JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Crystal Structures of Human DcpS in Ligand-free and m⁷GDP-bound forms Suggest a Dynamic Mechanism for Scavenger mRNA Decapping

Nan Chen[1], Martin A. Walsh[2], Yuying Liu[1], Roy Parker[3] and Haiwei Song[1,4*]

[1]*Laboratory of Macromolecular Structure, Institute of Molecular and Cell Biology 61 Biopolis Drive, Proteos Singapore 138673*

[2]*MRC France, CRG BM14 ESRF, B.P.220, F-38043 Grenoble CEDEX, France*

[3]*Department of Molecular and Cellular Biology and Howard Hughes Medical Institute University of Arizona, Tucson AZ 85721, USA*

[4]*Department of Biological Sciences, National University of Singapore, 14 Science Drive Singapore 117543*

***Corresponding author*

Eukaryotic cells utilize DcpS, a scavenger decapping enzyme, to degrade the residual cap structure following 3′-5′ mRNA decay, thereby preventing the premature decapping of the capped long mRNA and misincorporation of methylated nucleotides in nucleic acids. We report the structures of DcpS in ligand-free form and in a complex with m⁷GDP. apo-DcpS is a symmetric dimer, strikingly different from the asymmetric dimer observed in the structures of DcpS with bound cap analogues. In contrast, and similar to the m⁷GpppG–DcpS complex, DcpS with bound m⁷GDP is an asymmetric dimer in which the closed state appears to be the substrate-bound complex, whereas the open state mimics the product-bound complex. Comparisons of these structures revealed conformational changes of both the N-terminal swapped-dimeric domain and the cap-binding pocket upon cap binding. Moreover, Tyr273 in the cap-binding pocket displays remarkable conformational changes upon cap binding. Mutagenesis and biochemical analysis suggest that Tyr273 seems to play an important role in cap binding and product release. Examination of the crystallographic *B*-factors indicates that the N-terminal domain in apo-DcpS is inherently flexible, and in a dynamic state ready for substrate binding and product release.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: X-ray crystallography; protein structure; mRNA decay; mRNA decapping; scavenger mRNA decapping

Introduction

mRNA decay plays a key role in post-transcriptional regulation of gene expression. In eukaryotes, two major general mRNA decay pathways have been identified.[1] Both pathways are initiated with deadenylation of the 3′-poly(A) tail of mRNAs. In the 5′-3′ decay pathways, the 5′ cap structure can be removed by the Dcp1p/Dcp2p complex following deadenylation, thus exposing the 5′ end to 5′-3′ exoribonuclease activities. In the 3′-5′ decay pathways, deadenylation is followed by exosome-dependent 3′-5′ degradation of mRNA body. In addition, two specialized mRNA decay pathways exist to recognize and degrade aberrant mRNAs. In a process referred to as nonsense-mediated mRNA decay (NMD), transcripts with premature translation termination codons are degraded either by deadenylation-independent decapping (5′-3′ NMD), or by accelerated deadenylation and 3′-5′ exonucleolytic digestion by the exosome (3′-5′ NMD).[2–6] Moreover, mRNAs lacking translation termination codons are recognized and degraded rapidly 3′-5′ by the cytoplasmic exosome.[7,8]

Decapping in 5′-3′ mRNA decay pathways and removal of the residual cap structure generated by the 3′-5′ mRNA decay pathways require two different types of decapping enzymes. Hydrolysis of the cap structure in the 5′-3′ mRNA decay pathway requires the Dcp1/Dcp2 complex in yeast or Dcp2 alone in mammals.[1] Dcp2p, which contains a Nudix motif found in a class of pyrophosphatases,[9] catalyzes the release of m⁷G diphosphate (m⁷GDP) and monophosphate-terminated mRNA (pRNA) in a metal-dependent manner.[10–12] Interestingly, Dcp2p is an RNA-binding protein that

prefers substrates longer than 25 nucleotides.[13,14] The RNA-binding properties of Dcp2 prevent it from degrading the residual cap structure produced by 3′-5′ degradation and restrict its activity to the capped mRNA only.

Degradation of the residual cap structure of 3′-to-5′ exonucleolytic degradation is carried out by DcpS, the scavenger decapping enzyme identified as hDcpS in human and Dcs1p in yeast.[15,16] DcpS proteins are members of the HIT family of pyrophosphatases and use a histidine triad to carry out catalysis in a metal-independent manner, releasing m^7G monophosphate (m^7GMP) and diphosphate-terminated mRNA (ppRNA).[15–17] Structural analysis has revealed that HIT proteins exist as a homodimer through formation of a continuous ten-stranded antiparallel β sheet, with each HIT protomer containing an active site and nucleotide-binding pocket that coordinates the pyrophosphate bond with respect to the three histidine residues of the catalytic HIT motif.[18–20] Interestingly, DcpS is unable to decap substrates longer than ten nucleotides,[16] thereby preventing DcpS from prematurely decapping mRNAs not targeted for degradation. DcpS may also act in the 5′-3′ mRNA decay pathway by converting m^7GDP, a product released by Dcp2p to m^7GMP, thereby preventing misincorporation of methylated nucleotides in nucleic acids.[21] Further support that DcpS is implicated in the 5′-3′ mRNA decay comes from the observation that DcpS is susceptible to m^7GDP competition.[16]

The crystal structures of human DcpS with substitution of the active-site His277 with asparagine, in complex with either m^7GpppG or m^7GpppA have been reported recently.[22] The structures show that DcpS with bound cap analogue is an asymmetric dimer that simultaneously creates an open non-productive DcpS–cap complex and a closed productive DcpS–cap complex, which differs from the open complex by a 30 Å movement of an N-terminal domain. Combined with mutagenesis and biochemical analysis, an autoregulatory mechanism has been proposed to explain how premature decapping of mRNA can be avoided by blocking the conformational changes that are required to form a closed productive active site capable of cap hydrolysis. These differences suggest that DcpS undergoes significant conformational changes during its cycle of substrate binding, hydrolysis and product release.

To gain a better understanding of the molecular mechanisms by which DcpS functions, we determined the crystal structures of human DcpS in ligand-free form (apo-DcpS) and in complex with m^7GDP. apo-DcpS is a symmetric dimer different from the asymmetric dimer observed in the structures of hDcpS with bound cap analogues. The structure of hDcpS with bound m^7GDP is similar to that of hDcpS with the bound cap analogue. Structural comparison combined with mutagenesis suggests a dynamic mechanism for substrate binding and catalysis.

Results and Discussion

Structure determination

The full-length human DcpS consisting of amino acid residues 1–337 was expressed in *Escherichia coli* and purified to homogeneity. For multiwavelength anomalous dispersion (MAD) phasing, a mutant protein in which two Leu residues (Leu206 and Leu317) were mutated to Met was used to obtain the SeMet-substituted crystals. Since this mutant DcpS retained the same decapping activity as the wild-type protein (data not shown), we designated this mutant DcpS protein as apo-DcpS. Crystals of apo-DcpS belong to space group $P2_12_12$ and contain two molecules per asymmetric unit (AU). m^7GDP–DcpS complex crystals were obtained by soaking the native DcpS crystals in a buffer containing m^7GDP, belong to space group $P2_12_12_1$ and contain four molecules per AU. The structures of apo-DcpS and the m^7GDP–DcpS complex were solved by the molecular replacement method using the coordinates of the m^7GpppG–DcpS complex as the search model.[22] Difference Fourier maps clearly showed the bound nucleotides (Figure 1(d)). The electron densities for two of the m^7GDP molecules are well defined, while the β-phosphate groups are disordered for the other two m^7GDP molecules; thus, these two m^7GDP molecules are modeled as m^7GMP plus a phosphate group. In the structure of apo-DcpS, residues 1–39, 71–76, 286–293 and 337 for protomer A, and residues 1–39, 70–77, 286–292 and 337 for protomer B were not modeled, as these regions are not visible in the electron density map and are assumed to be disordered. In the structure of m^7GDP-DcpS, residues 1–39, 69–77, 111, and 337 for chain A, residues 1–39, 70–78, and 337 for chain B, residues 1–38, 68–76, and 337 for chain C, and residues 1–37, 71–74, 99–101, and 110 for chain D are assumed to be disordered, as interpretable electron densities are not observed for these regions. Statistics of structure determination and refinement are summarized in Table 1 (see Materials and Methods).

Overall structure description

apo-DcpS forms a symmetric homodimer with two protomers related by a 2-fold non-crystallographic symmetry in the AU (Figure 1(a) and (b)). As observed in the structure of the m^7GpppG–DcpS complex,[22] apo-DcpS contains two distinct domains, a dimeric C-terminal domain (residues 146–336) containing the putative active site and the HIT motif, and a swapped dimeric N-terminal domain (residues 40–145). Each of the two identical active sites in apo-DcpS is occupied by a phosphate ion, with its position close to the γ-phosphate group (P-P distance of 1.0 Å) of the bound cap analogue in the closed state of the m^7GpppG–DcpS complex. The N-terminal and C-terminal domains are connected covalently through a hinge linker,
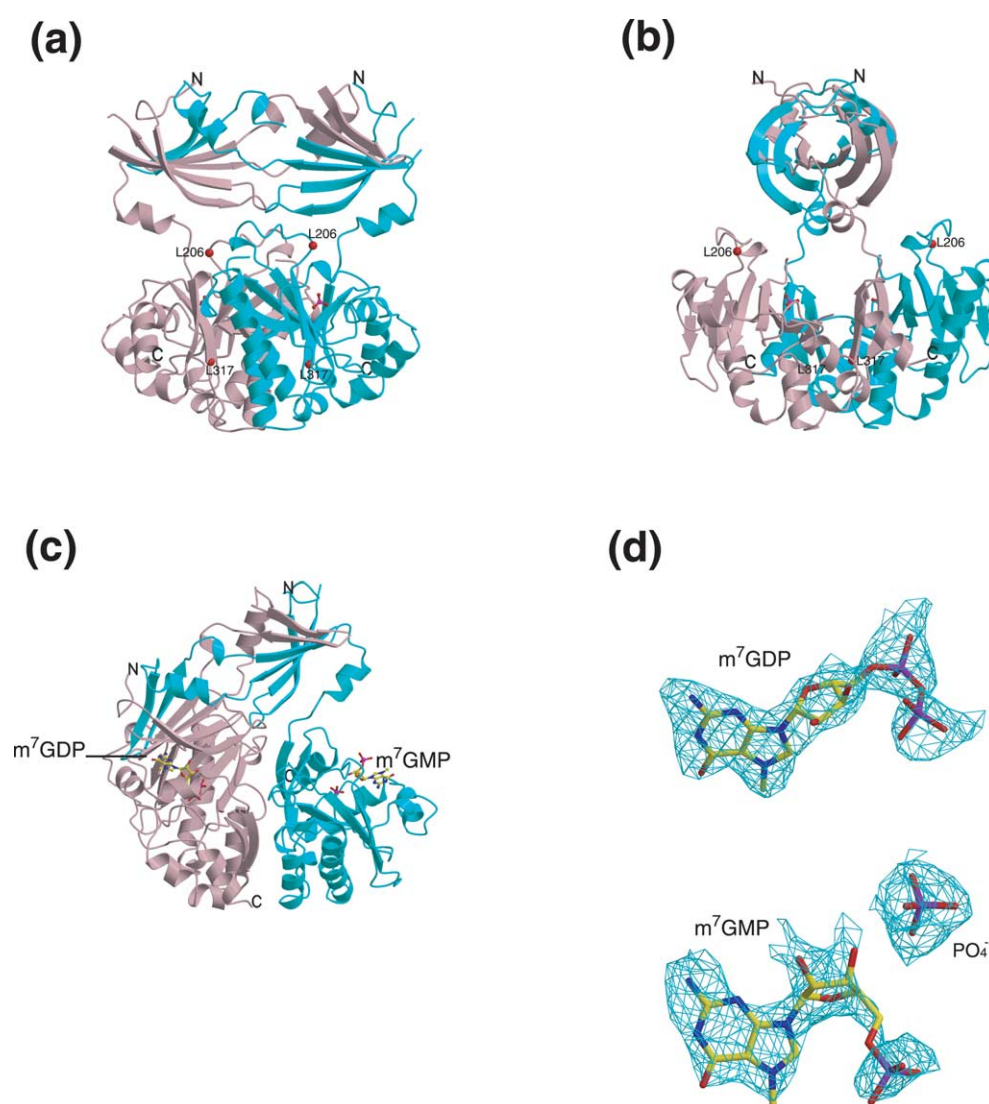
Figure 1. Structures of apo-DcpS and the m^{7}GDP–DcpS complex. (a) A ribbon diagram of the symmetric apo-DcpS dimer with protomer A colored pink and protomer B colored cyan. Two phosphate groups are shown as stick models. (b) Orthogonal view of the apo-DcpS structure in (a). (c) A ribbon diagram of the asymmetric DcpS dimer in the m^{7}GDP–DcpS complex. A phosphate ion and the m^{7}GDP and m^{7}GMP molecules are shown as stick models. The coloring scheme is as in (a). (d) Electron densities for bound m^{7}GDP, and m^{7}GMP plus a phosphate in the closed and open conformations of the m^{7}GDP–DcpS complex respectively. The $F_o - F_c$ σ_A-weighted map at 2.5 Å was calculated with phases from Dcps molecules only. Figures 1, 2, 3, and 4(a) and (b) were generated using MOLSCRIPT.[35]

which shows an identical conformation for each protomer.

The symmetric dimer of apo-DcpS is strikingly different from the asymmetric dimer observed in the structures of DcpS with bound cap analogue. The identical conformations for the hinge regions observed here are markedly different from those in the m^{7}GpppG–DcpS complex, wherein the hinge linker tilts nearly 90° toward the C-terminal domain in the closed state, whereas it adopts an extended conformation in the open state.[22] In addition, in the apo-DcpS dimer, there is no inter-domain contact between the N-terminal and C-terminal domains, whereas the N-terminal and C-terminal domains form an extensive interface in the closed conformation in the m^{7}GpppG–DcpS complex. The conformation of each protomer in apo-DcpS is reminiscent of the open conformation in DcpS with bound cap analogue wherein few contacts are observed between the N-terminal and C-terminal domains.

Like DcpS with bound cap analogue, m^{7}GDP–DcpS forms an asymmetric dimer (Figure 1(c)) with two asymmetric dimers related by a 2-fold non-crystallographic symmetry in the AU. Since no substantial difference is observed between the structures of the two m^{7}GDP–DcpS asymmetric dimer in the AU (r.m.s.d of 0.67 Å for all the equivalent C$^\alpha$ atoms), all subsequent analyses reported here use the AB dimer. Each of the active sites in two closed conformations of m^{7}GDP–DcpS is occupied by an m^{7}GDP molecule, while the active site in two open conformations is occupied by a m^{7}GMP molecule and a phosphate group. The

Table 1. Data collection and refinement statistics

	apo-DcpS[a]	m⁷GDP–DcpS complex
A. *Data collection*		
Wavelength (Å)	0.93	0.9997
Space group	$P2_12_12$	$P2_12_12_1$
Unit cell dimensions		
a (Å)	101.92	101.93
b (Å)	103.86	105.14
c (Å)	71.07	138.65
Resolution range (Å)	50.0–2.0	32.0–2.5
Unique reflections	51,633	51,097
Completeness (%)	100 (99.8)	98.9 (98.9)
R_{merge}[b] (%)	7.0 (40.0)	8.7 (26.0)
I/σ	22.3 (2.5)	7.5 (2.8)
B. *Refinement*		
Total atoms	5177	10,112
Resolution range (Å)	20.0–2.0	20.0–2.5
R_{work}[c] (%)	22.1	21.3
R_{free}[d] (%)	25.7	26.7
r.m.s deviations from ideal values		
Bond lengths (Å)	0.009	0.009
Bond angles (deg.)	1.14	1.34
Ramachandran plot		
Most-favored regions (%)	89.2	91.7
Disallowed regions (%)	0.1	0.4

Values in parentheses indicate the specific values in the highest-resolution shell.

  [a] The apo-DcpS structure was refined against a dataset collected from a SeMet-containing crystal.

  [b] $R_{merge} = \Sigma |I_j - \langle I \rangle| / \Sigma I_j$, where I_j is the intensity of an individual reflection i, and $\langle I \rangle$ is the average intensity of that reflection.

  [c] $R_{work} = \Sigma ||F_o| - |F_c|| / \Sigma |F_c|$, where F_o denotes the observed structure factor amplitude, and F_c denotes the structure factor amplitude calculated from the model.

  [d] R_{free} is as for R_{work} but calculated with 5.0% of randomly chosen reflections that are omitted from the refinement.

overall structures of the N-terminal and C-terminal domains are similar to each other in the structures of apo-DcpS and m⁷GDP–DcpS (mean pair-wise C^α r.m.s.d ∼0.75 Å) with the exception of the two hinge regions and the regions involved in cap binding.

Comparison between cap-bound and unbound DcpS

To gain insight into the DcpS structural changes upon cap binding, we compared the structures of apo-DcpS with both m⁷GDP–DcpS and m⁷GpppG–DcpS. The first point is that the individual dimeric domains are similar among these structures. Superposition of the N-terminal domains of apo-DcpS with those of m⁷GDP–DcpS (AB dimer) and m⁷GpppG–DcpS gives an r.m.s.d for equivalent C^α atoms of 0.85 Å and 0.76 Å, respectively (Figure 2(a)). When the dimeric C-terminal domains are superimposed, the r.m.s.d for C^α atoms is 0.73 Å and 0.90 Å for apo-DcpS *versus* m⁷GDP–DcpS and apo-DcpS *versus* m⁷GpppG-DcpS, respectively (Figure 2(b)).

Despite these similarities of the two domains, the relative orientation of the N-terminal domain with respect to the C-terminal domain in apo-DcpS is dramatically different from those in both m⁷GDP–DcpS and m⁷GpppG–DcpS. When the dimeric C-terminal domain of apo-DcpS is superimposed with that of m⁷GpppG–DcpS, the orientations of the N-terminal domains differ by 37° (Figure 2(c)). Similar results are obtained when the corresponding domains of apo-DcpS and m⁷GDP–DcpS are compared with each other. A thorough analysis of the intermolecular contacts ruled out the possibility that such large global conformational changes are induced by crystal packing. Given that the individual dimeric domains of DcpS are similar to each other in apo form and in complex with the cap analogue, the transition of symmetric to asymmetric dimer in DcpS results from the rigid-body movement of the N-terminal domain, presumably induced by cap binding. Comparison of m⁷GDP–DcpS with m⁷GpppG–DcpS showed that their overall structure and relative domain orientations are very similar between these two structures with pair-wise C^α r.m.s.d of 0.62 Å (Figure 2(c)). This similarity argues that binding of the m⁷Gpp portion of the cap analogue is necessary and sufficient to induce the orientation change of the N-terminal domain.

In addition to the large-scale, rigid-body movement of the N-terminal domain toward the C-terminal domain, binding of either m⁷GDP or m⁷GpppG leads to significant conformational changes in the cap-binding pocket (Figure 2(b)). Specifically, as shown in Figure 3(a), residues 170–189 and 204–217 in the closed state moved towards each other. Consequently, Trp175 and Leu206 (Met206 in the apo-DcpS) are displaced by 2.6 Å and 2.7 Å, respectively, from their respective positions in apo-DcpS, thus sandwiching the m⁷G base between them. Residues Glu185, Ile179, Asp205, and Lys207 undergo displacements of 1.5 Å, 1.5 Å, 1.5 Å and 3.2 Å, respectively, to interact with the bound cap. Ser272 and Tyr273 are flipped by 180° and 90°, respectively, to accommodate the bound cap (Figure 3(a) and see below). In the open state, residues 204–217 also moved towards the m⁷G base, with Leu206 in m⁷GDP–DcpS and m⁷GpppG–DcpS undergoing displacements of 1.9 Å and 2.5 Å, respectively (Figure 3(b)). Although residues 170–189 in the open state of m⁷GpppG–DcpS move towards the m⁷G base, as those observed in the closed state, the same region in m⁷GDP–DcpS moves back to its native positions with the largest positional shift of 3.5 Å confined to residue Leu171 (Figure 3(b)). Consequently, Trp175 is in an unfavorable stacking position against the m⁷G base. Such conformational changes of residues 170–189 in the open state of m⁷GDP–DcpS would make the cap binding less tightly and might play a role in facilitating product release. Furthermore, the conformations of Ser272 and Tyr273 in the open state of m⁷GDP–DcpS are similar to those in apo-DcpS but different from those in both the closed and open states of m⁷GpppG–DcpS (Figure 3(b)). Consistent with Ser272 being functionally important, mutation of
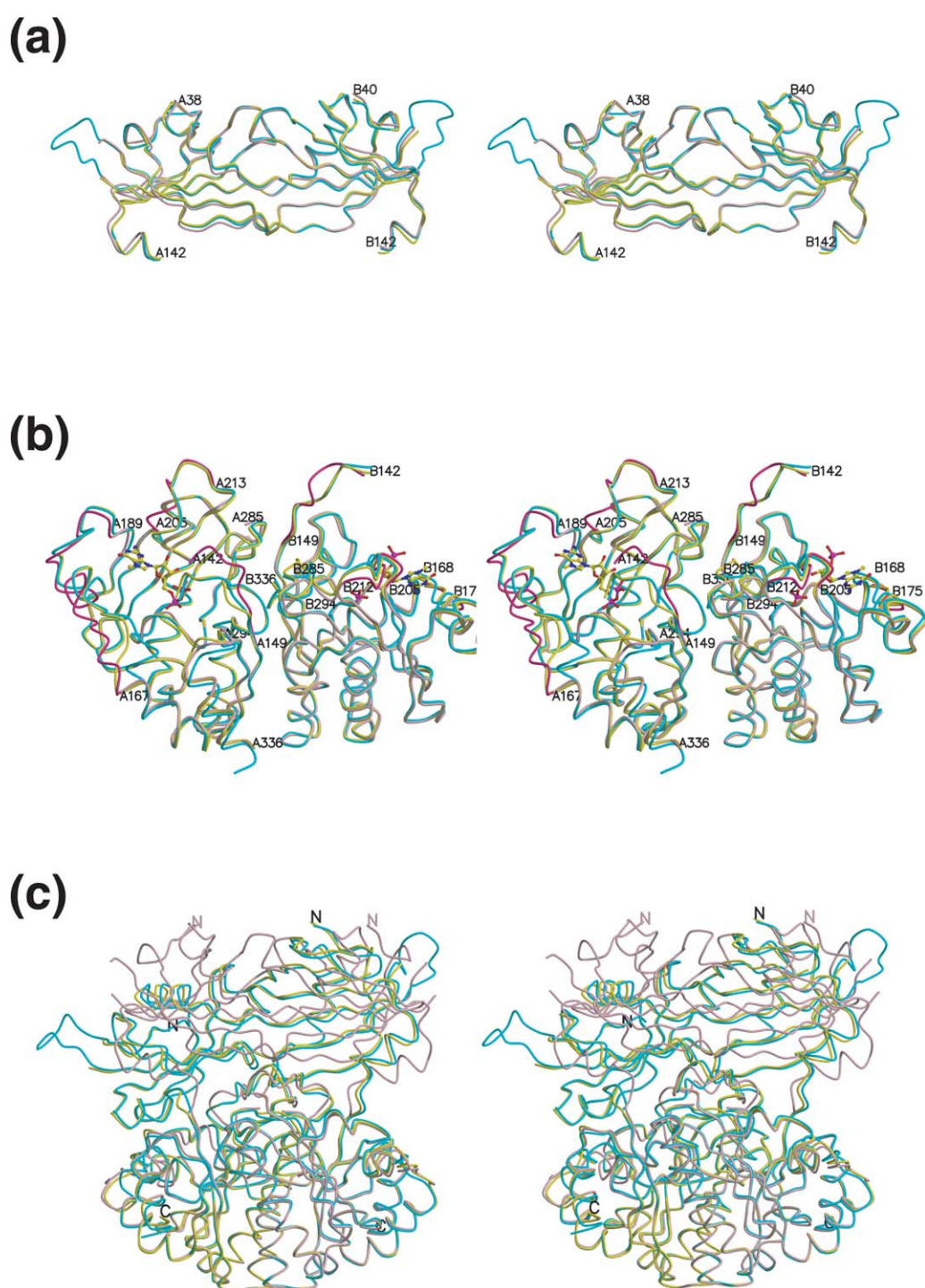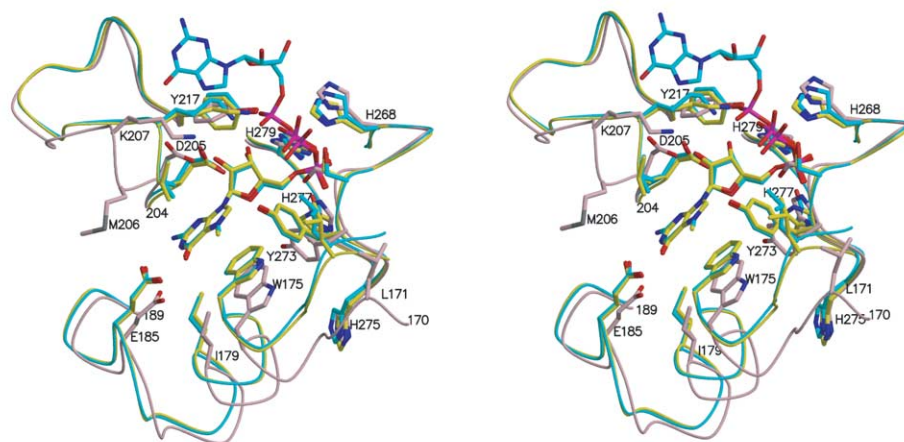
Figure 2. Comparison of apo-DcpS with m[7]GDP–DcpS and m[7]GpppG–DcpS. (a) Stereo view of superposition of the N-terminal domain-swapped dimer (residues 40–142) of apo-DcpS with those of m[7]GDP–DcpS and m[7]GpppG–DcpS. apo-DcpS is colored pink, m[7]GDP–DcpS is colored yellow and m[7]GpppG–DcpS is colored cyan. The same coloring scheme is used for (b) and (c). (b) Stereo view of superposition of the C-terminal dimeric domain (residues 142–336) of apo-DcpS with those of m[7]GDP–DcpS and m[7]GpppG–DcpS. Residues in apo-DcpS exhibiting large positional shifts upon cap binding are highlighted in magenta. m[7]GDP and m[7]GMP molecules in the m[7]GDP–DcpS complex are shown as stick models. (c) Stereo view of the C[α] traces of apo-DcpS, m[7]GDP–DcpS and m[7]GpppG–DcpS. The three structures were superimposed, aligning the C-terminal dimeric domain using all C[α] atoms. For clarity, the bound phosphate ions and nucleotides are removed. The view of apo-DcpS is the same as that in Figure 1(a).

Ser272 to Ala reduced the decapping activity by 30%.[22]

Other significant local conformational changes are observed in the hinge regions (residues 143–149) and the loop region (residues 285–294) connecting strands β12 and β13 defined in the complex of m[7]GpppG–DcpS.[22] In the apo-DcpS, both hinge regions adopt an extended conformation that is
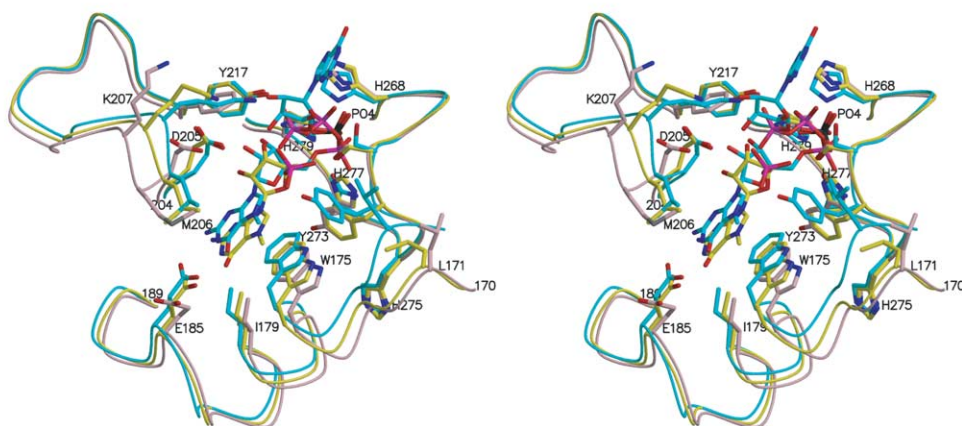
Figure 3. Comparison of the cap-binding pocket in apo-DcpS with those in m⁷GDP–DcpS and m⁷GpppG–DcpS. (a) Stereo view of superposition of the cap-binding pockets in the closed states of m⁷GDP–DcpS and m⁷GpppG–DcpS with that of apo-Dcps. (b) Stereo view of superposition of the cap-binding pockets in the open states of m⁷GDP–DcpS and m⁷GpppG–DcpS with that of apo-Dcps. The m⁷GMP, m⁷GDP m⁷GpppG molecules, the phosphate ion and residues involved in interactions with the nucleotides are shown as stick models. The coloring scheme is the same as that used in Figure 2.

different from those in both the closed and open states of DcpS with bound cap (Figure 2(b)). Residues 285–294 are located in the upper part of the dimer interface in the dimeric C-terminal domain. In the apo-DcpS, this region is disordered for both protomers, whereas it is ordered upon cap binding (Figure 2(b)), suggesting that this region may play a regulatory role in cap binding.

The role of Tyr273

Comparison of the structures suggests that Tyr273 plays an important role in cap binding. The structure of m⁷GpppG–DcpS showed that Tyr273 adopted the same conformation in both the closed and open states of DcpS, and makes van der Waals (VDW) contact with the m⁷G base.[22] Close

inspection of the cap-binding pocket among the structures of apo-Dcps, m⁷GDP–DcpS and m⁷GpppG–DcpS showed that Tyr273 displays remarkable conformational changes upon cap binding. Both Tyr273 residues in apo-DcpS pointed downwards, whereas they are flipped 90° and point upwards in both the closed and open states of m⁷GpppG–DcpS (Figures 3, 4(a) and (b)). Moreover, in the structure of m⁷GDP–DcpS, Tyr273 in the closed state adopts the same conformation as that in the m⁷GpppG–DcpS complex, whereas in the open state it displays the same conformation as that in apo-Dcps (Figures 3, 4(a) and (b)).

The conformational changes of Tyr273 in the structures of DcpS with, and without, bound ligands lead to different interactions with other
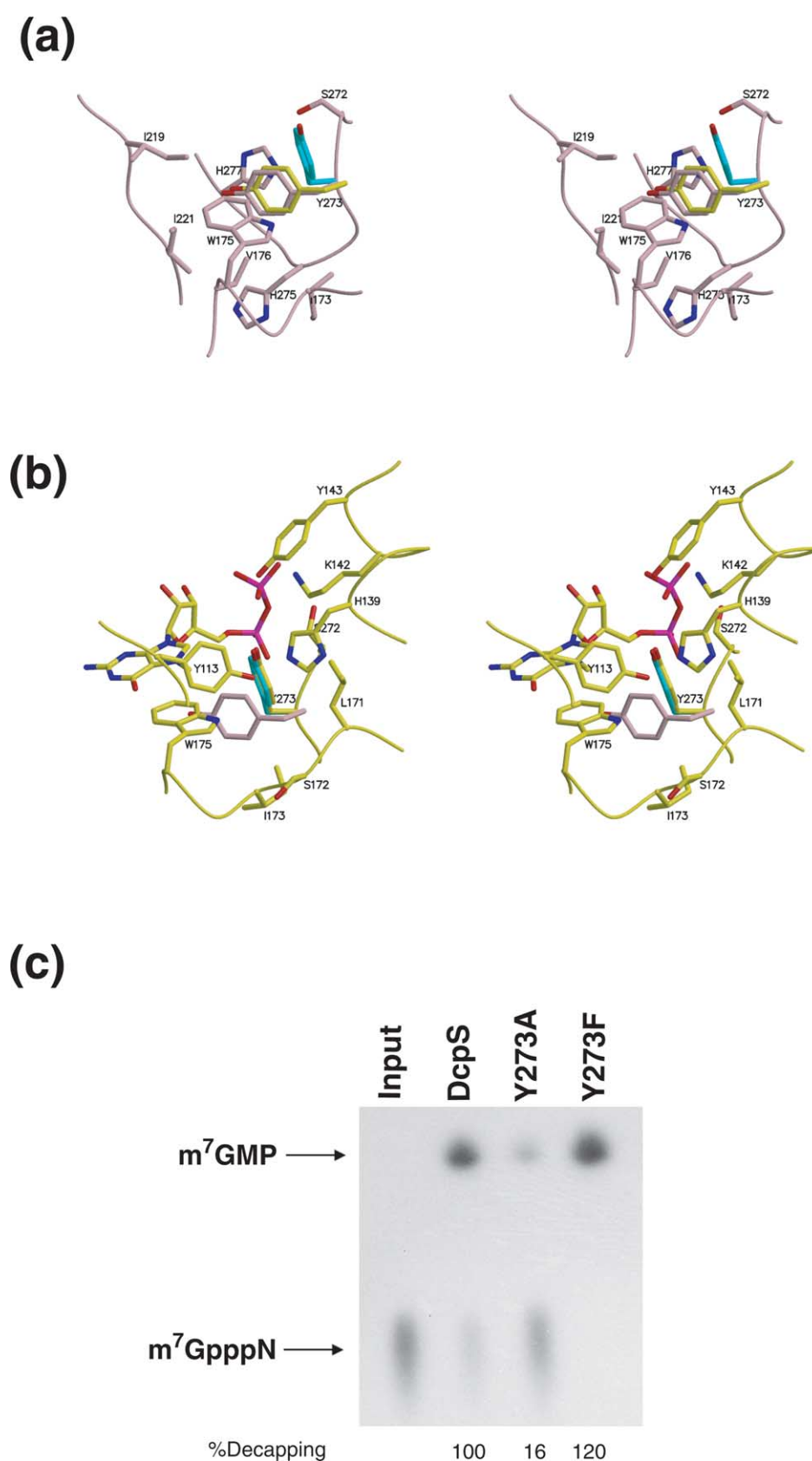
Figure 4. Conformational changes of Tyr273 and decapping assay of its mutants. (a) Stereo view of the conformation of Tyr273 in apo-DcpS and its interactions with other protein residues. Tyr273 residues in the open states of m^{7}GDP–DcpS and m^{7}GpppG–DcpS are colored yellow and cyan, respectively. (b) Stereo view of the conformation of Tyr273 in the closed state of m^{7}GDP–DcpS and its interactions with other protein residues. Tyr273 residues in the closed state of m^{7}GpppG–DcpS and in apo-DcpS are colored cyan and pink, respectively. (c) Decapping assays of wild-type DcpS or DcpS mutated on Tyr273. Reaction products were resolved by PEI-TLC developed in 0.3 M LiCl, 1 M formic acid. Each reaction contained 10 ng of the indicated protein.

amino acids. In the apo-DcpS and the open state of m^{7}GDP–DcpS, Tyr273 contacts with Ile221, Ile219, Trp175, Val176, and His277, residues exclusively from the C-terminal domain *via* hydrophobic, VDW and stacking interactions (Figure 4(a)). In the closed states of m^{7}GDP–DcpS and m^{7}GpppG–DcpS, Tyr273 makes hydrophobic interactions with residues Trp175, Leu171, VDW contacts with residues Tyr113 and His139 from the N-terminal, the m^{7}G base and residues Lys142 and Tyr143 from the hinge region (Figure 4(b)). In the open state of m^{7}GpppG–DcpS, the flipped Tyr273 makes hydrophobic interactions only with residues Trp175 and Leu171, due to its solvent exposure (Figure 3(b)).

To assess its functional role, Tyr273 was mutated to either Ala or Phe and tested for decapping activity. Substitution of Tyr273 with Ala resulted in a protein with activity equivalent to 16% of that of the wild-type enzyme, whereas mutation of Tyr273 to Phe generated a protein with 120% of the wild-type activity (Figure 4(c)), suggesting that Tyr273 plays an important role in cap binding and catalysis. Structural and mutational analysis of the m^{7}GpppG–DcpS complex suggested that the enzyme closure is a critical step during the catalytic cycle.[22] Examination of the structures of DcpS with bound cap showed that the flipped Tyr273 in the closed state interacts with residues from the N-terminal domain and the hinge region, as well as with the cap itself, thereby facilitating directly or indirectly the cap binding by contributing to the formation of the closed configuration (see above; and see Figure 4(b)). Mutation of Tyr273 to Ala would disrupt these interactions, therefore affecting cap binding and catalysis. Mutation of Tyr273 to Phe would have little effect on its interactions with other residues in the closed state, since these interactions are mediated by hydrophobic and VDW contacts in the closed state (Figure 4(b)), but would enhances its interaction with other amino acids in the predominantly hydrophobic environment in the open form of DcpS with bound m^{7}GMP or in apo-DcpS (Figure 4(a)), thereby facilitating the product release by helping the enzyme back to the native state after hydrolysis. Note that Tyr273 is conserved in the DcpS family and is replaced by Phe only in orthologues from both *Drosophila* and *Caenorhabditis elegans*.[22]

Coexistence of the substrate-bound and product-bound complexes in m^{7}GDP–DcpS

Examination of the open and closed states of m^{7}GDP–DcpS suggests that the closed complex represents a substrate complex, while the open complex represents a product-bound complex. As shown in Figure 3(a), in the closed configuration, m^{7}GDP adopts a conformation that is similar to the m^{7}Gpp portion of m^{7}GpppG in the closed state of the m^{7}GpppG–DcpS complex. Recognition and binding of m^{7}GDP is achieved in the same way as that observed in the closed state of the m^{7}GpppG–DcpS complex,[22] where the interactions of m^{7}G

base, ribose and the α and β-phosphate groups of m^{7}GDP with DcpS are mediated by the residues from both the N-terminal and C-terminal domains, and the hinge region. The active-site His277, which has been mutated to asparagine in the m^{7}GpppG–DcpS complex, is in an in-line attack position against the α-phosphate group of m^{7}GDP, with the distance of 3.43 Å between its N^ε atom and the α-phosphate group of m^{7}GDP (Figure 3(a)). This distance is similar to what is observed in other members of the HIT protein family.[18] His268 and His279, whose mutations inactivate DcpS in decapping assay,[22] are located at the positions similar to those in the closed productive complex of m^{7}GpppG-DcpS (Figure 3(a)). Moreover, structural comparison of DcpS with other HIT proteins has indicated that the active-site architectures are similar among these HIT proteins.[22] Taken together, these observations suggest that the closed state of m^{7}GDP–DcpS might be a substrate-bound complex, and DcpS is likely to use a catalytic mechanism similar to that used by other HIT proteins[18] for hydrolyzing the cap structure.

In contrast, in the open configuration, the conformation of m^{7}GMP is different from the m^{7}Gp portion of m^{7}GpppG in the open state of the m^{7}GpppG–DcpS complex, and may represent a product-bound form of the enzyme. In the open configuration of the m^{7}GpppG–DcpS complex, the m^{7}Gp portion is buried deeply in the nucleotide-binding pocket, whereas in the open configuration of the m^{7}GDP–DcpS structure, m^{7}GMP is more solvent-exposed with its α-phosphate group protruding into the solvent region and making no contact with the protein residues (Figure 3(b)). A phosphate ion is located in the same position as that in apo-DcpS in the nucleotide-binding pocket, with the distance of its phosphorus atom to that of the γ-phosphate group (corresponding to the α-phosphate group of m^{7}GDP) of m^{7}GpppG of 1.3 Å in the open form of the m^{7}GpppG–DcpS complex (Figure 3(b)). The distance between the N^ε atom of His277 and the phosphorus atom in the α-phosphate position of m^{7}GMP is nearly 10 Å, similar to the distance between the N^ε atom of His112 and the α-phosphate group of AMP observed in the product-bound structures of PKCI and FHIT.[18] Moreover, Trp175, which is involved in stacking interaction with the m^{7}G base, moves away from the m^{7}G base due to the large conformational change of residues 168–175 (Figure 3(b); and see above); thus, making less favorable stacking interactions between its aromatic ring and the m^{7}G base. Finally, Tyr273, which is exposed to the solvent region in the open state of the m^{7}GpppG–DcpS complex,[22] flips 90° and points to the bottom of the binding pocket with its conformation highly resembling that in apo-DcpS (see above). The solvent-exposed m^{7}GMP, the native-like conformation of Tyr273, the movement of Trp175, and the distant location between His277 and the α-phosphate group of m^{7}GMP, suggest that the open state of the m^{7}GDP–DcpS complex is probably

the product-bound complex in which the observed m^{7}GMP mimics the true product of scavenger decapping.

Inherent flexibility of the N-terminal domain in apo-DcpS

Dynamic protein motions play an important role in the catalytic function of all enzymes, and are implicated in events such as binding of substrates or cofactors, product release, or allosteric regulation.[23–26] These processes are often mediated by conformational changes in the active site, by hinge-bending motions, or by reorientation of protein domains or entire protein subunits. The symmetric to asymmetric transition of DcpS induced by cap binding implies that the decapping reaction catalyzed by DcpS may be a dynamic process involving substantial movements in the DcpS polypeptide chain.

It has long been recognized that the variations in crystallographic temperature factors (*B*-factors) reflect the amplitude of motion as it varies along the polypeptide chain, as well as between different liganded complexes of the enzyme. The averaged *B*-factors of the main-chain atoms for each residue for apo-DcpS and m^{7}GDP–DcpS have been calculated (data not shown). Strikingly, the whole N-terminal domain rather than a particular region in it has substantially higher *B*-factors than that of the C-terminal domain in apo-DcpS. In contrast, the *B*-factors differences between the N-terminal and C-terminal domains are much smaller for m^{7}GDP–DcpS. The averaged *B*-factors of all the main-chain atoms for the N-terminal and C-terminal domains in apo-DcpS are 46.2 Å^2 and 16.4 Å^2, respectively, whereas those for the N-terminal and C-terminal domains in m^{7}GDP–DcpS are 30.7 Å^2 and 14.5 Å^2, respectively. These results suggest that without bound ligand, the N-terminal domain is rather flexible and unstable, while the C-terminal domain is relatively stable, and binding of ligand to DcpS decreases the flexibility of the N-terminal domain. Consistent with this notion, the N-terminal domain is connected by the flexible hinge and has no contact with the C-terminal domain in apo-DcpS, whereas upon cap binding, the N-terminal domain rotates with respect to the C-terminal domain and contacts one side of the dimeric C-terminal domains, thus creating the closed and open states simultaneously. The proposed catalytic mechanism of DcpS requires that the closed and open states must be alternated in order to promote cap cleavage and release the product.[22] The alternating nature of this mechanism requires that the N-terminal domain must be flexible and in a dynamic state ready for substrate binding and product release.

Mechanistic implications

The set of structures available for DcpS now allows a model for the cycle of substrate binding, catalysis and product release to be proposed. In the absence of any substrate, DcpS is in a symmetric open state, with both active sites available for substrate binding. Upon interaction with a cap moiety in one of the active sites, that side of the protein adopts a closed state, which involves a large-scale rotation of the N-terminal head domain, and local structural changes in the cap-binding pocket to stabilize the substrate binding. Hydrolysis of the cap structure leads to conformational changes in the cap-binding pocket, similar to what is seen in the open configuration of the m^{7}GDP–DcpS structure. These changes would be predicted to weaken the interaction of the m^{7}GMP product with the enzyme and enhance product release.

It is currently unclear how the two active sites interact during cycles of cap structure hydrolysis by DcpS, and particularly how substrate binding in one site relates to product release from the other. The proposed catalytic mechanism[22] requires that the closed and open states are alternated *via* 30 Å conformational change; while one side closes to promote catalysis during cap cleavage, the other side is forced open to permit binding of a short capped mRNA. However, this model is unlikely, given that the intermediate state where both sites are closed would require substantial distortion of the N-terminal domains, which would presumably be energetically unfavorable. Moreover, the two sites do not appear to be connected mechanistically, since we do not observe substantial changes in the interface of N-terminal and C-terminal domains between apo-DcpS and the open states of both m7GDP-bound and m7GpppG-bound DcpS structures, which suggests that the two active sites act relatively independently of each other at a structural level. The flexible nature of the N-domain relative to the C-domain in the DcpS structures suggest a catalytic cycle wherein any transition between open and closed states on opposing sides of the dimer requires an intermediate form similar to the apo-DcpS structure, where both sites are in the open configuration. In this view, the binding of substrate to one site is coupled with hydrolysis and release of the product in the other site in a kinetic manner.

A key unresolved question concerns how access of the cap-binding site of DcpS is regulated. Such regulation is necessary to ensure that only residual cap structure substrate can be degraded, thereby preventing the premature decapping of the capped mRNA by DcpS. Biochemical studies have demonstrated that DcpS is a modular protein requiring both the core HIT fold at the C-terminal domain and the N-terminal domain for cap binding and hydrolysis.[27] The structures of DcpS with bound cap analogues led to a model wherein DcpS is unable to bind the large capped mRNA due to the increased entropy inherent to larger and more flexible RNA molecules, thus blocking substrate binding, closure, and enzyme activation.[22] However, more recently, it was shown that DcpS is capable of binding long capped mRNA with a K_d of

1.2 μM, but has 2500-fold lower capacity to hydrolyze the capped RNA relative to the cap structure substrate.[27] These results suggest that DcpS can bind the cap on larger mRNA molecules with reasonable affinity, but the enzyme closure and activation are probably restricted by a steric hindrance or increased entropy of larger RNA or both.

In order to gain insight into how RNA might interact with DcpS, we performed electrostatic potential mapping on the molecular surface of apo-DcpS. This revealed a prominent, positively charged patch located on the central channel of the enzyme (Figure 5). This channel is sufficiently wide for binding a single-strand mRNA molecule. Further experimental data are required to check whether this channel is involved in mRNA binding. One intriguing possibility is that a long RNA might bind in the putative RNA channel of both sites and, by doing so, prevent either site from closing completely and forming an active catalytic site. In this model, only mRNAs that are short enough to interact with only one site at a time would be substrates for DcpS.
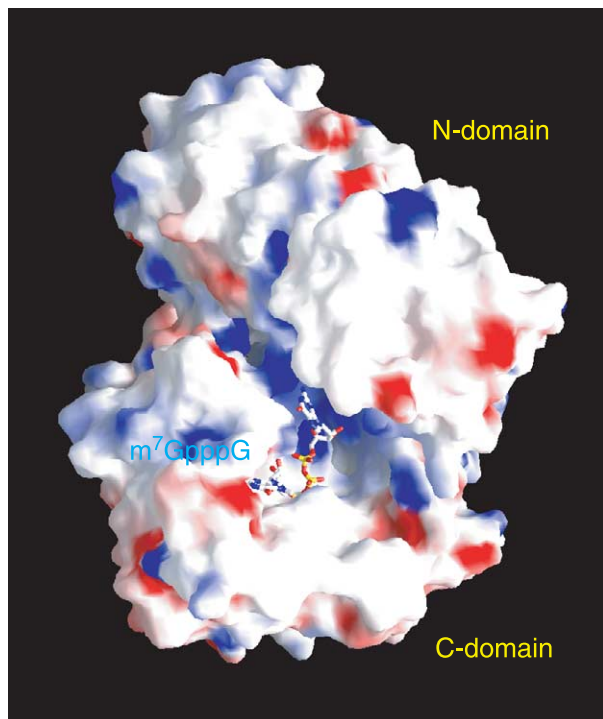

Figure 5. Solvent-accessible surface and electrostatic potential of apo-DcpS. The Figure reveals the positive electrostatic potential located in the channel between the N-terminal and C-terminal domains of apo-DcpS. m⁷GpppG of the m⁷GpppG–DcpS complex superimposed on apo-Dcps is shown as a stick model. This Figure was produced using GRASP.[36]

Materials and Methods

Protein expression and purification

The full-length DcpS (1-337) gene was cloned from human Marathon-Ready™ (Clontech) by PCR amplification and expressed as a glutathione-*S*-transferase (GST)-fusion protein in *E. coli*. Cells harboring the GST–DcpS fusion protein were lysed by incubation with lysozyme in a lysis buffer (20 mM Tris–HCl (pH 7.6), 500 mM NaCl, 2 mM DTT, 1 mM EDTA, 0.1 mM PMSF, 2 mM benzamide) followed by sonication. The clarified lysate was loaded onto a glutathione-Sepharose 4B column (Amersham). GST-fusion protein was eluted by glutathione and cleaved by PreScission protease (Amersham) overnight at 297 K. After desalting, cleaved protein was loaded onto a glutathione-Sepharose 4B column again and further purified by MonoQ™ (Amersham) anion-exchange chromatography and Hiload 20/60 Superdex75 (Amersham) gel-filtration column chromatography. Purified DcpS protein was eluted with 20 mM Tris–HCl (pH 7.6), 100 mM NaCl, 2 mM DTT. The protein was concentrated to ~2 mg/ml for crystallizations. Two Leu (Leu206 and Leu317) to Met mutations were introduced by using a QuickChange® Site-Directed Mutagenesis kit (STRATAGENE). SeMet-substituted mutant DcpS protein was purified using the same protocol as that used for wild-type DcpS, except the concentration of DTT was 10 mM.

Crystallization, data collection, and structure determination

Initial crystals of human DcpS were obtained from a condition consisting of 20% (w/v) PEG 8000, 0.05 M KH₂PO₄ by the hanging-drop, vapor-diffusion method at 293 K. By optimizing the conditions and seeding technique, crystals of native and SeMet-substituted mutant DcpS proteins suitable for data collection were obtained by mixing an equal volume of protein solution with a reservoir solution consisting of 25% (w/v) PEG 3350, 0.1 M KH₂PO₄, 0.05 M NaCl and 26% (v/v) glycerol. The m⁷GDP–DcpS crystals were obtained by soaking the native crystals in a stabilizing buffer containing 5 mM m⁷GDP for 12 hours. Before data collection, crystals were fished out directly from the mother liquor and fast-frozen in liquid nitrogen.

A SeMet single-wavelength anomalous diffraction (SAD) data set was collected at wavelength of 0.93 Å on beamline BM14UK at ESRF (Grenoble, France) using a MarCCD detector and processed with the HKL software package.[28] Data collection of m⁷GDP–DcpS was carried out on beamline BL42B at Spring-8 (Hyogo, Japan). Diffraction data were recorded on a MarCCD detector and processed with MOSFLM and CCP4 programs.[29] The structure of hDcpS in complex with m⁷GDP was solved using the molecular replacement method with Molrep,[30] using the structure of hDcpS in complex with m7GpppG[22] as a search model. Manual model rebuilding was performed using program O.[31] Crystallographic refinement was carried out using CNS.[32] The final round of refinement was done by REFMAC5.[33] The structure of the C-terminal dimeric domains of the m⁷GDP complex (residues A146–A336, andB146–B336) was then used as a search model for determining the structure of apo-DcpS by the molecular replacement method. A partial model of the N-domain structure of apo-DcpS was built automatically using Arp/wArp.[34] The rest of the model was built manually using O.[31] Crystallographic refinement was carried out with CNS[32]

and then REFMAC5.[33] The data collection and the final refinement statistics are summarized in Table 1.

Assay of decapping activity

RNA was transcribed *in vitro* by an AmpliScribe™ SP6 High Yield Transcription Kit (Clontech), using the control DNA in the kit as a template. RNA was cap-labeled using the vaccinia virus capping enzyme in the presence of [α-^{32}P]GTP and *S*-adenosyl-methionine (SAM) to label the first phosphate group relative to the terminal guanosine base (m7G*pppN-). The ^{32}P-labeled cap structure was generated by treating the cap-labeled RNA with one unit of nuclease S1 (Promega). The labeled cap analog was used as substrate for the decapping assays. Recombinant DcpS or its mutant derivatives were incubated with the labeled cap structure in IVDA buffer (10 mM Tris–HCl (7.5), 100 mM potassium acetate, 2 mM magnesium acetate, 2 mM DTT, 10 mM creatine phosphate, 1 mM ATP, 0.4 mM GTP, 0.1 mM spermine) for ten minutes at 37 °C. The reactions were stopped and extracted once with a 1 : 1 (v/v) mixture of phenol and chloroform. Products were resolved on PEI-cellulose TLC plates (Sigma). An aliquot of the reaction products was spotted onto the TLC plate and developed in 0.3 M LiCl, 1 M formic acid at room temperature. The plates were dried and exposed to Kodak BioMax film and quantified using a phosphorimager.

Protein Data Bank accession codes

The coordinates and structure-factor amplitudes for apo-DcpS and the m⁷GDP–DcpS complex have been deposited in the Protein Data Bank with accession codes 1XML and 1XMM, respectively.

References

1. Parker, R. & Song, H. (2004). The enzymes and control of eukaryotic mRNA turnover. *Nature Struct. Mol. Biol.* **11**, 121–127.
2. Muhlrad, D. & Parker, R. (1994). Premature translational termination triggers mRNA decapping. *Nature*, **370**, 578–581.
3. Cao, D. & Parker, R. (2003). Computational modeling and experimental analysis of nonsense-mediated decay in yeast. *Cell*, **113**, 533–545.
4. Mitchell, P. & Tollervey, D. (2003). An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3′–5′ degradation. *Mol. Cell.* **11**, 1405–1413.
5. Takahashi, S., Araki, Y., Sakuno, T. & Katada, T. (2003).

6. Lejeune, F., Li, X. & Maquat, L. E. (2003). Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. *Mol. Cell.* **12**, 675–687.
7. Frischmeyer, P. A., van Hoof, A., O'Donnell, K., Guerrerio, A. L., Parker, R. & Dietz, H. C. (2002). An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science*, **295**, 2258–2261.
8. van Hoof, A., Frischmeyer, P. A., Dietz, H. C. & Parker, R. (2002). Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science*, **295**, 2262–2264.
9. Bessman, M. J., Frick, D. N. & O'Handley, S. F. (1996). The MutT proteins or "Nudix" hydrolases, a family of versatile, widely distributed, "housecleaning" enzymes. *J. Biol. Chem.* **271**, 25059–25062.
10. Wang, Z., Jiao, X., Carr-Schmid, A. & Kiledjian, M. (2002). The hDcp2 protein is a mammalian mRNA decapping enzyme. *Proc. Natl Acad. Sci. USA*, **99**, 12663–12668.
11. Dunckley, T. & Parker, R. (1999). The DCP2 protein is required for mRNA decapping in *Saccharomyces cerevisiae* and contains a functional MutT motif. *EMBO J.* **18**, 5411–5422.
12. van Dijk, E., Cougot, N., Meyer, S., Babajko, S., Wahle, E. & Seraphin, B. (2002). Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. *EMBO J.* **21**, 6915–6924.
13. Piccirillo, C., Khanna, R. & Kiledjian, M. (2003). Functional characterization of the mammalian mRNA decapping enzyme hDcp2. *RNA*, **9**, 1138–1147.
14. Steiger, M., Carr-Schmid, A., Schwartz, D. C., Kiledjian, M. & Parker, R. (2003). Analysis of recombinant yeast decapping enzyme. *RNA*, **9**, 231–238.
15. Wang, Z. & Kiledjian, M. (2001). Functional link between the mammalian exosome and mRNA decapping. *Cell*, **107**, 751–762.
16. Liu, H., Rodgers, N. D., Jiao, X. & Kiledjian, M. (2002). The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases. *EMBO J.* **21**, 4699–4708.
17. Seraphin, B. (1992). The HIT protein family: a new family of proteins present in prokaryotes, yeast and mammals. *DNA Seq.* **3**, 177–179.
18. Lima, C. D., Klein, M. G. & Hendrickson, W. A. (1997). Structure-based analysis of catalysis and substrate definition in the HIT protein family. *Science*, **278**, 286–290.
19. Pace, H. C., Garrison, P. N., Robinson, A. K., Barnes, L. D., Draganescu, A., Rosler, A. *et al.* (1998). Genetic, biochemical, and crystallographic characterization of Fhit-substrate complexes as the active signaling form of Fhit. *Proc. Natl Acad. Sci. USA*, **95**, 5484–5489.
20. Brenner, C., Garrison, P., Gilmour, J., Peisach, D., Ringe, D., Petsko, G. A. & Lowenstein, J. M. (1997). Crystal structures of HINT demonstrate that histidine triad proteins are GalT-related nucleotide-binding proteins. *Nature Struct. Biol.* **4**, 231–238.
21. van Dijk, E., Le Hir, H. & Seraphin, B. (2003). DcpS can act in the 5′–3′ mRNA decay pathway in addition to the 3′–5′ pathway. *Proc. Natl Acad. Sci. USA*, **100**, 12081–12086.
22. Gu, M., Fabrega, C., Liu, S. W., Liu, H., Kiledjian, M. &

Lima, C. D. (2004). Insights into the structure, mechanism, and regulation of scavenger mRNA decapping activity. *Mol. Cell.* **14**, 67–80.

23. Schnell, J. R., Dyson, H. J. & Wright, P. E. (2004). Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 119–140.

24. Karplus, M. (2000). Aspects of protein reaction dynamics: deviations from simple behavior. *J. Phys. Chem. B*, **104**, 11–27.

25. Hammes, G. G. (1964). Mechanism of enzyme catalysis. *Nature*, **204**, 342–343.

26. Somogyi, B., Welch, G. R. & Damjanovich, S. (1984). The dynamic basis of energy transduction in enzymes. *Biochim. Biophys. Acta*, **768**, 81–112.

27. Liu, S. W., Jiao, X., Liu, H., Gu, M., Lima, C. D. & Kiledjian, M. (2004). Functional analysis of mRNA scavenger decapping enzymes. *RNA*, **10**, 1412–1422.

28. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.

29. CCP4 (Collaborative Computational Project No.4). (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.

30. Vagin, A. & Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *J. Appl. Crystallog.* **30**, 1022–1025.

31. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.

32. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al*. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.

33. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.

34. Perrakis, A., Morris, R. & Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol*, **6**, 458–463.

35. Kraulis, P. J. (1991). MOLSCRIPT:a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.

36. Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281–296.

Edited by J. Doudna