# X-Ray Crystal Structure of CutA from Thermotoga maritima at 1.4 Å Resolution

**ARTICLE** *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · NOVEMBER 2003

Impact Factor: 2.63 · DOI: 10.1002/prot.10585 · Source: PubMed



**9 AUTHORS**, INCLUDING:

Alex Savtchenko
Sanford Burnham Prebys Medical Discovery I…
**157** PUBLICATIONS **4,147** CITATIONS
SEE PROFILE

Roman Aleksander Laskowski
European Molecular Biology Laboratory
**99** PUBLICATIONS **30,908** CITATIONS
SEE PROFILE

Andrzej Joachimiak
Argonne National Laboratory
**400** PUBLICATIONS **16,455** CITATIONS
SEE PROFILE

Rongguang Zhang
**131** PUBLICATIONS **8,937** CITATIONS
SEE PROFILE

Published in final edited form as:
*Proteins*. 2004 January 1; 54(1): 162–165. doi:10.1002/prot.10585.

# X-Ray Crystal Structure of CutA From *Thermotoga maritima* at 1.4 Å Resolution

**Alexei Savchenko**[1], **Tatiana Skarina**[1], **Elena Evdokimova**[1], **James D. Watson**[4], **Roman Laskowski**[4], **Cheryl H. Arrowsmith**[1,3], **Aled M. Edwards**[1,2,*], **Andrzej Joachimiak**[5,*], and **Rong-guang Zhang**[5]

[1]Ontario Center for Structural Proteomics, University Health Network, Toronto, Ontario, Canada

[2]Banting and Best Department of Medical Research, Toronto, Ontario, Canada

[3]Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

[4]The Department of Crystallography, Birkbeck College, London

[5]Biosciences Division and Structural Biology Center, Argonne National Laboratory, Argonne, Illinois

## Introduction

The structure of the CutA protein from *Thermotoga maritima* (tmCutA) was determined at 1.4 Å resolution using the Se-Met multiwavelength anomalous diffraction (MAD) technique. This protein (TIGR annotation - TM1056, DNA bases 1,069,580–1,069,885) is conserved in numerous bacteria, archaea and eucarya, including plants and mammals (COG1324, Fig. 1). The CutA *Escherichia coli* homolog—CutA1 (35% ID) is involved in divalent cation homeostasis,[1] while the mammalian homolog—mCutA (40% ID) was found to be associated with cell surface acetylcholinesterase.[2] However, the biological function of the CutA proteins is yet to be determined.

The tmCutA monomer assumes a two-layer α/β α/β sandwich fold (β1β2α1β3β4α3β5β6α5) shared by a large number of proteins (CATH number 3.30.70.120). Two short (β2β5 and two long (β3β4) β-strands arranged in an antiparallel fashion β5↑β2↓β3↑β4↓ form a curved sheet. Three α-helices cover one side of the molecule (Fig. 2a). Each monomer is elongated with overall dimensions approximately 25 × 30× 60 Å (Fig. 2a). The monomers are assembled into a trimer with dimensions 55 × 55 × 35 Å through interaction of the edges of three β-strands (β3 on the top and β5β6 on the bottom) (Fig. 2b). In the trimer, the subunits are related by crystallographic three fold axes. The same oligomeric state was observed for tmCutA in solution by size exclusion experiments (see Methods) and appears to be functionally relevant. In the trimer, there is a cavity that is accessible from outside via three solvent accessible channels. The inner surface of the cavity is lined with several conserved residues including Cys35, which may serve as a potential metal binding site.

Structural homologues of tmCutA were identified using Dali search[3] and SSAP structure comparison program 4. The closest structural matches, with a Dali score of 9.4 Z, were the *E*.



*Correspondence to Aled M. Edwards, Banting and Best Department of Medical Research, 112 College St., Toronto, Ontario, Canada. aled.edwards@utoronto.ca. *Correspondence to: Andrzej Joachimiak, Biosciences Division and Structural Biology Center, Argonne National Laboratory, 9700 South Cass Avenue, Building 202, Argonne, Illinois 60439.
Alexei Savchenko's e-mail is savchenk@uhnres.utoronto.ca

*coli* signal transducer proteins GlnB (PII)[5] and GlnK[6] (Fig. 2c). These two proteins are 67% identical to each other but have insignificant sequence similarity (less than 15% ID) to tmCutA. GlnB and GlnK, which are also trimeric, are involved in maintaining nitrogen homeostasis[7–9] and have been shown to cooperatively bind ATP and α-ketoglutarate.[6,10] The ATP binding sites of GlnB and GlnK are located in clefts between the momoners,[11] but most of the conserved amino acids that mediate ATP binding in PII proteins are not found in tmCutA. The cavity on the side of tmCutA is formed by a number of conserved aromatic (Tyr45, Trp47, Tyr81, and Trp101) and charged residues (Asp54, Glu56, Glu82) (Figs. 1 and 2c), which makes this cleft a strong candidate for a conserved function.

In PII-type proteins, the two central β-strains β3 and β4 are joined by a flexible loop called the T-loop, which is composed of 17 amino acids. The T-loop plays a key role in protein-protein interactions with downstream effector proteins.[12] The β3 and β4 strands in tmCutA, while longer than the corresponding strands in the PII proteins (12 residues in the β3 strand of tmCutA compared with 9 in the β2 strand of PII proteins), are joined by only a two-amino acid (Lys48Gly49) turn and form a hairpin. Although the residues in the C-terminal extension of β3-strand in tmCutA proteins (Tyr45, Trp46, and Trp47, Fig. 1) are conserved among tmCutA homologs, the structure of this region in tmCutA is different than in PII proteins and thus this region of tmCutA may not participate in protein-protein interactions.

CutA is annotated as being involved in metal homeostasis. Several other structural homologs of tmCutA, including the PII proteins, metallochaperone Atx1 (5.9 Z score) and the metal-binding domain of the Menkes copper-transporting ATPase (4.5 Z score) have metal-binding properties.[13,14] These proteins coordinate the metal through a CXXC sequence motif,[15,16] but this motif is not found in CutA. Therefore, tmCutA may represent a new branch of this family that differs from PII and metal binding proteins. The biochemical experiments are under way to define possible ligands and functional partners of CutA proteins.

## Methods

The tmCutA gene was subcloned, expressed and its product purified and screened for crystallization as described previously.[17] Crystals for X-ray diffraction data collection were obtained from hanging drop vapor diffusion conditions containing 2 μl of Se-Met derivative of the protein plus 2 μl of 2 M ammonium sulphate, 0.1 M Tris HCl at pH 8.5 and 2% glycerol, over 2–5 days at 21°C. The crystals were flash frozen with crystallization buffer completed with 25% glycerol.

Diffraction data (Table I) were collected at beamline 19ID of the Advanced Photon Source, Argonne National Laboratory following the approach described earlier.[18] The three-wavelength inversed beam MAD data up to 1.4 Å were collected from one Se-Met labeled protein crystal at 100 K with 3 seconds exposure/1°/frame using a 100 mm crystal-to-detector distance. The total oscillation range was 90°, as predicted using strategy module within HKL2000 suite.[19] All data were processed and scaled with HKL2000 (Table I). The structure was determined by MAD phasing using CNS.[20] The initial model was built automatically using wARP,[21] and refined to 1.4 Å using CNS against the peak data. The final R factor was 0.218, and the free R was 0.241 (Table II). Final refinement permitted us to define the coordinates of all 101 amino acids of tmCutA, seven (ALYFMGH) N-terminal amino acids belonging to His6-tag and two (GS) C-terminal amino acids, which resulted from the cloning artifact.

HPLC size-exclusion chromatography was performed on a Superdex-75 and Superdex-200 columns (10 × 300 mm) pre-equilibrated with 10 mM HEPES pH 7.5, 0.5 M NaCl. The column was calibrated with cytochrome C (12.4 kDa), carbonic anhydrase (29 kDa), bovine serum albumin (66 kDa), alcohol dehydrogenase (150 kDa), β-amylase (200 kDa), and Blue Dextran

(2,000 kDa). A 25-μl of tmCutA protein sample at a 2 mg $ml^{-1}$ concentration or premixed with standard proteins was centrifuged at 14,000 rpm for 10 min before being injected into the column through a 20 μL injection loop. Filtration was carried out at 20°C at a flow rate of 1 ml $min^{-1}$. The eluted proteins were detected by measuring the absorbance at 280 nm.

## Acknowledgments

Atomic coordinates have been deposited in the Protein Data Bank (PDB) with PDB-ID 1KR4 and accession number RCSB015255. We wish to thank all members of the Structural Biology Center at Argonne National Laboratory for their help in conducting these experiments and S. Beasley for help in preparation of this manuscript. CHA is a CIHR Investigator.



Grant sponsor: National Institutes of Health; Grant Number: GM62414; Grant sponsor: U.S. Department of Energy, Office of Biological and Environmental Research, under contract W-31-109-Eng-38; Grant sposnor: the Ontario Research and Development Challenge Fund; Grant sponsor: National Institutes of Health; Grant Neumber: GM26290.

## REFERENCE

1. Fong ST, Camakaris J, Lee BT. Molecular genetics of a chromosomal locus involved in copper tolerance in Escherichia coli K-12. Mol Microbiol 1995;15:1127–1137. [PubMed: 7623666]
2. Perrier AL, Cousin X, Boschetti N, Haas R, Chatel JM, Bon S, Roberts WL, Pickett SR, Massoulie J, Rosenberry TL, Krejci E. Two distinct proteins are associated with tetrameric acetylcholinesterase on the cell surface. J Biol Chem 2000;275:34260–34265. [PubMed: 10954708]
3. Holm L, Sander C. Dali: a network tool for protein structure comparison. Trends Biochem Sci 1995;20:478–480. [PubMed: 8578593]
4. Orengo CA, Taylor WR. SSAP: sequential structure alignment program for protein structure comparison. Methods Enzymol 1996;266:617–635. [PubMed: 8743709]
5. Cheah E, Carr PD, Suffolk PM, Vasudevan SG, Dixon NE, Ollis DL. Structure of the Escherichia coli signal transducing protein PII. Structure 1994;2:981–990. [PubMed: 7866749]
6. Xu Y, Cheah E, Carr PD, van Heeswijk WC, Westerhoff HV, Vasudevan SG, Ollis DL. GlnK, a PII-homologue: structure reveals ATP binding site and indicates how the T-loops may be involved in molecular recognition. J Mol Biol 1998;282:149–165. [PubMed: 9733647]
7. Magasanik B. The regulation of nitrogen utilization in enteric bacteria. J Cell Biochem 1993;51:34–40. [PubMed: 8094391]
8. Ninfa AJ, Atkinson MR. PII signal transduction proteins. Trends Microbiol 2000;8:172–179. [PubMed: 10754576]
9. Arcondeguy T, Jack R, Merrick M. P(II) signal transduction proteins, pivotal players in microbial nitrogen control. Microbiol Mol Biol Rev 2001;65:80–105. [PubMed: 11238986]
10. Kamberov ES, Atkinson MR, Ninfa AJ. The Escherichia coli PII signal transduction protein is activated upon binding 2-ketoglutarate and ATP. J Biol Chem 1995;270:17797–17807. [PubMed: 7629080]
11. Xu Y, Carr PD, Huber T, Vasudevan SG, Ollis DL. The structure of the PII-ATP complex. Eur J Biochem 2001;268:2028–2037. [PubMed: 11277925]
12. Martinez-Argudo I, Contreras A. PII T-loop mutations affecting signal transduction to NtrB also abolish yeast two-hybrid interactions. J Bacteriol 2002;184:3746–3748. [PubMed: 12057972]
13. Gitschier J, Moffat B, Reilly D, Wood WI, Fairbrother WJ. Solution structure of the fourth metal-binding domain from the Menkes copper-transporting ATPase. Nat Struct Biol 1998;5:47–54. [PubMed: 9437429]

14. Rosenzweig AC, Huffman DL, Hou MY, Wernimont AK, Pufahl RA, O'Halloran TV. Crystal structure of the Atx1 metallochaperone protein at 1.02 A resolution. Structure Fold Des 1999;7:605–617. [PubMed: 10404590]

15. DiDonato M, Narindrasorasak S, Forbes JR, Cox DW, Sarkar B. Expression, purification, and metal binding properties of the N-terminal domain from the wilson disease putative copper-transporting ATPase (ATP7B). J Biol Chem 1997;272:33279–33282. [PubMed: 9407118]

16. Lutsenko S, Petrukhin K, Cooper MJ, Gilliam CT, Kaplan JH. N-terminal domains of human copper-transporting adenosine triphosphatases (the Wilson's and Menkes disease proteins) bind copper selectively in vivo and in vitro with stoichiometry of one copper per metal-binding repeat. J Biol Chem 1997;272:18939–18944. [PubMed: 9228074]

17. Zhang RG, Skarina T, Katz JE, Beasley S, Khachatryan A, Vyas S, Arrowsmith CH, Clarke S, Edwards A, Joachimiak A, Savchenko A. Structure of Thermotoga maritima stationary phase survival protein SurE: a novel acid phosphatase. Structure (Camb) 2001;9:1095–1106. [PubMed: 11709173]

18. Walsh MA, Evans G, Sanishvili R, Dementieva I, Joachimiak A. MAD data collection — current trends. Acta Crystallogr D Biol Crystallogr 1999;55:1726–1732. [PubMed: 10531522]

19. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997:307–326.

20. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998;54:905–921. [PubMed: 9757107]

21. Perrakis A, Harkiolaki M, Wilson KS, Lamzin VS. ARP/wARP and molecular replacement. Acta Crystallogr D Biol Crystallogr 2001;57:1445–1450. [PubMed: 11567158]

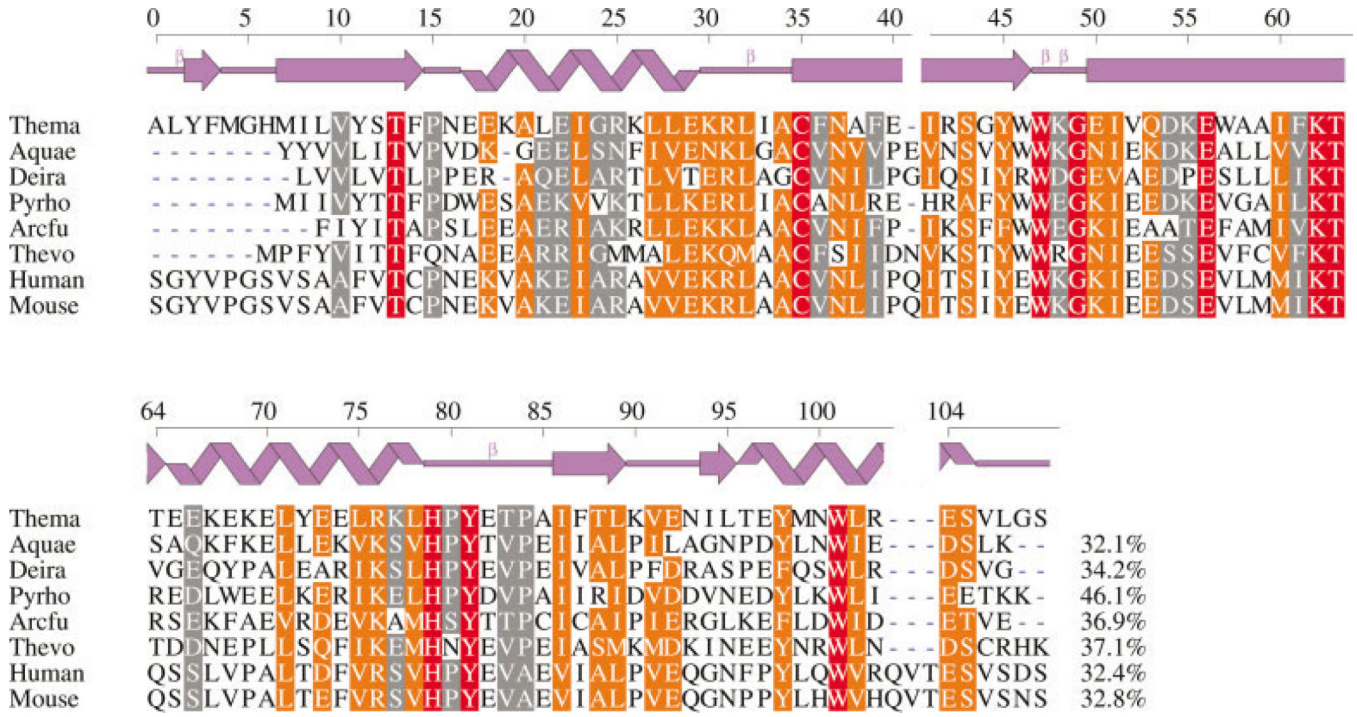


**Fig. 1.**

Multiple sequence alignment of CutA from *Thermotoga maritima* (Thema) against other periplasmic divalent cation tolerance proteins from bacteria (Aquae = *Aquifex aeolicus*, Deira = *Deinococcus radiodurans*) and archaea (Pyrho = *Pyrococcus horikoshii*, Arcfu = *Archaeoglobus fulgidus*, Thevo = *Thermoplasma volcanium*), together with one from Mouse and a brain acetylcholinesterase putative membrane anchor from Human. Sequence conservation is color coded from red for invariant residues to orange to grey to none. The secondary structure of the *T. maritima* CutA is schematically shown in purple above the alignment.

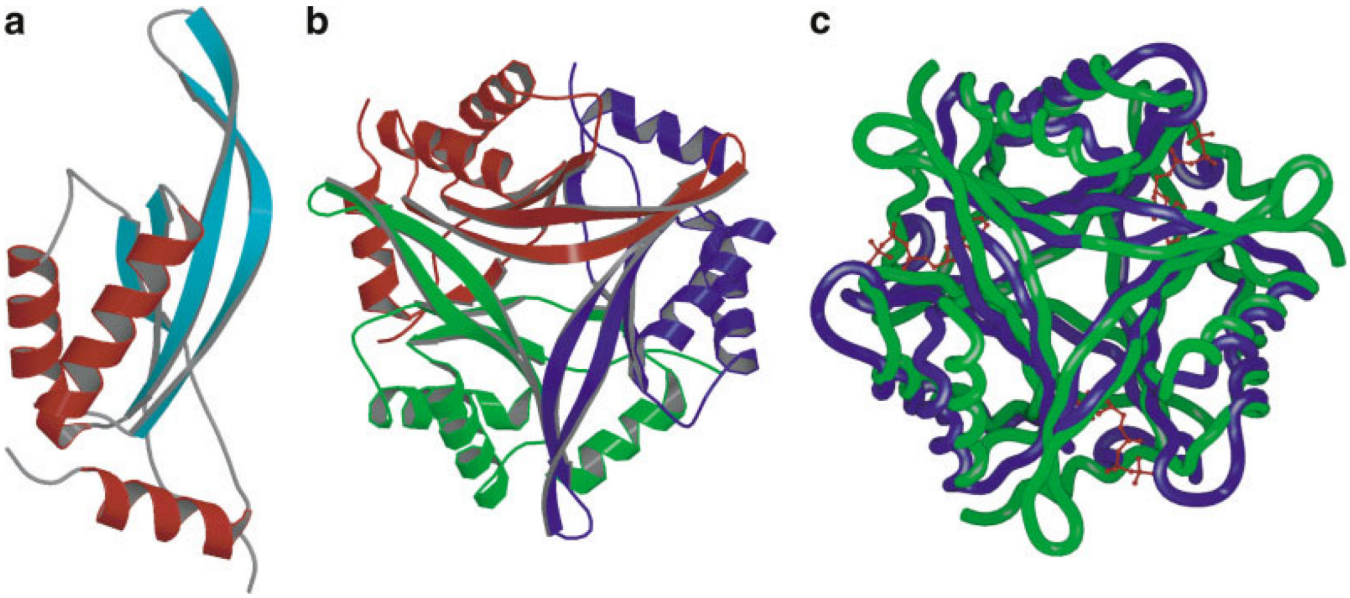


**Fig. 2.**

Ribbon representations of the tmCutA monomer (**a**), trimer (**b**) and superimposed structures of tmCutA and GlnK' complex with ATP (**c**). **a:** The α-helixes and β-strains are colored in red and blue, respectively. **b:** Each monomer in the tmCutA trimer is differentiated by color. **c:** The tmCutA is shown on green, GnlK is shown in blue and ATP molecules are colored in red.

## TABLE I

Summary of Crystal and MAD Data

| | |
|---|---|
| Unit cell | a = b = 52.237 Å, c = 33.846 Å α = β = 90° γ = 120° |
| Space group | P3 |
| MW (Da) (108 residues) | 12420 |
| Mol (AU) | 1 |
| SeMet (AU) | 3 |

| | MAD data collection | | |
|---|---|---|---|
| | **Edge** | **Peak** | **Remote** |
| Wavelength (Å) | 0.9795 | 0.9793 | 0.95354 |
| Resolution range (Å) | 1.4 | 1.4 | 1.4 |
| No. of unique reflections | 19614 | 19651 | 19451 |
| Completeness (%) | 95.9 | 96.1 | 95.1 |
| R merge (%) | 6.7 | 7.1 | 5.6 |

**TABLE II**

Crystallographic Statistics

| Phasing | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Centric | | Acentric | | All | | |
| Resolution range (Å) | FOM | Phasing power | FOM | Phasing power | Number | FOM | Phasing power |
| 30.0–1.4 | 0 | 0 | 0.5555 | 2.3552 | 17628 | 0.5555 | 2.3552 |
| Density modification | | | | | | | 0.80826 |

| Refinement | |
|---|---|
| Resolution range (Å) | 40–1.4 |
| No. of reflections | 19651 |
| σ cutoff | 0.0 |
| R-value (%) | 20.94 |
| Free R-value (%) | 24.08 (1571) |
| Rms deviations from ideal geometry | |
| bond length (1–2) (Å) | 0.006 |
| angle (°) | 1.2 |
| dihedral (°) | 21.8 |
| improper (°) | 0.66 |
| Number of atoms | |
| protein | 930 |
| water | 124 |
| Mean B-factor (Å$^2$) for all atoms | 15.3 |
| Ramachandran plot statistics (%) | |
| Residues in most favored regions | 99.2 |
| Residues in additional allowed regions | 0.8 |
| Residues in disallowed region | 0.0 |