Crystal Structure of Human Multiple Copies in T-Cell Lymphoma-1 Oncoprotein

Wolfram Tempel\textsuperscript{1}, Slav Dimov\textsuperscript{1}, Yufeng Tong\textsuperscript{1}, Hee-Won Park\textsuperscript{1,2}, Bum Soo Hong\textsuperscript{1*}

\textsuperscript{1} Structural Genomics Consortium, 101 College Street, MaRS South Tower, Toronto, Ontario, Canada

\textsuperscript{2} Department of Pharmacology, University of Toronto, Toronto, Ontario, Canada

Short title: Crystal structure of MCT-1

Key words: MCT-1; PUA domain; surface-entropy-reduced; cap structure; central pseudobarrel.

\textsuperscript{*}Address for Correspond:\textsuperscript{ }

Bum Soo Hong, Ph.D.
Structural Genomics Consortium
University of Toronto
MaRS South Tower, 101 College St, Room 735
Toronto, Ontario, M5G 1L7 Canada
E-mail: bumsso.hong@utoronto.ca
Tel (off): 416-946-3870
Fax: 416-946-0588
Abstract

Overexpression of multiple copies in T-cell lymphoma-1 (MCT-1) oncogene accompanies malignant phenotypic changes in human lymphoma cells. Specific disruption of MCT-1 results in reduced tumorigenesis, suggesting a potential for MCT-1-targeted therapeutic strategy. MCT-1 is known as a cap-binding protein and has a putative RNA-binding motif, the PUA-domain, at its C-terminus. We determined the crystal structure of apo MCT-1 at 1.7 Å resolution using the surface-entropy-reduced (SER) method. Notwithstanding limited sequence identity to its homologs, the C-terminus of MCT-1 adopted a typical PUA-domain fold that includes secondary structural elements essential for RNA recognition. The surface of the N-terminal domain contained positively charged patches that are predicted to contribute to RNA-binding.
Introduction

Gene amplification has been well studied as one of the critical genetic alterations in tumors and detected in a variety of human solid and hematological malignancies.\textsuperscript{1,2} Increase of the gene dosage by DNA amplification is generally accompanied by enhanced expression of genes contained within the amplified region. This principle is well exemplified by cellular oncogenes.\textsuperscript{2} Therefore, identification and characterization of oncogenes in amplified regions can provide important insights into tumorigenesis and cancer therapy.

More recently, genomic alteration through gene amplification has also been recognized as an important event in malignant lymphomagenesis.\textsuperscript{1,3} Multiple copies in T-cell lymphoma-1 (MCT-1) is an oncogene that was originally identified in a certain lymphoma cell line, where it was found amplified.\textsuperscript{4} Overexpression of MCT-1 in NIH 3T3 fibroblasts significantly shortened cell doubling time and promoted anchorage-independent growth \textit{in vitro}.\textsuperscript{4} Also, increased levels of MCT-1 protein were observed in a wide array of human lymphoma cells including aggressive non-Hodgkin’s lymphomas, and lymphoid cell lines overexpressing MCT-1 displayed increased growth rates and resistance to apoptosis.\textsuperscript{5,6} These findings suggest that the elevated level of MCT-1 protein is linked to cell transformation and proliferation, thereby being potentially implicated in human lymphomagenesis, although the exact molecular mechanism underlying this process remains elusive. MCT-1 is composed of 181 amino acid residues and contains a pseudouridine synthase and archaeosine transglycosylase (PUA) domain at its C-terminus.\textsuperscript{7,8} The PUA domain is a highly conserved RNA-binding motif and has mostly been found in tRNA- and rRNA-modifying enzymes,\textsuperscript{9} suggesting that MCT-1
may have RNA-binding capability. It has also been reported that MCT-1 binds to the cap structure $\text{m}^7\text{GpppN}$ (where N is any nucleotide) present at the 5′ end of all eukaryotic mRNAs and interacts with the translation machinery.\textsuperscript{7,10} In addition, MCT-1 protein recruits the density regulated protein (DENR/DRP), which contains an SUI1 domain involved in initiation site selection during translation.\textsuperscript{7} Taken together, these results support the idea that the oncogenic activity of MCT-1 would target RNA translation initiation/regulation. Meanwhile, it has been reported that the specific disruption of MCT-1 diminishes the malignant phenotype of human lymphoma cells,\textsuperscript{11,12} raising the possibility of developing novel therapeutic strategies that target MCT-1. In contrast to the considerable amount of data elucidating the physiological and cellular functions of oncogenic MCT-1, atomic level structural information providing valuable guidance for drug development has not yet been reported.

Here, we determined the crystal structure of human full-length MCT-1 at 1.7Å resolution. The MCT-1 crystals used for structure determination were obtained by using a surface-entropy-reduced mutant. The resulting high resolution crystal structure represents the first structural illumination of the human MCT-1 protein and is expected to expand overall knowledge about this oncoprotein containing a putative RNA recognition motif.

**Materials and Methods**

Cloning and mutagenesis
The full length cDNA encoding MCT-1 protein was cloned into the expression vector pNIC-CH (EF199843) to create a C-terminal His 6 tag (AHHHHHHH) fusion, and the resulting plasmid was utilized to make a series of site-directed MCT-1 mutants. The mutants were designed based on a surface-entropy reduction method through the use of the SER-prediction (SERp) server (http://services.mbi.ucla.edu/SER) (see “Results and Discussion”) and generated by using the Quick Change mutagenesis kit (Stratagene, La Jolla, CA, USA). All constructs used in this study were fully sequenced and then transformed into *E. coli* BL21(DE3) for overexpression.

Protein expression and purification

The *E. coli* transformants were initially cultured overnight in 50-ml LB-kanamycin media at 37 °C and then transferred into a LEX bioreactor system (Harbinger Biotechnology and Engineering Corp., Markham, Ontario, Canada) containing TB-kanamycin media for large-scale culture and protein expression. When the *A*₆₀₀ reached 3.0-4.0, the culture was cooled down to 18 °C and subsequently induced with 1.0 mM IPTG overnight at 18 °C. Cells were harvested by centrifugation and the pellets were stored at -80 °C until use. Frozen cell pellets were resuspended in nickel binding buffer (10 mM HEPES pH7.5, 500 mM NaCl, 5 mM imidazole, 10 % glycerol, 2.5 mM TCEP) containing protease inhibitor cocktail (EDTA-free, Roche Applied Science, USA) and resuspended cells were mechanically lysed with a microfluidizer (Microfluidizer Processor, M-110EH, USA). The cell lysate was then centrifuged to remove insoluble material, and the supernatant was loaded onto a DEAE-cellulose (DE52, Whatman, MA, USA) anion-exchange resin followed by a nickel-NTA agarose column (Qiagen, MD,
USA). Bound proteins were eluted with a buffer consisting of 10 mM HEPES pH7.5, 500 mM NaCl, 300 mM imidazole, 10% glycerol and 2.5 mM TCEP. The eluted sample was then dialyzed against 10 mM HEPES pH7.5, 10% glycerol and 2.5 mM TCEP to remove salts, and subjected to cation-exchange chromatography using a HiTrap SP column (GE Healthcare, NJ, USA) previously equilibrated with the same buffer used for dialysis. Protein was eluted with a linear gradient of 0-500 mM NaCl, and further purified by size exclusion chromatography on a Superdex 75 16/60 column (GE Healthcare, NJ, USA) previously equilibrated with a buffer consisting of 10 mM HEPES pH 7.5, 200 mM NaCl, 10% glycerol and 2.5 mM TCEP. The peak fractions containing MCT-1 protein were pooled, concentrated to 30 mg/ml and stored at -80 °C before crystallization. The purity and identity of the purified protein were confirmed by SDS-PAGE and mass spectroscopy. Selenomethionine (SeMet)-derivatized protein was purified in the same manner as the native protein.

Crystallization, data collection and structure determination

Crystallization was performed by the vapor diffusion method in sitting drops at 18 °C. A 1 µL aliquot of native protein sample was mixed with 1 µL reservoir buffer containing 2.5 M (NH₄)₂SO₄ and Bis-Tris propane pH7.0, and crystals grew in 3-4 days to a maximum size of 0.3 mm x 0.4 mm x 0.4 mm. Crystallization of SeMet-derivatized protein sample was carried out in the same manner as native protein, and selenium-containing crystals were obtained in 2.5M (NH₄)₂SO₄, Bis-Tris propane pH7.0 and 20% PEG3350. All crystals were cryoprotected in a 50:50 mixture of Paratone-N and mineral oils, and flash-frozen in liquid nitrogen for data collection.
Diffraction data were collected at the Advanced Photon Source, Argonne National Laboratory (Argonne, Illinois, USA) and reduced with programs of the XDS suite. The triclinic SeMet derivative crystal structure was phased using single wavelength anomalous diffraction and the SHELX program suite. An initial model of the triclinic structure was traced by the program BUCCANEER and preliminarily refined using REFMAC and COOT. The resultant preliminary homododecameric coordinates were then used as a search model for molecular replacement solution of the native hexagonal crystal structure. The program MOLREP placed one dodecamer in the asymmetric unit of the hexagonal lattice. After further refinement with REFMAC, model rebuilding with COOT and validation on the MOLPROBITY server, the coordinates for the final model were deposited in the Protein Data Bank under accession code 3R90.

Results and Discussion

Crystallization and structure determination

Full-length MCT-1 was expressed in *E.coli* as a recombinant protein possessing a C-terminal His-tag and exhibited relatively high solubility (about 15-20 mg per liter of culture). However, despite extensive efforts, our initial trials to crystallize purified MCT-1 were not successful. The truncated constructs containing only the C-terminal PUA domain (residues 93-181) of MCT-1 were also generated, but showed either no or very low solubility and marked precipitation during protein concentration. Therefore, as an alternative method for MCT-1 crystallization, surface entropy reduction (SER) was employed to induce epitopes that favor the formation of crystal contacts. To validate
whether the potential mutation sites in full-length MCT-1 predicted using the SERp server were located on the protein surface, the crystal structure of hypothetical protein APE0525 (PDB id 2CX0, 25% sequence identity to MCT-1) served as a homology model.\(^8\) All three SER mutants were designed through the process described above and generated for crystallization trials (Supporting Information Table SI). Although all MCT-1 mutants exhibited similar solubility to that of wild-type protein (data not shown), only one SER mutant containing triple alanine mutations (SER mutant no. #2 : E137A, K139A, Q140A, referred to as MCT-1\(^{\text{SER}}\)) was successfully crystallized in a form suitable for high resolution structural studies (Table I). Initial attempts at molecular replacement using the crystal structure of APE0525 as a search model, however, failed to find a convincing MCT-1\(^{\text{SER}}\) solution. Therefore, SeMet-derivatized MCT-1\(^{\text{SER}}\) was prepared, crystallized, and used to obtain the initial model. The final model was refined using native data to a resolution of 1.7 Å (Table I).

Crystallographic analysis indicated that twelve MCT-1\(^{\text{SER}}\) monomers were arranged as two ring-shaped hexamers in the asymmetric unit [Fig. 1(A)]. All twelve MCT-1\(^{\text{SER}}\) monomers in the asymmetric unit were mutually superimposable with average root mean square deviations (rmsd) of less than 0.9 Å. The triple alanine mutation sites were sequentially located in the C-terminal PUA domain of MCT-1 [Fig. 1(B)] and our MCT-1\(^{\text{SER}}\) model revealed that the mutated residues were clustered in the loop connecting the \(\beta8\) and \(\beta9\) strands (referred to as loop \(\beta8-\beta9^{\text{Ala}}\)) [Fig. 1(C,D)]. A close examination of the loop \(\beta8-\beta9^{\text{Ala}}\) region shows how these triple alanine mutations contributed to the crystallization of MCT-1\(^{\text{SER}}\). According to our present model, the loop \(\beta8-\beta9^{\text{Ala}}\) region was directly involved in the formation of reciprocal hydrophobic interactions with a
second MCT-1\textsuperscript{SER} molecule, related by a noncrystallographic 2-fold axis, wherein the C-alpha backbone of loop $\beta 8$-$\beta 9$\textsuperscript{Ala} on one molecule was contacted by the hydrocarbon side chains of Lys-99, Ile-102 and Lys-103 lined up in helix $\alpha 5$ on the other molecule, and \textit{vice versa} [Fig. 1(C)]. In the wild-type protein, however, the residues corresponding to substituted alanines in MCT-1\textsuperscript{SER}, namely Glu-137, Lys-139 and Gln-140, would have limited such hydrophobic interactions between protein molecules in solution because of the steric hindrance caused by their relatively bulky and flexible side chains. Consequently, our interpretation is that these hydrophobic contacts mediated by alanine mutations played a critical role in promoting the formation of MCT-1\textsuperscript{SER} crystals.

**Biophysical characterization**

Recently, it has been reported that human MCT-1 exists in a monomeric form in solution.\textsuperscript{8} Our investigation using analytical gel filtration chromatography reveals that the purified MCT-1\textsuperscript{SER} protein was eluted as a single, symmetrical peak and its elution profile overlapped with that of wild-type protein (Supporting Information Fig. S1(A), left panel). In addition, the molecular mass of MCT-1\textsuperscript{SER} (containing an N-terminal methionyl and a C-terminal His-tag) based on the standard curve with known proteins was estimated to be approximately 19.4 kDa (Supporting Information Fig. S1(A), right panel), close to the theoretical molecular mass of the monomeric form, 21.3 kDa. These data suggest that the hydrophobic contacts through the loop $\beta 8$-$\beta 9$\textsuperscript{Ala} as observed in crystals were a serendipitous artifact occurring under the present crystallization condition. Finally, to assess whether the triple alanine mutations gave rise to structural alterations, circular dichroism (CD) analysis was performed to measure the overall
secondary structure content and the melting temperature ($T_m$) of both MCT-1$^{\text{SER}}$ and wild-type. Their CD spectra were quite similar to each other (Supporting Information Fig. S1(B)), while the temperature scan revealed that the SER mutant transforms at a lower temperature ($T_m$ of MCT-1$^{\text{SER}}$ = 40.1 °C and $T_m$ of wild-type = 44.6 °C). Taken together, all these results strongly indicate that the triple alanine mutations did not perturb the overall folding of wild-type MCT-1, albeit slightly affecting its thermal stability.

Overall structure

The overall structure of MCT-1$^{\text{SER}}$ showed two globular and compact domains (designated herein as “N” and “C”) separated by two short linker peptides, and each domain was composed of a mixture of four $\alpha$-helices and five $\beta$-sheets, respectively [Fig. 1(D)]. N-domain (residues 1-92) has a five-stranded antiparallel $\beta$-sheet ($\beta$1-$\beta$5) with three $\alpha$-helices ($\alpha$2-$\alpha$4) on one side and two $\alpha$-helices ($\alpha$1 and $\alpha$8) on the other side. Helix $\alpha$8, as shown in Fig. 1(B), was assigned to the C-terminus based on its amino acid sequence. In our model, however, helix $\alpha$8 was tightly packed into the concave side of the N-terminal $\beta$-sheets [Fig. 1(D)], thereby enhancing interdomain contacts between N- and C-domains of MCT-1$^{\text{SER}}$. As anticipated, the C-domain (residues 93-181) of MCT-1$^{\text{SER}}$ adopted a highly conserved PUA fold [Fig. 2], although there was a slight structural discordance in the central pseudobarrel when compared to that of a typical PUA domain. Namely, while six $\beta$-strands contribute to the formation of the central pseudobarrel of a canonical PUA motif, one of the corresponding six $\beta$-strands in our refined model did not fully satisfy the expected interatomic distances for a $\beta$-sheet, and was consequently
assigned as a loop (residues 120-125). Both apical sides of the pseudobarrel were flanked by helices \( \alpha 5 \) and \( \alpha 7 \), respectively [Fig. 1(D)]. Structural details of MCT-1\(^{\text{SER}} \) are discussed below.

Structural comparison

We utilized a broad range of structural and functional homologs to derive possible cellular functions and targets of MCT-1. The structure of the N-domain is of particular interest because of its low sequence homology among known proteins. A detailed comparison with available structures in the Protein Data Base was carried out using a Dali search (http://ekhidna.biocenter.helsinki.fi/dali_server/). The results revealed that each N-domain from two top-ranked protein candidates, PH0734 from *Pyrococcus horikoshii* (PDB code: 3D79, Z-score = 7.8, rmsd = 2.5 Å, sequence identity = 27 %) and APE0525 from *Aeropyrum pernix* (PDB code: 2CX1, Z-score = 7.8, rmsd = 2.7 Å, sequence identity = 22 %), exhibits significant structural homology to that of MCT-1\(^{\text{SER}} \) (Supporting Information Fig. S2(A)). Both archaeal proteins, however, are currently classified into hypothetical proteins with unknown function,\(^8,22\) consequently making it difficult to deduce the functional role for the N-domain. Nonetheless, human MCT-1 shared significant similarities with these proteins such as average molecular weight of approximately 20 kDa and a PUA-domain harboring a C-terminal helical extension like helix \( \alpha 8 \) (Supporting Information Fig. S2(A)), suggesting that all three proteins may be closely related in their biological roles.
In contrast to the N-domain, a wealth of information regarding the PUA-domain, corresponding to C-domain of MCT-1, has been accumulated through structural studies of PUA-RNA complexes. Results from a Dali search also indicated that the top-ranked structural homologs to the C-domain of MCT-1\(^{\text{SER}}\) fall mostly into three PUA-containing proteins; Cbf5 from *Pyrococcus furiosus* (PDB code: 2HVY, Z-score = 11.3, rmsd = 1.2 Å, sequence identity = 25 %), TruB from *Thermotoga maritima* (PDB code: 1R3E, Z-score = 10.8, rmsd = 1.3 Å, sequence identity = 13 %), and ArcTGT from *P. horikoshii* (PDB code: 1J2B, Z-score = 10.5, rmsd = 1.7 Å, sequence identity = 15 %), in which the respective PUA-domains directly interacted with RNA molecules. Using these structural homologs, we investigated whether the amino acid residues involved in RNA-binding are also conserved in the PUA-domain of MCT-1. Structure-based sequence alignment revealed that a very small portion of the residues interacting with RNA molecules was conserved in the PUA-domain of MCT-1 [Fig. 1(B)]. In the case of Cbf5 PUA-domain, the closest homolog of that of MCT-1, ten amino acid residues contributed to major interactions with tRNA molecule\(^{23}\) and of these, only two residues matched the sequence of MCT-1 PUA-domain (Gly-266 and Lys-332 of Cbf5 correspond to Gly-108 and Lys-178 of MCT-1, respectively). It has been pointed out that PUA-domains make multiple contacts with RNA molecules mainly through i) a glycine-containing α5-β7 loop and ii) strand β10 (the labels of secondary structure elements described here follow those of MCT-1\(^{\text{SER}}\)), while sequence variations around these secondary structure elements contribute to unique recognition with RNA molecules.\(^{9}\) These structural features were also conserved on our MCT-1\(^{\text{SER}}\) model [Fig. 2]. Therefore we speculate that MCT-1 may have a different RNA-binding specificity or affinity from its homologs. Meanwhile,
both the C2- and C3 (PUA motif)-domains of ArcTGT were structurally related with the N-terminus and C-terminal PUA-domain in MCT-1\textsuperscript{SER}, respectively, in spite of low sequence similarity (15%). In the crystal structure of the ArcTGT-tRNA complex, surface-exposed basic residues of the C2-domain conferred an additional tRNA-binding interface through the electrostatic interactions between their positively charged side chains and the negatively charged phosphate backbone of RNA, thereby enhancing the specificity and overall interaction between tRNA and C3-domain.\textsuperscript{24} Likewise, analysis of the electrostatic surface potential revealed that there are several protruding clusters of positively charged residues on the N-domain of our MCT-1\textsuperscript{SER} model (Supporting Information Fig. S2(B)), providing a glimpse into their potential roles in forming a binary complex with RNA.

Meanwhile, recent reports suggest that MCT-1 interacts with the mRNA cap structure through its PUA-domain and recruits DENR.\textsuperscript{7,8} In an effort to obtain structural information regarding this, we have tried co-crystallization of both MCT-1\textsuperscript{SER} and wild-type with the cap structure m\textsuperscript{7}GpppN, but have so far failed to yield the complex structure. According to available structural information about cap-binding proteins, cap-binding pockets commonly consist of i) two aromatic residues for stacking interaction with the m\textsuperscript{7}G base and ii) basic and acidic residues interacting with the negatively-charged tri-phosphates and the cationic purine moiety of m\textsuperscript{7}G base, although overall structures of these proteins are markedly different.\textsuperscript{25} In the investigation based on this concept, the sulfate ion coordinated between N- and PUA-domains on MCT-1\textsuperscript{SER} model attracted our attention (Supporting Information Fig. S2(B)) since sulfate ions are known to frequently occupy the phosphate-binding sites in crystals of nucleotide-binding
proteins.\textsuperscript{26} It seems premature, however, to relate this region to a putative cap-binding site because there were no outstanding solvent-exposed aromatic residues for m\textsuperscript{7}G base. The interaction between the MCT-1 and the cap m\textsuperscript{7}GpppN has been previously verified by immunoblotting\textsuperscript{7,8}, but detailed kinetic data are not yet available. In our preliminary isothermal titration calorimetry (ITC) experiments, no measurable binding of the cap to the MCT-1 was detected (data not shown). Taken together, structural analyses of our apo MCT-1 along with ITC data lead us to assume that the cap binding-pocket is structurally less conserved in MCT-1, and the association between MCT-1 and the cap may be weak or accompanied by no enthalpy changes. Alternatively, the cap recognition of MCT-1 may be fully achieved through interaction with an additional binding partner such as DENR.

In conclusion, we have determined the high resolution crystal structure of MCT-1 oncoprotein using an SER mutant. Biophysical investigation showed that despite triple alanine mutations, MCT-1\textsuperscript{SER} has similar structural contents to the wild-type protein, thus providing a valid model for the investigation of the structure-function relationships of wild-type MCT-1 protein. Although our structural data is currently limited to the apo-protein structure, comparison with known structural homologs reveals that the structural properties important for direct interaction with RNA molecules are well conserved throughout both the N-domain and C-terminal PUA-domain of MCT-1\textsuperscript{SER} despite low sequence similarity. Recent extensive biochemical and structural studies on eIF4E oncoprotein, a notable cap-binding protein, constitute a significant breakthrough in developing cancer therapeutic agents such as cap-mimicking inhibitors or small molecules that disrupt the interaction with eIF4E-binding proteins.\textsuperscript{27,28} Likewise, it
would be interesting to verify the cap-binding site on MCT-1 and elucidate the role of DENR in forming a complex with the cap molecule in the future. These studies are also considered a significant step, not only in enhancing our overall understanding of translation regulation involving MCT-1 but also in evaluating its potential as a drug target.

Acknowledgments

We thank Dr. Charles M. Deber and Ms. Mira Glibowicka (Hospital for Sick Children, Toronto) for generously providing access to CD equipment and Dr. Guillermo Senisterra for interpretation of CD data. We also thank Dr. Jinrong Min and Mr. Scott Hughes for valuable discussions and corrections. We appreciate the crystal structure review and valuable comments by Dr. Amy K. Wernimont. This work was supported by the Structural Genomics Consortium, a registered charity (number 1097737) that receives funds from the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, Genome Canada through the Ontario Genomics Institute, GlaxoSmithKline, Karolinska Institutet, the Knut and Alice Wallenberg Foundation, the Ontario Innovation Trust, the Ontario Ministry for Research and Innovation, Merck & Co., Inc., the Novartis Research Foundation, the Swedish Agency for Innovation Systems, the Swedish Foundation for Strategic Research, and the Wellcome Trust. Diffraction data were collected on Advanced Photon Source (APS) beam lines at the Structural Biology Center and GM/CA-CAT, which has been funded in whole or in part by the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Sciences (Y1-GM-
1104). Use of the APS was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.
Reference List


Figure legends

Figure 1. Crystal structure of human MCT-1SER. (A) Ribbon diagrams of MCT-1SER monomers in asymmetric unit. Twelve MCT-1SER monomers are represented in different colors. All structural figures were generated using PyMol (http://www.pymol.org). (B) Structure-based sequence alignment between the PUA-domains of human MCT-1SER and homologs. The secondary structure elements of P.furiosus Cbf5 (NCBI accession number NP_579514, PDB code : 2HVY) and human MCT-1SER (GenBank accession number BAA86055, PDB code : 3R90) are placed on the top and the bottom of the alignment, respectively, and are compared to two additional homologs, P.horikoshii ArcTGT (NCBI accession number NP_143020) and T.maritime TruB (NCBI accession number NP_228665). Conserved residues are depicted in white on a red background. Physicochemically conserved residues are depicted in red. Overall conserved regions are framed in blue. The blue upward triangles highlight the residues involved in the significant interactions between P.furiosus Cbf5 PUA and RNA. The red circle indicates the site of each substituted alanine residue. The alignment was generated with ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and was printed using the ESPript 2.1 software package (http://escript.ibcp.fr/ESPript/ESPript/). (C) Zoomed view of the substituted triple alanine residues. The site where the substituted triple alanine residues face each other between two MCT-1SER monomers (colored in deep-yellow and cyan colors, respectively) is indicated by a black rectangle on the left and shown in close-up stereo view on the right. Substituted alanines are highlighted in red. Secondary structure elements and residues from the second monomer are marked with a single prime. (D) Ribbon diagram of overall MCT-1SER fold. The N-domain and C-terminal PUA domain
are represented in cyan and deep-red, respectively. Secondary structure elements of helices and strands are labeled.

Figure 2. PUA-domains of MCT-1 \(^{\text{SER}}\) and homologs. The name of each protein is given at the bottom of the PUA domain structure. Secondary structure elements important for RNA-binding are shown in yellow and labeled in the MCT-1 \(^{\text{SER}}\) model.
Table I. Data collection and refinement statistics of MCT-1<sup>SER</sup>.

<table>
<thead>
<tr>
<th></th>
<th>SeMet derivative</th>
<th>Native</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beam line</td>
<td>23ID</td>
<td>19ID</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.97944</td>
<td>0.97901</td>
</tr>
<tr>
<td>Number of oscillation frames (width, °)</td>
<td>360 (1.0)</td>
<td>360 (0.5)</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>78.58, 87.17, 92.19</td>
<td>115.59, 115.59, 157.79</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>80.36, 85.03, 89.12</td>
<td>90.00, 90.00, 120.00</td>
</tr>
<tr>
<td>Space group</td>
<td>P1</td>
<td>P3&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td>Resolution (outer shell) (Å)</td>
<td>40-1.85 (1.90-1.85)</td>
<td>30-1.70 (1.74-1.70)</td>
</tr>
<tr>
<td>Unique HKLs</td>
<td>390682 (28837)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>259245 (19130)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>95.2 (94.8)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>99.9 (100.0)</td>
</tr>
<tr>
<td>Rsym (%)</td>
<td>7.3 (64.6)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>8.4 (86.9)</td>
</tr>
<tr>
<td>⟨I/σI⟩</td>
<td>8.4 (1.4)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>13.3 (1.9)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>2.0 (2.0)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>5.6 (5.6)</td>
</tr>
<tr>
<td>PDB code</td>
<td>3R90</td>
<td></td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>28.4-1.7</td>
<td></td>
</tr>
<tr>
<td>HKLs used / &quot;free&quot;</td>
<td>254062 (5105)&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Rwork / Rfree (%)</td>
<td>20.2 / 23.5</td>
<td></td>
</tr>
<tr>
<td>Number of atoms / mean B-factor (Å)&lt;sup&gt;2&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>17596 / 21.6</td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>1374 / 26.0</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>131 / 39.0</td>
<td></td>
</tr>
<tr>
<td>RMSD bonds (Å) / angles (°)</td>
<td>0.013 / 1.3</td>
<td></td>
</tr>
<tr>
<td>Ramachandran allowed / preferred (%)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>100.0 / 98.1</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Bijvoet mates scaled separately.
<sup>b</sup> selected in thin resolution shells with program SFTOOLS (B. Hazes, Univ. of Alberta).
<sup>c</sup> Wilson B-factor: 19.8 Å<sup>2</sup> from TRUNCATE.
<sup>d</sup> Reference 29
Figure 1

A

Top

Side

90°

B

C

D

180°
Figure 2
Supplementary Table SI. List of MCT-1 SER mutants

<table>
<thead>
<tr>
<th>SER mutant no.</th>
<th>Residues no.</th>
<th>Crystallization</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>K46A, K47A</td>
<td>( \neg^a )</td>
</tr>
<tr>
<td>#2</td>
<td>E137A, K139A, Q140A</td>
<td>( +^b )</td>
</tr>
<tr>
<td>#3</td>
<td>E7A, K8A, E9A</td>
<td>( \neg )</td>
</tr>
</tbody>
</table>

\(^a\) : not crystallized  
\(^b\) : crystallized
Supplementary Figure S1

Figure S1. Biophysical characterization of MCT-1<sup>SER</sup> protein. All experiments described below were performed using protein samples that contained an N-terminal methionyl and
a C-terminal His-tag. (A) The elution profiles of both the wild-type (blue) and its SER mutant MCT-1$^{\text{SER}}$ (pink) proteins, determined by Superdex 200 10/300 analytical column (GE Healthcare) equilibrated with the same buffer as used for size exclusion chromatography, were superimposed. Arrows indicate elution volumes of standard proteins (left panel). A calibration curve was prepared using standard proteins (670 kDa, thyroglobulin; 158 kDa, gammaglobulin; 44 kDa, ovalbumin; 17 kDa, myoglobin, marked as blue-filled circles) (Bio-Rad) and utilized to determine the molecular mass of MCT-1$^{\text{SER}}$ (indicated by a hollow circle) (right panel). The logarithm of the molecular mass was plotted versus $K_{av}$ that was calculated for each protein from the equation $K_{av} = (V_e - V_o)/(V_t - V_o)$, where $V_e =$ elution volume for the protein, $V_o =$ column void volume, and $V_t =$ total bed volume.$^1$ (B) CD was measured with a Jasco J-810 spectropolarimeter using a 0.05-cm pathlength cell, and proteins were diluted to a final concentration of 20 $\mu$M in a buffer containing 5 mM HEPES pH 7.5 before measurements. At fixed temperature (20 °C), wavelength scans were recorded three times between 200 nm and 260 nm with a band width of 1 nm and a data pitch of 0.1 nm and averaged automatically. Protein spectra were then base line-corrected by subtracting buffer spectra. Time courses of temperature-induced conformational changes were followed by continuously monitoring at 222 nm. The sample cell was heated to 70 °C with a heat rate of 0.5 °C/min. Far ultraviolet CD spectra of the wild-type (blue) and its SER mutant MCT-1$^{\text{SER}}$ (pink) were superimposed after background subtraction and shown on the right.
Figure S2. (A) Superimposed C-alpha traces of *P. horikoshii* PH0734 (orange), *A. pernix* APE0525 (magenta) and MCT-1\textsuperscript{SER} (green) models. N-domain and C-terminal PUA
domain are indicated, respectively, and a helix corresponding to the α8 of MCT-1SER in each protein model is shown in yellow color. (B) Distribution of the electrostatic potential on the surface of MCT-1SER. The electrostatic potentials were calculated in the range of $-10 \ kT/e$ (red, negative potential) to $+10 \ kT/e$ (blue, positive potential) by APBS software$^2$ and mapped to the solvent-accessible surface, where $k$ denotes the Boltzmann’s constant, $T$ is the temperature in Kelvin, and $e$ is the charge of an electron. The intensity of color is proportional to the local potential. N-domain and C-terminal PUA domain are indicated as in (A). The top view of N-domain is shown on the right. The sulfate ion captured between two domains is represented as a space-filling model.

**References**
