Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structures of apo and GTP-bound molybdenum cofactor biosynthesis protein MoaC from *Thermus thermophilus* HB8

Shankar Prasad Kanaujia,[a] Jeyaraman Jeyakanthan,[b] Noriko Nakagawa,[c,d] Sathyaramya Balasubramaniam,[a] Akeo Shinkai,[c] Seiki Kuramitsu,[c,d] Shigeyuki Yokoyama[e,f] and Kanagaraj Sekar[a]*

[a]Bioinformatics Centre (Centre of Excellence in Structural Biology and Bio-computing), Indian Institute of Science, Bangalore 560 012, India, [b]Department of Bioinformatics, Alagappa University, Karaikudi 630 003, Tamilnadu, India, [c]RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan, [d]Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan, [e]RIKEN Systems and Structural Biology Center, Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan, and [f]Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Correspondence e-mail: sekar@physics.iisc.ernet.in

Received 5 March 2010
Accepted 21 May 2010

PDB References: MoaC, apo, 3jqj; 3jqk; GTP-bound, 3jqm.

The first step in the molybdenum cofactor (Moco) biosynthesis pathway involves the conversion of guanosine triphosphate (GTP) to precursor Z by two proteins (MoaA and MoaC). MoaA belongs to the *S*-adenosylmethionine-dependent radical enzyme superfamily and is believed to generate protein and/or substrate radicals by reductive cleavage of *S*-adenosylmethionine using an Fe–S cluster. MoaC has been suggested to catalyze the release of pyrophosphate and the formation of the cyclic phosphate of precursor Z. However, structural evidence showing the binding of a substrate-like molecule to MoaC is not available. Here, apo and GTP-bound crystal structures of MoaC from *Thermus thermophilus* HB8 are reported. Furthermore, isothermal titration calorimetry experiments have been carried out in order to obtain thermodynamic parameters for the protein–ligand interactions. In addition, molecular-dynamics (MD) simulations have been carried out on the protein–ligand complex of known structure and on models of relevant complexes for which X-ray structures are not available. The biophysical, structural and MD results reveal the residues that are involved in substrate binding and help in speculating upon a possible mechanism.

1. Introduction

Molybdenum cofactor (Moco) biosynthesis is an evolutionarily conserved pathway in archaea, eubacteria and eukaryotes, including humans. Moco consists of a mononuclear molybdenum coordinated by the dithiolene moiety of a family of tricyclic pyranopterin compounds (Chan *et al.*, 1995). It is considered to be an essential component that is required by enzymes that catalyze diverse key reactions in the global cycles of carbon, nitrogen and sulfur (Kisker *et al.*, 1997; Mendel & Bittner, 2006). A genetic deficiency of these enzymes leads to various autosomal recessive diseases with severe neurological symptoms that may even lead to death in early childhood (Johnson *et al.*, 1989). Five loci (*moa*, *mob*, *mod*, *moe* and *mog*) have been identified to be involved in Moco synthesis by chlorate-resistance screening in *Escherichia coli* (Rajagopalan & Johnson, 1992). Of these loci, *moa* and *moe* are required for the initial steps of Moco biosynthesis, which involve the synthesis of molybdopterin (MPT; Nohno *et al.*, 1988; Rivers *et al.*, 1993). The gene products of *moaA* and *moaC* (MOCS1A and MOCS1B in humans, Cnx2 and Cnx3 in plants) are involved in the conversion of GTP to cPMP (precursor Z; Wuebbens & Rajagopalan, 1993; Hanzelmann *et al.*, 2002, 2004), which is further converted to MPT by MoaD–MoaE protein complexes together with the MoeB protein (Rivers *et al.*, 1993; Pitterle & Rajagopalan, 1989, 1993). Similar types of processes have also been observed in the

biosynthesis of folate, riboflavin and biopterin, in which a guanosine derivative serves as an initial biosynthetic precursor. However, in contrast, Moco synthesis involves the rearrangement of the guanosine C8 atom as the first carbon of the precursor Z side chain. MoaA belongs to the *S*-adenosylmethionine-dependent radical enzymes and catalyzes the formation of protein and/or substrate radicals by reductive cleavage of *S*-adenosylmethionine with the help of a [4Fe–4S] cluster (Menendez *et al.*, 1996; Sofia *et al.*, 2001; Hanzelmann *et al.*, 2004; Hanzelmann & Schindelin, 2004, 2006). The GTP molecule first binds to MoaA and an intermediate compound (formamidopyrimidine-type; FPT) is generated which is subsequently used by MoaC. Furthermore, MoaC catalyzes the release of pyrophosphate from the MoaA-generated intermediate compound and the formation of the cyclic phosphate of precursor Z (Hanzelmann & Schindelin, 2006).

Ligand-free crystal structures of MoaC from *Escherichia coli* (*Ec*MoaC; PDB code 1ekr; Wuebbens *et al.*, 2000), *Pyrococcus horikoshii* (*Ph*MoaC; PDB code 2ekn; N. K. Lokanath, K. J. Pampa, T. Kamiya & N. Kunishima, unpublished work), *Sulfolobus tokodaii* (*St*MoaC; PDB code 2ohd; Yoshida *et al.*, 2008) and *Geobacillus kaustophilus* (*Gk*MoaC; PDB code 2eey; N. K. Lokanath, K. J. Pampa, T. Kamiya & N. Kunishima, unpublished work) are available. However, no structural study of a ligand-bound form is available in the literature. To this end, we have determined three crystal structures of MoaC from a highly thermophilic eubacterium (*Thermus thermophilius* HB8; *Tt*MoaC): two apo forms (space groups $P2_1$ and $R32$) and a GTP-bound form (space group $C222_1$). The ligand-bound form provides the first structural evidence of the binding of a 5′-GTP molecule to MoaC. In addition, isothermal titration calorimetry (ITC) experiments have been carried out to support the findings from the crystallographic results. Furthermore, molecular-dynamics (MD) simulations have been carried out on both known and modelled protein–ligand complexes to corroborate the above results. Thus, the present study should enhance the existing knowledge of the Moco biosynthesis pathway, particularly the first step, which is not clearly understood.

2. Materials and methods

2.1. Protein purification, crystallization and data collection

The cloning, expression, purification, crystallization and data collection of the ligand-free protein has been described previously (Kanaujia *et al.*, 2007). The crystallization of the GTP-bound form of *Tt*MoaC was carried out using the conditions that were used for the native form. Protein solution (0.6 m*M*) was incubated overnight with 5′-GTP at a final concentration of 10 m*M* before crystallization. A droplet consisting of 2 µl protein solution and 2 µl precipitant solution was equilibrated against 200 µl reservoir solution [0.1 *M* phosphate–citrate buffer pH 4.2, 25%(*v*/*v*) 1,2-propanediol, 5%(*w*/*v*) PEG 3350 and 10%(*v*/*v*) glycerol]. Crystals appeared within a week. The intensity data for the GTP-bound crystal were collected at 100 K using the home source with a MAR 345 imaging-plate detector mounted on a Rigaku RU-300 generator (operated at 40 kV and 80 mA). The data were processed and scaled using the *HKL* suite (Otwinowski & Minor, 1997). Details of the data-collection statistics are given in Table 1.

2.2. Structure solution, refinement and validation

The structures of the apo crystal forms ($P2_1$ and $R32$) were solved by molecular-replacement calculations using the program *Phaser* (McCoy *et al.*, 2007). The three-dimensional atomic coordinates of *Ec*MoaC (PDB code 1ekr) were used as the search model (53% sequence identity). A total of 5% of the reflections were kept aside for the calculation of $R_{\rm free}$ (Brünger, 1992). In both the apo forms difference electron density (up to 12σ in the $|F_{\rm o} - F_{\rm c}|$ map) appropriate for a phosphate ion was observed in the active site. However, water molecules were first located and added from difference electron-density maps with the criterion of peak heights greater than 2.8σ. Subsequently, phosphate ions were also modelled and refined. Details of the refinement statistics are given in Table 1.

A similar approach to that described above was used to refine the GTP-bound form. Preliminary calculations (Matthews coefficient of 1.99 Å^3 Da^{-1}, solvent content of 38.1%) suggested the presence of nine subunits in the asymmetric unit (Matthews, 1968). In the initial stage of refinement, clear difference electron density (up to 8σ in the $|F_{\rm o} - F_{\rm c}|$ map) appropriate for triphosphates was observed in the active site. Water molecules were first located and fitted into the model to improve the electron density for the bound ligand molecules. Subsequently, 5′-GTP molecules were added to the model and refined. The topology parameters for 5′-GTP were generated using the *HIC-UP* webserver (Kleywegt, 2007). The refinement statistics are given in Table 1. The molecular-modelling program *Coot* (Emsley & Cowtan, 2004) was used to display the electron-density maps for model fitting and adjustment. All atoms were refined with unit occupancies. Refinement was carried out using the program *CNS* v.1.2 (Brünger *et al.*, 1998). Simulated-annealing OMIT maps were calculated to correct or to check the final protein models. The programs *PROCHECK* (Laskowski *et al.*, 1993) and *MolProbity* (Chen *et al.*, 2010) were used to check and validate the quality of the final refined models. The atomic coordinates and structure factors of both the apo forms (PDB codes 3jqj and 3jqk) and the GTP-bound form (PDB code 3jqm) have been deposited in the RCSB Protein Data Bank (Berman *et al.*, 2000).

2.3. Isothermal titration calorimetry (ITC)

All the ITC experiments were performed using a VP-ITC MicroCalorimeter (MicroCal Inc., Northampton, Massachusetts, USA) at 293 K. In each experiment, purified *Tt*MoaC protein solution was dialyzed against 20 m*M* Tris–HCl buffer pH 8.0, 0.15 *M* NaCl for 12 h with three changes. The ligand solutions were prepared in the final dialyzed protein buffer. The sample cell (volume 1.4 ml) was filled with 75 µ*M* purified

Table 1
X-ray data and refinement statistics for free and GTP-bound forms of *Tt*MoaC.

Values in parentheses are for the highest resolution shell.

	Form I (PDB code 3jqj)	Form II (PDB code 3jqk)	Form III (PDB code 3jqm)
Wavelength (Å)	1.0	0.97243	1.5418
Temperature (K)	100	100	100
Space group	$P2_1$	$R32$	$C222_1$
Unit-cell parameters (Å, °)	a = 64.81, b = 109.84, c = 115.19, β = 104.9	$a = b$ = 106.58, c = 59.25	a = 69.93, b = 111.57, c = 311.42
Resolution range (Å)	50–1.90 (1.97–1.90)	50.0–1.75 (1.81–1.75)	30.0–2.5 (2.59–2.50)
Observed reflections	667221	288658	309014
Unique reflections	121501 (12072)	13115 (1288)	41061 (3786)
Completeness (%)	99.9 (99.7)	99.9 (100)	96.1 (89.6)
V_M (Å^3 Da^{-1})	1.94	1.90	1.99
Solvent content (%)	36.7	35.5	38.1
$I/\sigma(I)$	23.9 (5.3)	57.4 (14.6)	22.8 (3.1)
R_{merge}† (%)	4.5 (17.2)	8.0 (17.7)	8.7 (53.0)
Multiplicity	5.5 (5.1)	22.0 (21.3)	7.5 (7.2)
R_{work}/R_{free} (%)	18.8/21.9	19.9/21.8	20.2/27.0
Protein model			
No. of subunits in ASU	12	1	9
Protein atoms	13230	1078	9883
Water molecules	1181	132	427
Phosphate ions	12	1	—
GTP molecules	—	—	9
Others	22	2	39
Deviations from ideal geometry			
Bond lengths (Å)	0.005	0.004	0.007
Bond angles (°)	1.3	1.3	1.4
Dihedral angles (°)	23.4	23.2	23.1
Improper angles (°)	0.75	0.75	0.92
Average B factors (Å^2)			
Protein atoms	23.2	24.4	53.3
Water molecules	36.8	38.6	52.1
Phosphate ions	34.8	19.4	—
GTP molecules	—	—	83.0
Others	45.6	54.2	84.9
Ramachandran plot (%)			
Favoured	93.1	93.6	87.9
Allowed	6.9	6.4	12.1

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I(hkl)$ is the intensity of reflection hkl, $\sum_{hkl}$ is the sum over all reflections and $\sum_i$ is the sum over i measurements of reflection hkl.

*Tt*MoaC protein solution. The ligand concentrations in the ITC syringe (volume 298 µl) were 1 m*M*. Thus, the ITC experiments were performed under conditions in which the *C* value ($K_b \times$ Mt, where K_b and Mt represent the binding constant and the enzyme concentration, respectively) was greater than 1. Titrations were performed by a stepwise addition of small volumes (7 µl) of ligand solutions from the stirred syringe (307 rev min^{-1}) into the sample cell. A time interval of 180 s was used between successive injections. The values of the change in binding enthalpy (ΔH_b), binding constant (K_b) and binding stoichiometry (n) for the titration were determined by nonlinear least-squares fitting of the data using the program *Origin* 7.0. The change in entropy (ΔS) was obtained using the equation $\Delta G_b = \Delta H_b - T\Delta S$, where $\Delta G_b = -RT\ln K_b$; the parameters R and T represent the gas constant and the absolute temperature (K), respectively.

2.4. Molecular-dynamics (MD) simulations

MD simulations were performed using the package *GROMACS* v.3.3.3 running on parallel processors (Berendsen *et al.*, 1995; Lindahl *et al.*, 2001). The widely distributed AMBER all-atom force-field ports for the *GROMACS* suite were used (Duan *et al.*, 2003; Sorin & Pande, 2005). During MD simulations, the crystallographic water molecules were removed from the protein models. A cubic box was generated using the *editconf* module of *GROMACS* with a criterion that the minimum distance between the solute and the edge of the box was at least 0.75 nm. The protein models were solvated with an SPC (simple point charge) water model using the *genbox* program available in the *GROMACS* suite. H atoms were added to the ligand molecules using the *PRODRG* web server (Schüttelkopf & van Aalten, 2004). The parameters derived from AMBER03 (Case *et al.*, 2006) were used to generate ligand topologies, which were further converted to *GROMACS* format using a Perl script (`amb2gmx.pl`). Furthermore, the partial charges of the ligands were optimized using the *ab initio* program *Gaussian*03 (Frisch *et al.*, 2004). Chloride ions (in the range 10–37 m*M*) were used to neutralize the overall charge of the system where required. Energy minimization was performed using the conjugate-gradient and steepest-descent methods with the frequency of the latter at 1 in 1000 with a maximum force cutoff of 1 kJ mol^{-1} nm^{-1} for convergence of minimization. Subsequently, solvent equilibration by position-restrained dynamics for 10 ps was carried out. Simulations utilized the NPT ensembles with Parrinello–Rahman isotropic pressure coupling (τ_p = 0.5 ps) to 100 kPa and Nose–Hoover temperature coupling (τ_t = 0.1 ps) to 300 K. Long-range electrostatics were computed using the Particle Mesh Ewald (PME) method (Darden *et al.*, 1993) with a cutoff of 1.2 nm. A cutoff of 1.5 nm was used to compute the long-range van der Waals interactions. Bond lengths were constrained using the *LINCS* algorithm (Hess *et al.*, 1997). MD simulations were performed for a time period of 10 ns for all the structures discussed in the present study. However, analyses were performed for a time period of the last 9 ns. The protein–ligand interaction energies were calculated using the equation

$$E_{\text{protein-ligand}} = (E_{\text{protein-ligand}})_{\text{elec}} + (E_{\text{protein-ligand}})_{\text{vdw}},$$

where $E_{\text{protein-ligand}}$ denotes the interaction energy between the protein and the ligand and 'elec' and 'vdw' denote the electrostatics and van der Waals components of the energy,

respectively. The relative interaction energies among different ligands were obtained using the formula

$$\Delta\Delta E = E_{\text{protein-gtp}} - E_{\text{protein-ligand}},$$

where $E_{\text{protein-gtp}}$ is the interaction energy between the protein and the 5′-GTP molecule and $E_{\text{protein-ligand}}$ is that of the other ligands considered in the MD simulations.

2.5. Structural analysis

Invariant water molecules were identified using the *3dSS* server (Sumathi *et al.*, 2006). Most of the analyses of the MD simulations were performed using the *GROMACS* tools and locally developed Perl scripts. The freely available *PDB Goodies* server (Hussain *et al.*, 2002) was used to renumber the residues and to analyze the temperature factors. The figures were generated using the program *PyMOL* (DeLano Scientific LLC; http://www.pymol.org). Graphs were prepared using *Xmgr* (Paul J. Turner, Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, Oregon, USA). Structures were superposed using the program *ALIGN* (Cohen, 1997). Hydrogen bonds were calculated using the program *HBPLUS* (McDonald & Thornton, 1994). A donor–hydrogen–acceptor angle of greater than or equal to 120° and donor–acceptor distance of less than or equal to 3.5 Å were used as criteria for delineating hydrogen bonds. The solvent-accessible surface area of the invariant water molecules was computed using the program *NACCESS* (Hubbard & Thornton, 1993) with a probe radius of 1.4 Å. Water molecules with an accessible surface area of less than or equal to 2.5 Å^2 were considered to be internal/buried. The normalized temperature factor (B_i') for all the invariant water molecules was calculated using the formula $B_i' = (B_i - \langle B\rangle)/\sigma(B)$, where B_i is the B factor of each atom, $\langle B\rangle$ is the mean B factor and $\sigma(B)$ is the standard deviation of the B factors. The structure-based sequence alignment was generated using the program *MUSTANG* (Konagurthu *et al.*, 2006). The secondary-structural elements for the protein were assigned using the program *DSSP* (Kabsch & Sander, 1983). Electrostatic potentials were calculated using the module *APBS* (Baker *et al.*, 2001) plugged into *PyMOL*.

3. Results and discussion

3.1. The overall structure

All three forms of the crystal structure of *Tt*MoaC were solved by the molecular-replacement method using the atomic coordinates of *Ec*MoaC (PDB code 1ekr). The asymmetric units of the apo forms ($P2_1$, form I; $R32$, form II) and the GTP-bound form ($C222_1$, form III) contained 12, one and nine subunits, respectively. The refinement statistics of all the three crystal structures are given in Table 1. The final refined model in all three forms lacked the first ten residues at the N-terminus and three residues at the C-terminus. Each monomer of MoaC contains a $\alpha+\beta$ structure and is composed of a four-stranded antiparallel β-sheet with three helices, α_1, α_2 and α_3, located on the same side of the β-sheet (Fig. 1*a*). In addition, there is a short 3_{10}-helix (residues 90–92). MoaC belongs to the ferredoxin-like ($\beta\alpha\beta\beta\alpha\beta$) fold, with the insertion of a helix ($\beta\alpha\alpha\beta\beta\alpha\beta$). It is hexameric (~42% surface area buried), being made up of three dimers (Wuebbens *et al.*, 2000). A total of 22 and 18% of the monomer-accessible surface area is buried upon dimerization and trimerization, respectively. Furthermore, each monomer of the hexamer contacts another three subunits (two from the trimeric subunits and one from the dimeric subunit), similar to *Ec*MoaC. Each dimer of *Tt*MoaC is stabilized by 11 intersubunit hydrogen bonds, compared with the eight hydrogen bonds in *Ec*MoaC (Wuebbens *et al.*, 2000). Each trimer and hexamer of *Tt*MoaC has 36 and 115 intersubunit hydrogen bonds, respectively. The solvation-free (SF) energy gain

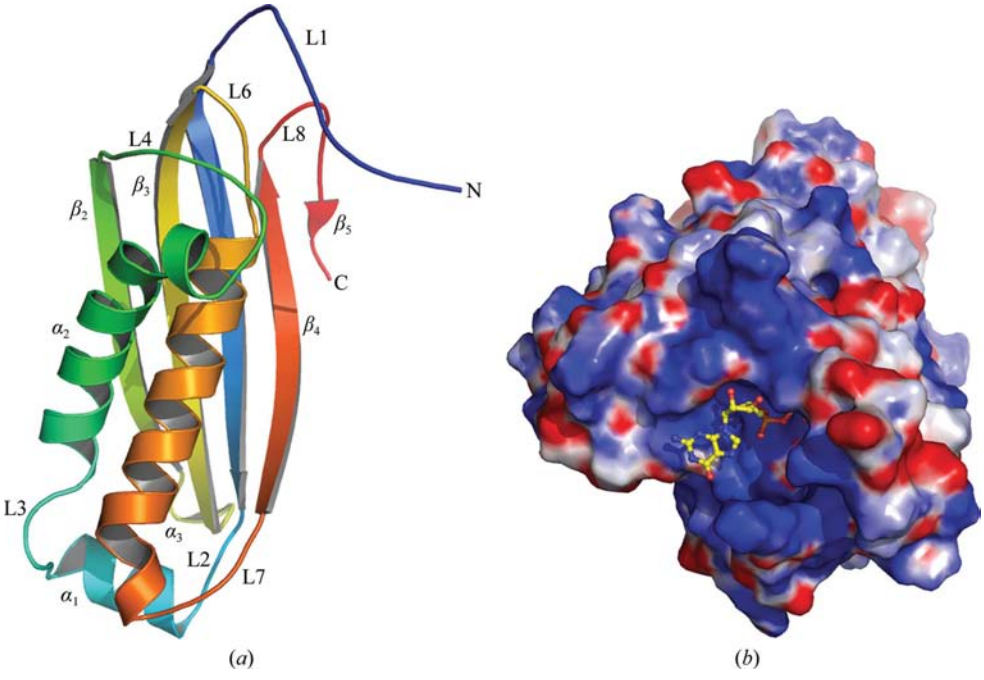


Figure 1
(*a*) Cartoon representation of the overall tertiary structure of the *Tt*MoaC monomer. The secondary-structural elements assigned using the program *DSSP* (Kabsch & Sander, 1983) and the terminal ends are labelled for clarity. (*b*) The electrostatic potentials calculated using the program *APBS* (Baker *et al.*, 2001) of the dimeric subunits of the protein molecule. Surface electrostatic potentials that are less than $-10kT$, neutral and greater than $10kT$ are displayed in red, white and blue, respectively. The bound 5′-GTP molecule is shown as a ball-and-stick model.

Table 2
Isothermal titration calorimetry data for the binding of 5′-GTP, 5′-GDP and 5′-GMP to *Tt*MoaC.

1 kcal = 4.186 kJ.

	T (K)	n	K_b (M^{-1}) $\times 10^4$	ΔH_b (kcal mol^{-1})	$T\Delta S$ (kcal mol^{-1} K^{-1})	ΔG_b (kcal mol^{-1})
GTP	293	0.4 ± 0.08	2.250 ± 0.38	−11.05 ± 2.59	−5.21	−5.84
GDP†	293	0.4 ± 0.79	0.468 ± 0.21	−9.37 ± 20.0	−4.69	−4.68
GMP†	293	1.0 ± 24.29	0.146 ± 1.02	−4.74 ± 131.4	−0.50	−4.24

† The values corresponding to 5′-GDP and 5′-GMP are approximate, as the iterations of the nonlinear curve fitting were never saturated.

upon formation of the assembly was predicted to be −522.8 kJ mol^{-1} using the *PISA* web server (Krissinel & Henrick, 2007). The SF energies for the phosphate bound in the apo and GTP-bound structures show an increase (of approximately 210 and 125 kJ mol^{-1}, respectively) compared with that of the ligand-free hexamer.

3.2. Active-site geometry

The active site of MoaC is located at the dimer interface and is composed of residues located in the six loops L1, L3, L4, L5, L6 and L8 and helix α_3, as observed in *Ec*MoaC (Wuebbens *et al.*, 2000). The residues Lys19, Arg24, Lys49, Gly50, His75, Thr107, Gly108, Glu110, Met111, Glu112, Asp126, Met127, Lys129, Lys145, Gly147 and Gly148 make up the active site of *Tt*MoaC. In addition, residue Lys65 (located in helix α_2 of the third subunit) is involved in the formation of the active site. Interestingly, Lys65 N$^{\zeta}$ is tri-coordinated to the main-chain carbonyl atoms of the conserved residues Ile71, Pro72 and Cys74. Furthermore, the surface-charge distribution in *Tt*MoaC is uniform; however, the active site is positively charged owing to the presence of basic residues (Fig. 1*b*).

3.3. Results of ITC experiments

ITC experiments were also carried out to ensure the binding of 5′-GTP to *Tt*MoaC. The ITC results revealed a dissociation constant of 44.4 ± 8 µ*M* and a binding stoichiometry of 0.4 ± 0.1 sites per monomer for 5′-GTP molecules (Fig. 2 and Table 2). In addition, the compounds 5′-GDP and

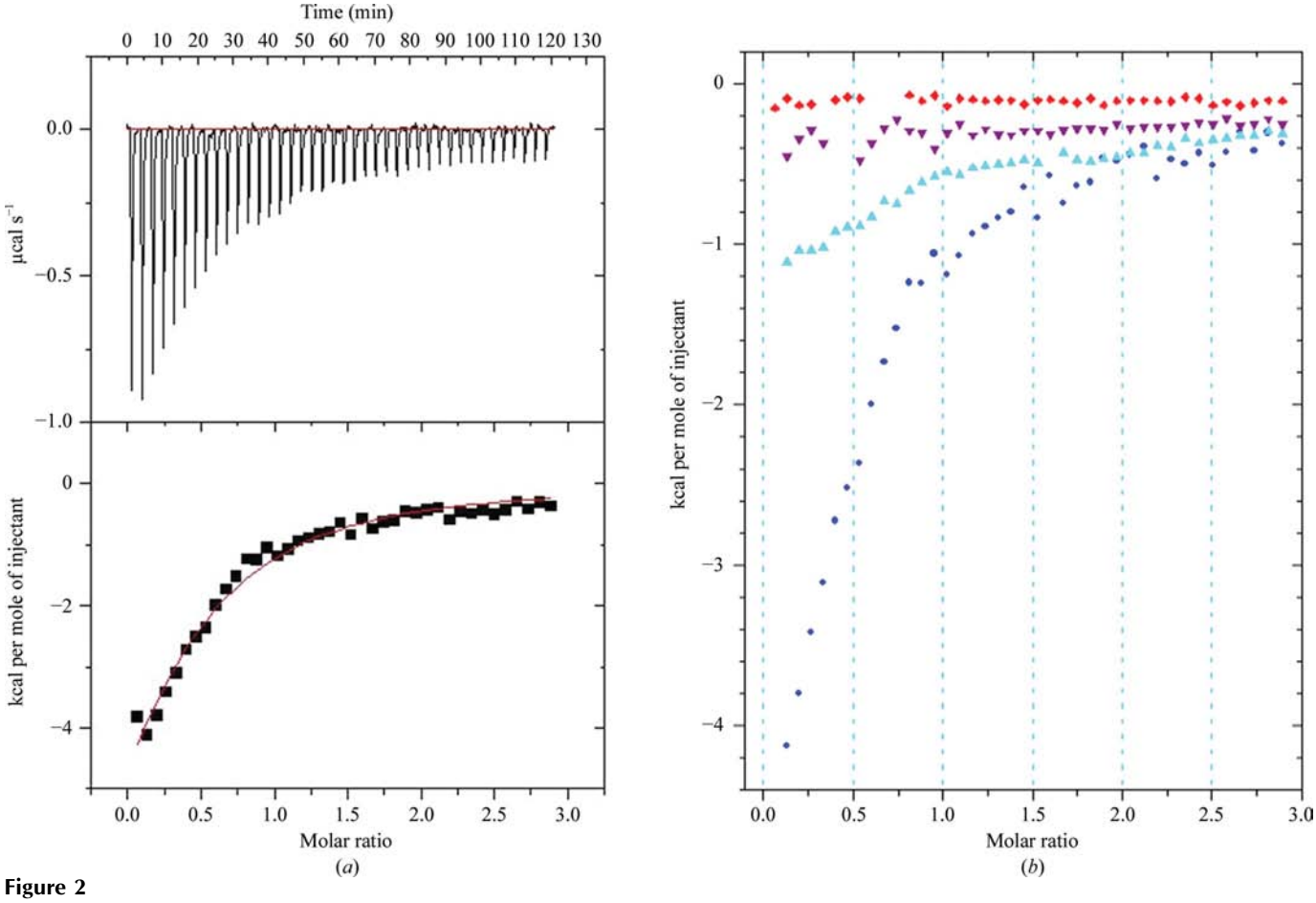


Figure 2
Isothermal titration calorimetry for the binding of 5′-GTP to *Tt*MoaC. (*a*) The upper panel shows the heat change elicited upon successive injections of 5′-GTP into *Tt*MoaC. The lower panel shows the binding isotherm as a function of the molar ratio of 5′-GTP to *Tt*MoaC. A theoretical curve was fitted to the integrated data using a single-site model. (*b*) The relative binding isotherm as a function of the molar ratio of ligands (5′-GTP, blue filled dots; 5′-GDP, cyan triangles; 5′-GMP, pink inverted triangles; dialyzed buffer used as a control, red squares) to *Tt*MoaC are shown. 1 kcal = 4.186 kJ.

5′-GMP that were also used for ITC experiments showed weak binding (fivefold and 15-fold weaker) compared with that of 5′-GTP (Fig. 2 and Table 2). However, the binding parameters for 5′-GDP and 5′-GMP are approximate as the iterations of curve fitting were never saturated. Comparison of GTP binding by MoaC and MoaA suggests that the binding of 5′-GTP to *Tt*MoaC is weaker (by ∼150-fold) compared with that to MoaA (Hanzelmann & Schindelin, 2006). These results, together with the crystal structure of GTP-bound *Tt*MoaC, suggest that GTP is the true substrate of MoaA and not of MoaC. Thus, it can be concluded that the substrate molecules of both the MoaA and MoaC proteins share a common motif (a triphosphate terminal group).

3.4. The phosphate ion and GTP-binding site

Both of the apo crystal structures of *Tt*MoaC contained a phosphate ion (present in the precipitant solution) bound in the active site of the protein molecule. It was found that the phosphate ions in the apo forms were located at the position of P^{γ} of GTP in the complex structure. The residues Lys49, Cys74, His75, Asp126 and Lys129 are involved in hydrogen bonding to the phosphate ions. However, hydrogen bonding to Cys74 is observed in only two subunits. In addition, two water molecules are observed to be coordinated to the phosphate ion in the form I structure. In form II, the residue Lys49 is located a little too far away to make a hydrogen bond to the phosphate ion. Instead, three water molecules are hydrogen bonded to the phosphate ion. Thus, the binding of phosphate ions in the active site of the ligand-free forms provides a possible clue to the binding of a molecule with terminal phosphate groups. The GTP-bound crystal structure revealed that the hydrogen-bonding interactions primarily contributed by the phosphate group stabilize the GTP molecule. The residues interacting with GTP are Val47, Lys49, Asp126 and Lys129 (from one subunit) and Cys74, His75 and Thr107 (from the other subunit of the dimer) and three water molecules (Fig. 3). However, interactions with Val47 and Cys74 are not observed in all subunits located in the asymmetric unit.

3.5. Other molecules bound in the active site

In addition to phosphate ions, glycerol (GOL) molecules and acetate (ACT) ions were observed in the active sites of the form I and II structures, respectively. Interestingly, difference electron density (up to 4.7σ in the $|F_o - F_c|$ map) in addition to the 5′-GTP was observed in the active site of form III. Based on the ingredients used in the crystallization, a citrate (FLC) ion was fitted. It is hydrogen bonded to residues Arg24, Glu110, Lys145, Lys149, Lys150 and, in some subunits, the N1 atom of 5′-GTP (Fig. 3). However, the interactions with Lys149 and Lys150 are perturbed in most of the subunits. The average closest distance between GTP and FLC is approximately 3.32 Å. Thus, these observations confirm that a longer

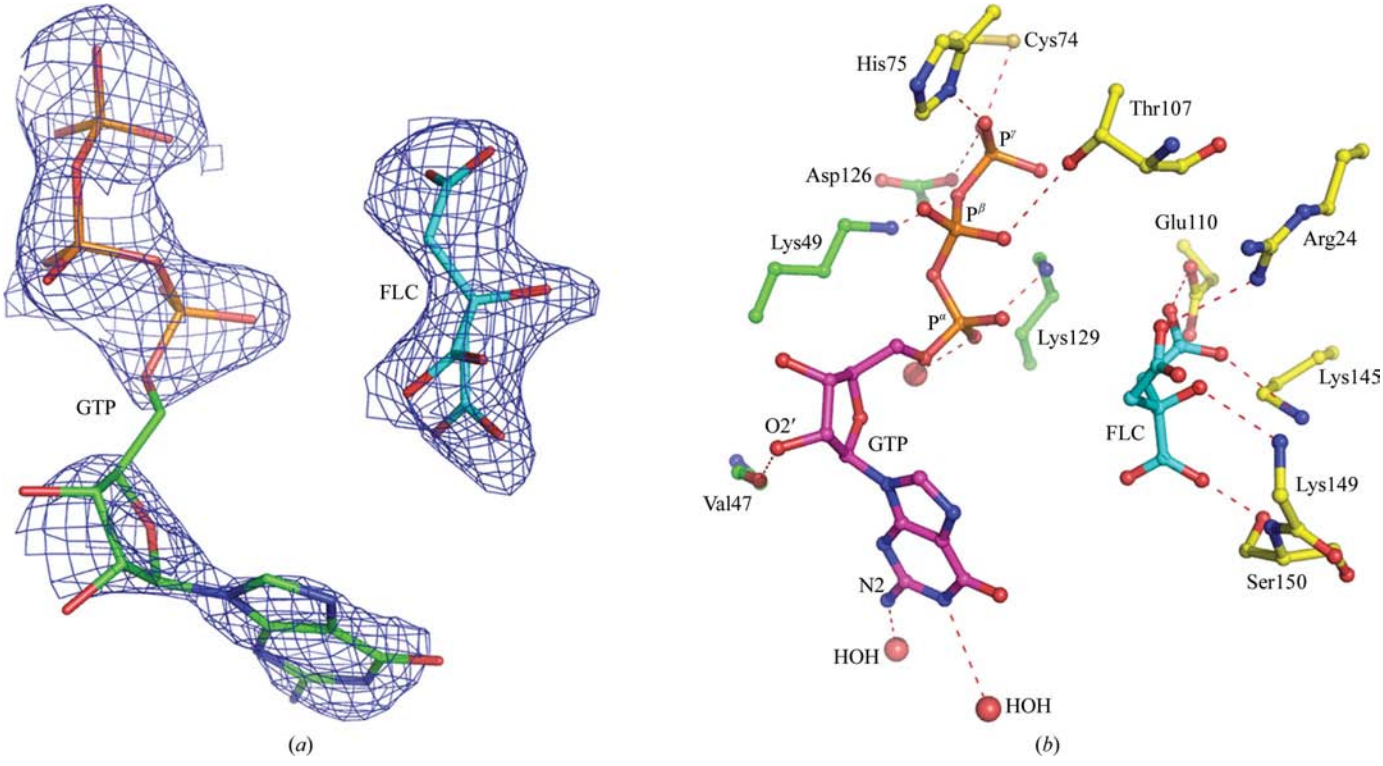


Figure 3
(*a*) The unbiased difference electron-density ($|2F_o - F_c|$) map for 5′-GTP and FLC (citrate) contoured at 0.8σ. The electron densities for both molecules are shown prior to their addition to the model. Both molecules are shown as ball-and-stick models. (*b*) The hydrogen-bond interactions to 5′-GTP and FLC at the dimeric interface of *Tt*MoaC. The residues involved in hydrogen-bond interactions are shown as ball-and-stick models in different colours for each subunit. The water molecules are shown in spheres.

Table 3
Invariant water molecules in the crystal structure of *Tt*MoaC.

1†	2‡	3§	4¶	5††	6‡‡
IW1	817	Thr25 $O^{\gamma1}$, Thr27 $O^{\gamma1}$, His144 $N^{\delta1}$	−0.7	1.6	0.97
IW2	818	Leu37(I) N, Asp126 $O^{\delta2}$, Met127 S^{δ}	−1.0	27.6	0.99
IW3	820	Leu62 O, HOH690 O	−0.6	13.7	0.68
IW4	822	Cys74 S^{γ}, His75 O, Glu112 $O^{\varepsilon2}$, HOH1185 O	0.3	1.5	0.73
IW5	823	Thr25 O, Val109 N	−0.6	0.0	0.83
IW6	824	Val14 N, Leu73 O, PO4 169 O4, HOH861 O	−0.2	14.0	1.00
IW7	827	Val54 O, Gln57 $O^{\varepsilon1}$, HOH849 O	−0.5	9.1	0.03
IW8	830	Thr100 O, HOH1153 O	0.0	3.3	1.00
IW9	832	Asp126 O, Ala130 N, HOH1178 O	−0.5	0.2	0.89
IW10	834	Leu88 N, HOH857 O	0.3	25.0	0.87
IW11	837	Val86 N, Val86 O, HOH840 O, HOH1044 O	−0.1	9.6	0.69
IW12	840	Leu53 O, HOH837 O, HOH846 O, HOH849 O	1.0	1.6	0.46
IW13	841	Asp69 $O^{\delta1}$, HOH845 O, HOH867 O	0.5	33.0	1.00
IW14	849	HOH827 O, HOH840 O	1.1	27.1	0.21
IW15	856	Val47 N, Lys49 N, Met127 O, HOH862 O	−0.2	0.4	0.77
IW16	862	Gly45 O, Gly48 N, Gly50 N, HOH1036 O	0.1	7.5	0.53

† Invariant water numbering scheme. ‡ The water-molecule number in the form I crystal structure of *Tt*MoaC. § Hydrogen-bond interactions observed in the crystal structure. ¶ Average normalized *B* factor (Å^2) calculated using the subunits from the apo crystal structure of *Tt*MoaC. †† Average solvent-accessible surface area (Å^2). ‡‡ Average occupancy computed from the MD calculations.

molecule such as FPT (an MoaA-generated intermediate compound) would tightly bind to MoaC.

3.6. Changes owing to substrate binding in the active site

The overall C^{α}-atom superposition of the apo and GTP-bound crystal structures of *Tt*MoaC shows an r.m.s.d. of 0.4 Å, indicating no significant change in the overall tertiary structure of the protein. However, the C-terminal loop regions (residues 148–151) deviate more (average r.m.s.d. of 1.7 Å). Furthermore, electron density for these residues is not clear in the GTP-bound structure. This may be a consequence of the binding of the 5′-GTP molecule in the active site. It should be noted that the guanosine moiety of 5′-GTP is close to loop L8 and shows high flexibility, causing local structural changes (Fig. 4).

3.7. Invariant water molecules

To study the role of water molecules, a total of 13 crystallographically independent subunits from apo forms of *Tt*MoaC were used to identify invariant water molecules. The identification of invariant water molecules was carried out in a similar way as performed previously in our laboratory (Kanaujia & Sekar, 2009). Water molecules in a pair of subunits were considered to be equivalent if they were less than or equal to 1.8 Å apart when the subunits, together with their hydration shell, were superposed on each other and if they have at least one common interaction with the protein molecule. Water molecules that are equivalent in all possible pairs among the subunits considered are termed invariant. A total of 16 invariant water molecules (Table 3) were identified.

3.8. Plasticity of *Tt*MoaC

The 22 copies of the *Tt*MoaC subunit in the two apo structures and the complex structure provide a database for exploring the relatively rigid and flexible structural features of the protomer. The analysis was performed using the program *ESCET* (Schneider, 2004). The program *ESCET* categorizes the molecule into conformationally invariant and variable regions by automated analysis of pairs of error-scaled difference distances (Cruickshank, 1999) of an ensemble of conformers (*e.g.* crystal structures from different crystal forms or molecules related by noncrystallographic symmetry). While using *ESCET*, a parameter σ is employed to divide a subunit into rigid and flexible regions and the value of σ is calculated from the error estimate using the atomic coordinates of the structures being compared. In the present calculation, the parameter σ was chosen so as to have roughly 69% and 31% [these values were derived using all the structures (approximately 11 671) belonging to the same class, fold, architecture and molecular topology as MoaC available in the Protein Data Bank] of residues in the invariant and variable regions, respectively (Schneider, 2004). Thus, of 143 consistent residues, 55 are predicted to be conformationally invariant, including all/most of the residues in the three helices. Interestingly, more than 90% of the β-sheet residues are predicted to be in an intermediate state. Most of the loops belong to the flexible region, with the exception of L3 and L5 (Fig. 5*a*). Comparison of the *ESCET* analysis and the *B* factors of each crystallographically independent subunit in all three struc-

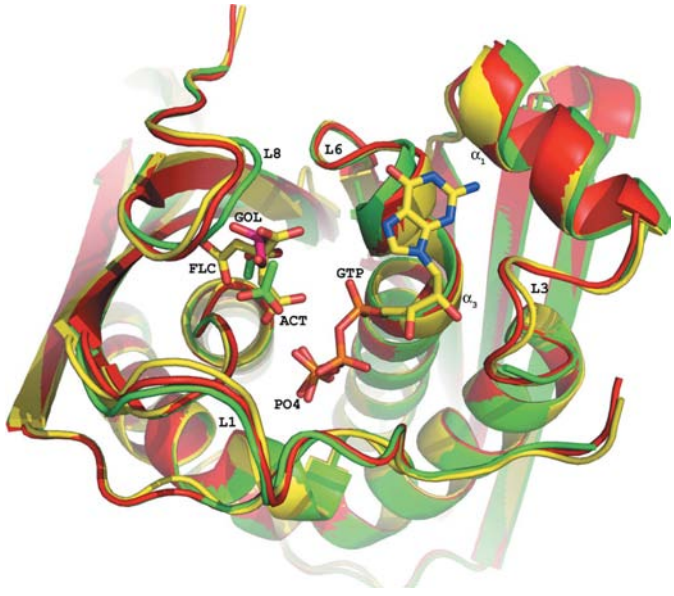


Figure 4
Active-site superposition of the two apo ($P2_1$, red; $R32$, green) and the complex (yellow) crystal structures. The phosphate (PO4) ions observed in the apo forms of the crystal structures are also shown. The glycerol (GOL) molecules and acetate (ACT) and citrate (FLC) ions observed in the crystal structures are also shown for comparison.

tures shows a significant difference (Fig. 5*b*). Although both of the methods categorize the loops as highly flexible, there is a difference for α-helices and β-strands. According to *ESCET* analysis most of the β-strands fall into an intermediate state, whereas *B*-factor analysis shows them to be rigid. A similar difference is found in the case of α-helices (Fig. 5*b*).

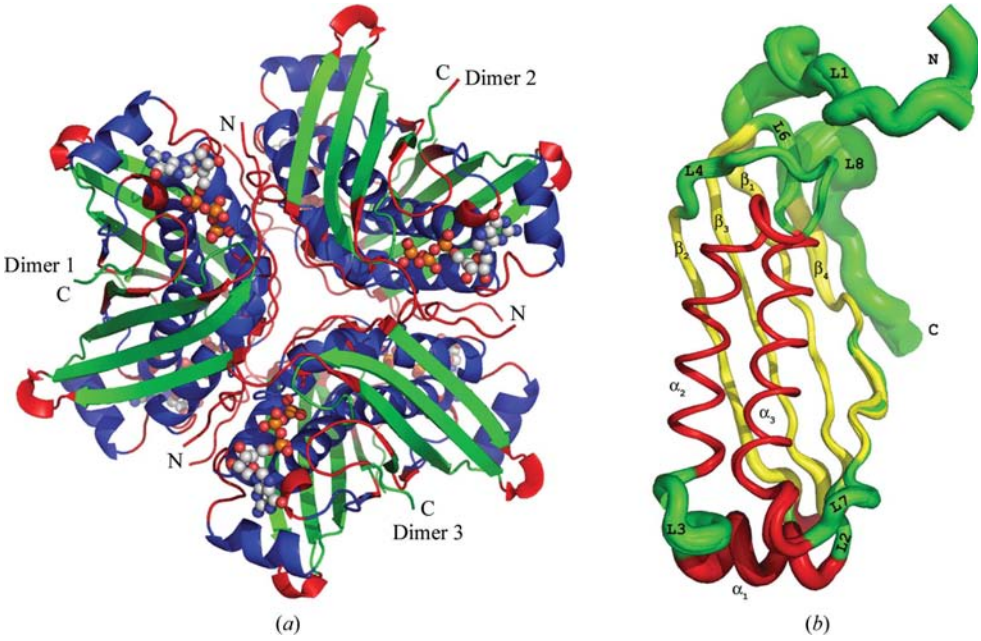


Figure 5
(*a*) Cartoon representation of structurally invariant and flexible regions of *Tt*MoaC. The invariant and flexible regions are shown in blue and red, respectively. The residues coloured green correspond to an intermediate state. The ligand (5′-GTP) molecules are shown as ball-and-stick models. (*b*) *B*-factor analysis of all 22 crystallographically independent subunits across all three crystal structures. The segments of each subunit are thickened according to their *B*-factor values. Segments with low *B* factors are thinner than those with high *B* factors. The secondary-structural elements α-helices, β-strands and loops are shown in red, yellow and green, respectively.

3.9. Comparison with MoaC from other sources

A structure-based sequence alignment of *Tt*MoaC, *Ec*MoaC, *Ph*MoaC, *St*MoaC and *Gk*MoaC is shown in Fig. 6(*a*). Although the overall sequence identity among them is low (21%), pairwise structure-based sequence alignment of these proteins shows that *Tt*MoaC has the most similarity to *Ec*MoaC and *Gk*MoaC (∼48%); that with *Ph*MoaC and *St*MoaC is only ∼35%. In addition, Fig. 6(*a*) shows that almost a quarter of the residues (36 of 146) are highly conserved among the species. A total of five residues (of these 36), Lys49, His75, Thr107, Asp126 and Lys129, are important for substrate binding (Fig. 3) and another three residues, Arg24, Lys133 and Lys145, are

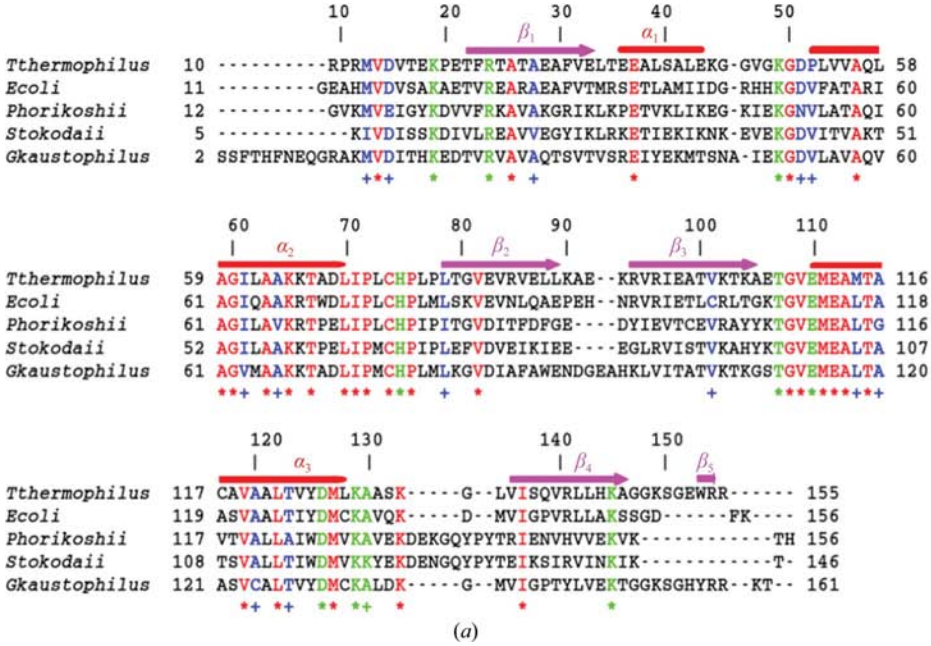
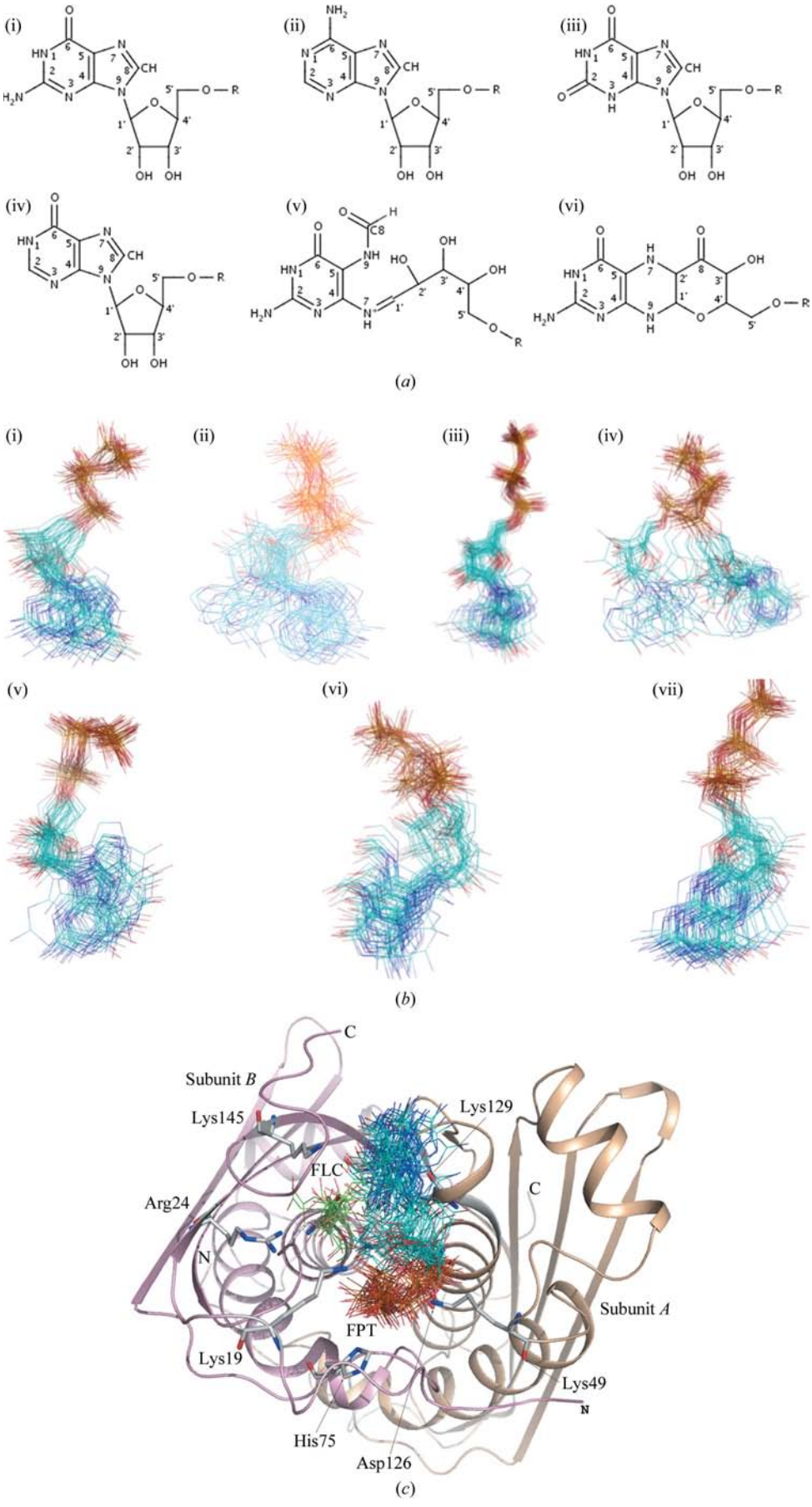


Figure 6
(*a*) Structure-based sequence alignment of MoaC proteins using the program *MUSTANG* (Konagurthu *et al.*, 2006). The secondary-structural elements are shown for *Tt*MoaC. Highly conserved and semi-conserved residues are shown in red and blue and are marked by symbols (* and +, respectively) at the bottom of the alignment. Residues that are possibly involved in the catalytic mechanism of the protein are shown in green. (*b*) An overall structural superposition of individual subunits of *Tt*MoaC (red), *Ec*MoaC (green), *Ph*MoaC (blue), *St*MoaC (yellow) and *Gk*MoaC (orange) is shown for comparison.



involved in binding citrate ions. The remaining 28 residues may play a role in stabilizing the overall tertiary structure. Gly50, which is part of the GTP-binding motif (GKG), is located in the active site of the molecule. Four residues, Ala56, Gly60, Ala63 and Leu65, are mainly involved in the oligomerization of the protein molecule. In addition, the LIP-*X*CHP motif (residues 70–76) and the residues Leu122 and Ile137 are involved in dimerization of the protein molecule.

The tertiary structures of *Ec*MoaC, *Ph*MoaC, *St*MoaC and *Gk*MoaC are similar to those of *Tt*MoaC, with r.m.s.d.s of 1.0, 1.2, 1.0 and 1.0 Å, respectively. Some minor deviations are observed in the region corresponding to the α_1 helix and loop L3 (Fig. 6*b*). In addition, amino acids at the N- and C-termini that are absent in MoaCs from other species are observed in the case of *Gk*MoaC. In the case of the *Ph*MoaC and *St*MoaC proteins there is an insertion of seven residues in loop L6. Moreover, the C-terminal loop L8 could only be traced in the case of *Tt*MoaC. It is interesting to recall that this loop shows structural changes upon substrate binding.

Figure 7
(*a*) Schematic diagrams of the basic units of the ligands (i) 5′-GTP, (ii) 5′-ATP, (iii) 5′-XTP, (iv) 5′-ITP, (v) FPT and (vi) PBT. (*b*) The ligand conformations accessed at each 100 ps during MD are shown for (i) 5′-GTP, (ii) 5′-ATP, (iii) 5′-XTP, (iv) 5′-ITP, (v) GTPWF, (vi) FPT and (vii) PBT. (*c*) The ensembles (45 in total) of FPT generated every 100 ps from the trajectories obtained from the MD simulation. Each conformation of FPT accessed during the MD simulation is shown at the dimeric interface of *Tt*MoaC. In addition, nine citrate (FLC) ions (one from each of the nine subunits of the complex structure) observed in the crystal structure of GTP-bound *Tt*MoaC are shown in green. Each subunit of a dimer is coloured differently. The active-site residues involved in substrate and FLC binding are also displayed and labelled.

Table 4
Broad average parameters for the *Tt*MoaC–ligand interactions derived from calculations.

1†	2‡	3§	4¶
GTP	−236.1 (12.4)	—	5.5 (0.6)
GDP	−202.3 (7.1)	−33.8	5.2 (0.5)
GMP	−151.2 (5.1)	−84.9	3.3 (0.4)
ATP	−212.9 (5.2)	−23.2	5.0 (0.5)
ADP	−167.4 (5.2)	−68.7	3.7 (0.4)
AMP	−132.6 (3.7)	−103.5	3.1 (0.4)
XTP	−233.9 (7.6)	−2.2	5.6 (0.5)
XDP	−199.2 (4.1)	−36.9	5.1 (0.4)
XMP	−123.5 (4.8)	−112.6	3.1 (0.4)
ITP	−250.4 (5.9)	14.3	5.7 (0.5)
IDP	−199.3 (5.1)	−36.8	5.0 (0.4)
IMP	−127.8 (4.9)	−108.3	3.2 (0.4)
GTPWF	−253.0 (5.5)	16.9	6.0 (0.6)
FPT	−263.3 (6.5)	27.2	7.6 (0.6)
PBT	−235.0 (6.4)	−1.1	6.1 (0.6)

† Ligands used for the calculations. ‡ Interaction energies ($E_{\text{protein–ligand}}$; kcal mol^{-1}). Standard deviations (SD) are given in parentheses. 1 kcal = 4.186 kJ. § The interaction-energy difference ($\Delta\Delta$E) between GTP and the respective ligands (kcal mol^{-1}). 1 kcal = 4.186 kJ. ¶ The average number of hydrogen bonds calculated from ensembles generated using the MD simulations. Standard deviations (SD) are given in parentheses.

4. Results of MD simulations

4.1. General features

A total of 16 simulations (15 protein–ligand complexes and one protein) were carried out, each for 10 ns. Based on the GTP-bound crystal structure of *Tt*MoaC (present study), 11 different ligands were modelled in the active site to study interactions involving triphosphate, diphosphate and monophosphate groups (Fig. 7*a* and Table 4). All the simulations (for protein–ligand complexes) were carried out in the presence of FLC as observed in the GTP-bound crystal structure. Thus, simulations of GTP without FLC (GTPWF) and two probable intermediate compounds (FPT and PBT) were also performed (Fig. 7*a*). The conformations accessed during the MD simulations by ligands containing triphosphate groups are shown in Fig. 7(*b*). This suggests that the phosphate groups of the ligands are rigid compared with the sugar and base rings.

4.2. Energetics

The protein–ligand interaction energies given in Table 4 are so large as to render the actual values somewhat meaningless. However, the fact that the interaction energies calculated from the MD simulations led to a difference ($\Delta\Delta E$) in favour of the correct ligand type is in itself satisfying. An analysis of these results revealed that ligands containing triphosphate groups are more favourable compared with diphosphates and monophosphates. The results of the ITC experiments corroborate the above conclusion. Interestingly, ligands containing triphosphates (5′-GTP, 5′-ATP and 5′-XTP) show similar interaction energies, whereas 5′-ITP shows a slightly higher interaction energy with a greater number of hydrogen bonds (Table 4). Furthermore, it was observed that FPT shows better binding compared with the other molecules considered in the present study. Moreover, it is observed that some of the atoms of FPT occupy the positions of FLC during the MD simulations (Fig. 7*c*). Thus, it may be suggested that a molecule containing a triphosphate group and an open sugar ring (similar to FPT) is a better substrate for MoaC. However, given the range of standard deviations, no clear distinction between the different substrates can be made except for the NTP, NDP and NMP trend.

4.3. Protein dynamics

The residue-wise root-mean-square fluctuation (RMSF) in C$^{\alpha}$ positions averaged over all the simulations, together with

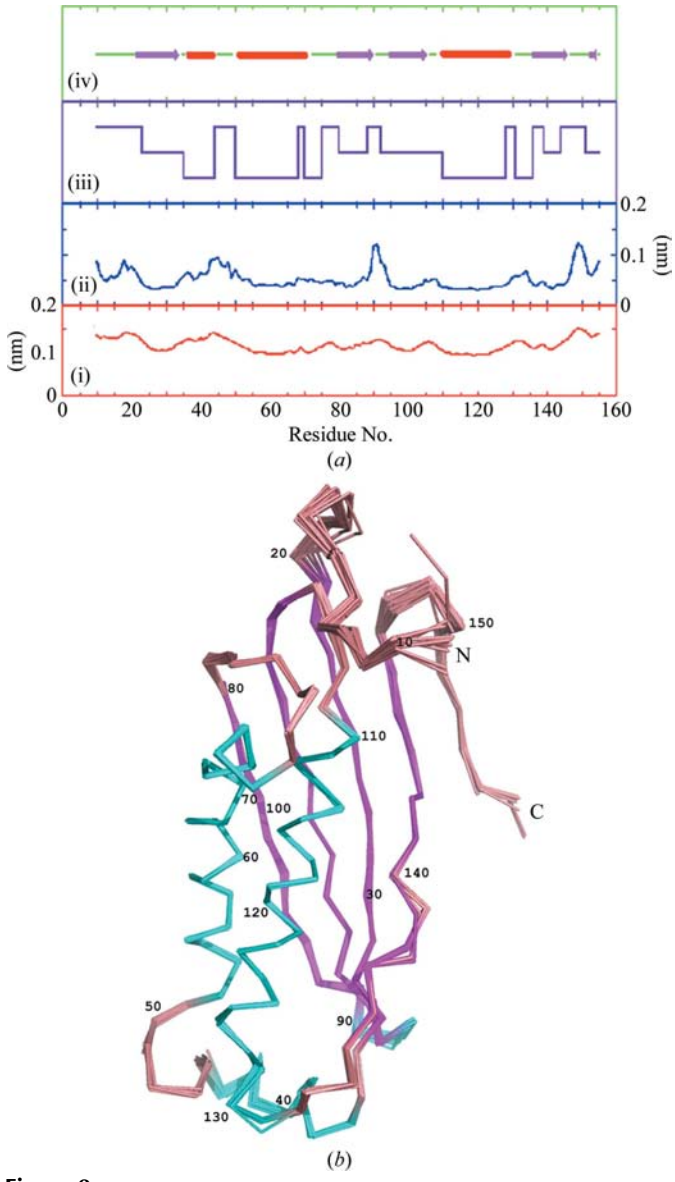


Figure 8
(*a*) Different representations of the plasticity of *Tt*MoaC. (i) The average RMSFs calculated from the crystal structures of *Tt*MoaC. (ii) The average RMSFs computed using the ensembles generated from MD simulations. (iii) Relatively rigid (bottom line), intermediate (middle line) and flexible (upper line) regions of the subunit. (iv) The secondary-structural elements of *Tt*MoaC. (*b*) Overlay of all 22 crystallographically independent subunits from all three crystal structures shown as a ribbon diagram. The secondary-structural elements α-helices, β-strands and loops are shown in cyan, magenta and brown, respectively.

the average atomic displacement derived from *B* factors obtained from crystallographic studies, is shown in Fig. 8(*a*). The different indicators presented in Fig. 8(*a*) provide valuable insights into the plasticity and dynamics of the protein molecule. The regions 19–23, 40–50, 89–92 and 147–152 are highly flexible. Of these regions, 40–50 and 147–152 are involved in substrate binding. Interestingly, the protein dynamics obtained from the crystallographic *B* factors and MD simulations differ in several regions of the subunit (Fig. 8*b*).

4.4. Role of invariant water molecules

The location of invariant water molecules identified from crystal structures, together with their hydrogen-bond interactions, is shown in Fig. 9. Their normalized *B* factors, solvent-accessible surface areas and occupancies computed from the MD calculations are provided in Table 3. Of the 16 invariant water molecules, seven (IW2, IW4, IW5, IW6, IW9, IW15 and IW16) are located in the vicinity of the active site and show a low average normalized *B* factor and a high occupancy (≥70%), with the exception of IW16, as computed using the ensembles generated during MD simulations. In addition, most of them are buried (Table 3). Interestingly, three of them (IW2, IW4 and IW9) are hydrogen bonded to the highly conserved residues His75 and Asp126 that are crucial for substrate binding (Fig. 3). Another water molecule, IW6, makes a hydrogen bond to the phosphate ion. Furthermore, two water molecules, IW15 and IW16, seem to stabilize the active-site loop L3. A set of five water molecules, IW7, IW10, IW11, IW12 and IW14, are involved in a water bridge near the active site and are located on the protein surface (average solvent accessibility of 14 Å^2) and are flexible, with a high average normalized *B* factor (Table 3). As expected, most of them have low occupancy as computed from the MD calculations, with the exception of IW10 (Table 3). Another four water molecules, IW1, IW3, IW8 and IW13, located on the surface of the protein molecule show low normalized *B* factors and high occupancy (Table 3). However, the role of these water molecules is not clear.

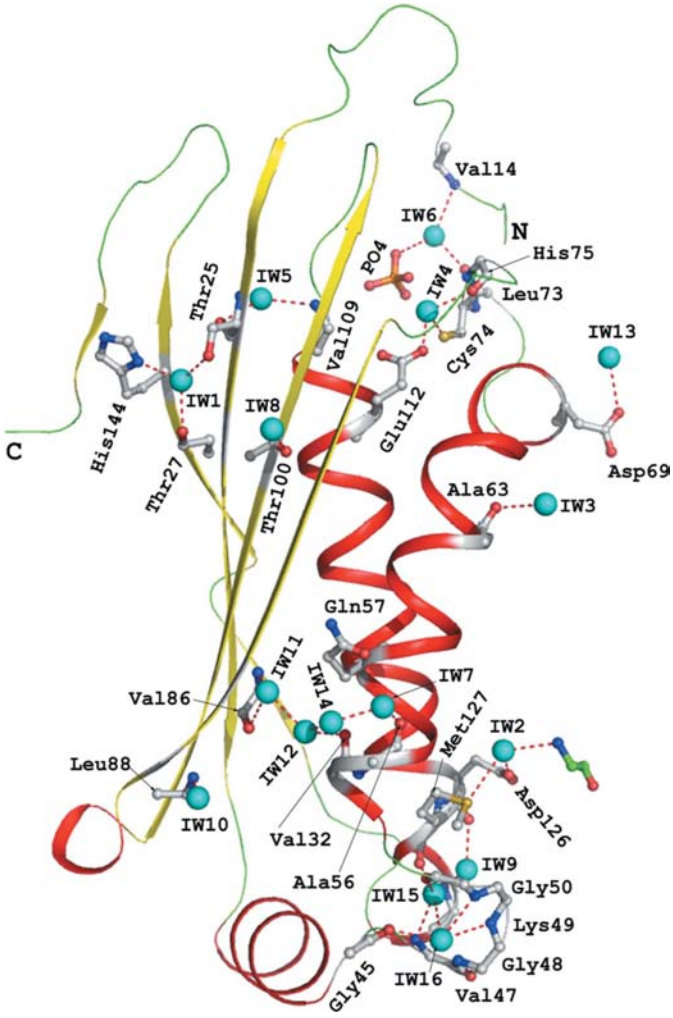


Figure 9
The hydrogen-bond interactions of invariant water molecules. The invariant water molecules are shown in cyan. Only the polar groups of the interacting residues of the protein molecule are shown for clarity. One of the residues, the C atoms of which are shown in green, is from the other monomer of the dimer.

5. A possible mechanism of the first step of Moco biosynthesis

Based on the previous studies of MoaA (Hanzelmann & Schindelin, 2006) and MoaC (Wuebbens *et al.*, 2000) and the present work, it may be suggested that the intermediate compound (FPT) generated by MoaA is the most potent substrate molecule for MoaC. However, the possibility of another compound (PBT) being the substrate of MoaC cannot be neglected. Thus, two possible sets of mechanisms are proposed here. Firstly, in the case where FPT is the substrate of MoaC, precursor Z (the final stable compound in the first step of the Moco biosynthesis pathway) can be generated in two ways (Fig. 10). In the second case, the ring formation of FPT molecule is completed first and the resulting compound (PBT) may play the role of the substrate of MoaC (Fig. 10). However, the interaction energies computed using MD simulations suggest that the first case is more favourable.

6. Conclusions

The crystallographic study of apo and GTP-bound crystal structures of MoaC from *T. thermophilus* coupled with ITC experiments and MD simulations provides insight into the substrate binding, structure dynamics and a possible mechanism. The GTP-bound crystal structure reveals that residues Lys49, His75, Asp126 and Lys129 are critical for the activity of the protein molecule. Together with the interaction energies calculated from MD simulations, the ITC results provide insight into the differentiation of the molecules binding to the protein molecule. They suggest that molecules with triphosphates are more potent for binding to MoaC. The study of the plasticity of the protein molecule reveals that all the α-helices, which are conserved among MoaC from different species, are highly rigid. In addition, most of the

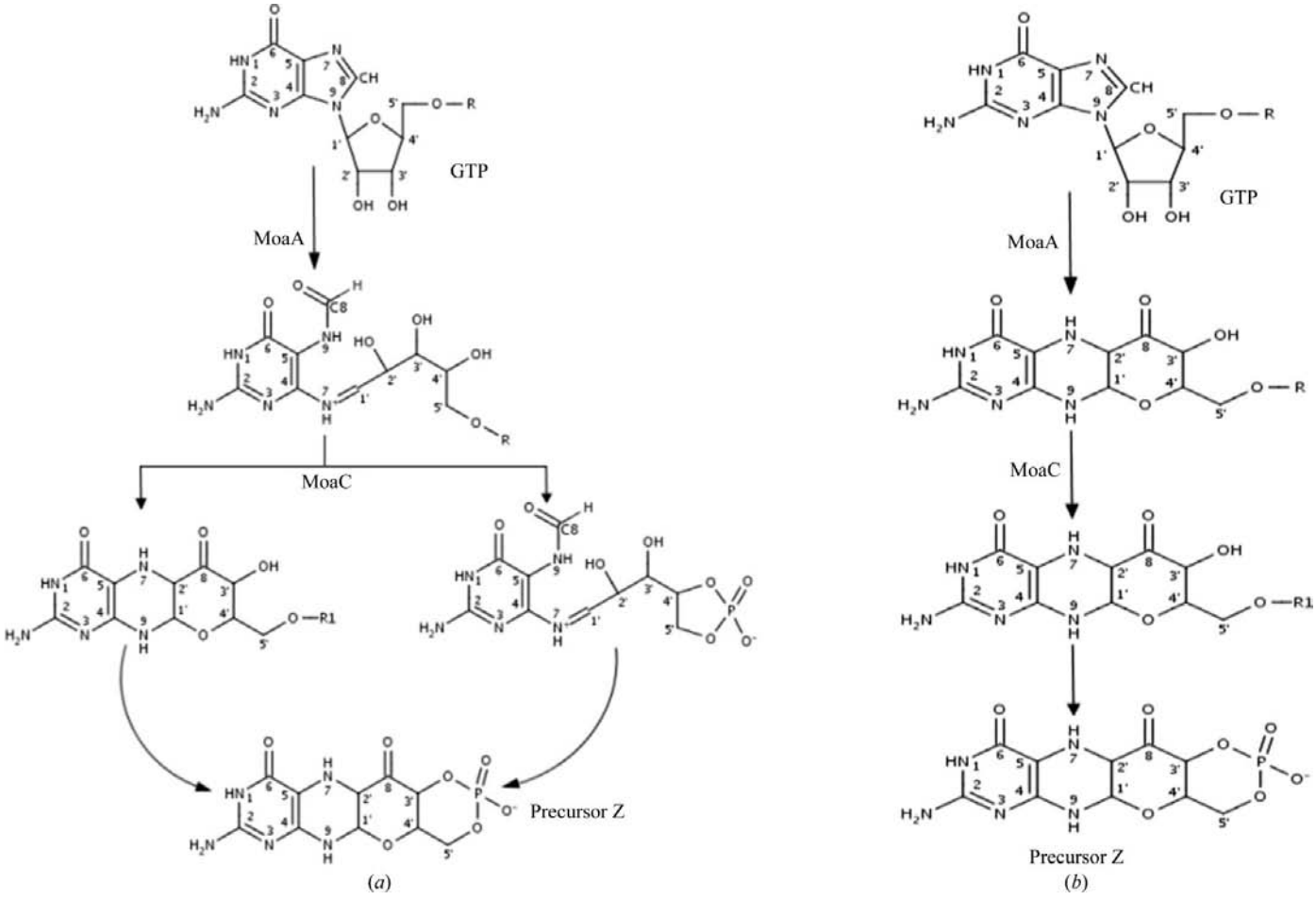


Figure 10
Schematic diagram of possible mechanisms proposed for the first step of the Moco biosynthesis pathway involving two probable substrate molecules (*a*) FPT and (*b*) PBT for MoaC (see text for details). R and R1 denote triphosphate and monophosphate groups, respectively.

residues involved in the β-sheet are flexible. In addition, 16 invariant water molecules were identified, some of which were located in the vicinity of the active site. Interestingly, the water molecules IW2, IW4 and IW9 may play a functional role in the catalytic activity of the protein molecule. The interaction energies obtained from MD simulations for the protein–ligand complexes revealed no clear distinction between the different substrates except for the NTP, NDP and NMP trend. In addition, these results support the crystallographic and ITC results.

KS and SPK thank the Bioinformatics Centre and the Interactive Graphics-Based Molecular Modelling Facility. The intensity data for the GTP-bound crystal were collected using the X-ray Facility for Structural Biology at the Molecular Biophysics Unit (MBU), Indian Institute of Science (IISc) supported by the Department of Science and Technology (DST) and the Department of Biotechnology (DBT), Government of India. KS and SPK thank Professor R. Varadarajan, MBU, IISc for his kind help during the ITC experiments. JJ thanks beamlines BL26B1 and BL26B2 of SPring-8 for excellent facilities and assistance and thanks the Department of Bioinformatics, Alagappa University, Karaikudi, Tamilnadu, India. This work was supported by the RIKEN Structural Genomic/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses and Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.
Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. (1995). *Comput. Phys. Commun.* **91**, 43–56.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Case, D. A. *et al.* (2006). *AMBER* 9. University of California, San Francisco, USA.
Chan, M. K., Mukund, S., Kletzin, A., Adams, M. W. & Rees, D. C. (1995). *Science*, **267**, 1463–1469.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.

Cohen, G. E. (1997). *J. Appl. Cryst.* **30**, 1160–1161.
Cruickshank, D. W. J. (1999). *Acta Cryst.* D**55**, 583–601.
Darden, T., York, D. & Pedersen, L. (1993). *J. Chem. Phys.* **98**, 10089–10092.
Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J. & Kollman, P. (2003). *J. Comput. Chem.* **24**, 1999–2012.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Frisch, M. J. *et al.* (2004). *Gaussian*03. Gaussian Inc., Wallingford, Connecticut, USA.
Hanzelmann, P., Hernandez, H. L., Menzel, C., Garcia-Serres, R., Huynh, B. H., Johnson, M. K., Mendel, R. R. & Schindelin, H. (2004). *J. Biol. Chem.* **279**, 34721–34732.
Hanzelmann, P. & Schindelin, H. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 12870–12875.
Hanzelmann, P. & Schindelin, H. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 6829–6834.
Hanzelmann, P., Schwarz, G. & Mendel, R. R. (2002). *J. Biol. Chem.* **277**, 18303–18312.
Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. (1997). *J. Comput. Chem.* **18**, 1463–1472.
Hubbard, S. J. & Thornton, J. M. (1993). *NACCESS.* Department of Biochemistry and Molecular Biology, University College London.
Hussain, A. S. Z., Shanthi, V., Sheik, S. S., Jeyakanthan, J., Selvarani, P. & Sekar, K. (2002). *Acta Cryst.* D**58**, 1385–1386.
Johnson, J. L., Wuebbens, M. M., Mandell, R. & Shih, V. E. (1989). *J. Clin. Invest.* **83**, 897–903.
Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.
Kanaujia, S. P., Ranjani, C. V., Jeyakanthan, J., Baba, S., Chen, L., Liu, Z.-J., Wang, B.-C., Nishida, M., Ebihara, A., Shinkai, A., Kuramitsu, S., Shiro, Y., Sekar, K. & Yokoyama, S. (2007). *Acta Cryst.* F**63**, 27–29.
Kanaujia, S. P. & Sekar, K. (2009). *Acta Cryst.* D**65**, 74–84.
Kisker, C., Schindelin, H. & Rees, D. C. (1997). *Annu. Rev. Biochem.* **66**, 233–267.
Kleywegt, G. J. (2007). *Acta Cryst.* D**63**, 94–100.
Konagurthu, A. S., Whisstock, J. C., Stuckey, P. J. & Leskh, A. M. (2006). *Proteins*, **64**, 559–574.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lindahl, E., Hess, B. & van der Spoel, D. (2001). *J. Mol. Model.* **7**, 306–317.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
McDonald, I. K. & Thornton, J. M. (1994). *J. Mol. Biol.* **238**, 777–793.
Mendel, R. R. & Bittner, F. (2006). *Biochim. Biophys. Acta*, **1763**, 621–635.
Menendez, C., Siebert, D. & Brandsch, R. (1996). *FEBS Lett.* **391**, 101–103.
Nohno, T., Kasai, Y. & Saito, T. (1988). *J. Bacteriol.* **170**, 4097–4102.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pitterle, D. M. & Rajagopalan, K. V. (1989). *J. Bacteriol.* **171**, 3373–3378.
Pitterle, D. M. & Rajagopalan, K. V. (1993). *J. Biol. Chem.* **268**, 13499–13505.
Rajagopalan, K. V. & Johnson, J. L. (1992). *J. Biol. Chem.* **267**, 10199–10202.
Rivers, S. L., McNairn, E., Blasco, F., Giordano, G. & Boxer, D. H. (1993). *Mol. Microbiol.* **8**, 1071–1081.
Schneider, T. R. (2004). *Acta Cryst.* D**60**, 2269–2275.
Schüttelkopf, A. W. & van Aalten, D. M. F. (2004). *Acta Cryst.* D**60**, 1355–1363.
Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F. & Miller, N. E. (2001). *Nucleic Acids Res.* **29**, 1097–1106.
Sorin, E. J. & Pande, V. S. (2005). *Biophys. J.* **88**, 2472–2493.
Sumathi, K., Ananthalakshmi, P., Roshan, M. N. A. M. & Sekar, K. (2006). *Nucleic Acids Res.* **34**, W128–W134.
Wuebbens, M. M., Liu, M. T., Rajagopalan, K. & Schindelin, H. (2000). *Structure*, **8**, 709–718.
Wuebbens, M. M. & Rajagopalan, K. V. (1993). *J. Biol. Chem.* **268**, 13493–13498.
Yoshida, H., Yamada, M., Kuramitsu, S. & Kamitori, S. (2008). *Acta Cryst.* F**64**, 589–592.