Supplementary material

The structures of the CutA1 proteins from *Thermus thermophilus* and *Pyrococcus horikoshii*: characterization of metal binding sites and metal-induced assembly.

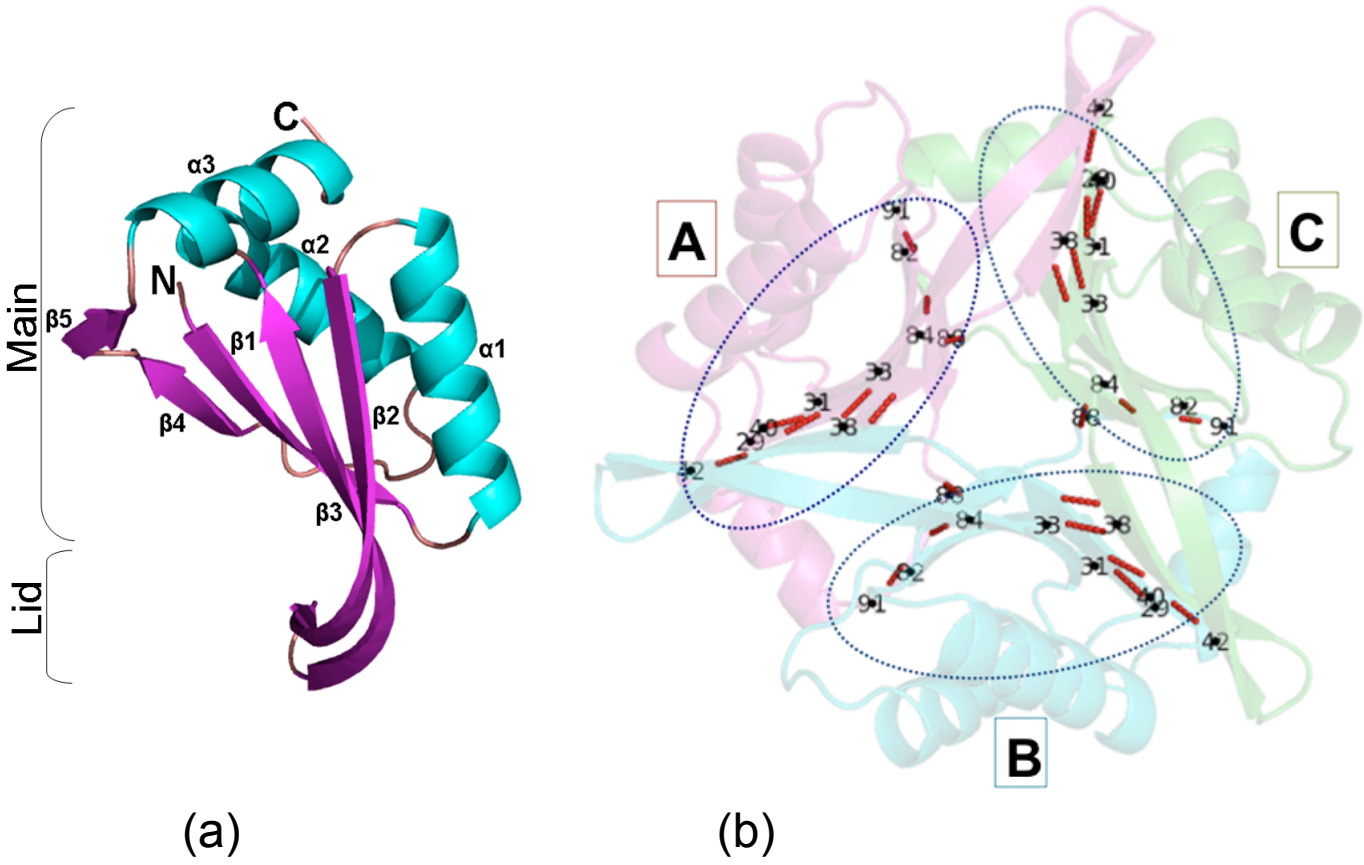
Bagautdin Bagautdinov[a,b*]

[a]Japan Synchrotron Radiation Research Institute (JASRI/SPring-8), 1-1-1 Kouto, Sayo, Hyogo 679-5198, Japan. [b]RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan

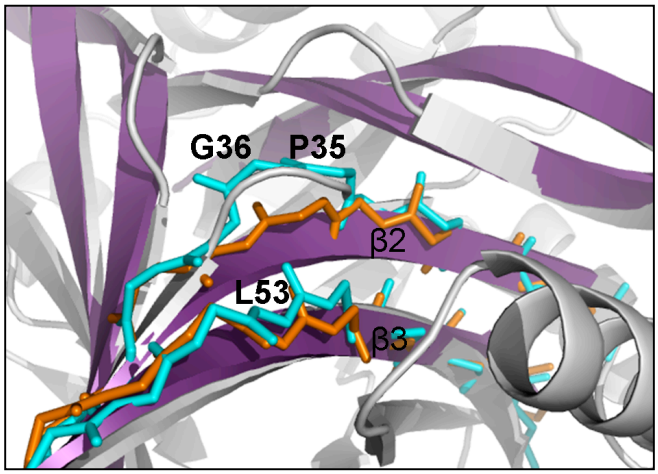
Correspodence e-mail: bagautdi@spring8.or jp

Supplementary Information Content:

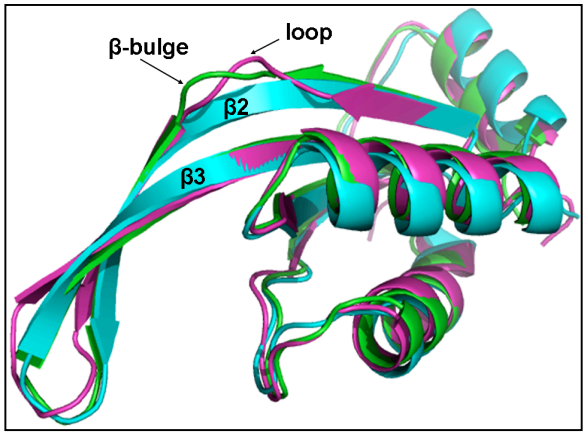
1. Supplementary Figures

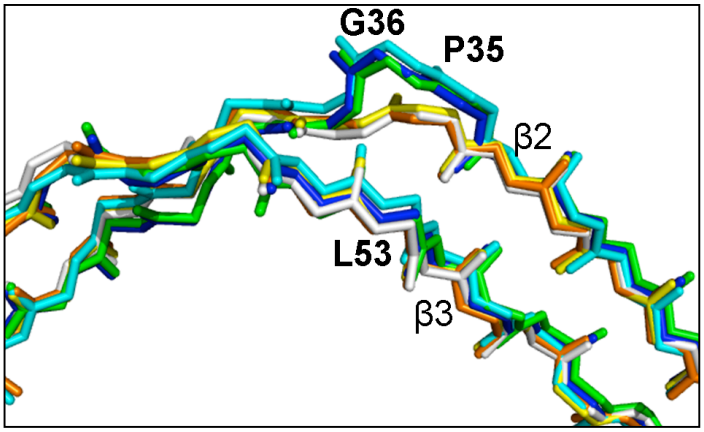


Supplementary Figure S1. (a) Ribbon model of the apo-*Tt*CutA1 protomer structure with indication of the main and lid parts. The α-helices and β-strains are colored in cyan and magenta, respectively. The total molecule has an elongated shape with overall dimensions of 26 x 37 x 50 Å. (b) The apo-*Tt*CutA1 trimer viewed down crystallographic threefold axis. Each protomer is colored differently. The red doted-lines represent the backbone inter-subunit N-H···O=C bonds. The H-bond residue numbers refer to the *Tt*CutA1 protein. The blue doted-circles represent the 7-stranded β-sheets: $(\beta_3\beta_2)_B(\beta_2\beta_3\beta_1\beta_4)_A(\beta_5)_C$, $(\beta_3\beta_2)_C(\beta_2\beta_3\beta_1\beta_4)_B(\beta_5)_A$, and $(\beta_3\beta_2)_A(\beta_2\beta_3\beta_1\beta_4)_C(\beta_5)_B$. In the functional biological unit, three protomers assemble into a trimer resembling a flattened barrel with an overall size of 30 x 50 x 55 Å.



(a)

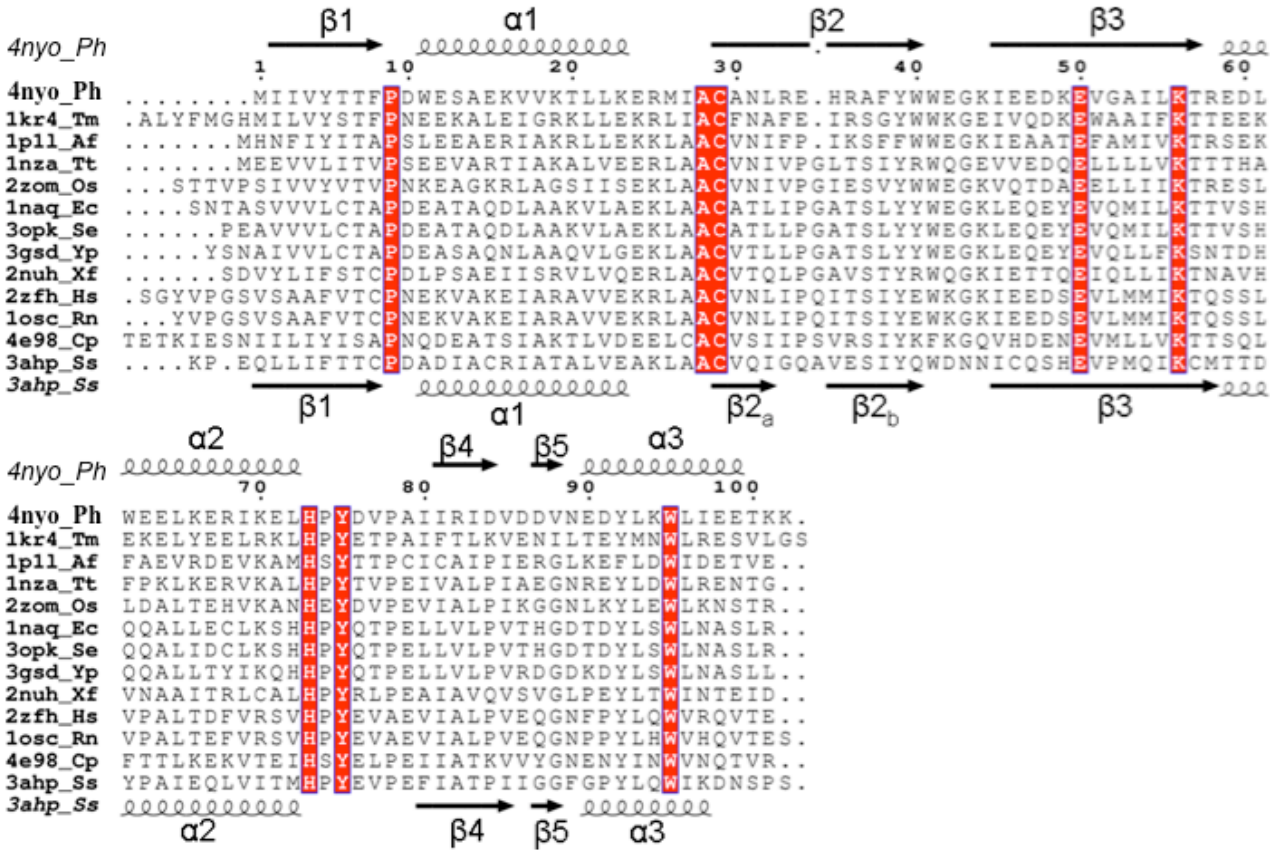




(b) (c)

Supplementary Figure S2. The AW β-bulge region of CutA1. (a) Superposed ribbon representations of *Tt*CutA1 (gray) and *Ph*CutA1 (magenta) of the trimmer tops. The amino acid backbones are shown in cyan (*Tt*CitA1) and in brown (*Ph*CutA1). For clarity the side chains are not shown. The bulge on β_2 of *Tt*CutA1 clearly deviates from the regular arrangement in *Ph*CutA1. (b) Superposition of the monomer structures (ribbon models) of *Tt*CutA1 (green), *Ph*CutA1 (cyan) and psychrotrophic *Ss*CutA1-*Shewanella sp*. Sib1 (magenta). The β_2 strands of *Tt*CutA1 and *Ss*CutA1 enzymes have the irregularity as bulge and loop, respectively. (c) Superposition of the bulged monomers of *Tt*CutA1 (cyan), *Hs*CutA1 (*Homo sapiens*) (blue), *Ec*CutA1 (green) and non-bulged monomers of *Ph*CutA1 (brown), *Tm*CutA1 (Thermotoga maritima) (yellow) and *Af*CutA1 (*Archaeoglobus*

fulgidus) (gray). Their main-chains are shown in sticks. The bulge residues of *Tt*CutA1 are labeled. Despite the presence of considerable differences on the β-sheets, other parts adopt very similar conformation.



Supplementary Figure S3. Structure-based sequence alignment of CutA1. 4nyo_*Ph-Pyrococcus horikoshii*; 1kr4_*Tm-Thermotoga maritima*; 1pll_*Af*C-*Archaeoglobus fulgidus*; 1nza_*Tt-Thermus thermophilus*; 2zom_*Os-Oryza sativa* subsp.; 1naq_*Ec-Escherichia coli*; 3opk_*Se-Salmonella enterica* subsp.; 3gsd_*Yp-Yersinia pestis*; 2nuh_*Xf-Xylella fastidiosa*; 2zfh_*Hs-Homo sapiens*; 1osc_*Rn-Rattus norvegicus*; 4e98_*Cp-Cryptosporidium parvum*; 3ahp_*Ss-Shewanella sp*. Sib1. CutA1 from a different organisms share identical structures while their amino-acid sequences show significant differences.

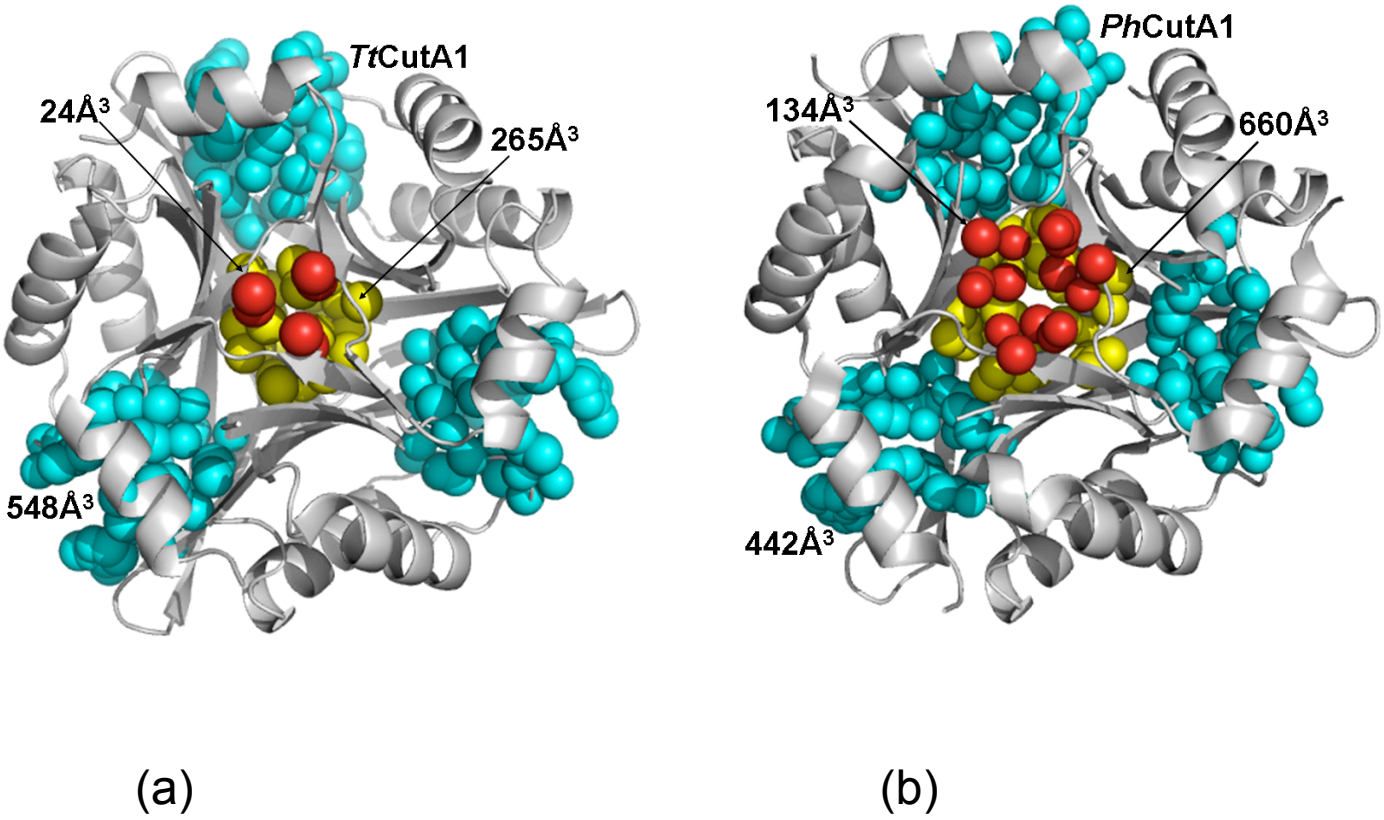
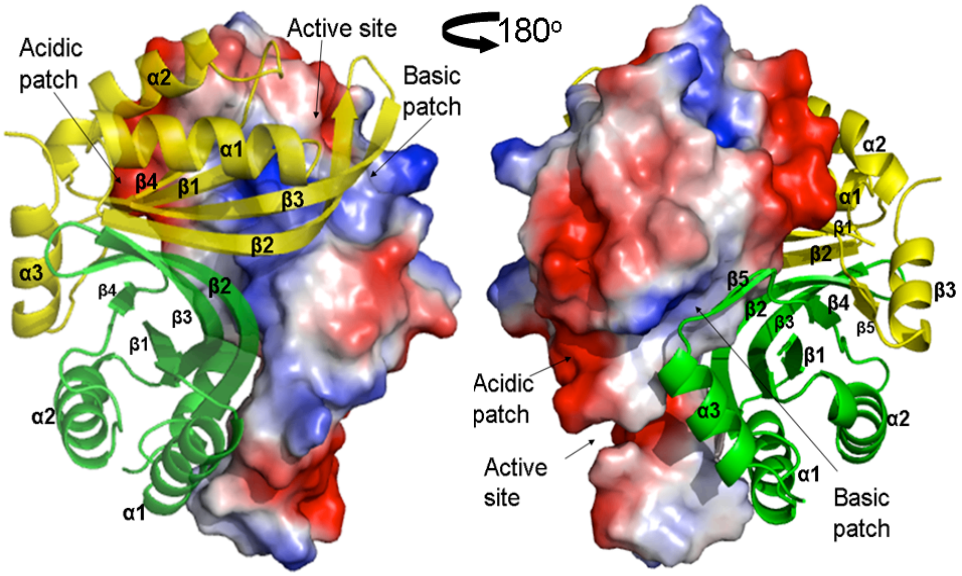
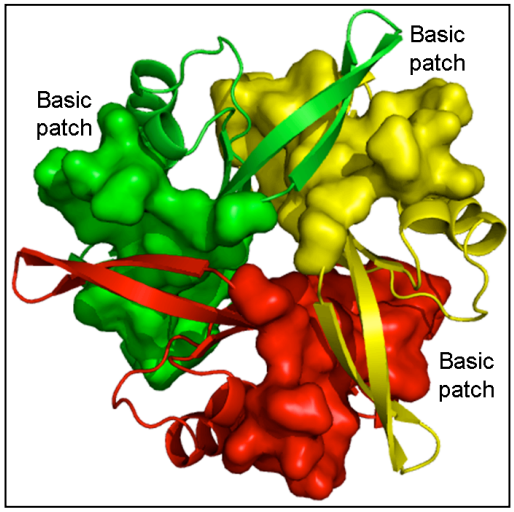


Figure 4. *CASTp* clefts of *Tt*CutA1 (a) and *Ph*CutA1 (b).

The interface pockets are highlighted using spheres presenting all C^{α} atoms of residues surrounding the cleft. Color scheme: side clefts, cyan; central cleft, yellow; bottom entry to the central cleft, red. The overall cleft volumes are presented and the trimeric structures are drawn in gray ribbons. A solvent probe of radius 1.4 Å was used for calculations.



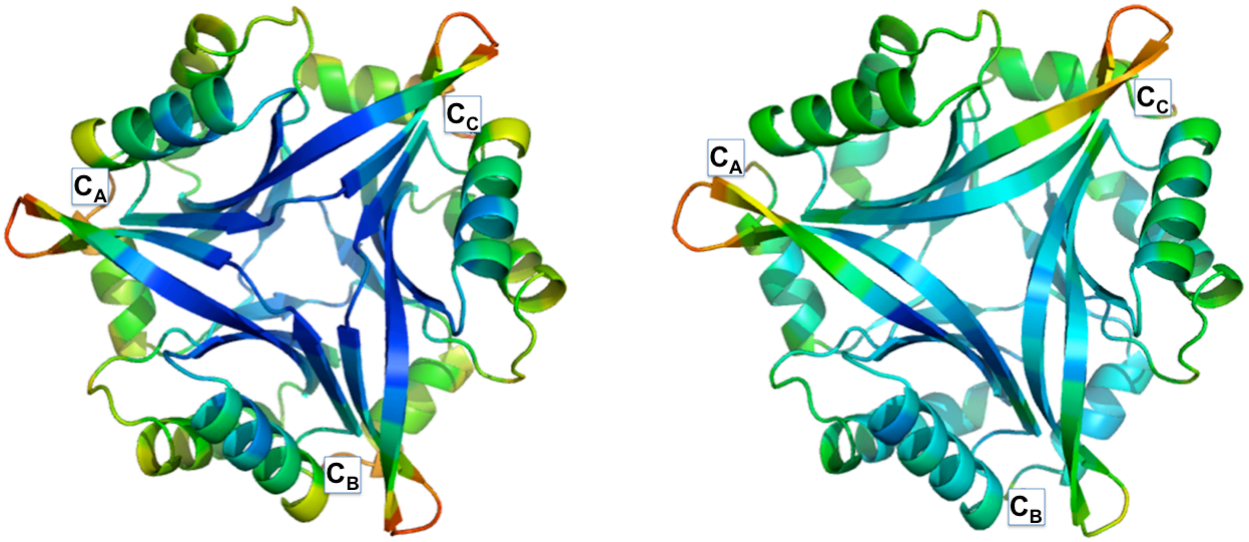
(a)



(b)

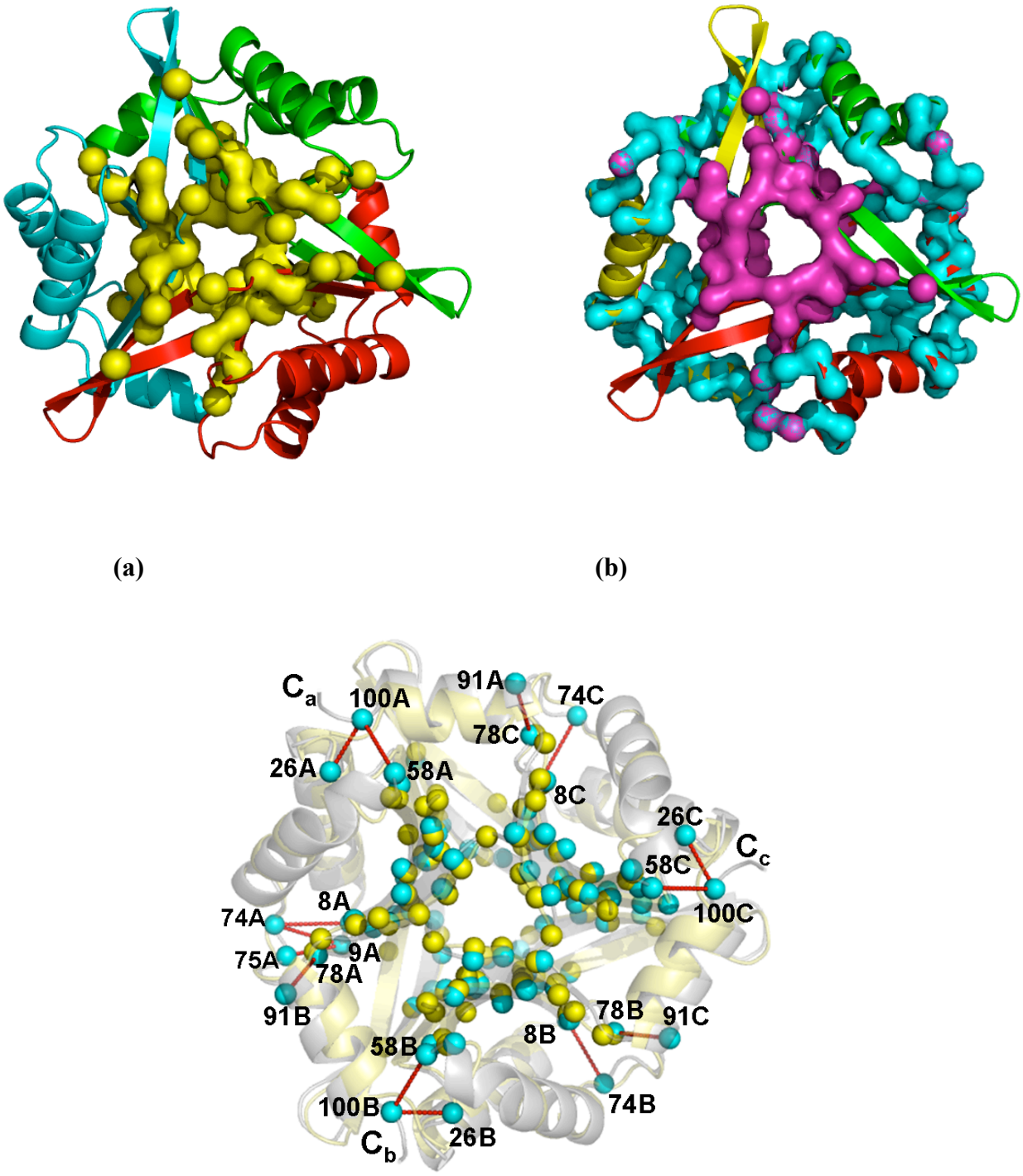
Supplementary Figure S5. Electrostatic and surface complementary of the CutA1 trimer. (a) Surface representations of the isolated *Ph*CutA1 subunit mapped by electrostatic potential at neutral pH; red, blue and gray patches refer to the negatively, positively and neutrally charged regions, respectively. The other subunits are shown as ribbons with highlighted secondary structures. Two views rotated by 180° about the vertical are shown. The molecular concave surface allows each protomer pairs strong dimerization and provides a location for third one. The molecular surface area of each CutA1 molecule is

~6438 Å^2, in which ~35% (~2279 Å^2) of the surface area facing the other two molecules is buried. (b) Surface representation of basic patches of the *Ph*CutA1 subunits. They overlapped by generally acidic $\beta_2\beta_3$–lids shown as ribbons. Each protomer interacts with the other two in two regions having opposite charges, the main part basic patch overlaps with $\beta_2\beta_3$-lid acidic region of the other protomer while its own $\beta_2\beta_3$-lid overlaps with the main part of a third protomer. The presence of contacts between negative and positive surface regions from different subunits stabilizes the trimer.



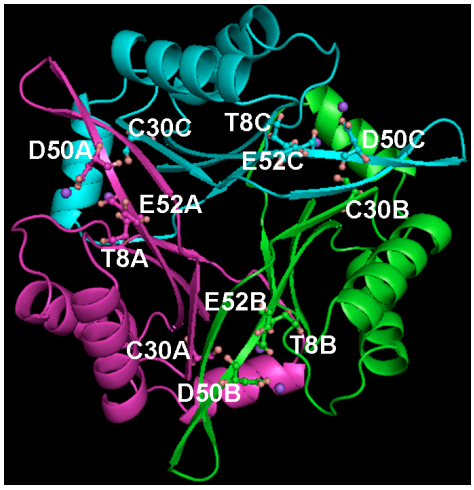
(a) **(b)**

Supplementary Figure S6. The *Tt*CutA1 (a) and *Ph*CutA1 (b) are color-coded by *B* factor from dark blue for low *B*=10 A^2 to red for high *B*=50 A^2. The C-terminus of each subunit polypeptide chain is highlighted. They indicate that the barrel area is well ordered in contrast to $\beta_2\beta_3$-lid, loops and helices which are relatively inherently mobile and characterized by relatively higher *B* factors.

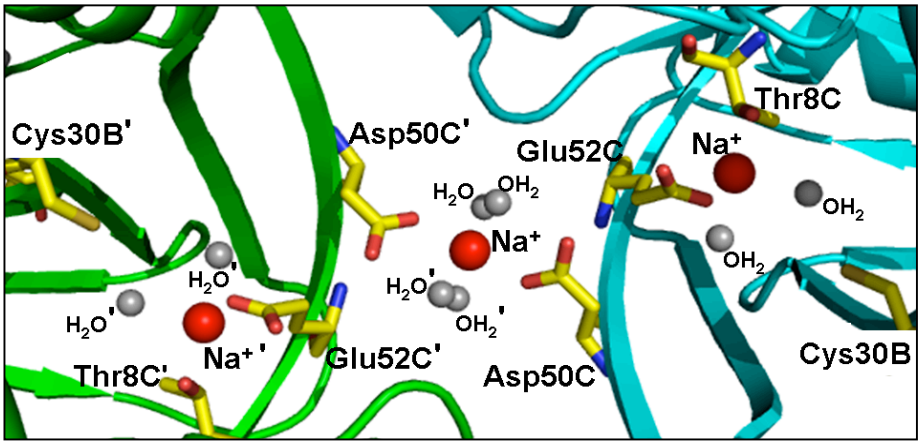


Supplementary Figure S7. Stabilization centers (SC) and SC clusters in *Tt*CutA1 and *Ph*CutA1. The trimers are shown in ribbon representation and chains colored differently. The SC residues and SC cluster residues are shown at the C^{α} positions and indicated by

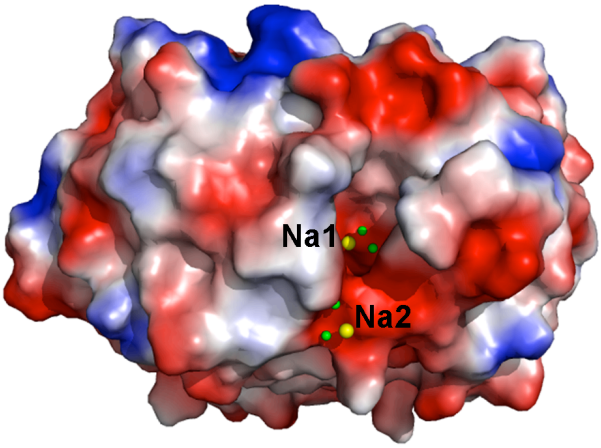
spheres and surfaces, respectively. (a) Distribution of SC clusters within the *Tt*CutA1 structure (yellow). The SC cluster in *Tt*CutA1 mainly occupied central part. (b) Distribution of SC-clusters in the *Ph*CutA1 structure: outer and core clusters are shown in cyan and magenta, respectively. The outer SC clusters of the *Ph*CutA1 trimer are overlapped with the inner core and form interconnected SC cluster. (c) Stabilization centers over the *Tt*CutA1 and *Ph*CutA1 trimers. The structures are superposed and ribbons are drawn in yellow and gray for *Tt*CutA1 and *Ph*CutA1, respectively. Allocation of the C^{α} positions for the SC residues are indicated by spheres and shown in the yellow and cyan colors for *Tt*CutA1 and *Ph*CutA1, respectively. The external-core SC-pairs of *Ph*CutA1 are connected by red lines and the pair's residues are numbered. The C-terminus in the model of each subunit is highlighted. The stabilization-centre (SC) residues and the structural clusters with dense networks of cooperative interactions in the both CutA1 proteins were found mainly in the core β-sheets region. But, *Ph*CutA1 presents extra SC pairs on the outer α-helices and loops some of which form outer-core SC contacts, anchor residues of different parts of the trimer contributing to its closure. Expanding of SC over the entire structure may promote holding the *Ph*CutA1 protein intact for an extended period and/or temperature, potentially to be functional at elevated temperatures. Thus, partial structural and amino acidic modifications in *Ph*CutA1 compared with the *Tt*CutA1 homologue expand and synchronize the densely interacting clusters that are vital for protein stability. For this, *Ph*CutA1 adapts residues (mainly charged amino acid residues) that enhance and improve local interactions generally close to the surface of the protein to exclude any affect on the 3D architecture and function.



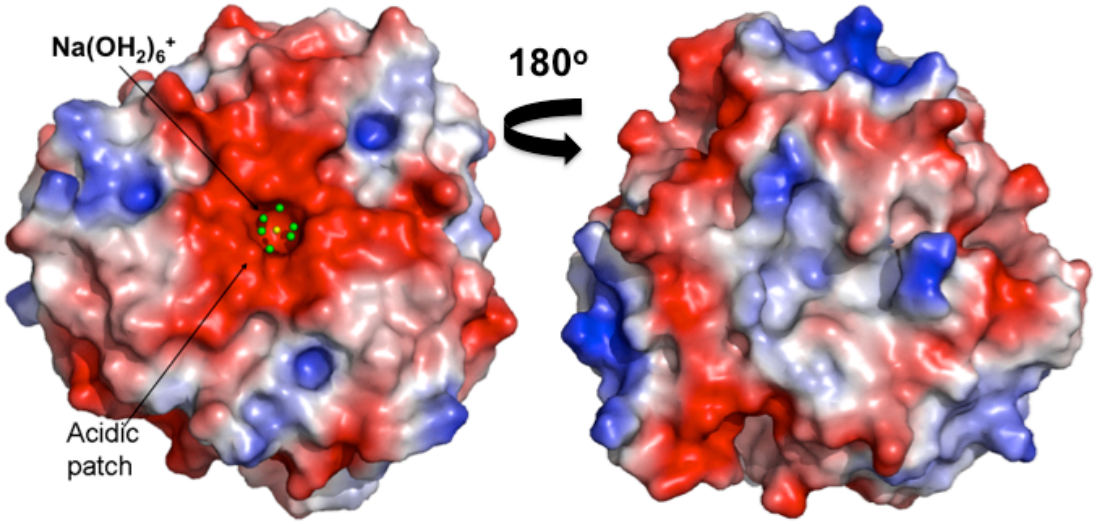
Supplementary Figure S8. The *Tt*CutA1 trimer with individual protomers are drawn in a differ colors (A-magenta, B-green, C-cyan). The residues interacting with sodium ion at Na1 position are highlighted in stick mode. The Cys30 on N-terminal of β_2 of each protomer acts in the cavity formed with triad residues of Thr8, Asp50 and Glu52 of other protomer, while itself triad residues forms cavity with Cy30 of third one.



Supplementary Figure S9. Sodium ions binding in *Tt*CutA1. Na^+ (depicted as red sphere) located inside of the trimer lateral putative active sites close to Glu52, Thr8, Cys30 while other one at the trimer-trimer border interacts with symmetrically related Asp50. The active residues are shown as sticks and labeled. The water molecules H-bonded with Na^+ are shown as small grey balls. Different trimers are depicted in a different color.



(a)



(b)

Supplementary Figure S10. Electrostatic potential at neutral pH distributed on the molecular surface of the isolated Na^+-*Tt*CutA1 (a) and Na^+-*Ph*CutA1 (b) trimers. The basic (dark blue) acidic (red) and hydrophobic (white) regions of surface are shown. The metal ions are shown in yellow balls and the bounded water molecules are shown in green small balls. The Na^+-*Tt*CutA1 trimer is viewed toward to the lateral cleft, and Na^+-*Ph*CutA1 trimer viewed along the crystallographic threefold axis toward to the trimer bottom (with liganded $Na(OH_2)_6^+$) and top, respectively. The negative charges surrounding the clefts are appropriate for the binding of metal ions. In Na^+-*Ph*CutA1, the metal ions are associated with six well-ordered water molecules forming an octahedral coordination complex.

2. Supplemental tables

Supplementary Table S1. Mainchain-mainchain hydrogen-bonding interaction between the A, B and C protomers of the *Tt*CutA1 trimer.

Maximum distance cutoff between contact atoms: 3.5 Å.

Protein atom				Distance (Å)
Donor		Acceptor		
Val31A	N	Ile40 C	O	2.81
Ile 33A	N	Thr38C	O	2.77
Thr38A	N	Ile33B	O	3.11
Ile40A	N	Val31B	O	2.86
Arg42A	N	Ala29B	O	3.01
Ala84A	N	Glu89B	O	2.65
Ala88A	N	Ala84C	O	2.84
Asn91A	N	Ile82C	O	2.70
Val31B	N	Ile40A	O	2.81
Ile33B	N	Thr38A	O	2.76
Thr38B	N	Ile33C	O	3.11
Ile40B	N	Val31C	O	2.86
Arg42B	N	Ala29C	O	3.01
Ala84B	N	Glu89C	O	2.65
Ala88B	N	Ala84A	O	2.84
Asn91B	N	Ile82A	O	2.70
Val31C	N	Ile40B	O	2.81
Ile33C	N	Thr38B	O	2.77
Thr38C	N	Ile33A	O	3.12
Ile40C	N	Val31A	O	2.86
Arg42C	N	Ala29A	O	3.01
Ala84C	N	Glu89A	O	2.65
Ala88C	N	Ala84B	O	2.84
Asn91C	N	Ile82B	O	2.70

Supplemental Table S2. Mainchain-mainchain hydrogen-bonding interaction between the A, B and C subunits of the *Ph*CutA1 trimer.

Maximum distance cutoff between contact atoms: 3.5 Å.

Protein atom				Distance (Å)
Donor		Acceptor		
Ala30A	N	Phe38C	O	3.05
Leu32A	N	Arg36C	O	2.85
Arg36A	N	Leu32B	O	2.77
Phe38A	N	Ala30B	O	3.03
Trp40A	N	Ala28B	O	3.29
Arg82A	N	Asp87B	O	2.82
Asp86A	N	Arg82C	O	2.93
Asp87A	N	Arg82C	O	3.46
Asn89A	N	Ile80C	O	3.26
Ala30B	N	Phe38A	O	2.86
Leu32B	N	Arg36A	O	2.85
Arg36B	N	Leu32C	O	2.82
Phe38B	N	Ala30C	O	2.93
Trp40B	N	Ala28C	O	3.11
Arg82B	N	Asp87C	O	2.77
Asp86B	N	Arg82A	O	2.98
Ala30C	N	Phe38B	O	2.80
Leu32C	N	Arg36B	O	2.79
Arg36C	N	Leu32A	O	2.76
Phe38C	N	Ala30A	O	2.90
Arg82C	N	Asp87A	O	2.86
Asp86C	N	Arg82B	O	3.00
Asp87C	N	Arg82B	O	3.45
Asn89C	N	Ile80B	O	3.27

Supplementary Table S3. β-bulges in CutA1.

The β-bulge region residues in positions X (on the normal strand) and 1 and 2 (on the bulged strand) are presented. The bulge type is described using letter A as the CutA1 bulges involve antiparallel β-strands, the second letter can be C (classic) or W (wide). Classic bulges (AC) and wide bulges (AW) both involve an extra residue (at position 1 or 2) on one β relative to X residue on neighboring strand. The AC β-bulges are conserved in the CutA1 structures where conserved Ala and Cys on β_2 form structurally conserved close pair of backbone H-bonds to the Lys residue on β_3. With the anti-parallel strand arrangement the interacting residues are aligned directly opposite each other to produce strong H-bonds and the small size of the Ala and Cys side chains minimizes repulsive interactions. As result, the AC H-bonds distort the local extended β_3 chain from a classical flat β-sheet. Probably, the accentuated β_3 twisting inducted by the strong AC β-bulge contacts is important for functional positions of the metal-binding residues of CutA1.

	Residues			Bulge
	X	1	2	type
1OSC	*Rattus norvegicus*			
	Val62	Pro44	Gln45	AW
	Lys67	Ala 38	Cys39	AC
2ZFH	*Homo sapiens*			
	Val119	Pro101	Gln102	AW
	Lys124	Ala95	Cys96	AC
2ZOM	*Oryza sativa*			
	Glu61	Pro43	Gly44	AW
	Lys66	Ala37	Cys38	AC
1NAQ	*Escherichia coli*			
	Val62	Pro44	Gly45	AW

	Lys67	Ala38	Cys39	AC
1NZA	*Thermus thermophilus*			
	Leu53	Pro35	Gly36	AW
	Lys58	Ala29	Cys30	AC
3GSD	*Yersinia pestis*			
	Val69	Pro51	Gly52	AW
	Lys74	Ala45	Cys46	AC
3OPK	*Salmonella enterica subsp*			
	Val65	Pro47	Gly48	AW
	Lys70	Ala41	Cys42	AC
4E98	*Cryptosporidium parvum*			
	Val65	Pro47	Ser48	AW
	Lys70	Ala41	Cys4	AC
2NUH	*Xylella fastidiosa*			
	Ile56	Pro38	Gly38	AW
	Lys61	Ala32	Cys33	AC
3AHP	*Shewanella sp*. SIB1			
	Lys62	Ala33	Cys34	AC
1J2V	*Pyrococcus horikoshii*			
	Lys56	Ala28	Cys29	AC
1KR4	*Thermotoga maritima*			
	Lys62	Ala34	Cys35	AC
1P1L	*Archaeoglobus fulgidus*			
	Lys57	Ala29	Cys30	AC

Supplementary Table S4. Cleft volumes (A^3) in CutA1

	Central	Lateral
The structures with AW β-bulge		
*Ec*CutA1	345	888
*Tt*CutA1	256	547
*Os*CutA1	119	492
*Rat*CutA1	98	646
*Hs*CutA1	95	640
The hyperthermophilic CutA1 structures with no AW β-bulge		
*Ph*CutA1:	551	422
*Af*CutA1:	401	320
*Tm*CutA1:	379	425
The psychrophilic CutA1		
*Ss*CutA1	913	521

Supplementary Table S5. Stabilization center (SC) pairs in *Tt*CutA1 and *Ph*CutA1 with secondary structure positions.

*Tt*CutA1	*Ph*CutA1
A-B	A-B
β1 Glu2-Thr60 β3	β1 Ile2-Thr57 β3
β1 Glu3-Thr59 β3	β1 Ile2-Arg58 β3α2
β1 Glu3-Thr59 β3	β1 Ile2-Glu59 β3α2
β1 Glu3-Leu85 β4	β1 Ile2-Ile83 β4
β1 Glu3-Pro86 β4	β1 Ile2-Asp84 β4
β1 Val4-Val57 β3	β1 Ile3-Arg82 β4
β1 Val4-Lys58 β3	β1 Ile3-Ile83 β4
β1 Val4-Leu85 β4	β1 Val4-Ile54 β3
β1 Val5-Leu56 β3	β1 Val4-Ile81 α2β4
β1 Val5-Val57 β3	β1 Tyr5-Ala53 β3
β1 Leu6-Leu55 β3	β1 Tyr5-Ile54 β3
β1 Leu6-Leu56 β3	β1 Tyr5-Ile80 β4
β1 Leu6-Ile82 β4	β1 Phe8-Pro74 α2β4
β1 Val9-Glu52 β3	β1α1 Pro9-Pro74 α2β4
β1 Val9-Leu53 β3	α1β2 Leu26-Thr100 α3
β2 Asn32-Leu55 β3	α1-β2 Ala28-Lys56 β3
β2 Ile33-Leu54 β3	α1β2 Cys29-Leu55 β3
β2 Ile33-Leu55 β3	α1-β2 Cys29-Lys56 β3
β2 Val34- Leu53 β3	β2 Ala30-Leu55 β3
β2 Val34-Leu54 β4	β2 Arg33-Val51 β3
β2 Gly36-Leu53 β3	β2 Arg33-Gly52 β3
β2 Ser39-Glu49 β3	β2 Glu34-Glu50 β3
A-B	β2 Glu34-Val51 β3
β2 Thr38-Asn32 β2	β2 Glu34-Gly52 β3
β2 Ser39-Asn32 β2	β2 His35-Glu50 β3
β2 Pro80-Asn91 β5	β2 His35-Val51 β3
β2 Ile82-Asn91 β5	β2 Ala37-Glu47 β3
β2 Val83-Glu89 β5	β3α2 Arg58-Thr100 α3

β2 Val83-Gly90 β5	A-B
β2 Ala84-Ala88 β5	β2 Arg36-Asn31 β2
β2 Ala84-Glu89 β5	β2 Ala37-Asn31 β2
β2 Leu85-Ile87 β5	β4 Arg82-Asp86 β5
A-C	A-C
β2 Asn32-Thr38 β2	β2 Asn31-Arg36 β2
β2 Asn32-Ser39 β2	β2 Asn31- Ala37 β2
β5 Ile87-Leu85 β4	β4β6 Asp86-Arg82 β4
β5 Ala88-Ala84 β4	α3 Asp91-Pro78 α2β4
β5 Glu89-Val83 β4	
β5 Glu89-Ala84 β4	
β4 Gly80-Val83 β4	
β5 Asn91-Pro80 β4	
β5 Asn91-Ile82 β4	

Supplementary Table S6. Sodium chelation in *Tt*CutA1
Distance cut-off: 2.1-2.9 Å

	Chelating residue	Chelating atom	Bond length (Å)	occ.	B metal	B donor
Na1 (104)	Thr8	OG1	2.29	1.0	14.7	16.17
	Glu52	OE1	2.42	1.0		19.57
	Glu52	OE2	2.41	1.0		23.21
	HOH10	O	2.40	1.0		29.35
	HOH20	O	2.53	1.0		38.41
Na2 (105)	Asp50	OD1	2.22	0.5	3.4	15.98
	HOH1	O	2.36	0.5		15.37
	HOH2	O	2.39	0.5		25.69

Supplementary Table S7. Sodium chelation in *Ph*CutA1
Distance cut-off: 2.1-2.9 Å

Divalent ion	Chelating residue	Chelating atom	Bond length (Å)	B metal	B donor
Na	HOH1	O	2.48/2.42	38.72/39.39	29.04/42.34
	HOH2	O	2.21/2.38		27.67/29.62
	HOH3	O	2.29/2.36		34.57/34.29
	HOH4	O	2.31/2.32		30.22/19.38
	HOH5	O	2.49/2.44		29.97/23.04
	HOH6	O	2.20/2.31		29.74/26.89
	Asp86D	OD1	HOH1	2.59/2.71	
	Asp86E	OD1	HOH2	2.67/2.70	
	Asp84E	O	HOH3	2.39/2.51	
	Asp86E	OD2	HOH3	2.58/2.81	
	Asp84F	O	HOH4	2.46/2.48	
	Asp86F	OD2	HOH4	2.65/2.49	
	Asp86F	OD1	HOH5	2.87/2.65	
	Asp84D	O	HOH6	2.49/2.52	
	Asp86D	OD2	HOH6	2.73/2.60	