6
R <sub>merge</sub> (%)	6.2(30.0)	3.5(10.5)	11.7(33.0)
<b>Refinement statistics</b>			
Resolution range (Å)	40-2.4	40-2.3	40-2.3
Reflection used in refinement	25793	13222	12336
R <sub>work</sub>	0.202	0.205	0.174
R <sub>free</sub>	0.251	0.247	0.238
No. of water molecules	226	114	300
RMSD bond length (Å)	0.0062	0.0059	0.0057
RMSD bond length (angle)	1.203	1.151	1.135
<b>Ramachandran statistics</b>			
Most favored region (%)	93.6	92.5	94.0
Additional allowed region (%)	6.4	7.5	6.0
<b>PDB code</b>			
	3AH6	3AA9	3AA8

Table 2.

SPMP scores of *Ec*CutA1 for the four elements, in the order of worst to best.  $\Delta G_{SPMP}$ ,  $\Delta G_{SP}$ ,  $\Delta G_{Hyd}$ ,  $\Delta G_{LC}$ , and  $\Delta G_{BR}$  represent scores of the wild-type amino acid, side-chain packing, hydration, local structure, and back-bone-side-chain repulsion, respectively. Negative values of  $\Delta G$  means "stabilization". The unit is kJ/mol of a monomer.

order	position	wild-type	SS <sup>a</sup>	burial <sup>b</sup>	$\Delta G_{SPMP}$	$\Delta G_{SP}$	$\Delta G_{Hyd}$	$\Delta G_{LC}$	$\Delta G_{BR}$	rank <sup>c</sup>	fitness of amino acids <sup>d</sup>
1	67	Lys	E	8	2.63	-1.21	2.26	1.30	0.29	17	FIYVLCMWHTSAGQRNKPDE
2	11	Ser	E	9	1.76	-0.42	1.09	1.00	0.08	13	<b><u>V</u>CIA</b> MLYFTPWHSGQNERDK
3	73	Gln	H	6	1.42	1.59	0.25	-0.75	0.33	14	I <b><u>V</u></b> MLCTFSHAWYGQKERNDP
4	59	Glu	E	5	1.25	-0.50	-0.04	1.71	0.08	15	<b><u>L</u></b> HFWYCRMVIT <b><u>Q</u></b> KSEPNADG
5	45	Gly	E	5	1.21	0.21	0.00	1.00	0.00	19	MLVTC <b><u>Q</u></b> YWSFHREPAKNGD
6	86	Tyr	C	2	0.96	0.00	2.01	-1.05	0.00	12	<b><u>D</u></b> NTSEKQRHCAYWMIFGLVP
7	60	Tyr	E	2	0.92	-0.67	2.47	-0.96	0.08	13	<b><u>T</u></b> ESQKRHNGVPDYMFIALWC
8	61	Glu	E	6	0.88	-1.76	0.63	1.71	0.29	12	W <b><u>T</u></b> H <b><u>V</u></b> YCFSAINEQRPDGM <b><u>L</u></b> K
9	96	Val	C	6	0.88	-0.67	-0.13	1.46	0.21	9	CSTGNADPVHMEW <b><u>I</u></b> QLYFRK
10	90	Glu	C	6	0.75	-0.71	0.71	0.59	0.17	18	PRWYFHGACKQVTSIM <b><u>N</u></b> ELD

<sup>a</sup>SS: secondary structure determined by DSSP<sup>15</sup>. Structures other than  $\alpha$  helix (H) and  $\beta$  strand (E) are denoted as coil (C).

<sup>b</sup>burial: degree of residue burial of the site (1: exposed to water, 9: buried in the protein).

<sup>c</sup>rank: fitness rank position of the wild-type amino acid at the position out of 20 amino acids.

<sup>d</sup>fitness of amino acid: 20 amino acids sorted from the most compatible one to the most incompatible one. Bold and bold-underlined letters denote the wild-type and substituted residues in this study, respectively.

Table 3. SPMP evaluation at the mutation sites of the three solved mutant structures and the wild-type one. The total SPMP values for E61V and S11V/E61V show average of 3 subunits and its deviation. Those of S11V are the average of 6 subunits in an asymmetric unit. Negative values of  $\Delta G$  mean "stabilization". The values of the four elements show only average values of 3 subunits.

proteins	position	$\Delta G_{SPMP}$	$\Delta G_{SP}$	$\Delta G_{Hyd}$	$\Delta G_{LC}$	$\Delta G_{BR}$	rank
S11V	11Val	-6.16 $\pm$ 0.55	-3.23	-1.53	-1.49	0.10	1
	61Glu	0.91 $\pm$ 0.48	-2.03	0.63	1.71	0.59	15
E61V	11Ser	1.21 $\pm$ 0.04	-0.18	0.99	0.32	0.08	12
	61Var	-2.09 $\pm$ 0.08	-1.24	-0.24	-1.30	0.68	7
S11V/E61V	11Val	-5.57 $\pm$ 0.79	-2.58	-1.55	-1.50	0.06	1
	61Val	-1.02 $\pm$ 0.92	-0.06	-0.15	-1.30	0.49	7
wild-type	11Ser	1.76	-0.42	1.09	1.00	0.08	13
	61Glu	0.88	-1.76	0.63	1.71	0.29	12

Table 4. Thermodynamic parameters of unfolding for the wild-type and mutant *Ec*CutA1 proteins.

Proteins	$T_d$ (°C)*	$\Delta G^0_{H_2O}$ (kJ/mol)*	midpoint (M)*	m (slope) (kJ/mol M)*	$\Delta\Delta G^0_{H_2O}$ (kJ/mol)**
<i>wild-type</i>	89.9 ± 0.5	165.2 ± 10.2	2.72 ± 0.01	-35.2 ± 2.1	0
S11V	105.0 ± 0.8	205.9 ± 8.8	3.31 ± 0.00	-41.2 ± 1.3	40.7
S11A	106.4 ± 0.2	185.9 ± 8.1	3.06 ± 0.00	-38.3 ± 1.2	20.7
S11I	102.7 ± 0.8	ND	ND	ND	
G45Q	82.1 ± 1.2	ND	ND	ND	
E59L	100.2 ± 0.3	177.3 ± 10.9	2.96 ± 0.01	-36.1 ± 2.7	12.1
E59K	104.2 ± 0.1	181.6 ± 9.5	3.01 ± 0.01	-37.2 ± 1.9	16.4
E59Q	98.3 ± 0.7	ND	ND	ND	
Y60T	91.9 ± 2.1	ND	ND	ND	
E61V	103.4 ± 0.4	179.5 ± 9.5	3.29 ± 0.01	-33.3 ± 1.5	14.3
E61T	98.6 ± 0.2	178.4 ± 10.0	2.99 ± 0.01	-34.5 ± 1.2	13.2
E61H	94.3 ± 1.7	ND	ND	ND	
Q73V	94.0 ± 2.3	147.9 ± 7.3	3.40 ± 0.02	-23.0 ± 1.4	-17.3
Y86D	76.8 ± 0.1	138.5 ± 2.6	2.67 ± 0.01	-25.9 ± 0.7	-26.7
S11V/E61T	112.3 ± 0.1	201.3 ± 17.3	3.48 ± 0.01	-34.7 ± 1.7	36.1
S11V/E61V	113.5 ± 0.4	245.0 ± 20.1	3.80 ± 0.01	-44.7 ± 2.0	79.8
E59K/E61V	104.3 ± 0.3	188.1 ± 19.2	3.24 ± 0.01	-36.6 ± 2.8	22.9
S11V/E61T/Q73V	113.6 ± 0.6	202.4 ± 15.2	4.20 ± 0.01	-30.0 ± 1.4	37.2
S11V/E61V/Q73V	116.5 ± 0.7	258.9 ± 42.2	4.67 ± 0.01	-34.8 ± 2.5	93.7
S11V/E59K/E61T	112.4 ± 0.3	209.6 ± 16.2	3.48 ± 0.02	-39.8 ± 2.1	44.4

\* $T_d$  values were determined from the DSC experiments at pH 9.0 and the other values were from the GuHCl denaturation experiments at pH 8.0.

\*\*The difference in  $\Delta G^0_{H_2O}$  between the wild-type and mutant.

Positive values of  $\Delta G$  and  $\Delta\Delta G$  mean stabilization. ND means "not determined".

Table 5. Hydrophobic energy of substitution residues of the three mutant proteins evaluated by changes in the ASA values and changes in secondary propensity due to mutations.

proteins	position	$\Delta G_{HP}$ (kJ/mol)*				$\Delta\Delta G_{HP}$ (kJ/mol)**	$\Delta\Delta G_{Pro}$ (kJ/mol)***
		A	B	C	average values		
S11V	Val11	15.9	15.6	16.3	15.9 $\pm$ 0.3	8.6(25.8)	2.0(5.9)
	Glu61	8.1	8.8	8.4	8.4 $\pm$ 0.2		
E61V	Ser11	6.8	7.5	7.5	7.3 $\pm$ 0.3	8.3(24.9)	2.7(8.2)
	Val61	17.0	16.2	16.9	16.7 $\pm$ 0.6		
S11V/E61V	Val11	17.5	15.7	15.7	16.3 $\pm$ 0.8	18.0(54.0)	4.7(14.0)
	Val61	16.5	17.6	18.0	17.4 $\pm$ 0.6		

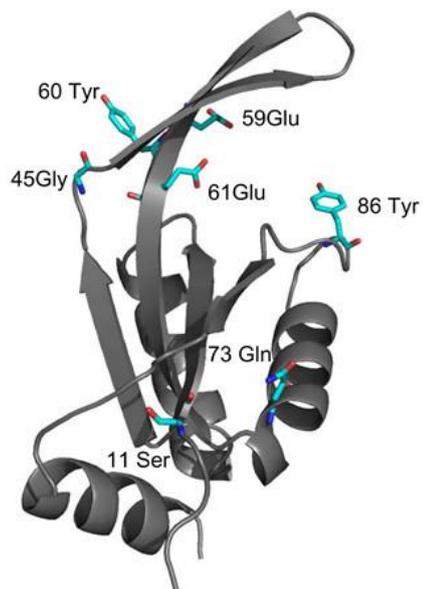
\*A, B, and C in a header represent unfolding Gibbs energy of hydrophobic interaction ( $\Delta G_{HP}$ ) at a given residue of each subunit.

\*\* $\Delta\Delta G_{HP}$  represents difference in  $\Delta G_{HP}$  due to mutations. The values in parentheses represent  $\Delta\Delta G_{HP}$  per a trimer.

\*\*\*Change in propensity to form a secondary structure ( $\Delta G_{Pro}$ ) due to mutations. The values in parentheses represent energies per a trimer.

Positive value means stabilization.

(A)



(B)

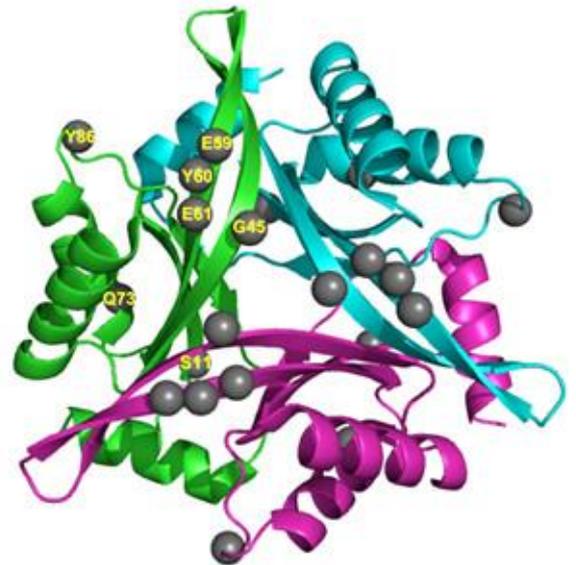


Fig. 1. The crystal structure of the wild-type *EcCutA1*. (A) Monomer structure of the wild-type *EcCutA1* (1NAQ). The labeled names represent the substituted residues. (B) Trimer structure of the wild-type *EcCutA1* (1NAQ). The different colors represent the different subunits. Black circles denote the substituted residues.

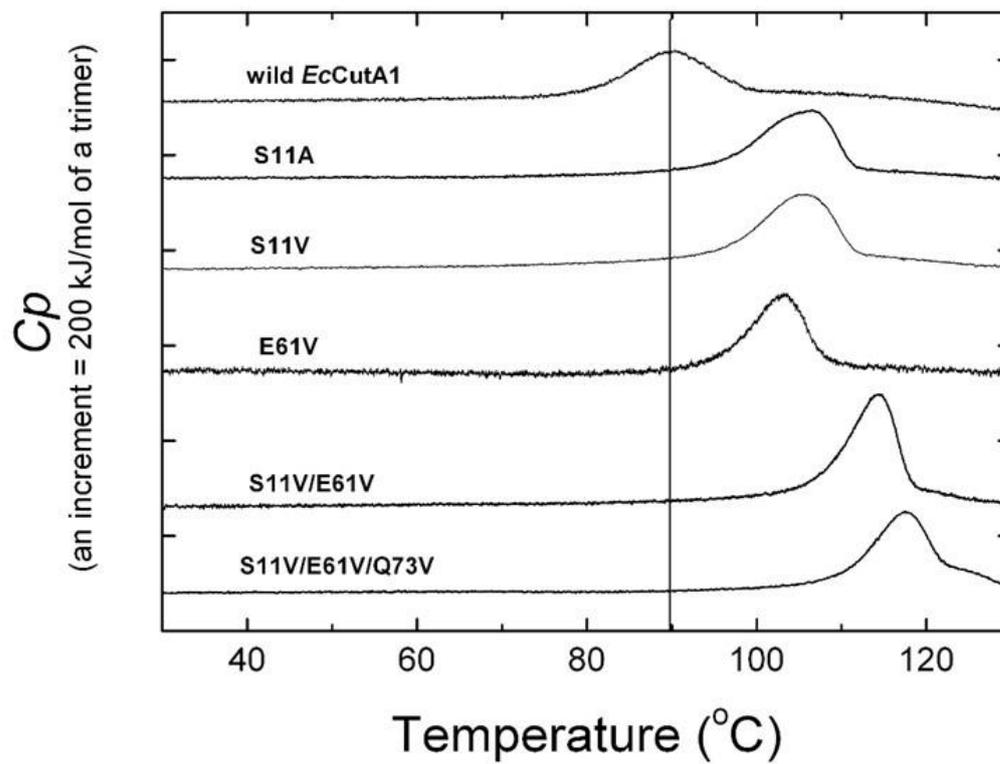


Fig. 2. Typical DSC curves of the wild-type and mutant *EcCutA1* at pH 9.0. The scan rate of each DSC curve was 60 °C/h. The perpendicular line at 89.9 °C shows the  $T_d$  value of the wild-type protein.

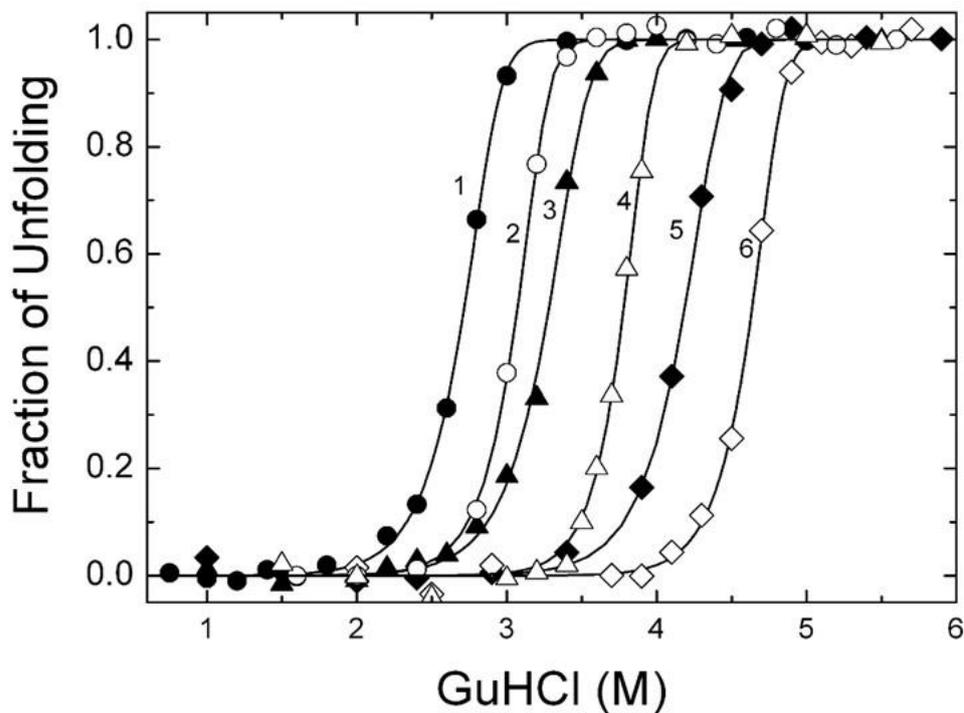


Fig. 3. Typical refolding curves of the wild-type and mutant *EcCutA1* at pH 8.0. Numbers 1-6 show the refolding plots of the wild-type, S11A, E61V, S11V/E61V, S11V/E61T/Q73V, and S11V/E61V/Q73V, respectively. These data represent the refolding points of each protein after a 1-day incubation at 37 °C as a function of the GuHCl concentration. The solid curves were obtained by fitting of the refolding data to equation 5 to obtain the unfolding  $\Delta G^0_{\text{H}_2\text{O}}$  values in Table 4.

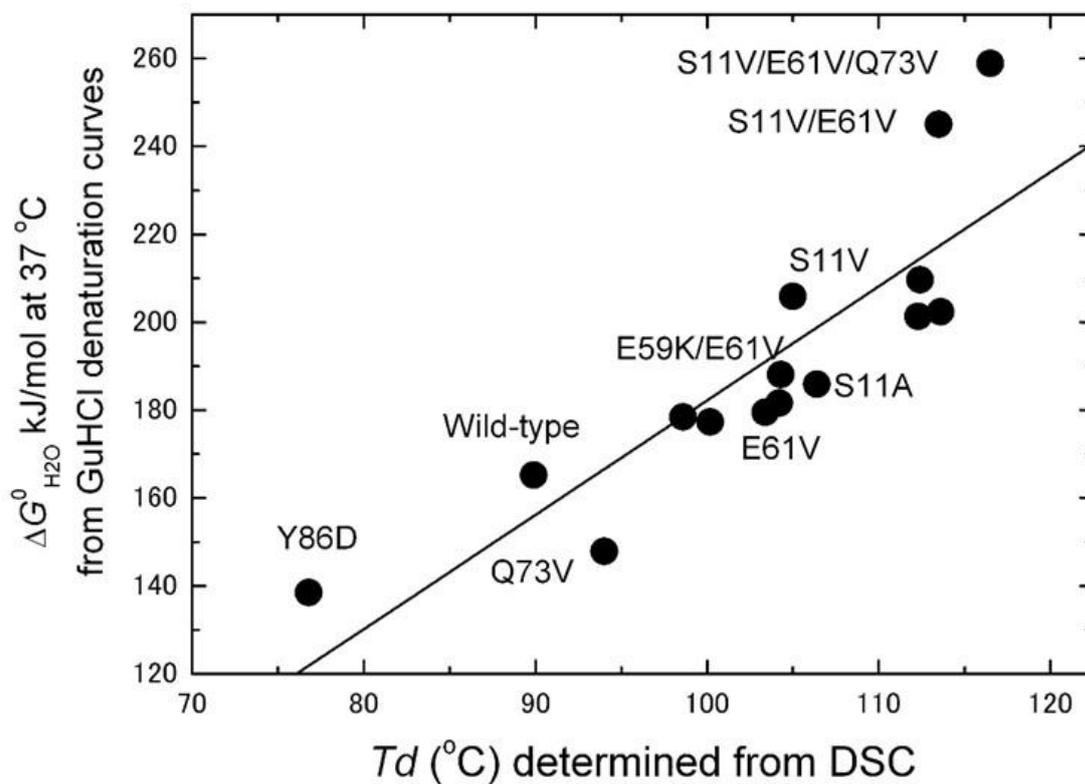


Fig. 4. The relationships between  $T_d$  determined from DSC and  $\Delta G^{\circ}_{\text{H}_2\text{O}}$  at 37  $^{\circ}\text{C}$  from the GuHCl denaturation experiments. All data were obtained from Table 4.

#### 4-5. References.

1. Ota, M., Kanaya, S., and Nishikawa, K. (1995) Desk-top analysis of the structural stability of various point mutations introduced into ribonuclease H. *J. Mol. Biol.* **248**, 733-738.
2. Ota, M. and Nishikawa, K. (1997) Assessment of pseudo-energy potentials by the best-five test: a new use of the three-dimensional profiles of proteins. *Protein Eng.* **10**, 339-351.
3. Ota, M., Isogai, Y., and Nishikawa, K. (2001) Knowledge-based potential defined for a rotamer library to design protein sequences. *Protein Eng.* **14**, 557-564.
4. Pace, C. N., Vajdos, F., Fee, G., Grimsley, G., and Gray, Y., (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411-2423.
5. Sawano, M., Yamamoto, H., Ogasahara, K., Kidokoro, S-i., Katoh, S., Ohnuma, T., Katoh, E., Yokoyama, S., and Yutani, K. (2008) Thermodynamic basis for the stabilities of three CutA1s from *Pyrococcus horikoshii*, *Thermus thermophilus*, and *Oryza sativa*, with unusually high denaturation temperatures. *Biochemistry* **47**, 721-730.
6. Backmann, J., Schafer, G., Wyns, L., and Bonisch, H. (1998) Thermodynamics and kinetics of unfolding of the thermostable trimeric adenylate kinase from the archaeon *Sulfolobus acidocaldarius*. *J. Mol. Biol.* **284**, 817-833.
7. Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30**, 1022-1025.
8. Brünger, A.T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst. D* **54**, 905-921.
9. Emsley, P. and Cowton, K. (2004) Coot: model-building tools for molecular graphics. *Acta Cryst. D* **60**, 2126-2132.
10. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291.

11. Funahashi, J., Takano, K., and Yutani, K. (2001) Are the parameters of various stabilization factors estimated from mutant human lysozymes compatible with other proteins? *Protein Eng.* **14**, 127-134.
12. Connolly, M. L. (1993) The molecular surface package. *J. Mol. Graphics* **11**, 139-141.
13. Arnesano, F., Banci, L., Benvenuti, M., Bertini, I., Calderone, V., Mangani, S., and Viezzoli, M. S. (2003) The evolutionarily conserved trimeric structure of CutA1 proteins suggests a role in signal transduction. *J. Biol. Chem.* **278**, 45999-46006.
14. Nishikawa, K. and Matsuo, Y. (1993) Development of pseudoenergy potentials for assessing protein 3-D-1-D compatibility and detecting weak homologies. *Protein Eng.* **6**, 811-820.
15. Kabsch, W. and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577-2637.
16. Kyte, J. and Doolittle, R. F., (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105-132.
17. Chou, P. Y. and Fasman, G. D. (1978) Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol.* **47**, 45-148.

## **Chapter 5: Stabilization of *Escherichia coli* CutA1 by introduction of charged residues**

### ***5-1. Introduction.***

The tertiary structures of proteins, which are vital for their physiological functions, are related to amino-acid sequences and stabilized by thermodynamic rules. To elucidate the mechanisms of protein folding and protein stabilization, it is critically important to obtain the thermodynamic parameters of protein denaturation as a function of temperature. One difficulty for studying the stabilization mechanism of proteins with denaturation temperatures above 100 °C is that the heat denaturation of proteins is usually irreversible at temperatures higher than 80°C<sup>1-8</sup>, because under these conditions proteins generally aggregate after heat denaturation. Thus, the thermodynamic features of protein stabilization at temperatures above 100 °C are not well understood. One important question in this field is whether the hydrophobic interactions that make the largest contributions to protein stability<sup>9-13</sup> still occur at temperatures above 100 °C<sup>14-15</sup>. Next, it is necessary to perform thermodynamic analyses of salt bridges in proteins that have denaturation temperatures above 100 °C, because many proteins from hyperthermophiles seem to be stabilized by an abundance of ion pairs formed by charged residues<sup>8, 16-26</sup>. Thermodynamics of protein denaturation at temperatures over 100 °C is also essentially important for the rational design of hyperthermostable proteins that would be highly useful for industrial and bio-technological processes.

*Ec*CutA1 and its mutants are not heat-reversible<sup>27</sup>. However, other CutA1 proteins,

such as *Tt*CutA1 and *Os*CutA1, which have fewer cysteine sulfhydryl (SH) groups, exhibit remarkable heat reversibility<sup>28</sup>. *Ec*CutA1 has three SH groups per subunit. Therefore, we designed an SH-free *Ec*CutA1 mutant, referred to as *Ec*CutA1\_0SH (Cys16→Ala, Cys39→Ala, Cys79→Ala), with the aim of achieving high heat reversibility.

In this study, we used the *Ec*CutA1\_0SH protein, which has excellent heat reversibility, as a template to design thermostabilized mutants. First, we constructed hydrophobic mutants from *Ec*CutA1\_0SH, which were effective in our previous work<sup>27</sup>. Then, to achieve a hyperthermostability comparable to that of *Ph*CutA1, we designed ionic mutants in which charged residues were introduced into a hydrophobic mutant (*Ec*CutA1\_0SH\_S11V/E61V) by substitution. We describe our use of these *Ec*CutA1 mutants to assess the thermodynamic characteristics of proteins at temperatures over 100 °C, in regard to both hydrophobic and ion-ion interactions.

## ***5-2. Experimental methods.***

### **Mutagenesis, expression, and purification of CutA1 mutants from *E. coli*.**

The mutagenesis, expression, and purification of CutA1 mutants from *E. coli* were performed as described<sup>27</sup> with minor modifications. The homogeneity and identity of the purified samples were assessed using SDS-PAGE. The protein concentration was estimated from the absorbance at 280 nm, assuming  $E^{1\text{cm}}_{1\%} = 14.96$ , based on the number of aromatic amino acids<sup>30</sup>.

### **Differential scanning calorimetry (DSC) experiments.**

To measure the changes in stability due to mutations, DSC was performed using a scan rate of 60 °C/h on a VP-capillary DSC platform (Microcal, USA) for temperatures up to 130 °C at pressures below 60 psi, or a Nano-DSC 6300Y microcalorimeter (TA Instruments, USA) for higher temperatures up to 160 °C at a pressure of 88 psi. Protein concentrations were around 0.6 mg/ml in a 50 mM glycine buffer at pH 9.0 containing 2 mM EDTA or a 50 mM glycine buffer at pH 2.0-3.5. All samples were dialyzed against the buffers overnight at 4°C and then filtered through a membrane with 0.22- $\mu$ m pores. The denaturation temperature ( $T_d$ ) is the temperature at which the area of the denaturation enthalpy ( $\Delta H$ ) is 0.5. The  $T_d$  and  $\Delta H$  values in this study represent the averages for at least six experiments.

To measure the heat capacity of mutant proteins in their native states, the protein concentrations were adjusted to around 2.0 mg/ml in a 50 mM glycine buffer at pH 9.0. Two different scan rates, 60 and 200°C/h, were used. Each experiment comprised six cycles of reheating to the pre-denaturation temperatures: 95 °C and 110 °C for *Ec0VV* and *Ec0VV\_6*, respectively. The partial specific volumes for the calculation of heat capacity were estimated from the amino-acid composition of each mutant protein<sup>31</sup>.

### **5-3. Results.**

#### **Hydrophobic mutants of *EcCutA1* with no SH group.**

We constructed hydrophobic mutants with no SH groups (*EcCutA1\_0SH\_S11V*, *EcCutA1\_0SH\_E61V*, *EcCutA1\_0SH\_S11V/E61V*), which we expected to increase stability<sup>27</sup>. Fig. 1A shows typical DSC curves of *EcCutA1\_0SH* and its hydrophobic mutant without SH groups at pH 9.0. As shown in Fig. 1B, the reheating curve (second scan) of *EcCutA1\_0SH* agrees completely with the first scan. The other two proteins also exhibited good reproducibility. These results indicate that the removal of SH groups facilitates excellent reversibility of heat denaturation under these conditions and that we can reliably determine the denaturation enthalpies of these proteins. The denaturation temperature ( $T_d$ ) of *EcCutA1\_0SH* decreased by 4.3 °C, relative to that (89.9 °C) of *EcCutA1* with SH groups, whereas those of *EcCutA1\_0SH\_S11V* and *EcCutA1\_0SH\_E61V* were 103 and 101 °C, respectively, which were remarkably improved relative to the template. Furthermore, the  $T_d$  of a double mutant, *EcCutA1\_0SH\_S11V/E61V*, was 113 °C, which is 28 °C higher than that of the template (Table 1). Hereafter, *EcCutA1\_0SH\_S11V/E61V* is abbreviated as *Ec0VV*. These changes in stability due to the hydrophobic mutations were comparable to those observed in mutant proteins with SH groups<sup>27</sup>.

#### **Ionic mutants of *EcCutA1\_0SH\_S11V/E61V* (*Ec0VV*).**

To examine the thermodynamic parameters of stabilization by ion-ion interaction at temperatures over 100 °C, we constructed several mutant proteins containing substitutions with charged residues, using *Ec0VV* as a template. Ionic mutants, whose denaturation temperatures are improved and whose DSC curves are suitable for thermodynamic analysis, were selected from our pool of stock mutants (Table 1). Typical

DSC curves and reversibility curves are shown in Fig. 2 and Fig. S2, respectively. Although the  $T_d$  of a double mutant, *Ec0VV\_T17K/S48D*, was lower than that of the template, it was selected because the  $T_d$  was over 100 °C and higher than those of the original single mutants (*Ec0VV\_T17K* and *Ec0VV\_S48D*) (Table 1). In the case of *Ec0VV\_A39D/S48K/H72K/S82K/Q87K/T88R*, which is abbreviated as *Ec0VV\_6*, the DSC curve was suitable for analysis (Fig. 2), but the reversibility curve could not be properly obtained due to certain side reactions that occurred at high temperatures. The  $T_d$  of *Ec0VV\_6* was  $136.8 \pm 0.9^\circ\text{C}$ , improved by  $23.6^\circ\text{C}$  with the introduction of six charged residues.

In acidic pH, negatively charged residues of a protein should be protonated, leading to a decrease in conformational stability. In the case of CutA1 from *P. horikoshii*, which is stabilized by many ionic interactions, the  $T_d$  of  $148.5^\circ\text{C}$  at pH 7.0 is drastically reduced to  $75.6^\circ\text{C}$  at pH 2.5, whereas the  $T_d$  of CutA1 from *T. thermophilus* changes from  $112.8^\circ\text{C}$  at pH 7.0 to  $86.6^\circ\text{C}$  at pH 2.5<sup>28</sup>. To confirm the stabilization resulted from ionic interactions, the stabilities of ionic mutants were examined under acidic conditions at pH 2-3. The  $T_d$  values of the ionic mutants monotonically decreased as the pH was lowered, reaching a constant minimum at pH 2-2.5 (Table S1). We plotted the  $T_d$  shift ( $T_d$  value at pH 9.0 vs. pH 2.0-2.5) versus the  $T_d$  value at pH 9.0 for several ionic mutants (closed circles in Fig. S3). Clearly, the  $T_d$  shift became greater as  $T_d$  increased. These results suggest that the electrostatic interactions dominates the thermo-stabilization of the ionic mutant proteins.

### **Temperature dependence of denaturation enthalpy at higher temperatures.**

The denaturation heat capacity ( $\Delta C_p$ ) is generally assumed to be constant at temperatures below 80 °C<sup>32</sup>, but it gradually decreases at higher temperatures<sup>14</sup>. Therefore, it is important to elucidate the temperature function of  $\Delta C_p$  at higher temperatures. To this end, we measured the  $C_p$  values of *Ec0VV* in the native state by DSC at temperatures up to 95 °C (Y2 of Fig. S4A). Unfortunately, the temperature dependence of  $C_p$  values in the denatured state could not be determined experimentally due to the high reversibility of denatured *Ec0VV*. Alternatively, assuming that the heat-capacity contribution of amino-acid groups is additive, the heat capacity of proteins in the denatured state can be calculated from their amino-acid composition<sup>33</sup>. Y1 in Fig. S4A also shows the temperature function of the heat capacity of *Ec0VV* in the denatured state, which was estimated from its amino-acid composition using the parameters in Table II of Makhatadze and Privalov<sup>34</sup>. Next, we were able to estimate the temperature function of the denaturation heat capacity ( $\Delta C_p$ ) for *Ec0VV* from these native and denatured  $C_p$  values (Y3 of Fig. S4A). The temperature function obtained of  $\Delta C_p$  can be expanded around  $T_d$  as a second-order polynomial:

$$\Delta C_p(T) = A + B(T - T_d) + C(T - T_d)^2 \quad (1)$$

Then, the temperature functions of denaturation enthalpy ( $\Delta H$ ) and denaturation entropy ( $\Delta S$ ) can be calculated by the following equations (2) and (3), respectively.

$$\begin{aligned} \Delta H(T) &= \Delta H(T_d) + \int_{T_d}^T \Delta C_p(T') dT' \\ &= \Delta H(T_d) + A(T - T_d) + \frac{B}{2}(T - T_d)^2 + \frac{C}{3}(T - T_d)^3 \end{aligned} \quad (2)$$

$$\begin{aligned} \Delta S(T) &= \Delta S(T_d) + \int_{T_d}^T \frac{\Delta C_p(T')}{T'} dT' \\ &= \Delta S(T_d) + (A - BT_d + CT_d^2) \ln \frac{T}{T_d} + (B - CT_d)(T - T_d) + \frac{C}{2}(T - T_d)^2 \end{aligned} \quad (3)$$

Fig. 3A shows the temperature function of  $\Delta H$  for *Ec0VV*. The  $\Delta H$  values of *Ec0VV*

were higher than those at each denaturation temperature of other proteins (*EcCutA1\_0SH*, *EcCutA1\_0SH\_S11V*, and *EcCutA1\_0SH\_E61V*). If we assume that the temperature function of  $\Delta C_p$  is not largely affected by the constitution of the protein, this observation indicates that stabilization of hydrophobic mutants at residue positions 11 and 61 is mainly caused by enthalpic effects.

The temperature function of  $\Delta C_p$  for *Ec0VV\_6* was also determined using the native  $C_p$  values (Fig. S4B), which were directly measured up to 110 °C. Fig. 3B shows the temperature function of  $\Delta H$  for *Ec0VV\_6* and the denaturation enthalpy values at the denaturation temperatures of several ionic *Ec0VV* mutants. This figure indicates that the  $\Delta H$  value of *Ec0VV* is similar to those of *Ec0VV\_S110R* and *Ec0VV\_6* at each denaturation temperature, but remarkably higher than those of other mutants derived from *Ec0VV* by the further addition of charged residues.

The thermodynamic parameters of denaturation for *EcCutA1\_0SH* mutants at the denaturation temperature (113.2 °C) of *Ec0VV* are listed in Table 2. The  $\Delta G$  values, estimated using the  $\Delta C_p$  temperature function obtained from *Ec0VV*, agreed well with those from *Ec0VV\_6*, around the denaturation temperature of *Ec0VV*. Fig. 4 also shows the temperature functions of  $\Delta G$ ,  $\Delta H$ , and  $T\Delta S$  for *Ec0VV* and *Ec0VV\_6* over a larger temperature range (between 280 and 420K), indicating that  $\Delta G$  values of *Ec0VV\_6* are positive over a broad range of temperatures.

#### **5-4. Discussion.**

Forty years ago, Privalov *et al.*<sup>14,32</sup> developed highly qualified adiabatic differential

micro-calorimeters for determining the thermodynamic parameters of protein denaturation. They reported that specific characters of amino-acid residues disappear during protein unfolding near 110 °C, and that all the observed entropy originates from the increase in conformational freedom of the polypeptide upon unfolding, because the specific  $\Delta H$  and  $\Delta S$  of unfolding for several proteins intersect at a single point near 110 °C. On the other hand, the thermodynamics of transfer of hydrocarbons to water provides a model for the temperature dependence of the hydrophobic interaction in protein folding. Baldwin<sup>15</sup> examined the solution thermodynamics of several liquid hydrocarbons in water. He found that the extrapolated temperatures at which the transfer  $\Delta S$  reaches zero, around 112.8 °C, were similar for six hydrocarbons. The extrapolated temperature of the transfer  $\Delta H$  is 22.0 °C. This means that at 113 °C, the hydrophobic interaction changes from being entropy-driven at 22 °C to being enthalpy-driven at 113 °C, and the contribution of water to the entropy of protein unfolding (hydrophobic hydration) is removed<sup>15</sup>. Regarding these estimations, the heat capacity change ( $\Delta C_p$ ) is assumed to be constant against temperature. Later, Makhatadze and Privalov<sup>34, 35</sup> reported that the temperature at which  $\Delta S$  is zero approaches 145 °C when the decreasing nature of  $\Delta C_p$  against temperature is taken into account, because the  $\Delta C_p$  value of hydrocarbon hydration decreases with increasing temperature.

In this study, the heat capacities of the native states of *Ec0VV* and *Ec0VV\_6* could be directly measured by DSC up to 95 and 110 °C, respectively, and the temperature functions of their  $\Delta C_p$  values were estimated as shown in Fig. S4A and Fig. S4B, respectively.

### **Hydrophobic effects strongly contribute to $\Delta H$ at temperatures around 100 °C.**

The temperature function of  $\Delta H$  of *Ec0VV* is depicted in Fig. 3A. Using the same temperature function of  $\Delta C_p$ , the  $\Delta H$  values of *Ec0VV*, *EcCutA1\_0SH\_S11V*, *EcCutA1\_0SH\_E61V*, and *EcCutA1\_0SH* were estimated to be 1569, 1396, 1340, and 1175 kJ/mol at 113.2°C, respectively (Table 2). The increase in  $\Delta H$  (394 kJ/mol) of *Ec0VV* (*EcCutA1\_0SH\_S11V/E61V*) agrees well with the sum (386 kJ/mol) of the increases in  $\Delta H$  of *EcCutA1\_0SH\_S11V* (221 kJ/mol) and *EcCutA1\_0SH\_E61V* (165 kJ/mol). On the other hand, the  $T\Delta S$  values of *Ec0VV* (the red curve in Fig. 3A) were larger than the  $\Delta H$  value ( $=T\Delta S$  at  $T_d$ ) at each  $T_d$  of *EcCutA1\_0SH\_S11V*, *EcCutA1\_0SH\_E61V*, and *EcCutA1\_0SH*, indicating that *Ec0VV* is entropically unfavorable compared with other proteins. That is, *Ec0VV*, which contains hydrophobic substitutions for hydrophilic residues in the interior of a molecule, is mainly stabilized by the enthalpic gain upon substitutions, but is partly destabilized by the entropic loss, probably due to the disruption of the hydrophilic solvation in the denatured state upon substitutions. These results at high temperatures around 100 °C are contrary to the well-accepted belief that the entropic gain from hydrophobic solvation can account for the stabilization effect of hydrophobic substitutions at lower temperatures<sup>15</sup>. Whereas, the estimations from the hydration of amino acids<sup>35</sup> are generally consistent with our results. This is the first experimental evidence pertaining to the hydrophobic effects on protein stability, which is obtained by direct measurement at temperatures around 100 °C.

***Ec0VV\_6* substituted with six charged residues is stabilized by both enthalpic and entropic effects around 137 °C.**

Below 100 °C, the ion-ion interaction (salt bridge) is driven entirely by entropic effects due to the release of water strongly bound to the ions of charged residues<sup>36-38</sup>. The proteins substituted with single charged residues, *Ec0VV\_H72K*, *Ec0VV\_S82K*, *Ec0VV\_Q87K*, and *Ec0VV\_T88R*, are stabilized by electrostatic interactions (Fig. S3). The  $\Delta H$  values of all of these proteins were drastically decreased relative to  $\Delta H$  at each corresponding temperature on the temperature function of *Ec0VV\_6* (Fig. 3B). Because the changes in  $\Delta H$  upon these mutations are unfavorable for folding, the observed improvements in stability were caused by entropic effects due to the release of water at the charged residues that we introduced (electrostatic solvation). The thermodynamic analyses also clearly confirmed stabilization resulted from entropic effects (Table 2). The other single mutants, *Ec0VV\_A39D* and *Ec0VV\_S48K*, whose mutation sites were located in the interior of the molecule (Table 3), demonstrated drastically decreased  $T_d$  and  $\Delta H$  values. However, a double mutant containing both of these single mutants, *Ec0VV\_A39D/S48K*, in which Asp39 forms a strong salt bridge with Lys48, was stabilized by an increase in  $\Delta H$  relative to the single mutants (Fig. 3B), probably due to Coulomb's force resulting from salt bridge formation. A similar result was obtained in the other double mutant, *Ec0VV\_T17K/S48D* (Fig. 3B). The double mutant *Ec0VV\_Q87K/T88R* was also stabilized by an increase in  $\Delta H$  relative to *Ec0VV\_Q87K* and *Ec0VV\_T88R*. Because this double mutant does not have additional salt bridges, two individual thermo-stabilizations might work synergistically at around 120 °C to promote the desolvation of the ionic residues introduced, thereby reducing both the enthalpic loss and the entropic gain that are mutually attributed to the electrostatic solvation. In addition, the increase in  $\Delta H$  in this double-ion mutant might have been caused by a hydrophobic interaction due to the alkyl groups of Lys87 and Arg88.

*Ec0VV\_6*, with six additional charged residues, was stabilized by an increase in the  $\Delta H$  value relative to each of the six original ionic mutants (Fig. 3B). The increase in  $\Delta H$  might indicate that hydrophobic effects due to the alkyl groups of Lys or Arg and Coulomb's force still function effectively at these high temperatures. According to Coulomb's law, the strength of an electrical interaction is inversely proportional to the dielectric constant. The dielectric constant of water drops from 80 at 0 °C to 55 at 100 °C<sup>39</sup>. Furthermore, Elcock<sup>40</sup> found that increasing temperature decreases the electrostatic desolvation penalty incurred in forming a salt bridge, leading to an increase in salt bridge stabilization from his continuum solvation model<sup>40</sup>. The increase in  $\Delta H$  of *Ec0VV\_6* might result mainly from the high degree of desolvation of the ionic residues in the denatured state at around 137 °C, in addition to the other effects described above.

The temperature dependence of  $\Delta H$  for *Ec0VV\_6* has an intersection near  $T_d$  (113°C) of *Ec0VV*, as shown in Fig. 4, suggesting that the contribution of  $\Delta H$  to the stability of *Ec0VV\_6* with additional 6 charged residues becomes favorable in the temperature region above 113 °C, compared with those of *Ec0VV*. Furthermore, Fig. 4 shows that the increase in  $\Delta G$  of *Ec0VV\_6* results largely from the decrease in  $\Delta S$  of *Ec0VV\_6* when compared with the template *Ec0VV* at temperatures below 113 °C. That is, the stabilization due to ionic mutations results mainly from both the enthalpic gain from ion-ion interactions in the native state and the entropic gain from the water release of ionic residues in the denatured state at temperatures over 113 °C.

A mutant, *Ec0VV\_S110R*, whose substitution position is located in the C-terminus of  $\alpha$ -helix and almost buried (Table 3), was stabilized by an increase in  $\Delta H$  (Fig. 3B). In this case, due to local conformations, the decrease in  $\Delta H$  due to water release might be

suppressed by the effects of other stabilizing factors.

Table 1. Denaturation enthalpy of examined *EcCutA1\_0SH* mutants at denaturation temperatures.

	$T_d$ (°C)*			$\Delta H$ (kJ mol <sup>-1</sup> )*		
		±			±	
<i>EcCutA1_0SH</i>	85.6	±	0.3	870	±	21
<i>EcCutA1_0SH_S11V</i>	102.7	±	0.3	1302	±	21
<i>EcCutA1_0SH_E61V</i>	101.0	±	0.2	1228	±	15
<i>Ec0VV**</i>	113.2	±	0.2	1569	±	15
<i>Ec0VV_T17K</i>	107.2	±	0.3	1315	±	15
<i>Ec0VV_S48D</i>	105.4	±	0.2	1242	±	14
<i>Ec0VV_T17K/S48D</i>	112.0	±	0.1	1387	±	22
<i>Ec0VV_A39D</i>	105.4	±	0.5	1197	±	31
<i>Ec0VV_S48K</i>	112.2	±	0.9	1173	±	35
<i>Ec0VV_A39D/S48K</i>	118.3	±	0.7	1410	±	33
<i>Ec0VV_H72K</i>	118.4	±	0.4	1521	±	34
<i>Ec0VV_S82K</i>	116.9	±	0.5	1479	±	22
<i>Ec0VV_S82R</i>	117.1	±	0.5	1446	±	24
<i>Ec0VV_T88R</i>	117.6	±	0.6	1501	±	58
<i>Ec0VV_Q87K</i>	116.8	±	0.5	1424	±	13
<i>Ec0VV_Q87K/T88R</i>	122.4	±	0.6	1582	±	56
<i>Ec0VV_S110R</i>	117.3	±	0.4	1637	±	35
<i>Ec0VV_6***</i>	136.8	±	0.9	1739	±	59

\*Average value and its standard deviation of at least 6 data.

\*\**Ec0VV* represents *EcCutA1\_0SH\_S11V/E61V* mutant.

\*\*\**Ec0VV\_6* represents *Ec0VV\_A39D/S48K/H72K/S82K/Q87K/T88R* mutant.

Table 2. Thermodynamic parameters of denaturation for *EcCutA1\_0SH* mutants at the denaturation temperature (113.2 °C) of *Ec0VV*. The unit is kJ mol<sup>-1</sup>.

	$\Delta H$		$T\Delta S$		$\Delta G (= \Delta H - T\Delta S)$		$\Delta\Delta H$		$T\Delta\Delta S$	
	a*	b*	a	b	a	b	a	b	a	b
<i>EcCutA1_0SH</i>	1175	1254	1255	1336	-80	-82	-394	-315	-314	-233
<i>EcCutA1_0SH_S11V</i>	1396	1440	1434	1478	-38	-38	-173	-129	-135	-91
<i>EcCutA1_0SH_E61V</i>	1340	1377	1382	1419	-42	-42	-229	-192	-187	-150
<i>Ec0VV</i>	1569	1569	1569	1569	0	0	0	0	0	0
<i>Ec0VV_A39D</i>	1265	1289	1290	1315	-26	-26	-304	-280	-279	-254
<i>Ec0VV_S48K</i>	1181	1184	1184	1187	-3	-3	-388	-385	-385	-382
<i>Ec0VV_A39D/S48K</i>	1375	1358	1356	1340	18	18	-194	-211	-213	-229
<i>Ec0VV_H72K</i>	1485	1469	1465	1449	20	20	-84	-100	-104	-120
<i>Ec0VV_S82K</i>	1453	1442	1439	1428	14	14	-116	-127	-130	-141
<i>Ec0VV_T88R</i>	1470	1456	1453	1440	17	17	-99	-113	-116	-129
<i>Ec0VV_Q87K</i>	1398	1387	1385	1374	13	13	-171	-182	-184	-195
<i>Ec0VV_Q87K/T88R</i>	1524	1495	1488	1459	36	36	-45	-74	-81	-110
<i>Ec0VV_S110R</i>	1608	1595	1590	1578	17	17	39	26	21	9
<i>Ec0VV_6</i>	1638	1560	1540	1465	98	96	69	-9	-29	-104

\*a and b represent the calculated results using the temperature function of  $\Delta C_p$  obtained from *Ec0VV* and *Ec0VV\_6*, respectively.

Table 3. Burial rates of target residues of *Ec0VV* mutants. The burial rates (%) were estimated from the average of ASA values for nine structures during 40 ns MD at 300 K.

Mutants	Residues	All atoms		Charged atoms		location
<i>Ec0VV_6</i>	Asp39	100	± 1	99	± 2	beta 2
	Lys48	96	± 3	93	± 6	beta 2
	Lys72	80	± 5	55	± 18	alpha 2, N terminal
	Lys82	10	± 6	13	± 8	alpha 2, C terminal
	Lys87	10	± 14	-2	± 21	loop
	Arg88	55	± 20	40	± 34	loop
	Val11	98	± 1			beta 1
	Val61	97	± 3			beta 3
<i>Ec0VV_S110R</i>	Arg110	84	± 10	66	± 20	alpha 3, C-terminal

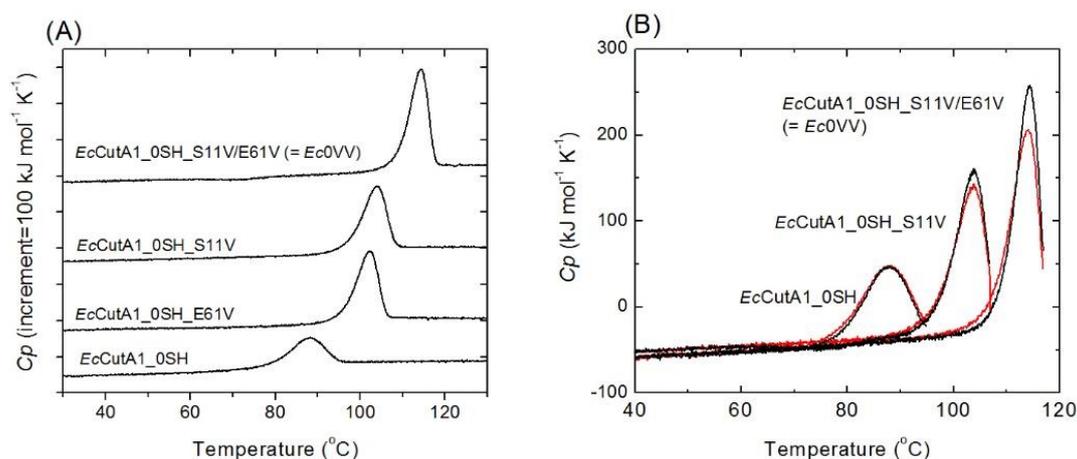


Fig. 1. DSC curves of hydrophobic *EcCutA1* mutants with no SH group at pH 9.0. (A) Typical DSC curves of four mutants: *EcCutA1\_0SH*, *EcCutA1\_0SH\_S11V*, *EcCutA1\_0SH\_E61V*, and *EcCutA1\_0SH\_S11V/E61V*. Scan rates were 60 °C/h. (B) Reversibility of heat denaturation of *EcCutA1\_0SH*, *EcCutA1\_0SH\_S11V*, and *EcCutA1\_0SH\_S11V/E61V*. The red curves of three proteins are the second runs of DSC, just after the cooling step of the first run (the black curves). Scan rates of both curves were 60 °C/h.

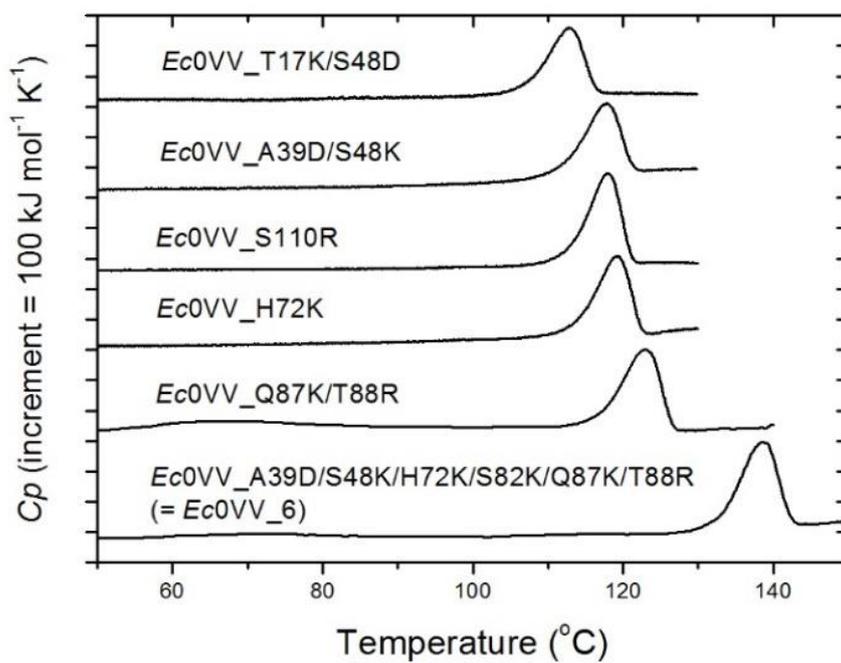


Fig. 2. Typical DSC curves of six ionic *EcCutA1* mutants from *Ec0VV* at pH 9.0. The six mutants are *Ec0VV\_T17K/S48D*, *Ec0VV\_A39D/S48K*, *Ec0VV\_S110R*, *Ec0VV\_H72K*, *Ec0VV\_Q87K/T88R*, and *Ec0VV\_6*. Scan rates were 60 °C/h.

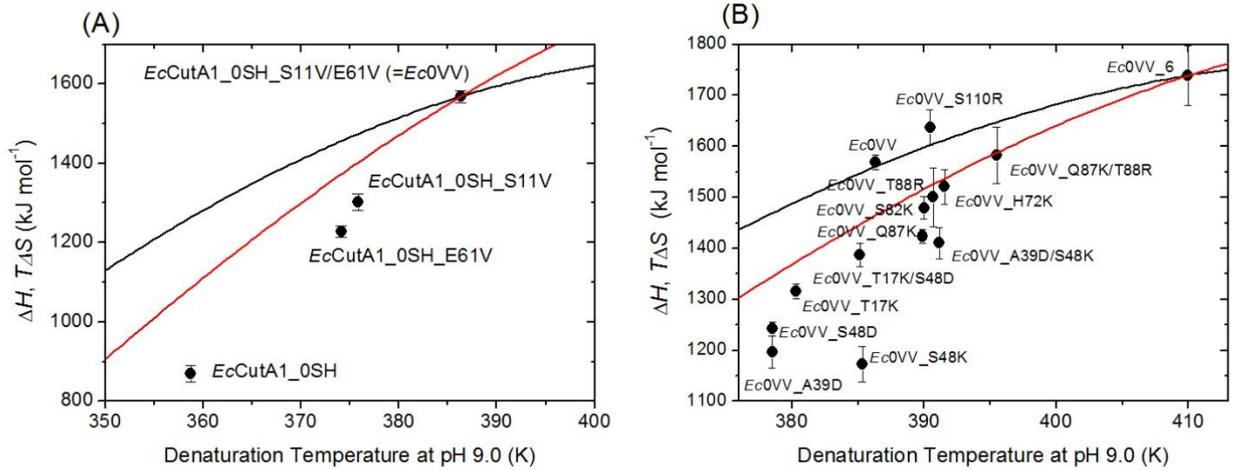


Fig. 3. Temperature dependence of  $\Delta H$  for hydrophobic *EcCutA1* mutants at pH 9.0. (A)  $\Delta H$  values (closed circles) of denaturation for *EcCutA1* mutants come from Table 1. The black curves represent the temperature function of  $\Delta H$  upon denaturation, using the temperature function of  $\Delta C_p$  for *Ec0VV* obtained from Y3 of Fig. S4B. The red curve represents  $T\Delta S$  of *Ec0VV*. In the case of *Ec0VV*, the parameters A, B, and C of  $\Delta C_p$  (in  $\text{kJ mol}^{-1} \text{K}^{-1}$ ) in equation (1) were calculated to be 7.61029,  $-0.26614$ , and  $-8.4434 \times 10^{-4}$ , respectively. (B) Temperature dependence of  $\Delta H$  for ionic *Ec0VV* mutants at pH 9.0.  $\Delta H$  values (closed circles) of denaturation for *EcCutA1* mutants come from Table 1. Black curves represent the temperature function of  $\Delta H$  upon denaturation, using the temperature function of  $\Delta C_p$  for *Ec0VV\_6* obtained from Y3 of Fig S4B. The red curve represents  $T\Delta S$  of *Ec0VV\_6*. In the case of *Ec0VV\_6*, the parameters A, B, and C of  $\Delta C_p$  (in  $\text{kJ mol}^{-1} \text{K}^{-1}$ ) in equation (1) were calculated to be 4.22658,  $-0.29835$ , and  $-10.0757 \times 10^{-4}$ , respectively.

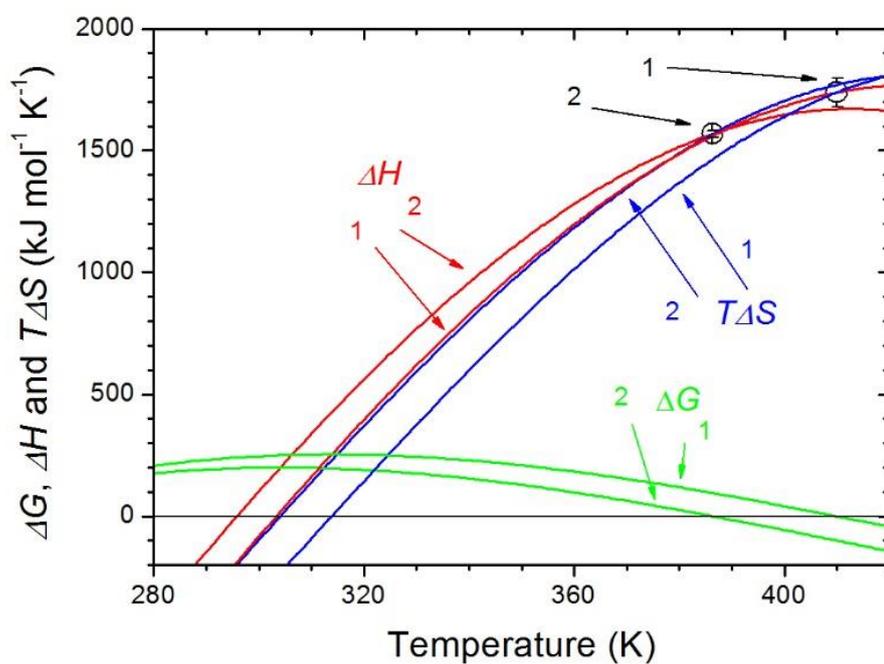


Fig. 4. Temperature functions of  $\Delta G$ ,  $\Delta H$ , and  $T\Delta S$  for *Ec0VV*, and *Ec0VV\_6* between 280 and 420 K. Temperature functions of  $\Delta H$  and  $\Delta S$  were obtained using equations (2) and (3), respectively, in which each temperature function of  $\Delta C_p$  was used for the calculation of *Ec0VV* and *Ec0VV\_6*. The green, red, and blue curves represent values of  $\Delta G (= \Delta H - T\Delta S)$ ,  $\Delta H$ , and  $T\Delta S$ , respectively. Numbers 1 and 2 represent *Ec0VV\_6* and *Ec0VV*, respectively. Open circles with error bars show the  $\Delta H$  value of each protein at the denaturation temperature, as indicated in the figure.

5-5. Supporting information.

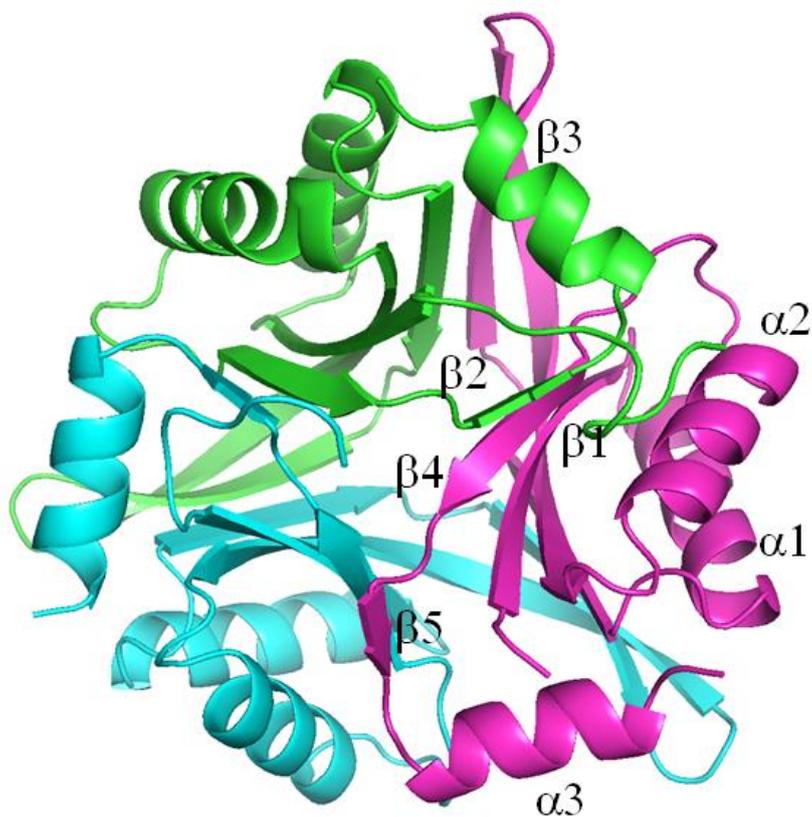


Fig. S1. The trimer crystal structure of *EcCutA1\_0SH* (PDB ID 4Y65). Different colors represent different chains.  $\alpha$  and  $\beta$  represent  $\alpha$  helix and  $\beta$  strand, respectively. Three N terminal residues of B subunit (cyan) and eight N terminal residues of C subunit (magenta) are missing in the crystal structure.

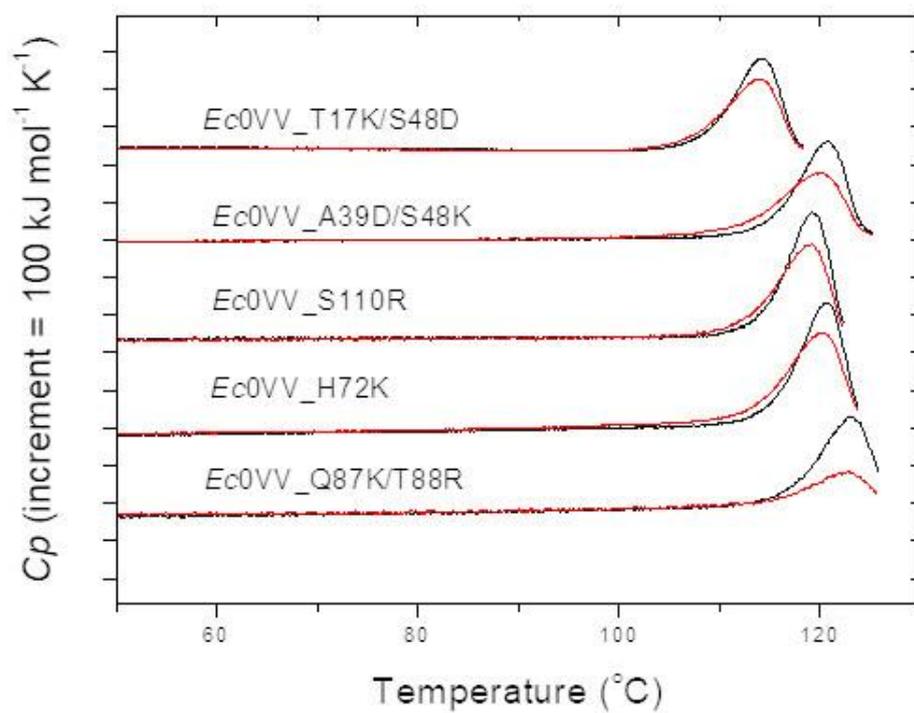


Fig. S2. Reversibility of the DSC curves of ionic mutants from *Ec0VV* at pH 9.0. Red curves are the second runs of DSC just after cooling of the first run (black curves). Scan rates of both curves were  $60 \text{ }^{\circ}\text{C/h}$ .

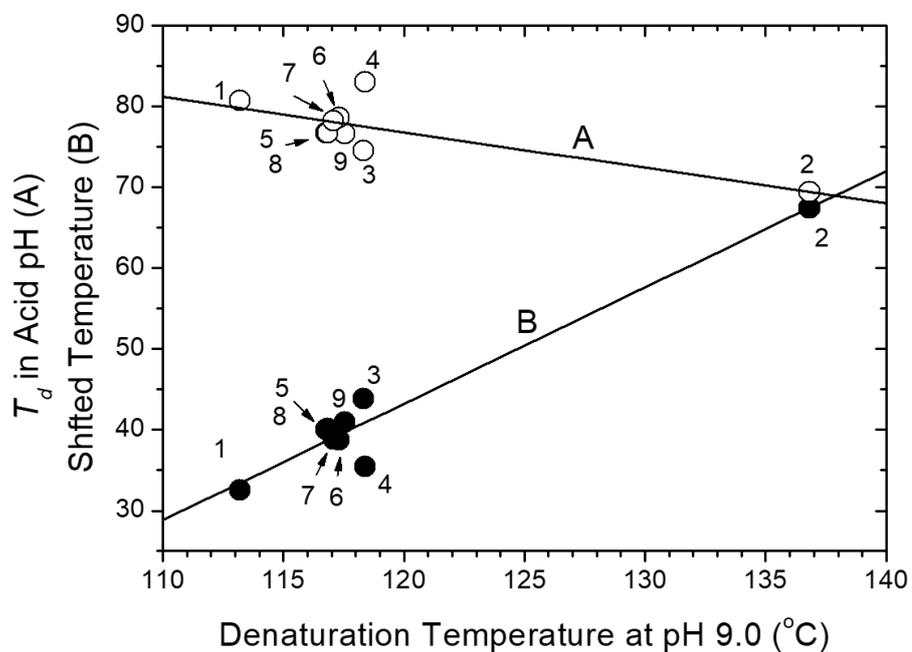


Fig. S3. Comparison between  $T_d$  values of ion mutants of *Ec0VV* at pH 9.0 and pH 2.0-2.5. Open circles represent average  $T_d$  values at pH 2.0, 2.25, and 2.5 for ionic mutants of *Ec0VV* (Table S1). Closed circles represent shifted temperatures, which are the differences between the  $T_d$  values at pH 9.0 (Table 1) and the average  $T_d$  values at pH 2.0-2.5 for the ionic mutants of *Ec0VV* (Table S1). Lines A and B represent linear regressions for open and closed circles, respectively. Numbers 1-9 represent mutant proteins of *Ec0VV*, *Ec0VV*\_6, *Ec0VV*\_A39D/S48K, *Ec0VV*\_H72K, *Ec0VV*\_S82K, *Ec0VV*\_S110R, *Ec0VV*\_S82R, *Ec0VV*\_Q87K, and *Ec0VV*\_T88R.

(A)

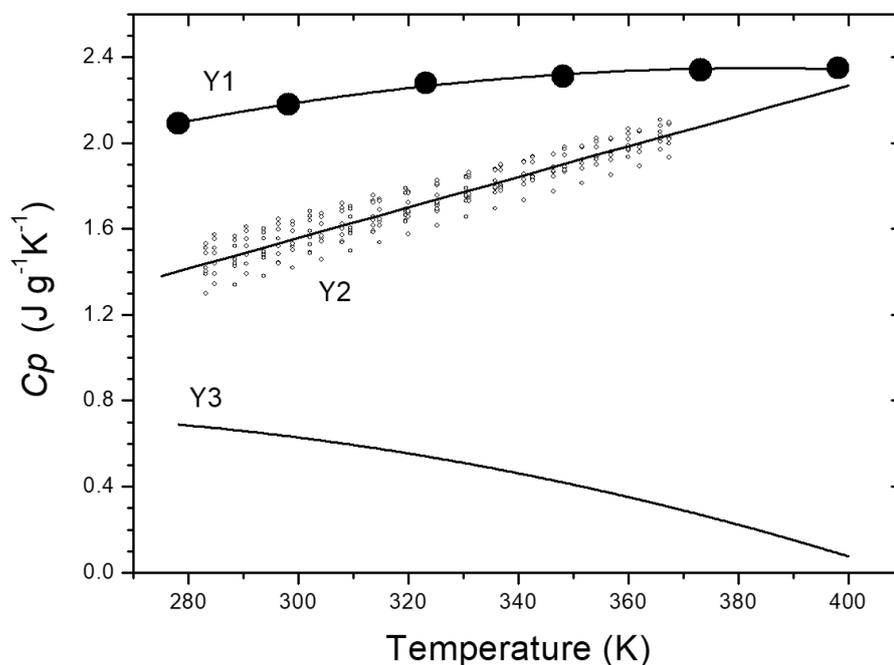
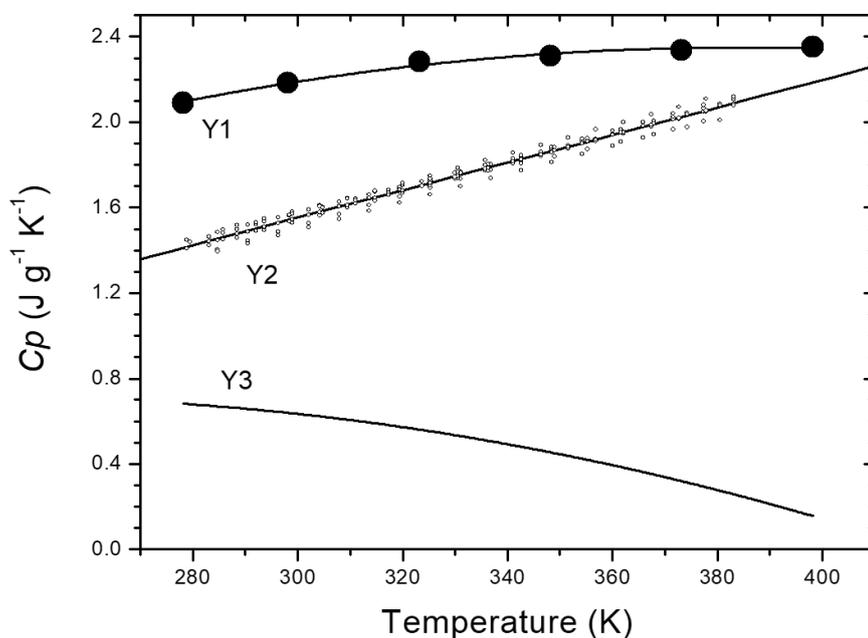


Fig. S4. (A) Temperature dependence of  $C_p$  for *Ec0VV* in the native and denatured states. Y1 represents the temperature dependence of  $C_p$  ( $\text{Jg}^{-1}\text{K}^{-1}$ ) for *Ec0VV* in the denatured state. Black closed circles in Y1 represent the heat capacity of *Ec0VV* in the denatured state, estimated from the amino-acid composition using the parameters in Table II of Makhatadze and Privalov<sup>34</sup>. The Y1 curve is the result of fitting a secondary expression to the data. Curve  $Y1 = -1.01674 + 1.753 \times 10^{-2}T - 2.282 \times 10^{-5}T^2$ , where  $T$  is the temperature in Kelvin. Y2 represents the heat capacity in the native state; small circles represent experimental data. Each experiment comprised the six times cycles of reheating to the pre-denaturation temperature. The data points (small circles) show all data in 3 time experiments of the liner regression obtained from each experiment. The liner line in Y2 is liner regression of all data shown in the figure. Line  $Y2 = -0.56817 + 0.709 \times 10^{-2}T$ . Y3 represents the temperature function of denaturation heat capacity,  $\Delta C_p$ , between the native and denatured states, i.e.,  $Y3 = Y1 - Y2$ . Thus,  $Y3 = -0.44857 + 1.044 \times 10^{-2}T - 2.282 \times 10^{-5}T^2$ .

(B)



(B) Temperature dependence of  $C_p$  for *Ec0VV\_6* in the native and denatured states. Y1 represents the temperature dependence of  $C_p$  ( $\text{Jg}^{-1}\text{K}^{-1}$ ) for *Ec0VV\_6* in the denatured state. Black closed circles and a curve in Y1 show the heat capacity of *Ec0VV\_6* in the denatured state, estimated as shown in the legend of Fig. S4A. Curve  $Y1 = -0.98373 + 1.735 \times 10^{-2}T - 2.25855 \times 10^{-5}T^2$ , where  $T$  is a temperature in Kelvin. Y2 represents the heat capacity in the native state, where small circles are experimental data. Each experiment comprised the six times cycles of reheating to the pre-denaturation temperature. The linear line in Y2 is linear regression of all data shown in the figure. Line  $Y2 = -0.38148 + 0.645 \times 10^{-2}T$ . Y3 represents  $\Delta C_p$ , between the native and denatured states, i.e.,  $Y3 = Y1 - Y2$ . Thus,  $Y3 = -0.60225 + 1.09 \times 10^{-2}T - 2.2586 \times 10^{-5}T^2$ .

Table S1. Denaturation temperature of *Ec0VV* mutants in the acidic region.

Mutants	pH						
	3.0	2.9	2.7	2.5	2.25	2.0	2.5-2.0*
<i>Ec0VV</i>	85.9	85.0	82.6	80.6	80.4	81.0	80.7
<i>Ec0VV_A39D/S48K</i>	85.6	82.0	77.3	75.1	73.6	74.9	74.5
<i>Ec0VVE_H72K</i>	88.7	87.2	84.5	83.4	82.5	83.2	83.0
<i>Ec0VV_S82K</i>	84.3	81.8	78.6	76.7	76.9	76.7	76.7
<i>Ec0VV_S82R</i>	83.8	82.2	79.2	78.1	78.6	78.0	78.2
<i>Ec0VV_Q87K</i>	83.2	81.1	78.2	76.6	76.7	76.9	76.7
<i>Ec0VV_T88R</i>	83.2	80.8	77.8	76.5	76.8	76.4	76.6
<i>Ec0VV_S110R</i>	84.2	82.4	80.1	78.5	78.3	79.0	78.6
<i>Ec0VV_6</i>	85.3	80.8	73.3	70.1	69.6	68.5	69.4

Each data represents average of two data.

The unit of data is °C.

\*Average value of pH 2.5, 2.25, and 2.0.

## 5-6. References.

1. Cavagnero, S., Debe, D. A., Zhou, Z. H., Adams, M. W., and Chan, S. I. (1998) Kinetic role of electrostatic interactions in the unfolding of hyperthermophilic and mesophilic rubredoxins. *Biochemistry* **37**, 3369-3376.
2. Pothekin, S.A., Ogasahara, K., and Yutani, K. (2000) Transition state of heat denaturation of methionine aminopeptidase from a hyperthermophile. *J. Therm. Anal. Calorim.* **62**, 111-122.
3. Wrba, A., Schweiger, A., Schultes, V., Jaenicke, R., and Závodszky, P. (1990) Extremely thermostable D-glyceraldehyde-3-phosphate dehydrogenase from the eubacterium *Thermotoga maritima*. *Biochemistry* **29**, 7584-7592.
4. Klump, H., Di Ruggiero, J., Kessel, M., Park, J. B., Adams, M. W., and Robb, F. T. (1992) Glutamate dehydrogenase from the hyperthermophile *Pyrococcus furiosus*. Thermal denaturation and activation. *J. Biol. Chem.* **267**, 22681-22685.
5. McAfee, J.G., Edmondson, S.P., Zegar, I., and Shriver, J.W. (1996) Equilibrium DNA binding of Sac7d protein from the hyperthermophile *Sulfolobus acidocaldarius*: fluorescence and circular dichroism studies. *Biochemistry* **35**, 4034-4045.
6. Pfeil, W., Gesierich, U., Kleemann, G. R., and Sterner, R. (1997) Ferredoxin from the hyperthermophile *Thermotoga maritima* is stable beyond the boiling point of water. *J. Mol. Biol.* **272**, 591-596.
7. Pappenberger, G., Schurig, H., and Jaenicke, R. (1997) Disruption of an ionic network leads to accelerated thermal denaturation of D-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima*. *J. Mol. Biol.* **274**, 676-683.
8. Ogasahara, K., Lapshina, Sakai, M., Izu, Y., Tsunasawa, S., Kato, I., and Yutani, K. (1998) Electrostatic stabilization in methionine aminopeptidase from hyperthermophile *Pyrococcus furiosus*. *Biochemistry* **37**, 5939-5946.
9. Yutani, K., Ogasahara, K., Sugino, Y., and Matsushiro, A. (1977) Effect of a single amino acid substitution on stability of conformation of a protein. *Nature* **267**, 274-275.
10. Yutani, K., Ogasahara, K., Tsujita, T., and Sugino, Y. (1987) Dependence of conformational stability on hydrophobicity of the amino acid residue in a series of

- variant proteins substituted at a unique position of tryptophan synthase  $\alpha$  subunit. *Proc. Natl. Acad. Sci. USA* **84**, 4441-4444.
11. Privalov, P. L. and Gill, S.J. (1988) Stability of protein structure and hydrophobic interaction. *Adv. Protein. Chem.* **39**, 191-234.
  12. Mukaiyama, A. and Takano, K. (2009) Slow unfolding of monomeric proteins from hyperthermophiles with reversible unfolding. *Int. J. Mol. Sci.* **10**, 1369-1385.
  13. Pace, C.N., Scholtz, J. M., and Grimsley, G.R. (2014) Forces stabilizing proteins. *FEBS Lett.* **588**, 2177-2184.
  14. Privalov, P. L. (1979) Stability of proteins: small globular proteins. *Adv. Protein. Chem.* **33**, 167-241.
  15. Baldwin, R. L. (1986) Temperature dependence of the hydrophobic interaction in protein folding. *Proc. Natl. Acad. Sci. USA* **83**, 8069-8072.
  16. Korndörfer, I., Steipe, B., Huber, R., Tomschy, A., and Jaenicke, R. (1995) The crystal structure of holo-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima* at 2.5 Å resolution. *J. Mol. Biol.* **246**, 511-521.
  17. Hennig, M., Darimont, B., Sterner, R., Kirschner, K., and Jansonius, J. N. (1995) 2.0 Å structure of indole-3-glycerol phosphate synthase from the hyperthermophile *Sulfolobus solfataricus*: possible determinants of protein stability. *Structure* **3**, 1295-1306.
  18. Ermler, U., Merckel, M. C., Thauer, R., and Shima, S. (1997) Formylmethanofuran: tetrahydromethanopterin formyltransferase from *Methanopyrus kandleri* - new insights into salt-dependence and thermostability. *Structure* **5**, 635-646.
  19. Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi M., and Pearl L. H. (1997) Crystal structure of the  $\beta$ -glycosidase from the hyperthermophilic archeon *Sulfolobus solfataricus*: resilience as a key factor in thermostability. *J. Mol. Biol.* **271**, 789-802.
  20. Vetriani, C., Maeder, D. L., Tolliday, N., Yip, K. S.-P., Stillman, T. J., Britton, K. L., Rice, D. W., Klump, H. H., and Robb, F. T. (1998) Protein thermostability above 100 °C: a key role for ionic interactions. *Proc. Natl. Acad. Sci. USA* **95**, 12300-12305.

21. Tahirov, T. H., Oki, H., Tsukihara, T., Ogasahara, K., Yutani, K., Ogata, K., Izu, Y., Tsunasawa, S., and Kato, I. (1998) Crystal structure of methionine aminopeptidase from hyperthermophile, *Pyrococcus furiosus*. *J. Mol. Biol.* **284**, 101-124.
22. de Bakker, P.I., Hünenberger, P.H., and McCammon, J. A. (1999) Molecular dynamics simulations of the hyperthermophilic protein Sac7d from *Sulfolobus acidocaldarius*: contribution of salt bridges to thermostability. *J. Mol. Biol.* **285**, 1811-1830.
23. Karshikoff, A. and Ladenstein, R. (2001) Ion pairs and the thermotolerance of proteins from hyperthermophiles: a "traffic rule" for hot roads. *Trends Biochem. Sci.* **26**, 550-556.
24. Yamagata, Y., Ogasahara, K., Hioki, Y., Lee, S. J., Nakagawa, A., Nakamura, M., Ishida, M., Kuramitsu, S., and Yutani, K. (2001) Entropic stabilization of the tryptophan synthase  $\alpha$ -subunit from a hyperthermophile, *Pyrococcus furiosus*: X-ray analysis and calorimetry. *J. Biol. Chem.* **276**, 11062-11071.
25. Robinson-Rechavi, M., Alibés, A., and Godzik, A. (2006) Contribution of electrostatic interactions, compactness and quaternary structure to protein thermostability: lessons from structural genomics of *Thermotoga maritima*. *J. Mol. Biol.* **356**, 547-557.
26. Tanaka T., Sawano M., Ogasahara K., Sakaguchi Y., Bagautdinov B., Katoh E., Kuroishi C., Shinkai A., Yokoyama S., and Yutani K. (2006) Hyper-thermostability of CutA1 protein, with a denaturation temperature of nearly 150 °C. *FEBS Lett.* **580**, 4224-4230.
27. Matsuura, Y., Ota, M., Tanaka, T., Takehira, M., Ogasahara, K., Bagautdinov, B., Kunishima, N., and Yutani, K. (2010) Remarkable improvement in the heat stability of CutA1 from *Escherichia coli* by rational protein design. *J. Biochem.* **148**, 449-458.
28. Sawano, M., Yamamoto, H., Ogasahara, K., Kidokoro, S., Katoh, S., Ohnuma, T., Katoh, E., Yokoyama, S., and Yutani, K. (2008) Thermodynamic basis for the stabilities of three CutA1s from *Pyrococcus horikoshii*, *Thermus thermophilus*, and *Oryza sativa*, with unusually high denaturation temperatures. *Biochemistry* **47**, 721-730.
29. Elcock, A. H., and McCammon, J. A. (1997) Continuum solvation model for studying protein hydration thermodynamics at high temperatures. *J. Phys. Chem.* **101**, 9624-9634.

30. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411-2423.
31. Durchschlag, H. (1986) Specific volumes of biological macromolecules and some other molecules of biological interest. *Thermodynamic Data for Biochemistry and Biotechnology*, ed Hinz H-J (Springer-Verlag, Berlin, Germany), pp. 45-128.
32. Privalov, P. L. and Khechinashvili, N. N. (1974) A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study. *J. Mol. Biol.* **86**, 665-684.
33. Privalov, P. L., and Makhatadze, G. I. (1990) Heat capacity of proteins II. Partial molar heat capacity of the unfolded polypeptide chain of proteins: protein unfolding effects. *J. Mol. Biol.* **213**, 385-391.
34. Makhatadze, G. I. and Privalov, P. L. (1995) Energetic of protein structure. *Adv. Protein Chem.* **47**, 307-425.
35. Makhatadze, G. and Privalov, P. L. (1994) Hydration effects in protein unfolding. *Biophys. Chem.* **51**, 291-309.
36. Kauzmann, W. (1959) Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* **14**, 1-63.
37. Fersht, A. R. (1972). Conformational equilibria in  $\alpha$  and  $\delta$  chymotrypsin. The energetics and importance of the salt bridge. *J. Mol. Biol.* **64**, 497-509.
38. Lam, S. Y., Yeung, R. C., Yu, T. H., Sze, K. H., and Wong, K. B. (2011) A rigidifying salt-bridge favors the activity of thermophilic enzyme at high temperatures at the expense of low-temperature activity. *PLoS Biol.* **9**, e1001027.
39. Pace, C. N. (2000) Single surface stabilizer. *Nat. Struct. Biol.* **7**, 345-346.
40. Elcock, A. H. (1998) The stability of salt bridges at high temperatures: Implications for hyperthermophilic proteins. *J. Mol. Biol.* **284**, 489-502.

## Chapter 6: Stabilization of *Escherichia coli* CutA1 by ion–ion interactions

### 6-1. Introduction.

We present the results of experiments in which substitutions have been made in the sequence of mesophilic CutA1 (*Ec*CutA1) to probe the role of charged residues in increasing the stability of this protein possibly to the level similar to that of its hyperthermophilic counterpart, *Ph*CutA1. To this end we have constructed many mutants of *Ec*CutA1 in which substitutions to charged residues were made. These charged mutations were incorporated in the background of *Ec*CutA1 mutant (*Ec*0VV)<sup>1,2</sup>. *Ec*0VV has additional two substitutions (S11V/E61V) at the buried positions and has higher  $T_d$  of 113.2 °C<sup>1,2</sup> relative to a cysteine-free *Ec*CutA1\_0SH that has  $T_d$  of 85.6 °C. Single and multiple ionic mutants of more than 100 different variants were measured experimentally by DSC. The highest stability of multiple mutants was a mutant substituted by 9 charged residues (Fig. 1) with  $T_d$  of 142.2 °C, which is close to the  $T_d = 148.5$  °C of hyperthermophilic *Ph*CutA1<sup>3</sup>.

The native protein structure is not static as depicted by x-ray crystallography, but dynamic. To account for structural flexibility of residues in the native state, we generated a structural ensemble using all-atom explicit solvent molecular dynamics (MD) simulations<sup>4</sup> at 300 K. This structural ensemble was used to evaluate the energy of ion-ion interactions. The obtained results showed that the denaturation temperatures linearly increase with the increment of the energy of ion-ion interactions for ionic mutant proteins,

suggesting that ion-ion interactions cumulatively contribute to the stabilization of a protein at temperatures of over 100 °C.

## **6-2. Experimental methods.**

### **Mutagenesis, expression, and purification of *EcCutA1* mutants.**

Hydrophobic cysteine-free mutant of *EcCutA1* (*EcCutA1\_C16A/C39A/C79A\_S11V/E61V*) was used as a template for all mutations. Mutagenesis, expression, and purification of mutants were performed as described<sup>2</sup>. All purified mutant proteins ran as single band in SDS-PAGE. Protein concentrations were determined using an absorption coefficient of  $E^{1\text{cm}}_{1\%} = 14.96$ , based on the number of aromatic amino acid<sup>5</sup>.

### **Differential scanning calorimetry (DSC) experiments.**

To measure the changes in stability due to mutations, DSC was performed using a scan rate of 60 °C/h on a VP-capillary DSC platform (Microcal, USA) for temperatures up to 130 °C at pressures below 60 psi, or a Nano-DSC 6300Y microcalorimeter (TA Instruments, USA) for higher temperatures up to 160 °C at a pressure of 88 psi. Protein concentration was kept at 0.6 mg/ml. All samples were dialyzed overnight at 4 °C against the 50 mM glycine buffer at pH 9.0 containing 2 mM EDTA and then filtered through a 0.22 µm membrane. The denaturation temperature,  $T_d$ , is defined as the temperature at which the area of the denaturation enthalpy is 0.5. When it seemed to be aggregation after heat denaturation, temperature of the peak maximum of DSC curves was estimated to be

$T_d$  (Table 1). Considering very high overall enthalpy of unfolding, such approximation does not lead to significant errors in  $T_d$ . The reported results are the average value of at least two independent experiments.

### **MD simulation of mutant *EcCutA1* proteins.**

Structures were prepared for MD simulations as follows. The missing residues in the coordinate file (PDB ID 4Y65) of a mutant, *EcCutA1\_C16A/C39A/C79A*, which are three N-terminal residues of B-subunit and eight N-terminal residues of C-subunit in a trimer complex, were added using QUANTA2000 (Accelrys Inc), using the coordinates of N-terminal residues of A-subunit as a template. The structure of mutant proteins was modeled using FoldX (<http://foldxsuite.crg.eu>)<sup>6</sup>.

MD simulations were performed using GROMACS software (ver. 4.5.5)<sup>7, 8</sup>. For the MD simulations the protein was placed in a cubic box with 1.2 nm between the protein and the box. Counterions were added to neutralize any net charge. The long-range electrostatic interactions were computed using the Particle-Mesh-Eward (PME) method<sup>9</sup>. The GROMOS 43a1 force field and spc/e water model<sup>10</sup> was employed. The system was weakly coupled to a heat bath by velocity rescaling<sup>11</sup> with a relaxation time of 0.1 ps. A Parrinello-Rahman barostat<sup>12</sup> was used to maintain a pressure constant at 1 atm with a relaxation time of 0.5 ps. Hydrogen atoms were constrained using LINCS<sup>13</sup>, and MD simulations were performed at 300 K with an integration time-step of 1 femtosecond (fs). Prior to the production run, energy minimization for 1000 steps was followed by increasing temperature from 50 K to 300 K in increments of 50 K, with 10,000 integration

steps at each temperature and a harmonic constraint of C-alpha atoms. Thereafter, the ensemble was equilibrated through four 100-picosecond (ps) cycles with gradually released harmonic constraints: 1000, 100, 10, and 1 kJ mol<sup>-1</sup> nm<sup>-2</sup>. The subsequent MD production stages for the *Ec*CutA1 mutants were carried out without any restraint at 300 K. The obtained MD trajectories were analyzed using GROMACS software. Fig. S2 shows of root mean square deviation (RMSD) of C $\alpha$  atoms as a function of simulation time for ionic mutants at 300 K. The RMSD values of majority of mutants reached constant values after 2-3 ns and for all at 8 ns. It should be noted that Sawle & Ghosh<sup>14</sup> using a different force field and water model, i.e. Amber99sb/tip3p, report convergence at much longer time scales.

### **Evaluation of the energy of ion-ion interactions for *Ec*0VV mutants.**

The computer algorithm, FoldX<sup>6</sup>, can quantitatively estimate the stabilization factors that are important for protein stability. FoldX is available via a web-interface at <http://foldxsuite.crg.eu>. The electrostatic energies due to ion-ion interactions between charged residues were calculated with “AllAtoms\_Electro” file in FoldX. The electrostatic energy in FoldX is calculated from a simple implementation of Coulomb’s law, in which the dielectric constant is scaled with the burial of the bond under consideration<sup>6</sup>. Structural snapshots from the MD simulations of mutants substituted by charged residues are picked up at every 2 ns after 8 ns run during 40ns (17 structures). The average computed energies from these structures are shown in Tables 2 and S1.

### **6-3. Results.**

#### **Thermostability of *Ec0VV* mutants substituted by charged residues.**

There are few reports addressing how charged residues contribute to the thermostability of a protein at the temperature of over 100 °C. Therefore, in order to examine the relationship between the ion-ion interactions and thermostability of a protein at over 100 °C, single and multiple ionic mutants of more than 100 were designed to improve the stability using various stabilization methods already reported<sup>6, 15-25</sup>. The thermostabilities were experimentally measured by DSC (Fig. 2). Out of 83 single ionic mutants, there were 37 mutants with increase in  $T_d$  ( $\Delta T_d > +0.5$  °C) and 31 mutants with a decrease in  $T_d$  ( $\Delta T_d < -0.5$  °C) (Table 1). The remaining 15 mutants showed only negligible changes in  $T_d$  ( $-0.5$  °C  $< \Delta T_d < +0.5$  °C). Multiple single or double (in the case of A39D/S48K) mutations that showed the highest increase in stability were combined to make *Ec0VV\_6* and *Ec0VV\_9* mutants of *EcCutA1*. These *EcCutA1* mutants showed the denaturation temperature close to that of *PhCutA1* (Table 1). The *Ec0VV\_6* mutant (six mutations A39D/S48K/H72K/S82K/Q87K/T88R) has  $T_d$  of  $136.8 \pm 0.9$  °C, a 24 °C increase in thermostability. The *Ec0VV\_9* mutant (three additional mutations Q25R/T101E/N108E relative to *Ec0VV\_6*) showed highest thermostability of  $142.2 \pm 0.9$  °C, which is an increase of 29.0 °C as compared to starting *Ec0VV* template.

#### **Evaluation of the energy of ion-ion interactions using the conformational ensemble obtained from MD simulations.**

Charged residues on the surface of a protein show a high degree of flexibility in water<sup>26, 27</sup>. Therefore, MD simulations of ionic mutants of *EcCutA1* were performed at 300 K to examine how charged residues introduced by mutations interact with other charges. Fluctuations of ion pairs can be estimated by analyzing distances between charged residues (see e.g. representative examples in Fig. 3).

The distance between C $\gamma$  atom of Asp39 in A-subunit and C $\epsilon$  atom of Lys67 in A-subunit of *Ec0VV\_6* remains constant  $2.2 \pm 0.2$  Å, suggesting that the fluctuations are rather small (Fig. 3a and Table S2B). In the case of Glu90 and Lys67 (Fig. 3b), there are fluctuations between two distances suggesting that it alternates between two isomers of the sidechain of Glu90 (Fig. S3). The distance between Lys5 and Glu4 (Fig. 3c) shows large fluctuations suggesting that a salt bridge between these two residues is formed only transiently. On the other hand, the inter subunit interaction between Arg88 in B-subunit and Asp39 in C-subunit (Fig. 3d) appears to undergo yet a different type of fluctuation. Other inter-subunit interactions between these residues were not detected in the range of less than 8.0 Å (Table S2B).

To evaluate the energy due to ion-ion interaction in water, the structural ensemble obtained from MD simulations was used to evaluate the energy of ion-ion interactions. The electrostatic energies of ion-ion interaction for each structure were estimated using FoldX energy function<sup>6</sup>. Table 2 shows the average values of the total computed energy of each ionic mutant protein. The difference in the energy of ion-ion interactions between ionic mutants and *Ec0VV* template shows a linear correlation with the difference in the denaturation temperatures between them, as shown in Fig. 4. These results suggest that

ion-ion interactions cumulatively contribute to the stabilization of a protein up to the temperatures near 140 °C.

#### **6-4. Discussion.**

##### **Role of charged residues in stabilization of *Ec0VV\_6*.**

The  $T_d$  of *Ec0VV\_6* (with 6 additional charged residues) was increased by 23.6 °C, and the energy of ion-ion interactions was improved by 91.1 kJ/mol (Table 2), as compared with those of the template, *Ec0VV*.

Table S1 shows the energy of ion-ion interaction at targeted charged residues of *Ec0VV* mutants. All pair residues interacting with the targeted residue are listed and discriminated between inter and intra interactions in a trimer structure. Arg88 among the targeted ion residues is mainly inter interacted with Asp102, Lys48 is strongly intra interaction with Asp39 and strongly intra interaction with Glu59 and Glu90, and Asp39 and Lys72 are mainly intra interacted with pair-residues. Table S2 also shows the average distance of their ion-ion interactions.

Although the mutant *Ec0VV\_A39D* was significantly less stable due to substitution of charged residue in the buried position (Table 1A), there was an increase in the stability of double mutant, *Ec0VV\_A39D/S48K*, even though *Ec0VV\_S48K* mutation is also destabilizing. Inspection of the structure and energetics of ion-ion interactions suggests that the stability increase of the double mutant, *Ec0VV\_A39D/S48K*, is probably due to the formation of a salt bridge between these two residues (Table 1B). Similar effects have

been observed previously in other model protein systems<sup>28</sup>. The computed interaction energy between Asp39 and Lys48 was similar in both the *Ec0VV\_6* and *Ec0VV\_9* mutants (Table S1A). Lys72 in *Ec0VV\_H72K* forms a strong salt bridge with carboxylate of the C-terminal Arg112 (Table S2B): this salt bridge also remains present in both *Ec0VV\_6* and *Ec0VV\_9* backgrounds (Table S1C).

Substitutions at adjacent residues, *Ec0VV\_Q87K* and *Ec0VV\_T88R* showed an increase in thermostability,  $\Delta T_d$ , by 3.6 and 4.4 °C, respectively. The  $\Delta T_d$  of a double mutant 9.2 °C, *Ec0VV\_Q87K/T88R* was somewhat more than the sum of stability increase of individual mutants. Similar effects have been previously reported for the mutations at the adjacent position in the sequence of ubiquitin and have been shown to be related to the electrostatic context into which such mutations are incorporated<sup>29</sup>.

### **Features of the stabilization of *Ec0VV\_9*.**

Three mutants, *Ec0VV\_E34R*, *Ec0VV\_E57R*, and *Ec0VV\_S110R*, showed an increase in  $T_d$  relative to the *Ec0VV* background (Table 2). However, these single site substitutions in the *Ec0VV\_6* did not lead to an increase in the stability or in the computed energy of ion-ion interactions (Table 2). On the other hand, when three substitutions, Q25R, T101E, and N108E, were simultaneously introduced into *Ec0VV\_6*, the  $T_d$  of its mutant protein, *Ec0VV\_9*, was increased by 5.4 °C. Furthermore, the computed energy of ion-ion interaction of *Ec0VV\_9* was less than that of *Ec0VV\_6*.

The ionic features of six charged residues substituted in *Ec0VV\_6* were hardly affected by the introduction of three charged residues in *Ec0VV\_9* (Table S1). The mutant

*Ec0VV\_Q25R* seems to be stabilized by increase in electrostatic energy (Table 2) and favorable ion-ion interaction of Arg25 remains in the *Ec0VV\_9* background (Table S1G). The introduction of Glu101 decreased the energy of ion-ion interactions in *Ec0VV\_T101E* and *Ec0VV\_9* (Table S1H), but might increase the stability due to strengthening of the helix dipole moment<sup>30-32</sup>. Glu101 is located in the N-terminal of  $\alpha$ -3 helix and the corresponding residue of CutA1 from hyperthermophile (*PhCutA1*) is also Glu. The computed energy of ion-ion interaction due to introduction of Glu108 was negligible in both *Ec0VV\_N108E* and *Ec0VV\_9* (Table 2 and Table S1I). The introduction of either negative (*Ec0VV\_N108E*) or positive (*Ec0VV\_N108K*, and *Ec0VV\_N108R*) residues at this position leads to an increase in  $T_d$  ( $\Delta T_d$  for *Ec0VV\_N108E*, *Ec0VV\_N108K*, and *Ec0VV\_N108R*, is 2.1, 2.8, and 2.3 °C, respectively). Considering that N108 is located in the helical region, the stabilization might be due to the increase in helix propensity of the residues as it is known that E, R and K have higher helical propensity than N<sup>33</sup>.

**Stabilization strategy of the CutA1 protein that improved up to the denaturation temperature of 142.2 °C.**

The average increase in stability due to single site substitutions was 2.2 °C and the highest increase was 5.2 °C for *Ec0VV\_H72K* (Table 1A). This suggests that it is necessary to introduce many charged residues to construct mutant proteins with  $T_d$  comparable to that of *PhCutA1* from hyperthermophile. The  $T_d$ 's of *Ec0VV\_6* and *Ec0VV\_9* were increased by 23.6 and 29.0 °C, respectively, relative to that of *Ec0VV* (Table 1B). As shown in Fig. 4, the increment of  $\Delta T_d$  for mutant proteins linearly

correlates with the difference in computed ion-ion interaction energy of *Ec0VV* and *Ec0VV* mutants. This suggest that the ion-ion interaction at multiple distant position in the native structure contribute additively to the stabilization of a protein even at the temperature of over 100 °C, although other factors in addition to ion-ion interactions such as helix-propensity and charge-helix dipole interactions also play a role in modulating the stability.

It is interesting to compare the number and distribution of charge residues and corresponding *CvP*-bias of the *Ec0VV\_9* mutant with that of hyperthermophilic *PhCutA1*. The *Ec0VV\_9* mutant contains fewer charged residues and has *CvP*-bias of 10%. This *CvP* is much higher than for mesophilic *EcCutA1* (-4%), much lower than for *PhCutA1* (32%) and more on par with the average value computed for thermophilic proteomes, 10-14%<sup>34</sup>. Moreover, the distribution of charges along the sequence between *Ec0VV\_9* and *PhCutA1* is also different (see Fig. 1). We have shown previously that ion-ion interactions make significant contribution to the stability of *PhCutA1*. The increase in stability of *Ec0VV\_9* relative to the template *Ec0VV* is also due to the optimization of the ion-ion interactions. This suggests that it is possible to use different arrangements of ionizable residues to achieve significant protein stabilization. Overall, our results suggest that introduction of multiple charged residues represents a useful strategy in engineering proteins with high denaturation temperatures of over 100 °C.

```

PhCutA1      -----MIIVYTTFPDWESAEKVVKTLLKERLIACAN-LREHRAFYWEGKIE
Ec0VV        MLDEKSSNTAVVVVLATAPDEATAQDLAAKVLAEKLAAAATLIPGATSLYYWEGKLE
Ec0VV_9      MLDEKSSNTAVVVVLATAPDEATARDLAAKVLAEKLAADATLIPGATKLYYWEGKLE

PhCutA1      EDKEVGAILKTREDLWEELKERIKELHPYDVPAIIRIDVDDVNEDYLKWLIEETKK
Ec0VV        QEYVVQMILKTTVSHQQALLEALKSHHPYQTPELLVLPVTHGDTDYLSWLNASLR-
Ec0VV_9      QEYVVQMILKTTVSKQQALLEALKKHHPYKRPELLVLPVTHGDEDYLSWLEASLR-

```

Fig. 1. Sequence alignment of *PhCutA1* with *Ec0VV* and *Ec0VV\_9*. Red and blue colors represent positively and negatively charged residues, respectively. Underbars represent the 9 charged residues newly substituted in *Ec0VV\_9*.

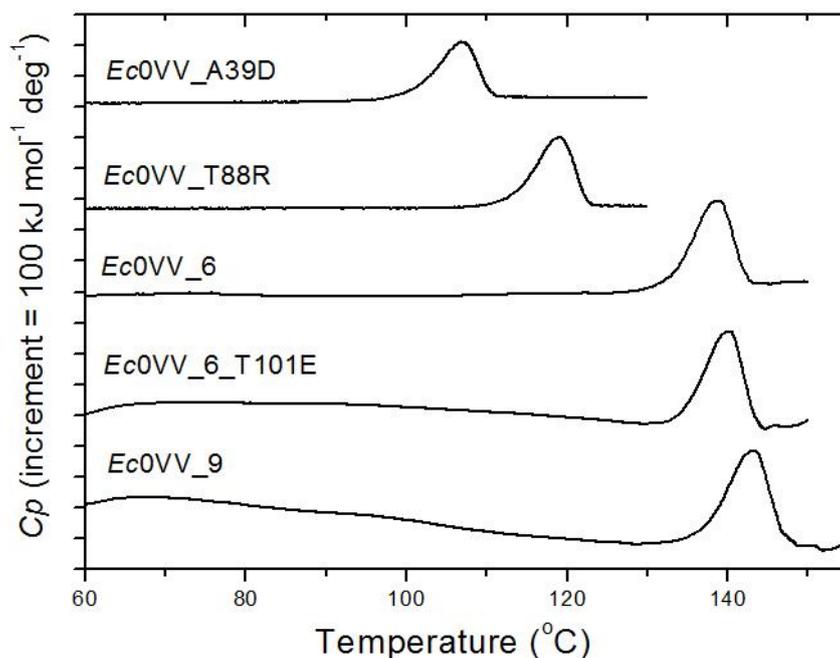


Fig. 2. Typical DSC curves of five ionic mutants of *Ec0VV* at pH 9.0. Six mutants are *Ec0VV\_A39D*, *Ec0VV\_T88R*, *Ec0VV\_6*, *Ec0VV\_6\_T101E*, and *Ec0VV\_9*. Scan rates were 60 °C/h. The curve of *Ec0VV\_6* is reported in reference<sup>2</sup>.

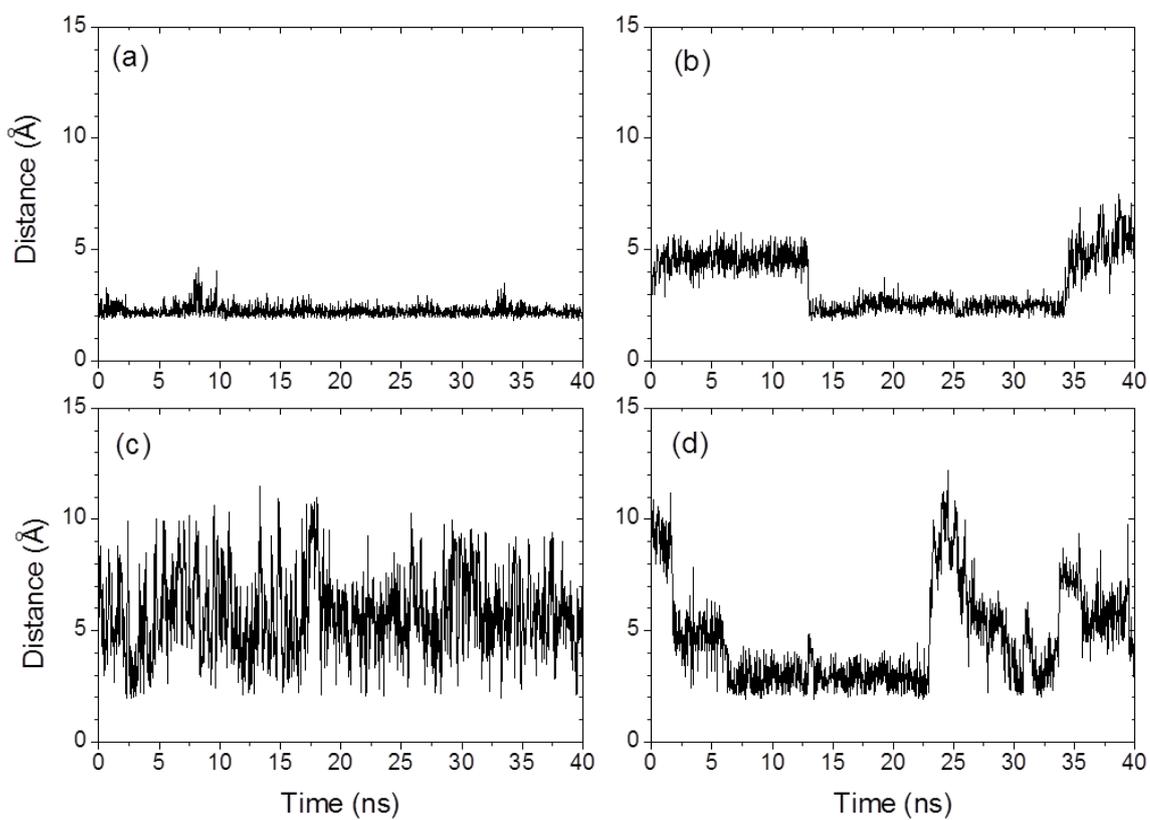


Fig. 3. MD trajectory of ion-ion interaction of *Ec0VV\_6* for 40 ns at 300 K. (a) Distance between  $C_{\gamma}$  atom of Asp39 in A-subunit and  $C_{\epsilon}$  atom of Lys67 in A-subunit. (b) Distance between  $C_{\delta}$  atom of Glu90 in A-subunit and  $C_{\epsilon}$  atom of Lys67 in B-subunit. (c) Distance between  $C_{\delta}$  atom of Glu4 in A-subunit and  $C_{\epsilon}$  atom of Lys5 in A-subunit. (d) Distance between  $C_{\delta}$  atom of Arg88 in B-subunit and  $C_{\gamma}$  atom of Asp39 in C-subunit.

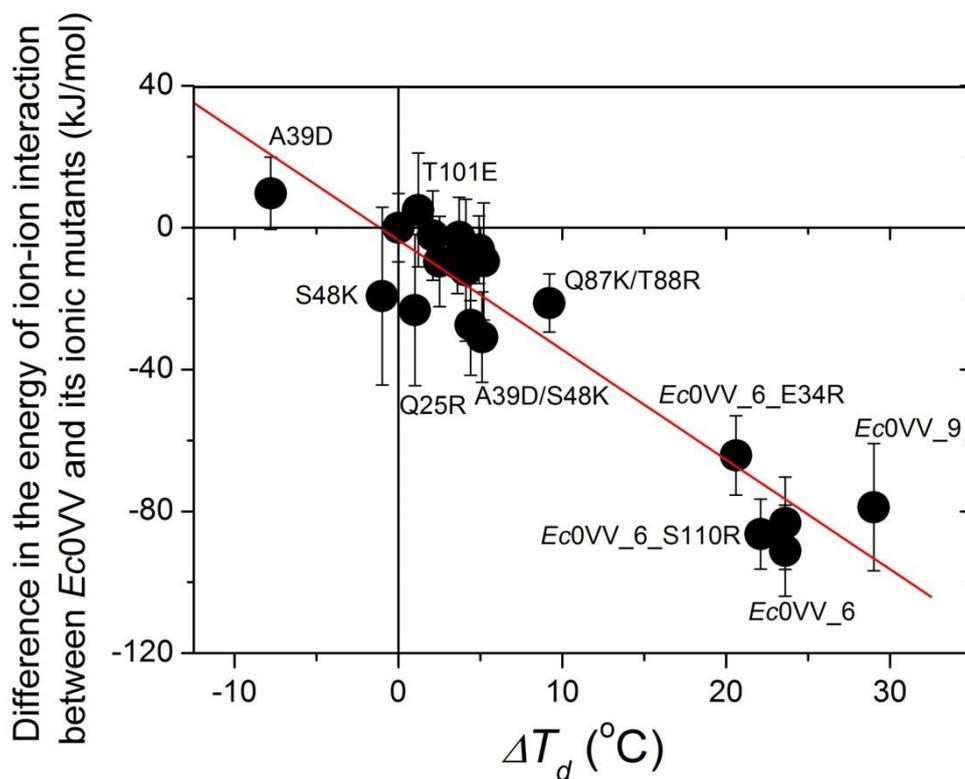


Fig. 4. Relationship of  $\Delta T_d$  and difference in the ionic interaction energy between *Ec0VV* and its ionic mutant proteins. Errors bar represent the standard deviation of 17 structures obtained by MD simulation at 300 K, red line represents a linear regression of 21 points:  $r = -0.94$  and  $p = 2.70 \times 10^{-10}$ .

Table 1.

Denaturation temperature ( $T_d$  in °C) of *Ec0VV* mutants substituted with charged residues at pH 9.0.

(A) Single mutants.

Mutants	$T_d$ (°C)		$\Delta T_d$ (°C)	S.S	ASA (%)
<i>Ec0VV</i> H72K <sup>+</sup>	118.4	± 0.4	5.2	$\alpha 2$	13
<i>Ec0VV</i> E57R**	118.1	± 0.3	4.9	$\beta 3$	46
<i>Ec0VV</i> H72R**	118.0	± 0.2	4.8	$\alpha 2$	13
<i>Ec0VV</i> E57K**	117.6	± 0.0	4.4	$\beta 3$	46
<i>Ec0VV</i> T88R <sup>+</sup>	117.6	± 0.6	4.4	L	30
<i>Ec0VV</i> S110R <sup>+</sup>	117.3	± 0.4	4.1	$\alpha 3$	9
<i>Ec0VV</i> S82R <sup>+</sup>	117.1	± 0.5	3.9	$\alpha 2$	100
<i>Ec0VV</i> Q87R	117.1	± 0.2	3.9	L	100
<i>Ec0VV</i> S82K <sup>+</sup>	116.9	± 0.5	3.7	$\alpha 2$	100
<i>Ec0VV</i> Q87K <sup>+</sup>	116.8	± 0.5	3.6	L	100
<i>Ec0VV</i> T88K**	116.1	± 0.2	2.9	L	30
<i>Ec0VV</i> N108K**	116.0	± 0.2	2.8	$\alpha 3$	57
<i>Ec0VV</i> S71R**	116.0	± 0.1	2.8	$\alpha 2$	100
<i>Ec0VV</i> E34K**	116.0	± 0.1	2.8	$\alpha 1$	75
<i>Ec0VV</i> S105R <sup>+</sup>	115.7	± 0.2	2.5	$\alpha 3$	70
<i>Ec0VV</i> S71K**	115.7	± 0.3	2.5	$\alpha 2$	100
<i>Ec0VV</i> E34R**	115.7	± 0.1	2.5	$\alpha 1$	75
<i>Ec0VV</i> H98D**	115.6	± 0.0	2.4	$\beta 4$	1
<i>Ec0VV</i> N108R	115.5	± 0.1	2.3	$\alpha 3$	57
<i>Ec0VV</i> T47K**	115.3	± 0.2	2.1	$\beta 2$	26
<i>Ec0VV</i> N108E	115.3	± 0.1	2.1	$\alpha 3$	57
<i>Ec0VV</i> A75R**	115.3	± 0.1	2.1	$\alpha 2$	41
<i>Ec0VV</i> Q25K**	115.2	± 0.0	2.0	$\alpha 1$	48
<i>Ec0VV</i> W52K	115.0	± 0.2	1.8	$\beta 2$	39
<i>Ec0VV</i> W52R	114.9	± 0.3	1.7	$\beta 2$	39
<i>Ec0VV</i> S105E	114.4	± 0.2	1.2	$\alpha 2$	70
<i>Ec0VV</i> Q58R**	114.4	± 0.4	1.2	$\beta 3$	52
<i>Ec0VV</i> T101E	114.4	± 0.3	1.2	$\alpha 3$	39
<i>Ec0VV</i> A109R <sup>+</sup>	114.2	± 0.1	1.0	$\alpha 3$	54
<i>Ec0VV</i> D26K	114.2	± 0.1	1.0	$\alpha 1$	59
<i>Ec0VV</i> Q25R	114.2	± 0.1	1.0	$\alpha 1$	48
<i>Ec0VV</i> S110K	114.2	± 0.2	1.0	$\alpha 3$	9
<i>Ec0VV</i> A33R	114.1	± 0.2	0.9	$\alpha 1$	49
<i>Ec0VV</i> Q58K	114.0	± 0.0	0.8	$\beta 3$	52
<i>Ec0VV</i> A33K	114.0	± 0.2	0.8	$\alpha 1$	49
<i>Ec0VV</i> S105K	113.9	± 0.1	0.7	$\alpha 3$	70
<i>Ec0VV</i> A22K**	113.9	± 0.0	0.7	$\alpha 1$	61
<i>Ec0VV</i> A29K	113.7	± 0.1	0.5	$\alpha 1$	32
<i>Ec0VV</i> A10R**	113.6	± 0.2	0.4	L	69
<i>Ec0VV</i> A29R	113.6	± 0.1	0.4	$\alpha 1$	32
<i>Ec0VV</i> T101R	113.5	± 0.1	0.3	$\alpha 3$	39
<i>Ec0VV</i> A79K**	113.4	± 0.2	0.2	$\alpha 2$	20
<i>Ec0VV</i> E53R	113.1	± 0.1	-0.1	L	76
<i>Ec0VV</i> L2E	113.1	± 0.1	-0.1	L	56
<i>Ec0VV</i> Y60E	113.1	± 0.0	-0.1	$\beta 3$	62
<i>Ec0VV</i> A10E	113.0	± 0.6	-0.2	L	69
<i>Ec0VV</i> A22R	112.9	± 0.0	-0.3	$\alpha 1$	61
<i>Ec0VV</i> T97E	112.9	± 0.0	-0.3	$\beta 4$	33
<i>Ec0VV</i> T69K**	112.8	± 0.0	-0.4	$\beta 3$	5
<i>Ec0VV</i> A10K	112.7	± 0.1	-0.5	L	69
<i>Ec0VV</i> E59K	112.7	± 0.0	-0.5	$\beta 3$	47
<i>Ec0VV</i> Q74R**	112.7	± 0.0	-0.5	$\alpha 2$	21
<i>Ec0VV</i> L2R	112.5	± 0.2	-0.7	L	56
<i>Ec0VV</i> A109E	112.4	± 0.2	-0.8	$\alpha 3$	54
<i>Ec0VV</i> S6R	112.3	± 0.1	-0.9	L	67
<i>Ec0VV</i> H83K	112.3	± 0.1	-0.9	L	41
<i>Ec0VV</i> A33E	112.3	± 0.2	-0.9	$\alpha 1$	49
<i>Ec0VV</i> A29E	112.3	± 0.1	-0.9	$\alpha 1$	32
<i>Ec0VV</i> Q74E	112.2	± 0.2	-1.0	$\alpha 2$	21
<i>Ec0VV</i> S48K <sup>+</sup>	112.2	± 0.9	-1.0	$\beta 2$	1
<i>Ec0VV</i> S7K	112.1	± 0.1	-1.1	L	35
<i>Ec0VV</i> S48R**	112.1	± 0.1	-1.1	$\beta 2$	1
<i>Ec0VV</i> Q25E	112.0	± 0.1	-1.2	$\alpha 1$	48
<i>Ec0VV</i> S71E	111.8	± 0.1	-1.4	$\alpha 2$	100
<i>Ec0VV</i> Q58E	111.6	± 0.1	-1.6	$\beta 3$	52
<i>Ec0VV</i> V70E	111.5	± 0.0	-1.7	L	45
<i>Ec0VV</i> A75E	110.5	± 0.3	-2.7	$\alpha 2$	41
<i>Ec0VV</i> V70R**	109.5	± 0.8	-3.7	L	45
<i>Ec0VV</i> A33D	109.5	± 0.1	-3.7	$\alpha 1$	49
<i>Ec0VV</i> A79E	109.2	± 0.2	-4.0	$\alpha 2$	20
<i>Ec0VV</i> K81R	109.1	± 0.1	-4.1	$\alpha 2$	49
<i>Ec0VV</i> Q73E	108.8	± 0.3	-4.4	$\alpha 2$	1
<i>Ec0VV</i> A39K	108.7	± 0.0	-4.5	$\beta 2$	0
<i>Ec0VV</i> Q73K**	108.1	± 0.1	-5.1	$\alpha 2$	1
<i>Ec0VV</i> T97K**	107.7	± 0.0	-5.5	$\beta 4$	33
<i>Ec0VV</i> T69D	107.5	± 0.2	-5.7	$\beta 3$	5
<i>Ec0VV</i> A39R**	107.0	± 0.1	-6.2	$\beta 2$	0
<i>Ec0VV</i> G45K**	106.3	± 0.4	-6.9	L	31
<i>Ec0VV</i> S48D <sup>+</sup>	105.4	± 0.2	-7.8	$\beta 2$	1
<i>Ec0VV</i> A39D <sup>+</sup>	105.4	± 0.5	-7.8	$\beta 2$	0
<i>Ec0VV</i> H72E	102.4	± 0.0	-10.8	$\alpha 2$	13
<i>Ec0VV</i> Q63K	102.2	± 0.5	-11.0	$\beta 3$	1
<i>Ec0VV</i> H72D	96.3	± 0.2	-16.9	$\alpha 2$	13

SS represents the secondary structure of the template mutant at a substituted residue. ASA represents accessible surface area of the template mutant at a substituted residue.

(B) Multiple mutants.

Mutants	$T_d$ (°C)			$\Delta T_d$ (°C)
		±		
<i>Ec</i> 0VV template*	113.2	±	0.2	0.0
<i>Ec</i> 0VV A39D/S48K*	118.3	±	0.7	5.1
<i>Ec</i> 0VV Q87K/T88R*	122.4	±	0.6	9.2
<i>Ec</i> 0VV Q87K/T88R/S110R **	120.8	±	0.4	7.6
<i>Ec</i> 0VV Q87K/T88R/S82K **	126.0	±	0.8	12.8
<i>Ec</i> 0VV Q87K/T88R/H72K**	127.5	±	0.5	14.3
<i>Ec</i> 0VV Q87K/T88R/H72K/S82K **	131.5	±	0.6	18.3
<i>Ec</i> 0VV_6*	136.8	±	0.9	23.6
<i>Ec</i> 0VV_6 E34R	133.8	±	0.0	20.6
<i>Ec</i> 0VV_6 S110R	135.3	±	0.0	22.1
<i>Ec</i> 0VV_6 E57R	136.8	±	0.1	23.6
<i>Ec</i> 0VV_6 T101E	139.3	±	0.3	26.1
<i>Ec</i> 0VV_6 Q25R/T101E	140.3	±	0.1	27.1
<i>Ec</i> 0VV_6 Q25R/T101E/N108E ( <i>Ec</i> 0VV_9)	142.2	±	0.2	29.0

The heating rate (scan rate) of DSC measurements was 60 °C/h.

These parameters are the average value of more than two data.

Positive values of  $\Delta T_d$  indicate the increase in stability due to mutations.

\*These data are reported in reference<sup>2</sup>.

\*\* $T_d$  is the temperature of peak center of DSC curves, because it seemed to be aggregated after heat denaturation.

*Ec*0VV6 represents *Ec*CutA1\_0SH\_S11V/E61V/A39D/S48K/H72K/S82K/Q87K/T88R mutant.

Table 2.

Total energy of ion-ion interactions of *Ec0VV* mutants using MD simulation data.

Mutants	Total energy of ion-ion interaction*			Difference**	$\Delta T_d$ (°C)
		±			
<i>Ec0VV</i> template	-62.7	±	9.6	0.0	0.0
<i>Ec0VV_Q25R</i>	-85.9	±	21.4	-23.2	1.0
<i>Ec0VV_E34R</i>	-72.2	±	12.8	-9.6	2.5
<i>Ec0VV_A39D</i>	-53.0	±	10.2	9.7	-7.8
<i>Ec0VV_S48K</i>	-82.0	±	25.1	-19.3	-1.0
<i>Ec0VV_E57K</i>	-73.0	±	10.3	-10.3	4.4
<i>Ec0VV_E57R</i>	-68.9	±	9.5	-6.2	4.9
<i>Ec0VV_H72K</i>	-72.2	±	16.5	-9.5	5.2
<i>Ec0VV_S82K</i>	-65.2	±	11.1	-2.6	3.7
<i>Ec0VV_Q87K</i>	-72.7	±	8.5	-10.1	3.6
<i>Ec0VV_T88R</i>	-90.1	±	14.2	-27.4	4.4
<i>Ec0VV_T101E</i>	-57.7	±	16.1	5.0	1.2
<i>Ec0VV_N108E</i>	-64.9	±	12.6	-2.2	2.1
<i>Ec0VV_S110R</i>	-74.7	±	20.0	-12.0	4.1
<i>Ec0VV_A39D/S48K</i>	-93.6	±	12.7	-30.9	5.1
<i>Ec0VV_Q87K/T88R</i>	-83.9	±	8.2	-21.2	9.2
<i>Ec0VV_6</i> ***	-153.8	±	12.9	-91.1	23.6
<i>Ec0VV_6_E34R</i>	-126.9	±	11.2	-64.2	20.6
<i>Ec0VV_6_E57R</i>	-146.0	±	13.0	-83.3	23.6
<i>Ec0VV_6_S110R</i>	-149.1	±	9.9	-86.4	22.1
<i>Ec0VV_9</i> ****	-141.5	±	17.9	-78.8	29.0

\*The total energy of ion-ion interactions are average values of energies obtained from 17 structures during 40ns MD simulation. The energy was calculated by "AllAtoms Electro" of FoldX.

\*\*Difference between the ion-ion interaction energies of mutants and *Ec0VV* Template.

\*\*\**Ec0VV\_6*: *Ec0VV\_A39D/S48K/H72K/S82K/Q87K/T88R*.

\*\*\*\**Ec0VV\_9*: *Ec0VV\_Q25R/A39D/S48K/H72K/S82K/Q87K/T88R/T101E/N108E*.

### 6-5. Supporting information.

Table S1. The ion-ion interaction energy at targeted charged residues of *Ec0VV* mutants. The energy of ion-ion interactions is an average value of energies obtained from 17 structures during 40ns MD simulation. The energy (kJ/mol) was calculated by a software, "AllAtoms Electro" in FoldX.

(A) Pair residues interacting with Asp39 in 3 mutants and its ionic energies

	mutants								
	<i>Ec0VV_A39D/S48K</i>			<i>Ec0VV_6</i>			<i>Ec0VV_9</i>		
	pairs		energy	pairs		energy	pairs		energy
Asp39	LYS48	inter	-14.4	LYS48	inter	-13.5	LYS48	inter	-13.3
	GLU53	inter	0.3	GLU53	inter	0.4	GLU53	inter	0.5
	LYS55	inter	-0.2	LYS55	inter	-0.3	LYS55	inter	-0.5
	GLU57	inter	1.2	GLU57	inter	1.6	GLU57	inter	1.5
	GLU59	inter	2.6	GLU59	inter	3.0	GLU59	inter	2.6
	LYS81	inter	-0.2	LYS81	inter	-0.3	LYS81	inter	-0.4
	HIS84	inter	0.0	HIS84	inter	-0.1			
				LYS87	inter	-0.5	LYS87	inter	-0.4
				ARG88	inter	-5.4	ARG88	inter	-2.5
	GLU90	inter	15.7	GLU90	inter	17.8	GLU90	inter	14.6
		Inter-sum	5.0			2.7			2.0
	GLU21	intra	0.0				GLU21	intra	0.0
							ARG25	intra	-0.8
	ASP26	intra	0.0						
	LYS35	intra	-0.1	LYS35	intra	-0.1	LYS35	intra	-0.2
	LYS67	intra	-22.7	LYS67	intra	-22.1	LYS67	intra	-20.9
	HIS72	intra	-0.1	LYS72	intra	0.0	LYS72	intra	0.0
	GLU90	intra	0.1						
				HIS98	intra	0.0			
	ASP100	intra	1.2	ASP100	intra	1.6	ASP100	intra	1.7
						GLU101	intra	0.0	
ASP102	intra	0.7	ASP102	intra	0.8	ASP102	intra	1.0	
						GLU108	intra	0.6	
ARG112	intra	0.0	ARG112	intra	-0.1	ARG112	intra	0.0	
	Intra-sum	-20.8			-20.0			-18.7	
Total sum		-15.8			-17.3			-16.7	

## (B) Pair residues interacting with Lys48 in 3 mutants and its ionic energies

	mutants									
	<i>Ec0VV_A39D/S48K</i>			<i>Ec0VV_6</i>			<i>Ec0VV_9</i>			
	pairs		energy	pairs		energy	pairs		energy	
Lys48	GLU21	inter	0.0				GLU21	inter	0.0	
							ARG25	inter	0.5	
							LYS35	inter	0.0	
	ASP39	inter	-14.4	ASP39	inter	-13.5	ASP39	inter	-13.3	
	LYS67	inter	5.4	LYS67	inter	3.8	LYS67	inter	3.6	
	ASP100	inter	-0.9	ASP100	inter	-0.9	ASP100	inter	-1.0	
							GLU101	inter	0.0	
	ASP102	inter	-0.6	ASP102	inter	-0.5	ASP102	inter	-0.6	
							GLU108	inter	0.0	
	ARG112	inter	0.0							
		Inter-sum		-10.4			-11.1			-10.8
	ASP20	intra	-0.4	ASP20	intra	-0.6	ASP20	intra	-0.6	
	GLU21	intra	-0.1	GLU21	intra	-0.2	GLU21	intra	-0.2	
							ASP26	intra	0.0	
	GLU53	intra	-0.2	GLU53	intra	-0.1	GLU53	intra	-0.2	
	LYS55	intra	0.2	LYS55	intra	0.2	LYS55	intra	0.5	
	GLU57	intra	-1.0	GLU57	intra	-1.3	GLU57	intra	-1.4	
GLU59	intra	-6.9	GLU59	intra	-10.1	GLU59	intra	-7.8		
LYS81	intra	0.2	LYS81	intra	0.3	LYS81	intra	0.3		
HIS84	intra	0.0	HIS84	intra	0.2					
						LYS87	intra	0.5		
						ARG88	intra	2.1		
GLU90	intra	-19.5	GLU90	intra	-19.9	GLU90	intra	-20.7		
	Intra-sum		-27.6			-28.1			-27.3	
Total sum			-38.0			-39.2			-38.2	

(C) Pair residues interacting with Lys72 in 3 mutants and its ionic energies

	mutants										
	<i>Ec0VW_H72K</i>			<i>Ec0VW_6</i>			<i>Ec0VW_9</i>				
	pairs		energy	pairs		energy	pairs		energy		
Lys72	N-terminal	inter	0.1	N-terminal	inter	0.3	N-terminal	inter	0.0		
	ASP3	inter	-0.2	ASP3	inter	-0.2	ASP3	inter	-0.2		
	GLU4	inter	-0.3	GLU4	inter	-0.2	GLU4	inter	-1.0		
	LYS5	inter	0.0	LYS5	inter	0.0	LYS5	inter	0.2		
	GLU53	inter	-0.7	GLU53	inter	-0.7	GLU53	inter	-0.6		
	Inter-sum			-1.0				-0.7			-1.6
	LYS30	intra	0.0	LYS30	intra	0.0	LYS30	intra	0.0		
	GLU34	intra	-1.5	GLU34	intra	-1.7	GLU34	intra	-2.0		
	LYS35	intra	1.3	LYS35	intra	1.5	LYS35	intra	1.7		
				ASP39	intra	0.0	ASP39	intra	0.0		
	GLU78	intra	-0.8	GLU78	intra	-0.8	GLU78	intra	-0.7		
				LYS82	intra	0.2	LYS82	intra	0.2		
	HIS83	intra	0.1	HIS83	intra	0.0	HIS83	intra	0.0		
							GLU108	intra	-0.8		
	C-terminal	intra	-12.5	C-terminal	intra	-16.2	C-terminal	intra	-14.5		
Intra-sum					-16.9				-16.1		
Total sum										-17.7	



## (E) Pair residues interacting with Lys87 in 4 mutants and its ionic energies

	mutants											
	<i>Ec0VW_Q87K</i>			<i>Ec0VW_Q87K_T88R</i>			<i>Ec0VW_6</i>			<i>Ec0VW_9</i>		
	pairs	energy		pairs	energy		pairs	energy		pairs	energy	
Lys87	N-terminal	inter	0.1	N-terminal	inter	0.0	N-terminal	inter	0.0	N-terminal	inter	0.1
	LYS35	inter	0.0	LYS35	inter	0.0	LYS35	inter	0.0	ARG25	inter	0.0
							ASP39	inter	-0.5	ASP39	inter	-0.4
	LYS67	inter	0.1	LYS67	inter	0.2	LYS67	inter	0.0	LYS67	inter	0.0
	ASP100	inter	-0.7	ASP100	inter	-0.5	ASP100	inter	-0.9	ASP100	inter	-0.8
										GLU101	inter	-0.2
	ASP102	inter	-1.4	ASP102	inter	-1.2	ASP102	inter	-2.3	ASP102	inter	-1.4
	ARG112	inter	0.0	ARG112	inter	0.0	ARG112	inter	0.0	ARG112	inter	0.0
		Inter-sum	-1.9			-1.5			-3.5			-2.6
	ASP20	intra	-4.3	ASP20	intra	-0.3	ASP20	intra	-0.4	ASP20	intra	-1.5
	GLU21	intra	-0.4	GLU21	intra	0.0	GLU21	intra	-0.1	GLU21	intra	-0.2
	ASP26	intra	-0.5	ASP26	intra	0.0	ASP26	intra	-0.1	ASP26	intra	-0.2
	LYS30	intra	0.2	LYS30	intra	0.1	LYS30	intra	0.0	LYS30	intra	0.2
	GLU34	intra	0.0	GLU34	intra	0.0	GLU34	intra	0.0	GLU34	intra	-0.1
							LYS48	intra	0.6	LYS48	intra	0.5
	GLU53	intra	-0.3	GLU53	intra	-0.6	GLU53	intra	-0.3	GLU53	intra	-0.3
	LYS55	intra	0.4	LYS55	intra	0.5	LYS55	intra	0.6	LYS55	intra	0.5
	GLU57	intra	-1.5	GLU57	intra	-4.0	GLU57	intra	-1.7	GLU57	intra	-1.2
	GLU59	intra	-3.5	GLU59	intra	-6.3	GLU59	intra	-3.4	GLU59	intra	-2.3
	GLU78	intra	-0.1	GLU78	intra	-0.1	GLU78	intra	-0.1	GLU78	intra	-0.3
	LYS81	intra	0.6	LYS81	intra	0.4	LYS81	intra	0.7	LYS81	intra	0.8
							LYS82	intra	0.1	LYS82	intra	0.4
	HIS83	intra	0.0	HIS83	intra	0.1	HIS83	intra	0.0			
	HIS84	intra	0.0				HIS84	intra	0.0			
				ARG88	intra	3.0	ARG88	intra	3.1	ARG88	intra	2.9
	GLU90	intra	-0.7	GLU90	intra	-0.8	GLU90	intra	-0.7	GLU90	intra	-0.8
		Intra-sum	-10.1			-8.1			-1.6			-1.7
Total sum		-12.0			-9.6			-5.1			-4.3	

(F) Pair residues interacting with Arg88 in 4 mutants and its ionic energies.

	mutants											
	<i>Ec0VV_T88R</i>			<i>Ec0VV_Q87K/T88R</i>			<i>Ec0VV_6</i>			<i>Ec0VV_9</i>		
	pairs	inter	energy	pairs	inter	energy	pairs	inter	energy	pairs	inter	energy
Arg88	N-terminal	inter	0.2	N-terminal	inter	0.4	N-terminal	inter	0.1	N-terminal	inter	0.3
										ASP3	inter	0.0
										GLU4	inter	-0.2
										LYS5	inter	0.0
	LYS35	inter	0.2	LYS35	inter	0.1	LYS35	inter	0.1	LYS35	inter	0.1
							ASP39	inter	-5.4	ASP39	inter	-2.5
	LYS67	inter	0.9	LYS67	inter	1.4	LYS67	inter	1.3	LYS67	inter	0.7
							HIS98	inter	0.0			
	ASP100	inter	-5.0	ASP100	inter	-5.5	ASP100	inter	-5.4	ASP100	inter	-7.8
										GLU101	inter	-1.9
	ASP102	inter	-16.0	ASP102	inter	-12.4	ASP102	inter	-13.8	ASP102	inter	-13.0
										GLU108	inter	-0.8
	ARG112	inter	0.5	ARG112	inter	0.6	ARG112	inter	1.0	ARG112	inter	0.7
	C-terminal	inter	0.0	C-terminal	inter	-0.1	C-terminal	inter	0.0			
		Inter-sum	-19.2			-15.4			-22.1			-24.4
	ASP20	intra	0.0	ASP20	intra	0.0	ASP20	intra	-0.1	ASP20	intra	-0.2
							ASP26	intra	0.0	ASP26	intra	-0.1
							LYS30	intra	0.1	LYS30	intra	0.2
										GLU34	intra	-0.1
							LYS48	intra	2.9	LYS48	intra	2.1
	GLU53	intra	-2.3	GLU53	intra	-1.7	GLU53	intra	-1.9	GLU53	intra	-1.3
LYS55	intra	0.9	LYS55	intra	0.8	LYS55	intra	0.9	LYS55	intra	0.9	
GLU57	intra	-2.0	GLU57	intra	-2.7	GLU57	intra	-1.8	GLU57	intra	-1.0	
GLU59	intra	-2.7	GLU59	intra	-2.7	GLU59	intra	-2.6	GLU59	intra	-1.7	
GLU78	intra	-0.2	GLU78	intra	-0.2	GLU78	intra	-0.2	GLU78	intra	-0.9	
LYS81	intra	2.5	LYS81	intra	2.7	LYS81	intra	2.5	LYS81	intra	3.8	
						LYS82	intra	0.3	LYS82	intra	0.6	
HIS83	intra	0.0							HIS83	intra	0.0	
						HIS84	intra	0.1				
						LYS87	intra	3.1	LYS87	intra	2.9	
GLU90	intra	-3.6	GLU90	intra	-5.5	GLU90	intra	-5.1	GLU90	intra	-4.9	
	Intra-sum	-7.5			-6.3			-1.8			0.2	
Total sum		-26.7			-21.7			-23.9			-24.2	

(G) Pair residues interacting with Arg25 in 2 mutants and its ionic energies

	mutants						
	<i>Ec</i> 0VV_Q25R			<i>Ec</i> 0VV_9			
	pairs		energy	pairs		energy	
Arg25	ASP20	inter	0.0				
	GLU21	inter	-0.2	GLU21	inter	-0.2	
				LYS48	inter	0.5	
	LYS55	inter	0.3	LYS55	inter	0.2	
	GLU57	inter	-2.6	GLU57	inter	-1.9	
	GLU59	inter	-1.5	GLU59	inter	-1.4	
				LYS87	inter	0.0	
	GLU90	inter	-0.6	GLU90	inter	-0.2	
		Inter-sum		-4.6			-3.1
	ASP20	intra	-2.6	ASP20	intra	-2.8	
	GLU21	intra	-5.2	GLU21	intra	-5.6	
	ASP26	intra	-5.2	ASP26	intra	-4.7	
	LYS30	intra	0.7	LYS30	intra	0.7	
	GLU34	intra	-0.1	GLU34	intra	-0.1	
	LYS35	intra	0.0				
	GLU59	intra	0.0	ASP39	intra	-0.8	
	LYS67	intra	0.2	LYS67	intra	0.2	
	HIS83	intra	0.1				
	GLU90	intra	0.0	GLU90	intra	0.0	
	Intra-sum		-12.1			-13.0	
Total sum			-16.6			-16.1	

## (H) Pair residues interacting with Glu101 in 2 mutants and its ionic energies

	mutants						
	<i>Ec0VV_T101E</i>			<i>Ec0VV_9</i>			
	pairs		energy	pairs		energy	
Glu101				LYS48	inter	0.0	
				GLU53	inter	0.0	
		GLU78	inter	2.6	GLU78	inter	2.4
		LYS81	inter	-2.3	LYS81	inter	-3.6
					LYS82	inter	-0.4
					LYS87	inter	-0.1
					ARG88	inter	-1.9
		GLU90	inter	0.0	GLU90	inter	0.2
		Inter-sum		0.3			-3.4
		N-terminal	intra	-11.2	N-terminal	intra	-6.1
		ASP3	intra	2.1	ASP3	intra	3.8
		GLU4	intra	1.1	GLU4	intra	2.7
		LYS5	intra	-0.2	LYS5	intra	-0.5
		LYS67	intra	0.0	ASP39	intra	0.0
					LYS67	intra	0.0
					HIS98	intra	-0.5
		ASP100	intra	5.5	ASP100	intra	6.3
		ASP102	intra	4.4	ASP102	intra	5.3
		ARG112	intra	-0.3	GLU108	intra	2.5
		C-terminal	intra	0.1	ARG112	intra	-0.4
	Intra-sum		1.5			13.3	
Total sum			1.7			9.9	

## (I) Pair residues interacting with Glu108 in 2 mutants and its ionic energies

	mutants						
	<i>Ec0VW_N108E</i>			<i>Ec0VW_9</i>			
	pairs		energy	pairs		energy	
Glu108				ASP3	inter	0.0	
	GLU4	inter	0.4	GLU4	inter	0.5	
	LYS5	inter	-0.2	LYS5	inter	-0.4	
				LYS48	inter	0.0	
	GLU53	inter	0.8	GLU53	inter	0.5	
	LYS81	inter	-0.1	LYS81	inter	-0.2	
				ARG88	inter	-0.8	
	GLU90	inter	0.0	GLU90	inter	0.0	
				HIS98	inter	0.0	
		Inter-sum		1.0			-0.3
		N-terminal	intra	-0.1	N-terminal	intra	-4.6
		ASP3	intra	0.2	ASP3	intra	0.6
		GLU4	intra	0.2	GLU4	intra	0.2
		LYS5	intra	-0.1	LYS5	intra	0.0
		GLU34	intra	0.0			
		LYS35	intra	-0.4	LYS35	intra	-0.4
					ASP39	intra	0.6
		LYS67	intra	-0.1	LYS67	intra	-0.2
		HIS72	intra	-0.5	LYS72	intra	-0.8
		GLU78	intra	0.0			
		HIS98	intra	0.0	HIS98	intra	-0.1
		ASP100	intra	0.9	ASP100	intra	1.4
					GLU101	intra	2.5
	ASP102	intra	1.3	ASP102	intra	1.5	
	ARG112	intra	-2.4	ARG112	intra	-3.9	
	C-terminal	intra	0.8	C-terminal	intra	0.1	
		Intra-sum				-3.2	
Total sum			0.9			-3.5	

Table S2. The average distance of ion-ion interaction in 40 ns MD trajectory at 300 K\*.

(A) *Ec*0VV template

ion pairs		inter or intra-subunit	AA or AB-subunit	BB or BC-subunit	CC or CA-subunit
N-terminal	Glu78	intra		2.9 ± 1.1	3.8 ± 0.8
Asp3	Lys5	intra	6.4 ± 1.2	5.7 ± 1.1	5.9 ± 1.0
Glu4	Lys5	intra	6.8 ± 2.1	5.1 ± 2.1	6.2 ± 1.9
Asp26	Lys30	intra	3.3 ± 1.8	3.4 ± 1.4	3.3 ± 1.0
Lys30	Glu34	intra		7.4 ± 2.0	7.8 ± 1.5
Glu34	Lys35	intra	7.0 ± 2.1		
Lys35	Glu50	inter		4.7 ± 0.3	
Glu53	Lys55	intra	8.0 ± 2.4		3.3 ± 4.1
Glu53	Arg112	inter	7.2 ± 3.8	6.3 ± 2.2	6.9 ± 3.8
Lys55	Glu57	intra	6.2 ± 2.0	5.0 ± 1.7	4.4 ± 2.0
Lys67	Glu90	intra	2.3 ± 0.4	3.9 ± 0.7	2.6 ± 0.4
Glu78	Lys81	intra	7.6 ± 2.0	5.5 ± 2.0	
Asp100	Lys81	inter	2.4 ± 0.3	2.8 ± 0.7	4.0 ± 1.1
Asp102	Lys81	inter	3.2 ± 0.7	4.2 ± 1.0	7.0 ± 1.8

\*The listed interactions show average values of favorable interactions less than 8.0 Å.

(B) *Ec0VV\_6* mutant

ion pairs		inter or intra-subunit	AA or AB-subunit	BB or BC-subunit	CC or CA-subunit
N-terminal	Glu78	inter		3.1 ± 1.1	
Asp3	Lys5	intra	6.2 ± 1.0	5.9 ± 1.2	5.9 ± 1.1
Glu4	Lys5	intra	5.8 ± 1.8	6.5 ± 2.2	6.2 ± 2.1
Asp26	Lys30	intra	3.3 ± 1.3	3.6 ± 1.7	3.7 ± 1.5
Lys30	Glu34	intra			7.4 ± 2.1
Glu34	Lys35	intra	7.2 ± 1.6	7.5 ± 1.6	
Lys35	Glu53	intra	5.8 ± 2.2	6.6 ± 2.6	6.5 ± 2.2
Asp39	Lys48	intrer	3.4 ± 0.7	4.1 ± 1.2	5.4 ± 0.8
Asp39	Lys67	intra	2.2 ± 0.2	2.5 ± 0.3	2.5 ± 0.4
Asp39	Arg88	inter		4.6 ± 2.2	7.6 ± 2.0
Lys48	Glu59	intra	3.9 ± 0.9	3.0 ± 0.7	4.8 ± 1.3
Lys48	Glu90	intra	2.6 ± 0.5	2.8 ± 0.5	3.1 ± 0.9
Glu53	Arg112	inter	5.5 ± 2.4	6.3 ± 2.8	7.2 ± 2.8
Lys55	Glu57	intra	5.2 ± 2.1	5.1 ± 2.1	5.8 ± 2.3
Glu59	Arg88	intra		7.1 ± 2.1	
Lys67	Glu90	inter	3.5 ± 1.2	3.6 ± 1.1	2.8 ± 0.6
Lys72	C-terminal	intra	2.6 ± 0.8	2.9 ± 1.1	3.1 ± 1.1
Glu78	Lys82	intra	4.6 ± 1.5	4.4 ± 1.3	4.5 ± 1.5
Lys81	Glu78	intra	7.0 ± 2.2	7.5 ± 2.0	8.0 ± 1.9
Lys81	Asp100	inter	2.7 ± 0.6	2.7 ± 0.8	2.6 ± 0.7
Lys81	Asp102	inter	4.2 ± 1.3	3.8 ± 1.1	3.8 ± 1.1
Arg88	Glu90	intra		5.1 ± 2.1	
Arg88	Asp100	inter	5.6 ± 1.6	6.2 ± 1.1	5.9 ± 1.0
Arg88	ASp102	inter	4.5 ± 2.0	6.1 ± 2.1	4.4 ± 1.7

\*The listed interactions show average values of favorable interactions less than 8.0 Å. *Ec0VV\_6* is the mutant *Ec0VV\_A39D/S48K/H72K/S82K/Q87K/T88R*.

(C) *Ec0VV\_9* mutant

ion pairs		inter or intra-subunit	AA or AB-subunit	BB or BC-subunit	CC or CA-subunit
N-terminal	Asp3	intra	7.9 ± 0.9		7.3 ± 0.9
N-terminal	Glu78	intra	4.3 ± 1.7	7.1 ± 1.2	
		inter	4.4 ± 1.0	4.0 ± 1.2	
N-terminal	Asp100	intra		7.3 ± 1.4	7.2 ± 1.5
		inter			7.6 ± 1.5
N-terminal	Glu101	intra	7.3 ± 2.8	4.0 ± 0.9	6.1 ± 2.3
		inter	6.4 ± 2.6	3.3 ± 0.6	6.7 ± 2.2
N-terminal	Glu108	intra	6.7 ± 4.7		
		inter	7.1 ± 4.4		
Asp3	Lys5	intra	6.4 ± 1.1	6.2 ± 1.1	5.8 ± 1.3
Glu4	Lys5	intra	5.8 ± 2.1	5.7 ± 2.2	7.1 ± 2.1
Asp20	Arg25	intra	7.9 ± 1.1	7.8 ± 1.1	7.5 ± 1.1
Glu21	Arg25	intra	4.6 ± 1.9	4.7 ± 2.3	5.3 ± 1.6
Arg25	Asp26	intra	7.7 ± 1.3	7.5 ± 1.2	7.5 ± 1.3
Asp26	Lys30	intra	3.4 ± 1.3	3.2 ± 1.0	3.4 ± 1.1
Lys30	Glu34	intra			7.4 ± 1.4
Glu34	Lys82	intra			7.9 ± 2.2
Lys35	Glu53	inter	5.2 ± 2.5	6.5 ± 2.5	5.5 ± 2.7
Asp39	Lys48	inter	4.6 ± 1.4	3.2 ± 0.7	3.9 ± 1.0
Asp39	Arg88	inter	7.7 ± 2.1	7.6 ± 1.5	
Lys48	Glu59	intra	3.6 ± 1.2	4.6 ± 1.6	4.6 ± 1.5
Lys48	Glu90	intra	2.4 ± 0.3	2.8 ± 0.7	2.4 ± 0.5
Glu53	Arg112	inter		7.0 ± 3.3	
Lys67	Asp39	intra	2.6 ± 0.4	2.5 ± 0.4	2.3 ± 0.3
Lys67	Glu90	inter	5.3 ± 0.8	2.7 ± 0.6	4.7 ± 1.3
Lys72	C-terminal	intra	2.9 ± 0.6	3.1 ± 0.7	3.4 ± 1.4
Glu78	Lys81	intra	5.4 ± 2.3	6.2 ± 2.1	6.4 ± 1.9
Glu78	Lys82	intra	4.8 ± 1.6	4.6 ± 1.5	4.8 ± 1.8
Lys81	Asp100	inter	2.8 ± 0.6	2.3 ± 0.4	3.0 ± 0.7
Lys81	Glu101	inter	5.5 ± 1.4	7.6 ± 1.1	6.4 ± 1.6
Lys81	Asp102	inter	4.6 ± 1.1	3.3 ± 0.9	5.4 ± 1.7
Arg88	Glu90	intra	4.8 ± 1.5		
Arg88	Asp100	inter	5.1 ± 1.5	5.3 ± 1.1	5.9 ± 1.6
Arg88	Asp102	inter	5.2 ± 1.9	3.6 ± 1.3	6.1 ± 2.9
Glu108	Arg112	intra	7.1 ± 3.6	5.8 ± 1.9	6.5 ± 3.4

\*The listed interactions show average values of favorable interactions less than 8.0 Å.

*Ec0VV\_9* is the mutant

*Ec0VV* Q25R/A39D/S48K/H72K/S82K/Q87K/T88R/T101E/N108E.

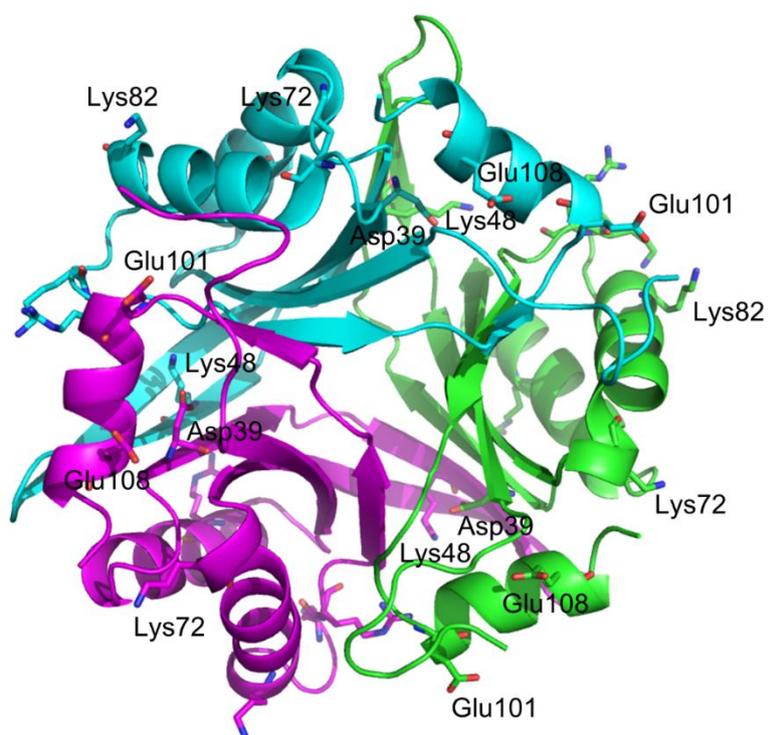


Fig. S1. The snapshot of the trimer structure of *Ec0VV\_9* at 20 ns of 300 K MD simulation. Sticks represent the 9 targeted charged residues. The different colors represent the different subunits.

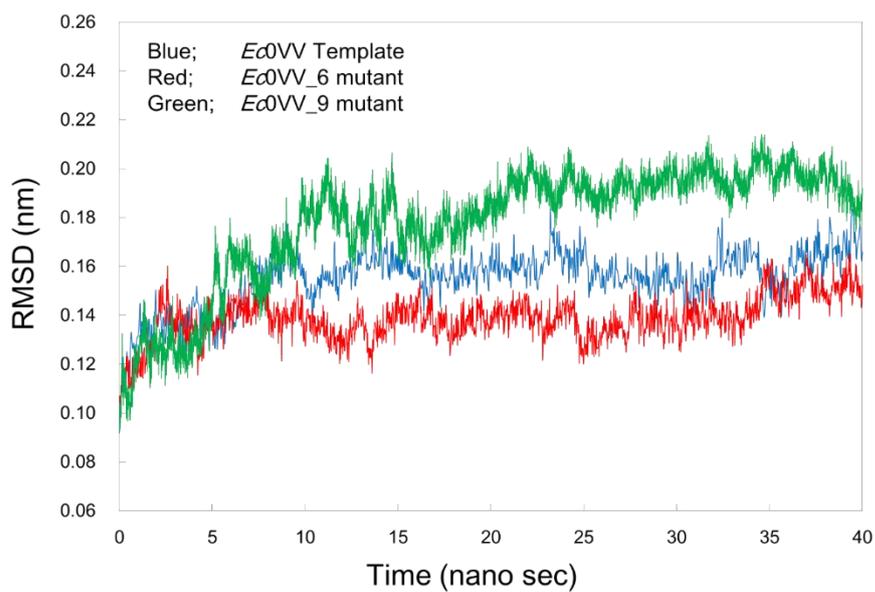


Fig. S2. MD trajectory of root mean square deviation (RMSD) of ionic mutant *EcCutA1* for 40ns at 300 K. Blue, red, and green represent *Ec0VV* template, *Ec0VV\_6* and *Ec0VV\_9*, respectively.

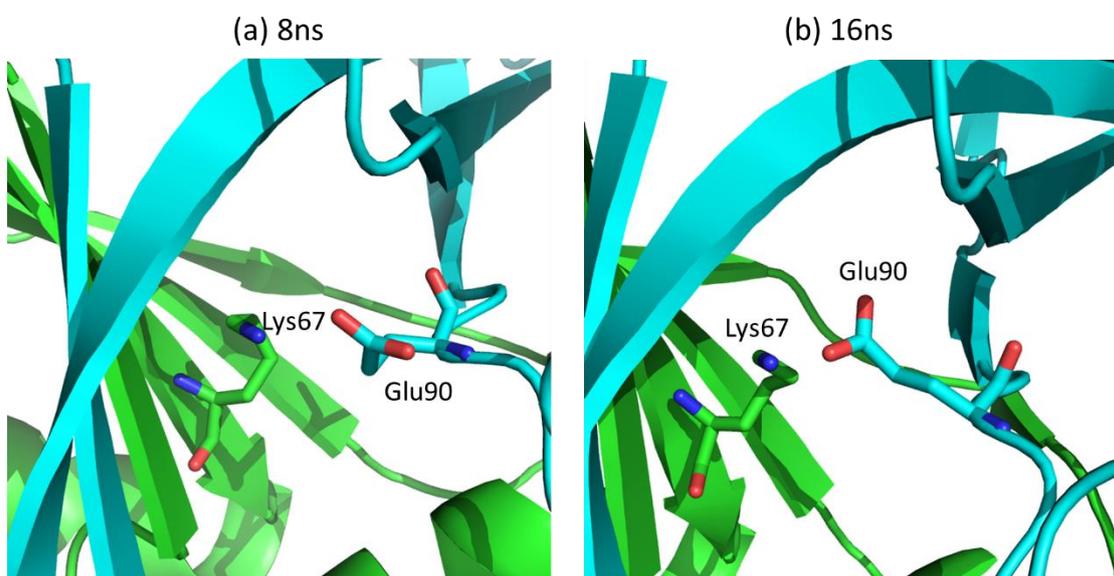


Fig. S3. The configuration of charged residues, Lys67 and Glu90 at (a) 8ns and (b) 16 ns of MD simulation of *Ec0VV\_6*.

## 6-6. References.

1. Matsuura, Y., Ota, M., Tanaka, T., Takehira, M., Ogasahara, K., Bagautdinov, B., Kunishima, N., and Yutani, K. (2010) Remarkable improvement in the heat stability of CutA1 from *Escherichia coli* by rational protein design. *J. Biochem.* **148**, 449-458.
2. Matsuura, Y., Takehira, M., Joti, Y., Ogasahara, K., Tanaka, T., Ono, N., Kunishima, N., and Yutani, K. (2015) Thermodynamics of protein denaturation at temperatures over 100 °C: CutA1 mutant proteins substituted with hydrophobic and charged residues. *Sci. Rep.* **5**, 1 5545.
3. Tanaka, T., Sawano, M., Ogasahara, K., Sakaguchi, Y., Bagautdinov, B., Katoh, E., Kuroishi, C., Shinkai, A., Yokoyama, S., and Yutani, K. (2006) Hyperthermostability of CutA1 protein, with a denaturation temperature of nearly 150 °C. *FEBS Lett.* **580**, 4224-4230.
4. Thomas A. S. and Elcock A. H. (2004) Molecular simulations suggest protein salt bridges are uniquely suited to life at high temperatures. *J. Am. Chem. Soc.* **126**, 2208-2014.
5. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411-2423.
6. Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. (2005) The FoldX web server: an online force field. *Nucleic Acids Res.* **33**, W382-W388.
7. Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* **4**, 435-447.
8. van der Spoel, D., Lindahl, E., Hess, B., van Buuren, A. R., Apol, E., Meulenhoff, P. J., Tieleman, D. P., Sijbers, A. L. T. M., Feenstra, K. A., van Drunen, R. and Berendsen, H. J. C. (2010) Gromacs User Manual version 4.5.4, [www.gromacs.org](http://www.gromacs.org)
9. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N·log (N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089-10092.
10. Berendsen, H. J. C., Grigera, J. R., and Straatsma, T. P. (1987) The missing term in effective pair potentials *J. Phys. Chem.* **91**, 6269-6271.

11. Bussi, G., Donadio, D., and Parrinello, M. (2007) Canonical sampling through velocity rescaling. *J.Chem. Phys.* **126**, 014101.
12. Parrinello, M. and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182-7190.
13. Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations. *J. Comp. Chem.* **18**, 1463-1472.
14. Sawle, L. and Ghosh, K. (2016) Convergence of Molecular Dynamics Simulation of Protein Native States: Feasibility vs Self-Consistency Dilemma. *Chem. Theory and Comput.* **12**, 861-869.
15. Vieille, C. and Zeikus, G. J. (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* **65**, 1-43.
16. Ota, M., Kanaya, S., and Nishikawa, K. (1995) Desk-top analysis of the structural stability of various point mutations introduced into ribonuclease H. *J. Mol. Biol.* **248**, 733-738.
17. Ota, M., Isogai, Y., and Nishikawa, K. (2001) Knowledge-based potential defined for a rotamer library to design protein sequences. *Protein Eng.* **14**, 557-564.
18. Sterner, R. and Liebl, W. (2001) Thermophilic adaptation of proteins. *Crit Rev Biochem Mol Biol.* **36**, 39-106.
19. Makhatadze, G. I., Loladze V. V., Ermolenko, D. N., Chen, X., and Thomas, S. T. (2003) Contribution of surface salt bridges to protein stability: guidelines for protein engineering. *J Mol Biol.* **327**, 1135-1148.
20. Permyakov, S. E., Makhatadze, G. I., Owenius, R., Uversky, V. N., Brooks, C. L., Permyakov, E. A., and Berliner, L. J. (2005) How to improve nature: study of the electrostatic properties of the surface of  $\alpha$ -lactalbumin. *Protein Eng. Des. Sel.* **18**, 425-433.
21. Sadeghi, M., Naderi-Manesh, H., Zarrabi, M., and Ranjbar, B. (2006) Effective factors in thermostability of thermophilic proteins. *Biophys. Chem.* **119**, 256-270.

22. Gribenko, A. V., Patel, M. M., Liu, J., McCallum, S. A., Wang, C., and Makhatadze G. I. (2009) Rational stabilization of enzymes by computational redesign of surface charge-charge interactions. *Proc. Natl. Acad. Sci. USA* **106**, 2601-2606.
23. Schweiker, K. L. and Makhatadze, G. I. (2009) A computational approach for the rational design of stable proteins and enzymes: optimization of surface charge-charge interactions. *Methods Enzymol.* **454**, 175-211.
24. Sawle, L. and Ghosh, K. (2011) How do thermophilic proteins and proteomes withstand high temperature? *Biophys. J.* **101**, 217-227.
25. Pace, N. C., Scholt, J. M., and Grimsley, G. R. (2014) Forces stabilizing proteins. *FEBS Lett.* **588**, 2177-2184.
26. Frauenfelder, H., Petsko, G. A., and Tsernoglou, D. (1979) Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature* **280**, 558-563.
27. Artymiuk, P. J., Blake C. C. F., Grace, D. E. P., Oatley, S. J., Phillips, D. C., and Sternberg, M. J. E. (1979) Crystallographic studies of the dynamic properties of lysozyme. *Nature* **280**, 563-568.
28. Karp, D. A., Stahley, M. R., and Bertrand, G.-M. E. (2010) Conformational consequences of ionization of Lys, Asp, and Glu buried at position 66 in staphylococcal nuclease. *Biochemistry* **49**, 4138-4146.
29. Loladze, V. V. and Makhatadze, G. I. (2011) Energetics of charge-charge interactions between residues adjacent in sequence. *Proteins* **79**, 3494-3499.
30. Sheridan, R. P., Levy, R. M., and Salemme, F. R. (1982)  $\alpha$ -Helix dipole model and electrostatic stabilization of 4- $\alpha$ -helical proteins. *Proc. Natl. Acad. Sci. USA* **79**, 4545-4549.
31. Sali, D., Bycroft, M., and Fersht, A. R. (1988) Stabilization of protein structure by interaction of  $\alpha$ -helix dipole with a charged side chain. *Nature* **335**, 740-743.
32. Aqvist, J., Luecke, H., Quioco, F. A., and Warshel, A. (1991) Dipoles localized at helix termini of proteins stabilize charges. *Proc. Natl. Acad. Sci. USA* **88**, 2026-2030.

33. Ermolenko, D. N., Richardson, J. M., and Makhatadze, G. I. (2003) Noncharged amino acid residues at the solvent-exposed positions in the middle and at the C terminus of the  $\alpha$ -helix have the same helical propensity. *Protein Sci.* **12**, 1169-1176.
34. Suhre, K. and Claverie, J.-M. (2003) Genomic correlates of hyperthermostability, an update. *J. Biol. Chem.* **278**, 17198-171202.

## Chapter 7: Conclusions.

There are few reports addressing how charged residues contribute to the thermostability of a protein at temperatures of over 100 °C.

In Chapter 2, analysis of the stability change of exhaustive mutation results indicate that (1) many Glu and Asp residues of *PhCutA1* seem to be essential for highly efficient interactions with positively charged residues and for generating high electrostatic energy, although they were forced to be partially repulsive to each other, (2) the changes in stability of mutant proteins with a  $T_d$  of around 140 - 150 °C were explained by considering factors important for protein stability and the structural features of mutant sites, and (3) these findings are useful for designing proteins that are stable at temperatures over 100 °C.

In Chapter 3, to compare the strengths of salt bridges among six combinations of force fields, we performed MD simulations using *PhCutA1*. The average strengths of the salt bridges for each positively charged residue did not differ greatly among force fields, but the strengths at specific sites within the structure depended sensitively on the force field used. In the case of the Gromos group, positively charged residues engaged in favorable interactions with many more charged residues than in the other force fields, especially in loop regions; consequently, the apparent strength at each site was lower.

In Chapter 4, to enhance the heat stability of *EcCutA1* ( $T_d = 90$  °C) so that it had comparable stability to *PhCutA1* ( $T_d = 150$  °C), first, we used the stability profile of mutant protein (SPMP). These analyses showed that (1) the stability of *EcCutA1* was remarkably improved by slight substitutions, even though the stability of the wild-type protein was considerably high, (2) remarkable improvements in stability were quantitatively explained based on the newly solved native structure, and (3) SPMP is a

powerful tool to examine substitutions that improve protein stability.

In Chapter 5, to evaluate the thermodynamics of protein denaturation at temperatures over 100 °C, we designed hydrophobic and ionic mutant *EcCutA1* proteins that showed a reversible heat denaturation. Thermodynamic analyses of these mutant proteins indicated that the hydrophobic mutants were stabilized by the accumulation of denaturation enthalpy ( $\Delta H$ ) with no entropic gain from hydrophobic solvation around 100 °C, and that the stabilization due to salt bridges resulted from both the increase in  $\Delta H$  from ion-ion interactions and the entropic effect of the electrostatic solvation over 113 °C. This is the first experimental evidence that has successfully overcome the typical technical difficulties.

In Chapter 6, to elucidate the contribution of charged residues to protein stabilization at temperatures of over 100 °C, we constructed many ionic *EcCutA1* mutants. To evaluate the energy of ion-ion interactions of the mutant proteins, we used the structural ensemble obtained by MD simulation at 300 K. The  $T_d$  of the ionic mutants linearly increased with the increments of the computed energy of the ion-ion interactions for ionic mutant proteins, even to temperatures near 140 °C, suggesting that ion-ion interactions cumulatively contribute to the stabilization of a protein at high temperatures.

In this thesis, a protein thermostabilization strategy using ion-ion interactions at temperatures of over 100 °C has been suggested. These findings have enabled us to provide a thermostable protein useful for protein sciences and biotechnological industries.

## List of publications.

### **Chapter 2.**

Y. Matsuura, M. Takehira, M. Sawano, K. Ogasahara, T. Tanaka, H. Yamamoto, N. Kunishima, E. Katoh, K. Yutani, Role of charged residues in stabilization of *Pyrococcus horikoshii* CutA1, which has a denaturation temperature of nearly 150 °C. *FEBS J.* **279**: 78-90 (2012)

### **Chapter 3.**

Y. Matsuura, Y. Joti, B. Bagautdinov, K. Yutani, Evaluating the strengths of salt bridges in the CutA1 protein using molecular dynamic simulations: a comparison of different force fields. *FEBS Open Bio* **9**: 1939-1956 (2019)

### **Chapter 4.**

Y. Matsuura, M. Ota, T. Tanaka, M. Takehira, K. Ogasahara, B. Bagautdinov, N. Kunishima, K. Yutani, Remarkable improvement in the heat stability of CutA1 from *Escherichia coli* by rational protein design. *J. Biochem.* **148**: 449-458 (2010)

### **Chapter 5.**

Y. Matsuura, M. Takehira, Y. Joti, K. Ogasahara, T. Tanaka, N. Ono, N. Kunishima, K. Yutani, Thermodynamics of protein denaturation at temperatures over 100 °C: CutA1 mutant proteins substituted with hydrophobic and charged residues. *Sci Rep.* **5**: 15545 (2015)

### **Chapter 6.**

Y. Matsuura, M. Takehira, GI. Makhatadze, Y. Joti, H. Naitow, N. Kunishima, K. Yutani, Strategy for stabilization of CutA1 proteins due to ion-ion interactions at temperatures of over 100 °C. *Biochemistry* **57**: 2649–2656. (2018)

## **Acknowledgements.**

The author would like to express his sincerest gratitude to Professor Dr. Kazufumi Takano of the Kyoto Prefectural University for giving him the opportunity to get doctor's degree and critical reading of this manuscript.

The author would like to express his grateful acknowledgment to Dr. Katsuhide Yutani, for his valuable suggestions and discussions throughout in this research.

The author wishes to express his special thanks to all of co-authors of his paper for helpful comments, useful advices and kind friendships.

Finally, the author wishes to express his gratitude to his parent, brother and wife for their encouragements and assistances.

Yoshinori Matsuura