

# Structure of the RNA Triphosphatase Domain of Mouse Capping Enzyme

A. Changela, A. Mondragón

*Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, IL, U.S.A.*

## Introduction

RNA triphosphatase is an essential mRNA processing enzyme that catalyzes the first step in 5' cap formation, the hydrolysis of the phosphate of nascent pre-mRNA to form a diphosphate end that is then capped with GMP by the enzyme RNA guanylyltransferase.<sup>1</sup> In yeast, the triphosphatase and guanylyltransferase activities reside in separate polypeptides encoded by different genes, whereas mammalian species contain a single bifunctional polypeptide (Mce1) composed of an N-terminal triphosphatase domain (residues 1-210) and a C-terminal guanylyltransferase domain (residues 211-597). The triphosphatase domain of metazoan and plant capping enzymes contains a HCxxxxR(S/T) motif that defines the cysteine phosphatase superfamily, which includes protein tyrosine phosphatases, dual specificity phosphatases, and phosphoinositide phosphatases. The finding of the HCxxxxR(S/T) motif in metazoan capping enzymes and the demonstration that the conserved cysteine is required for RNA triphosphatase activity<sup>2,4</sup> challenged the presumption that polypeptides containing the signature sequence necessarily act as protein phosphatases. The functional diversity of the cysteine phosphatase superfamily was further underscored by the discovery that the tumor suppressor PTEN dephosphorylates phospholipid substrates.<sup>5</sup>

In order to understand the catalytic mechanism of the RNA triphosphatase domain of a metazoan enzyme and to compare its structure to that of other proteins of the cysteine phosphatase superfamily, we solved the structure of the N-terminal domain of Mce1 using MAD phasing.

## Methods and Materials

Mce1 crystals were grown as described elsewhere.<sup>6</sup> They were transferred in a single step to crystallization solution containing 30% 2, 4-methylpentanediol for 2-3 minutes and then flash cooled in liquid nitrogen. In order to decrease crystal mosaicity, the method of macromolecular crystal annealing<sup>7</sup> was used. Flash-cooled crystals were removed from the nitrogen stream and transferred to 300  $\mu$ l of cryoprotectant for 3 minutes before being refrozen in the stream. All data were collected at 100K using synchrotron radiation. Se-Met MAD data<sup>8</sup> were measured at four wavelengths corresponding to the selenium absorption edge peak ( $\lambda_1$ ), inflection point ( $\lambda_2$ ), a high remote point ( $\lambda_3$ ), and a low remote point ( $\lambda_4$ ). MAD data were collected in 15° segments using the inverse beam technique. All data were processed using MOSFLM<sup>9</sup> and scaled using SCALA.<sup>10</sup> The crystals belong to space group C222<sub>1</sub> with unit cell dimensions of a=62.2 Å, b=98.9 Å, and c=71.8 Å and have one molecule in the asymmetric unit.

Four out of five possible selenium sites were identified using a heavy atom search in CNS<sup>11</sup> with data collected at the selenium absorption peak ( $\lambda_1$ ). The missing fifth site corresponded to the disordered N-terminal methionine. The four selenium sites were refined, and phases to 2.05 Å were calculated by treating the

MAD data as a special case of MIR using MLPHARE.<sup>10</sup> Phases were improved by density modification using DM.<sup>10</sup> The resulting electron density map clearly showed elements of protein secondary structure, and even solvent molecules could be identified. The excellent quality of the experimental map and the availability of a high resolution native data set enabled the use of ARP/wARP<sup>12</sup> to automatically trace ~90% of the polypeptide chain. Additional model building was carried out in O.<sup>13</sup> The model was refined with REFMAC<sup>10</sup> using data to 1.65 Å. ARP<sup>10</sup> was used to place water molecules into Fo-Fc difference density peaks greater than 3  $\sigma$  and that were within hydrogen bonding distances. No electron density was seen for the first 4 N-terminal residues or the last 12 C-terminal residues, and a small loop region was also found to be disordered. The final model contains residues 5-113, 119-198, and 212 water molecules. All residues are found within the most favored or allowed regions in the Ramachandran plot.

## Results

The structure of the triphosphatase domain of the mouse capping enzyme was solved to high resolution using MAD techniques. The 1.65 Å structure reveals a deep, positively charged active site pocket that can fit a 5' triphosphate end. The structure, together with biochemical and mutational results, shows that, despite sharing a HCxxxxR(S/T) motif, a phosphoenzyme intermediate, and a core  $\beta$  fold with other cysteine phosphatases, the mechanism of phosphoanhydride cleavage by mammalian capping enzyme is different in key respects from that used by protein phosphatases to hydrolyze phosphomonoesters. The most significant difference is the absence of a carboxylate general acid catalyst in RNA triphosphatase.

## Acknowledgments

Research was supported by the NIH and an NRSA Institutional Training Grant in Molecular Biophysics for A.C. Portions of this work were performed at the DuPont-Northwestern-Dow Collaborative Access Team (DND-CAT) Synchrotron Research Center at the Advanced Photon Source DND-CAT is supported by DuPont, Dow, and the NSF. Use of the APS was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-ENG-38.

## References

- 1 S. Shuman, *Prog. Nucleic Acid Res. Mol. Biol.* **66**, 1-40 (2000).
- 2 C.K. Ho, B. Schwer, and S. Shuman, *Mol. Cell Biol.* **18**, 5189-5198 (1998).
- 3 T. Takagi, C.R. Moore, F. Diehn, and S. Buratowski, *Cell* **89**, 867-873 (1997).
- 4 Y. Wen, Z. Yue, and A.J. Shatkin, *Proc. Natl. Acad. Sci. USA* **95**, 12226-12231 (1998).
- 5 T. Maehama and J.E. Dixon, *J. Biol. Chem.* **273**, 13375-13378 (1998).

- <sup>6</sup> A. Changela, C.K. Ho, A. Martins, S. Shuman, and A. Mondragón, *EMBO Journal*, in press.
- <sup>7</sup> J.M. Harp, D.E. Timm, and G.J. Bunick, *Acta Crystallogr. D* **54**, 622-628 (1998).
- <sup>8</sup> W.A.Hendrickson, *Science* **254**, 51-58 (1991).
- <sup>9</sup> A.G.W. Leslie, *Crystallographic Computing* (Oxford University Press, Oxford, 1990).
- <sup>10</sup> Collaborative-Computational-Project-4. *Acta Crystallogr. D* **50**, 760-763 (1994).
- <sup>11</sup> A.T. Brunger et al., *Acta Crystallogr. D* **54**, 905-921 (1998).
- <sup>12</sup> A. Perrakis, R. Morris, and V.S. Lamzin, *Nat. Struct. Biol.* **6**, 458-463 (1999).
- <sup>13</sup> T.A. Jones, J.Y. Zou, S.W. Cowan, and Kjeldgaard. *Acta Crystallogr. A* **47**, 110-119 (1991).