5-1-2012

Structure of the Catalytic Chain of Methanococcus Jannaschii Aspartate Transcarbamoylase in a Hexagonal Crystal Form: Insights into the Path of Carbamoyl Phosphate to the Active Site of the Enzyme

Jacqueline Vitali
Cleveland State University, j.vitali@csuohio.edu

Aditya K. Singh
Cleveland State University

Alexei S. Soaresb
Brookhaven National Laboratory

Michael J. Colaneri
SUNY College at Old Westbury

Publisher's Statement

Original Citation

Vitali, J., Singh, A.K., Soares, A.S. and Colaneri, M.J. (2012) Structure of the catalytic chain of Methanococcus jannaschii aspartate transcarbamoylase in a hexagonal crystal form: insights into the path of carbamoyl phosphate to the active site of the enzyme. Acta Crystallographica F68, 527-534

Repository Citation

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Structure of the catalytic chain of *Methanococcus jannaschii* aspartate transcarbamoylase in a hexagonal crystal form: insights into the path of carbamoyl phosphate to the active site of the enzyme

Jacqueline Vitali,[a]* Aditya K. Singh,[a] Alexei S. Soares[b] and Michael J. Colaneri[c]

[a]Department of Physics, Cleveland State University, Euclid Avenue at East 24th Street, Cleveland, OH 44115, USA, [b]Biology Department, Brookhaven National Laboratory, Upton, NY 11973, USA, and [c]Department of Chemistry and Physics, SUNY College at Old Westbury, Old Westbury, NY 11568, USA

Correspondence e-mail: j.vitali@csuohio.edu

Received 8 November 2011
Accepted 13 March 2012

PDB Reference: aspartate transcarbamoylase catalytic chain, 4ekn.

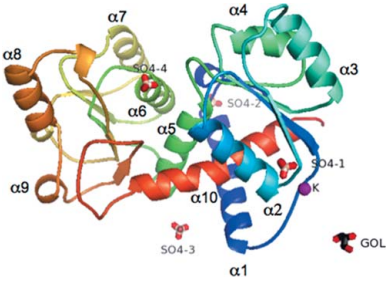
Crystals of the catalytic chain of *Methanococcus jannaschii* aspartate transcarbamoylase (ATCase) grew in the presence of the regulatory chain in the hexagonal space group $P6_322$, with one monomer per asymmetric unit. This is the first time that crystals with only one monomer in the asymmetric unit have been obtained; all known structures of the catalytic subunit contain several crystallographically independent monomers. The symmetry-related chains form the staggered dimer of trimers observed in the other known structures of the catalytic subunit. The central channel of the catalytic subunit contains a sulfate ion and a K^+ ion as well as a glycerol molecule at its entrance. It is possible that it is involved in channeling carbamoyl phosphate (CP) to the active site of the enzyme. A second sulfate ion near Arg164 is near the second CP position in the wild-type *Escherichia coli* ATCase structure complexed with CP. It is suggested that this position may also be in the path that CP takes when binding to the active site in a partial diffusion process at 310 K. Additional biochemical studies of carbamoylation and the molecular organization of this enzyme in *M. jannaschii* will provide further insight into these points.

1. Introduction

Aspartate transcarbamoylase (ATCase; EC 2.1.3.2) catalyzes the second step of *de novo* pyrimidine biosynthesis: the reaction between carbamoyl phosphate (CP) and aspartate to form *N*-carbamoyl-L-aspartate (CA) and inorganic phosphate (Jones *et al.*, 1955). It exists in different forms and molecular organizations in different organisms. In prokaryotes, the first three enzymes of the pathway, namely carbamoyl phosphate synthetase (CPSase), ATCase and dihydroorotase (DHOase), are commonly expressed separately. They function either independently, as in *Escherichia coli*, or form oligomeric complexes, as in *Thermus* ZO5 (Van de Casteele *et al.*, 1997) and *Aquifex aeolicus* (Purcarea *et al.*, 2003). In mammals, these three activities are part of the same polypeptide chain called CAD, which self-associates to form hexamers of 1.5 MDa (Evans & Guy, 2004).

There are three known forms of ATCase in prokaryotes. Type A1 ATCase is a dodecamer of six catalytic ATCase chains and six active DHOase chains as in *A. aeolicus* (Ahuja *et al.*, 2004) and *Thermus aquaticus* (Van de Casteele *et al.*, 1997). Type A2 complexes are similar to type A1 complexes except that the DHOase domain is inactive and fulfills only a structural role, as in *Pseudomonas aeruginosa* (Vickrey *et al.*, 2002). Type B enzymes form a dodecamer of six catalytic chains and six regulatory chains as in *E. coli* (Wiley & Lipscomb, 1968). Type C enzymes function as unregulated free trimers as in *Bacillus subtilis* (Brabson *et al.*, 1985). In all known ATCase enzymes the catalytic chains are active as homotrimers, in which the active sites are formed by residues from two subunits.

The structure and properties of the *E. coli* enzyme (type B) have been extensively studied (Hervé, 1989; Allewell, 1989; Lipscomb, 1992, 1994; England *et al.*, 1994). The holoenzyme has a dodecameric structure containing two trimers of catalytic chains linked by three regulatory dimers. The catalytic chains have two domains: the CP-binding and the aspartate-binding domains. The regulatory chains also have two domains: the nucleotide-binding and zinc-binding



domains. The mechanism of catalysis in the *E. coli* enzyme is sequential ordered, with CP binding first and conditioning the active site for binding of the second substrate, aspartate. The binding of aspartate triggers the closure of the two domains which is necessary for catalysis and the conformational change from a low-activity T state to a high-activity R state. CA is released first, followed by inorganic phosphate. *E. coli* ATCase is an allosteric enzyme that exhibits cooperativity for aspartate and heterotropic effects, being activated by ATP and inhibited by CTP. The large conformational differences between the crystal structures of unliganded ATCase (Stevens *et al.*, 1990*a*,*b*) and the *N*-phosphonacetyl-L-aspartate (PALA) liganded enzyme (Ke *et al.*, 1988; Jin *et al.*, 1999) have been proposed to define the structural differences between the T and R states.

The characterization of ATCase from the hyperthermophilic and barophilic archaeon *Methanococcus jannaschii* (Hack *et al.*, 2000) suggested that it has a molecular weight similar to that from *E. coli*. Kinetic analysis of *M. jannaschii* ATCase from cell-free extracts showed that it has limited homotropic cooperativity and little if any regulatory properties with ATP and CTP. Kinetic analysis of the *M. jannaschii* catalytic trimer showed hyperbolic kinetics with an activation energy similar to that of the *E. coli* trimer and with activity increasing with temperature. It is stable at 358 K.

We have previously determined the structure of the catalytic trimer of *M. jannaschii* ATCase in two crystal forms: monoclinic (Vitali *et al.*, 2008) and orthorhombic (Vitali & Colaneri, 2008). These studies and comparisons with *E. coli* ATCase (Stevens *et al.*, 1990*a*,*b*; Jin *et al.*, 1999; Beernink *et al.*, 1999; Endrizzi *et al.*, 2000) and the hyperthermophilic ATCases from *Pyrococcus abyssi* (Van Boxstael *et al.*, 2003) and *Sulfolobus acidocaldarius* (De Vos *et al.*, 2004) gave insights into the strategies for thermostabilization adopted by the *M. jannaschii* enzyme. An interesting feature of both structures was the vertical association of catalytic subunits in pairs into staggered dimers of trimers with a short separation distance of 33.7 Å. It was suggested that this arrangement may be present in the holoenzyme *in vivo* in the presence of the regulatory subunits. In order to further investigate this hypothesis, we attempted to cocrystallize the holoenzyme from stoichiometric mixtures of the catalytic and regulatory chains. However, we obtained crystals containing the same dimers of trimers without regulatory chains. The crystals diffracted to a higher resolution than in the earlier studies. The present work provides structural insight into how CP, an unstable metabolite, may be protected from thermal degradation in *M. jannaschii*, as well as the possible paths it may follow to reach the active site.

2. Materials and methods

2.1. Protein preparation and crystallization

The *M. jannaschii* ATCase catalytic trimer (NCBI NP_248590; UniProt Q58976) was prepared from *E. coli* strain EK1911 (Hack *et al.*, 2000), which has a deletion in the *pyrBI* region of the chromosome and contains plasmids pEK406 coding for the *M. jannaschii* ATCase catalytic chain and pSJS1240 (Kim *et al.*, 1998) coding for rare archaeal tRNAs. For the regulatory subunit (NCBI NP_248409; UniProt Q58801), we used an ATCase-deficient derivative of *E. coli* strain C600 (Van Boxstael *et al.*, 2003), into which we introduced a T7 RNA polymerase gene under lacUV5 control as a lambda lysogen using the λDE3 lysogenization kit from Novagen according to the manufacturer's instructions. Plasmids pEK407 (Hack *et al.*, 2000) and pSJS1240 (Kim *et al.*, 1998) were co-transformed into these cells. The genes for the *M. jannaschii* ATCase catalytic and regulatory chains are not associated with any tags in plasmids pEK406 and pEK407 (Hack *et al.*, 2000). The cells for each subunit were separately grown to an OD_{600} of ∼0.8 at 310 K in LB Lennox medium containing 100 mg ml^{-1} ampicillin and 100 mg ml^{-1} spectinomycin and were induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 1 m*M* for 5 h at 310 K. They were harvested by centrifugation at 4000 rev min^{-1} and stored at 193 K.

Before use, the frozen pellets were thawed and resuspended in breakage buffer as described by Hack *et al.* (2000), sonicated and centrifuged at 16 000 rev min^{-1} and the proteins were obtained from the cell-free supernatants. The purification of the catalytic subunit followed the procedure of Hack *et al.* (2000) and involved a 30% ammonium sulfate precipitation step, a heat step at 358 K for 15 min and chromatography using a Q-Sepharose Fast Flow anion-exchange column (HiPrep Q FF 16/10, GE Healthcare) and a phenyl Sepharose column [HiPrep Phenyl FF (high Sub) 16/10, GE Healthcare]. The purification of the regulatory subunit involved a 28% ammonium sulfate step, a heat step at 358 K for 5 min, a Q-Sepharose Fast Flow anion-exchange column (HiTrap Q FF, GE Healthcare) and an SP Sepharose Fast Flow cation-exchange column (Lab Pack from GE Healthcare). Both columns were pre-equilibrated with 50 m*M* Tris pH 9, 0.05 m*M* zinc acetate, 2 m*M* β-mercaptoethanol (BME) and were eluted with a stepped gradient of 0–0.5 *M* NaCl.

The regulatory and catalytic chains were mixed and were further purified using size-exclusion chromatography on a HiPrep 16/60 Sephacryl S-300 HR column (GE Healthcare) in 40 m*M* KH_2PO_4, 2 m*M* BME, 0.1 *M* KCl, 0.05 *M* zinc acetate pH 8.1. The fractions in which the subunits co-eluted in a molar ratio of 1:1 as apparent on SDS–PAGE gels were dialyzed three times in 50 m*M* Tris pH 8.3, 150 m*M* NaCl, 2 m*M* BME, 0.05 m*M* zinc acetate and concentrated to 11 mg ml^{-1} with centrifugal filters (Amicon Ultra) with a molecular-weight cutoff (MWCO) of 10 kDa. All centrifugations were carried out in a Beckman J2-21 centrifuge, and the chromatography for the catalytic subunit and the size-exclusion chromatography for the mixture were performed using an ÄKTAprime system (GE Healthcare).

Crystallizations were carried out at 295 K by the sitting-drop vapor-diffusion method in 24-well plates from BD Biosciences using micro-bridges from Hampton Research and 500 µl reservoir solutions. Initially, crystallization conditions were sought using the sparse-matrix Crystal Screen and Crystal Screen 2 from Hampton Research (Jancarik & Kim, 1991). The drops consisted of 1 µl reservoir solution and 1.3 µl complex solution. The best condition was condition No. 14 from Crystal Screen 2 (2.0 *M* ammonium sulfate, 0.2 *M* potassium sodium tartrate tetrahydrate, 0.1 *M* sodium citrate tribasic dihydrate pH 5.6). Nine further conditions produced interesting results. Condition No. 14 was further optimized for pH and concentration of potassium sodium tartrate tetrahydrate. The drops in the extrapolation screen contained 2 µl reservoir solution and 2.6 µl complex solution. The best condition involved reservoirs consisting of 2.0 *M* ammonium sulfate, 0.2 *M* potassium sodium tartrate tetrahydrate, 0.1 *M* Tris–HCl pH 7.5. The crystal used in this study was a hexagonal prism with approximate dimensions of 0.1 × 0.1 × 0.2 mm and was grown under these conditions.

2.2. X-ray data collection

Diffraction data were measured on the X12C beamline of the National Synchrotron Light Source at Brookhaven National Laboratory using an ADSC Q210 CCD detector. The temperature was 100 K, the wavelength was 1.1 Å and the crystal-to-detector distance was 200.00 mm. The cryoprotectant used was 25% glycerol.

Table 1
Data-collection and final refinement statistics.

Values in parentheses are for the highest resolution shell used in the refinement.

Data collection	
Space group	$P6_322$
Unit-cell parameters (Å)	$a = b = 96.96$, $c = 136.44$
Resolution range (Å)	50.0–2.5 (2.59–2.50)
Wavelength (Å)	1.1
No. of unique reflections	13471 (1314)
Multiplicity	3.3 (3.3)
Completeness (%)	98.0 (99.2)
Mean $I/\sigma(I)$	9.31 (1.62)
R_{merge}†	0.135 (0.825)
Final refinement	
Resolution range (Å)	42.0–2.5
No. of reflections	13469‡
No. of reflections in test set	1377
Contents of asymmetric unit	
No. of protein atoms	2460
No. of waters	145§
No. of sulfates/K^+/GOL	4/1/1¶
R_{work} (90% of data)	0.183
R_{free} (10% of data)	0.270
R_{all} (all data)	0.192
R.m.s.d. bond lengths (Å)	0.005
R.m.s.d. bond angles (°)	0.65
B factors (Å^2)	32.4
From Wilson plot	32.4
Mean, over all atoms	38.5
Mean, protein main chain	35.7
Mean, protein side chains	40.9
Ramachadran plot (%)	
Most favored	90.5
Additional allowed	8.8
Generously allowed	0.4††
Disallowed (%)	0.4‡‡

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ and $\langle I(hkl)\rangle$ are the observed intensity of measurement i and the mean intensity of the reflection with indices hkl, respectively. ‡ Two outliers identified by the program were excluded. § Three water molecules are located on crystallographic symmetry axes. ¶ All ligands are on crystallographic symmetry axes, except for SO4-2 and SO4-4, which are in general positions. GOL is threefold disordered around the crystallographic threefold axis. †† Only Ser128 is in this region and corresponds to poor electron density. ‡‡ Leu263 is in a non-accepted region, as is often the case for active-site residues. This residue is found in non-accepted regions in the PALA-liganded and unliganded *E. coli* catalytic subunit (Endrizzi *et al.*, 2000; Beernink *et al.*, 1999) and holoenzyme (Jin *et al.*, 1999; Stevens *et al.*, 1990*a*,*b*), in the PALA-liganded *P. abyssi* catalytic trimer (Van Boxstael *et al.*, 2003) and in the orthorhombic form of this enzyme (Vitali & Colaneri, 2008).

Oscillations were measured for 60 s each at 1.0° intervals in φ. The data were indexed, integrated and scaled using *HKL*-2000 (Otwinowski & Minor, 1997). The crystal was hexagonal and belonged to space group $P6_322$, with unit-cell parameters $a = b = 96.96$, $c = 136.44$ Å, one catalytic chain per asymmetric unit and a V_M of 2.25 Å^3 Da^{-1} (Matthews, 1968). Data statistics are summarized in Table 1.

2.3. Structure determination and refinement

The structure was solved using molecular replacement with the program *MOLREP* (Vagin & Teplyakov, 2010) in the *CCP*4 package (Winn *et al.*, 2011). The search model consisted of one catalytic chain of the monoclinic crystal form (PDB entry 2rgw, chain *D*; Vitali *et al.*, 2008). Refinement was carried out with the *PHENIX* suite of programs (Adams *et al.*, 2010) using simulated annealing with torsion-angle dynamics and was alternated with manual model building and rebuilding using *Coot* (Emsley & Cowtan, 2004). The *massage* algorithm in *phenix.reflection_file_converter* was used to treat negative intensities for refinement of the structure. The later stages of the refinement were carried out using data to 2.5 Å resolution, beyond the nominal resolution of 2.58 Å [$I/\sigma(I) = 2.0$], in order to take advantage of the information contained in the weaker reflections (Wang, 2010). There are four sulfate ions bound in the monomer: one on the crystallographic threefold axis (SO4-1), one on a crystallographic twofold axis (SO4-3) and two in general positions (SO4-2 and SO4-4). In addition, the structure has a K^+ ion on the threefold axis. Finally, there is a glycerol molecule from the cryoprotectant on this axis with threefold disorder around it. The K^+ ion was given an occupancy of one, as partial occupancy at a special position is taken care of internally in the program. All atoms of the other ligands in special positions were given partial occupancy based on the multiplicity of the special position to turn on the nonbonded symmetry exclusion for these ligands. Water molecules were added to the model in *phenix.refine* with default parameters, except that the $mF_o - DF_c$ map cutoff was 2.5σ and the largest hydrogen-bonding distance for water was taken to be 3.6 Å. A few waters were positioned manually after visual inspection of the electron-density maps. The program *PROCHECK* (Laskowski *et al.*, 1993) was used to assess the quality of the model. Final refinement statistics are provided in Table 1. A view of the asymmetric unit of the structure and its attached ligands is shown in Fig. 1. The final electron density contoured at 1.0σ is continuous for most main-chain and side-chain atoms, with the exceptions of residues 74–82 and 127–128 where it is weak and fragmented, the last two residues at the C-terminus where it is absent, and a few surface residues where it is very weak.

Refinement of the structure was initially carried out at 2.7 Å resolution. The sulfate and the K^+ ions were identified from strong regions of electron density in the $2mF_o - DF_c$ and $mF_o - DF_c$ maps. Their peak heights were in the range 1.0–4.0σ in the $2mF_o - DF_c$ map and 4.0–11.0σ in the $mF_o - DF_c$ map. These sites were chemically consistent with sulfate and K^+ sites in their environments and interactions and the ions were present in the crystallization medium. However, there was additional electron density along the hydrogen bonds of sulfate SO4-4 to Arg164 N at lower contour levels of less than 1.0σ in the $2mF_o - DF_c$ maps. In addition, the side chain of Arg164 was visible in these maps at a contour level of less than 0.8σ. The inclusion of data to 2.5 Å resolution in the refinement and the use of averaged kick maps (Pražnikar *et al.*, 2009) established the map interpretation in this region by disconnecting the sulfate peak from the main-chain electron density and improving the side-chain density for Arg164. Fig. 2 shows the electron density for the SO4-1, K^+ and SO4-4 sites as well as their environments at the end of the analysis. For computation of these maps, the ions and the residues interacting

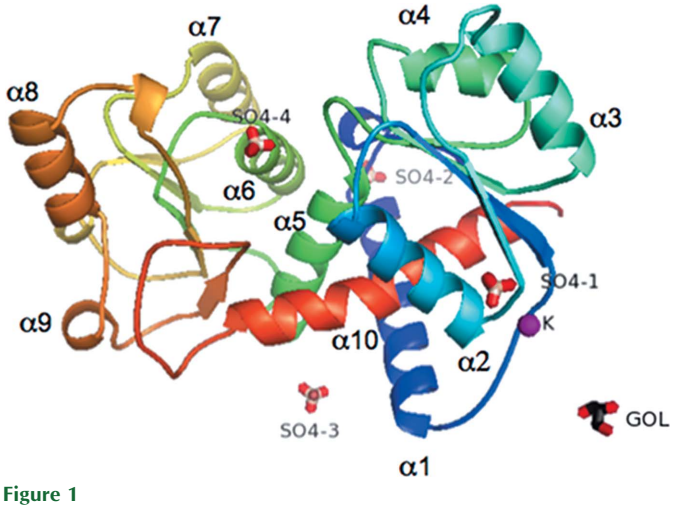


Figure 1
Ribbon representation of the catalytic chain, illustrating the positions of the ligands. Colors are from blue at the N-terminus to red at the C-terminus. Helices are labeled according to Vitali *et al.* (2008).

with them were removed from the model, which was then refined using simulated annealing with torsion-angle dynamics.

2.4. Model analysis

Hydrogen bonds were calculated with *HBPLUS* (McDonald & Thornton, 1994) using donor–acceptor distances of less than 3.6 Å, hydrogen–acceptor distances of less than 2.5 Å and associated angles of greater than 90°. Salt bridges between two charged groups correspond to distances of less than 4.0 Å. The *Protein Interfaces, Surfaces and Assemblies* (*PISA*) service at the European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html; Krissinel & Henrick, 2007) was used to compute buried surface areas. Structure superpositions were performed with *LSQMAN* (Kleywegt, 1996).

Planar angles between the CP-binding and the aspartate-binding domains were computed by a modification of the method of Williams *et al.* (1998) using the angle between the geometric centers of the two domains and a hinge point. The geometric centers of the CP-binding and the aspartate-binding domains of *M. jannaschii* ATCase were computed from the C^{α} atoms of residues 1–131 and 147–280, respectively. The hinge point was taken as the C^{α} atom of residue 137.

The global association of two catalytic subunits in a complex is described by the distance between their geometric centers and the torsional angle between the individual chains of the two subunits around the axis defined by this line. The geometric centers were computed from the C^{α} atoms of residues 1–131 and 147–280.

Figures were prepared with *PyMOL* (http://www.pymol.org). The central channel of the catalytic subunit was illustrated using the *CAVER* plugin (Petrek *et al.*, 2006). The electrostatic surfaces were calculated using *APBS* (Baker *et al.*, 2001) through the *PDB2PQR* web portal (Dolinsky *et al.*, 2004) at neutral pH and zero ionic strength with the AMBER force field (Case *et al.*, 2005). The dielectric constant was set to 2.0 for the protein and 78.0 for the solvent.

3. Results and discussion

3.1. Description of the structure

The catalytic chain (Fig. 1) is similar to other known structures of the catalytic chain of *M. jannaschii* ATCase (Vitali *et al.*, 2008; Vitali & Colaneri, 2008), with r.m.s.d.s between corresponding C^{α} atoms in the range 0.39–0.56 Å. There is a variation in the planar angle between the CP-binding and aspartate-binding domains among the known structures. The planar angle of 124.5° in the present structure is comparable to the planar angles of 122.9–125.2° in the orthorhombic form (Vitali & Colaneri, 2008), but is smaller by ∼5° than the planar angles in the monoclinic form (average of 129.5°; Vitali *et al.*, 2008). It is likely that this variation reflects the flexibility and reactivity of the catalytic chain.

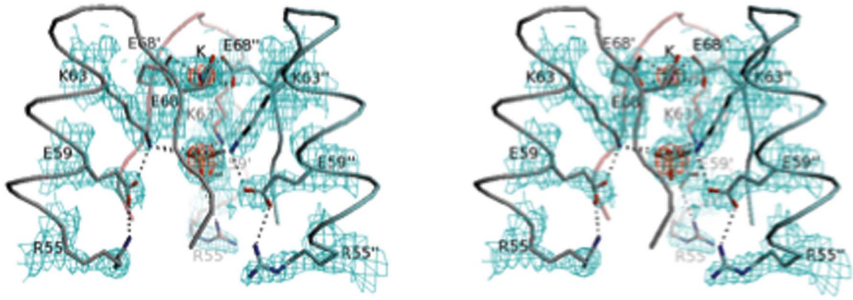


Figure 2
(*a*) A stereo pair illustrating the interactions involving the ions inside the central channel. The color scheme is as follows. C atoms are shown in silver for C1 (main molecule), salmon for C2 (′) and cyan for C3 (″). O atoms are shown in red, N atoms in blue and S atoms in wheat. The purple sphere is the K^{+} ion. The C^{α} backbone is shown as a cartoon tube. Primes and double primes are included in the residue names of C2 and C3 to emphasize that these chains are related to chain C1 by the threefold axis. For clarity, amino-acid names use one-letter codes in this figure. For calculation of the electron-density maps, the structure was refined using torsion-angle simulated annealing with the ions and residues of the channel omitted from the model. The teal electron density is a $2mF_o - DF_c$ map at 1.2σ and the red electron density is an $mF_o - DF_c$ map at 7.5σ. (*b*) Environment of sulfate ion SO4-4. A stereo pair. There are three salt bridges between the sulfate O atoms and the guanidino N atoms of Arg226 and three hydrogen bonds involving the amide N atoms of Arg164 and Thr165. For the calculation of the electron-density maps, the structure was refined using torsion-angle simulated annealing with the atoms of the sulfate ion and the residues interacting with it omitted from the model. The teal electron density is a $2mF_o - DF_c$ map at 0.8σ.

The catalytic chain in the asymmetric unit makes contacts with symmetry-related chains to form the catalytic trimer and the staggered dimer of trimers observed in other known structures of the catalytic subunit of *M. jannaschii* ATCase. However, the threefold symmetry of the catalytic trimer and the 32 symmetry of the dimer of trimers in the present structure are formed by crystallographic symmetry operations, whereas in the other structures the symmetry of these complexes is noncrystallographic. This is the first time that a single catalytic chain has been observed in the asymmetric unit; all other characterized crystals of the catalytic subunit contained multiple copies. Even though the catalytic and regulatory subunits were mixed in approximately the exact stoichiometric ratio, they did not cocrystallize. The situation is similar to that for DHOase from *A. aeolicus* (Martin *et al.*, 2005). A mesh representation of the hexamer that includes the names of the chains and the corresponding equivalent positions is shown in Fig. 3.

PISA predicts that the hexameric complex is a stable quaternary structure for this enzyme. This prediction is consistent with our observations since the hexameric species persists in different crystalline environments. However, previous size-exclusion chromatography studies have shown that the catalytic subunits exist as isolated trimers in Tris solution (Hack *et al.*, 2000). The possibility that the association that we observe in the crystalline state may occur at high concentrations of the protein and/or in the presence of ammonium sulfate was tested with dynamic light scattering (Vitali & Colaneri, 2008). These studies were consistent with the formation of hexamers but were inconclusive as the solutions showed high polydispersity. It was suggested that the hexameric species may be part of the holoenzyme *in vivo* in the presence of the regulatory subunits (Vitali & Colaneri, 2008).

The vertical association of the catalytic subunits in the hexamer shows some flexibility in the rotation around the axis connecting their geometric centers at a constant intersubunit vertical separation of 33.7 Å in the several crystal forms. It is more eclipsed in the present structure: by 4° from the hexamers in the orthorhombic form and by 8° from the hexamer in the monoclinic form. The global torsional angles C1—C4, C1—C6 and C1—C5 between the chains of the two catalytic subunits are −37, −157 and 83° in the present structure compared with −40, −160 and 80° and −41, −161 and 80° in the orthorhombic form and −44, −165 and 76° in the monoclinic form, respectively.

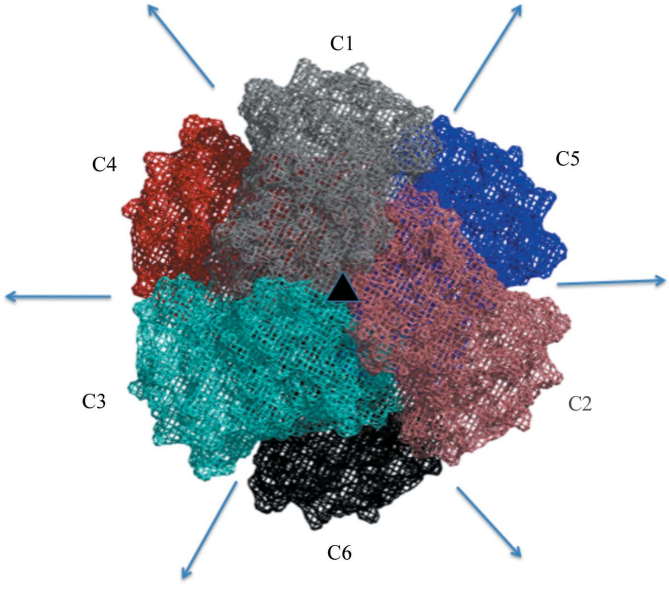


Figure 3
Mesh representation of the dimer of trimers looking down the crystallographic threefold axis. The arrows indicate the twofold axes and the triangle at the center indicates the threefold axis. Ligands are not shown in this figure. Catalytic chains C1, C2 and C3 comprise the top trimer and catalytic chains C4, C5 and C6 comprise the bottom trimer. The equivalent positions corresponding to the chains are C1 = (x, y, z), C2 = $(1 - y, x - y, z)$, C3 = $(-x + y + 1, -x + 1, z)$, C4 = $(x, x - y, -z + 1/2)$, C5 = $(-x + y + 1, y, -z + 1/2)$, C6 = $(-y + 1, -x + 1, -z + 1/2)$. C1–C4 have the smallest global angular separation and C1–C5 the next smallest. Colors: silver, C1; salmon, C2; cyan, C3; red, C4; blue, C5; black, C6.

The central channel of the catalytic subunit (Fig. 2*a*) contains a sulfate ion, SO4-1, and a K^+ ion on the crystallographic threefold axis that relates the three monomers as well as several waters. The sulfate ion is located at the center of the CP-binding domains of the three monomers. One of its O atoms is along the crystallographic threefold, while the other three are related by it. As in the previous structures (Vitali *et al.*, 2008; Vitali & Colaneri, 2008), the sulfate ion is involved in an extended ion-pair network with all three monomers of its subunit through charged residues of the α2 helix that point into the central channel: Lys63, Glu59 and Arg55. These residues form salt bridges with each other in each chain and Lys63 from each chain directly makes salt bridges to two of the sulfate O atoms around the threefold axis. The K^+ ion is located 6.0 Å away from the sulfate ion towards the top of the dome-shaped subunit. It is coordinated directly by the three Glu68 carboxylates related by the threefold in a bidentate mode. The $K^+\cdots$OE1 and $K^+\cdots$OE2 distances, of 2.8 and 2.7 Å, respectively, are close to the mean 2.9 Å for coordination of K^+ with carboxylates of Glu from the MESPEUS database (Hsin *et al.*, 2008) and structures to 2.0 Å resolution (http://tanna.bch.ed.ac.uk/). The entrance to the central channel, at the top of the dome-shaped subunit, has a glycerol molecule on the threefold axis threefold disordered around it (not shown).

The active site has a sulfate ion, SO4-4, near Arg164 (Fig. 2*b*). This sulfate is involved in three salt bridges to Arg226 and three hydrogen bonds to the amide N atoms of Arg164 and Thr165.

3.2. Structural insights into the path of CP to the active site of the enzyme

One intriguing question in hyperthermophilic organisms is how unstable metabolites such as CP, which is a key intermediate in both pyrimidine and arginine biosynthesis, are preserved from thermal degradation. CP has a half-life for thermal decomposition of less than 2 s at 373 K (Legrain *et al.*, 1995) and decomposes to the toxic cyanate, a promiscuous alkylating agent (Allen & Jones, 1964). In contrast, the half-life of CP at 310 K is 5 min. Therefore, these organisms must have a mechanism or mechanisms for protecting CP from thermal degradation.

It is likely that binding of CP to the active site of *M. jannaschii* ATCase stabilizes CP against thermal decomposition. The stereochemistry of binding in the active site of *M. jannaschii* ATCase is expected to be similar to that in *E. coli* ATCase (Wang *et al.*, 2008) as the residues involved in this interaction are conserved between the two systems. Furthermore, enzymatic studies and quantum-mechanics/molecular-mechanics calculations have shown that the stereochemistry of binding in *E. coli* ATCase precludes thermal decomposition by inhibiting the Allen–Jones pathway (Allen & Jones, 1964). The question that then remains is what is the stabilized path that CP takes to reach the active site once it is synthesized by CPSase.

Substrate channeling is prominent for CP in the pyrimidine and arginine pathways of hyperthermophilic organisms. It has been demonstrated in *Thermus* ZO5 (Van De Casteele *et al.*, 1997),

A. aeolicus (Purcarea *et al.*, 2003), *P. furiosus* (Massant & Glansdorff, 2005) and *P. abyssi* (Purcarea *et al.*, 1999). In all of these systems the corresponding enzymes form transient or short-lived complexes as opposed to stable stoichiometric complexes. Even so, the efficiency of intermediate transfer in such systems may be quite high. It is possible that a similar mechanism operates in *M. jannaschii* even though there are no kinetic data to support this hypothesis at present.

The presence of the two sulfate ions, one in the central channel (SO4-1) and one in the active site near Arg164 (SO4-4), suggests two possible routes that CP may follow to the active site. The chemical properties of sulfate and phosphate are sufficiently similar that all locations found to bind sulfate may also be viewed as potential phosphate-binding sites. In a few structures where experiments have been performed with both sulfates and phosphates, the two groups exploit the same residues for binding even though the details of the geometry of binding may differ for the two systems (Copley & Barton, 1994). It may be noted that in the present structure it is possible to model a CP molecule in the central channel at the position of sulfate SO4-1. In addition, the structure of *E. coli* ATCase in the presence of two CP molecules in the active site (EcATCase-2CP; Wang *et al.*, 2005; PDB entry 1za2) features one CP in the regular Ser52 position (Ser51 in *M. jannaschii* numbering) and the second weakly bound near Arg167 (Arg164 in *M. jannaschii* numbering), ∼9 Å away from the first CP (Fig. 4). A superposition of EcATCase-2CP on the present structure (Fig. 4) shows that the position of the second CP in EcATCase-2CP corresponds to sulfate SO4-4 near Arg164 in the present structure. Fig. 5 shows the electrostatic potential as calculated by *APBS* mapped onto the surface in the three active sites. The potential is positive between the two sulfates and Ser51, indicating electrostatically favorable paths for CP from either position to the active site (Ser51). A similar approach was followed by Ramón-Maiques *et al.* (2010) in their study of *Enterococcus faecalis* carbamate kinase, in which bound sulfate ions in the active site of the enzyme were considered to mimic the phosphate group of CP.

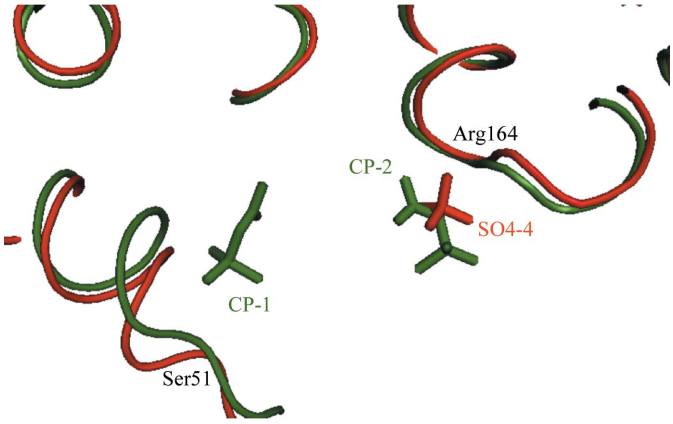


Figure 4
Superposition of the present structure with EcATCase-2CP (PDB entry 1za2) showing the ligands in the active site. The present structure is shown in red and EcATCase-2CP is shown in green. Note that the second CP in EcATCase-2CP (CP-2) is near SO4-4 in the present structure. The positions of Ser51 and Arg164 in *M. jannaschii* ATCase are marked.

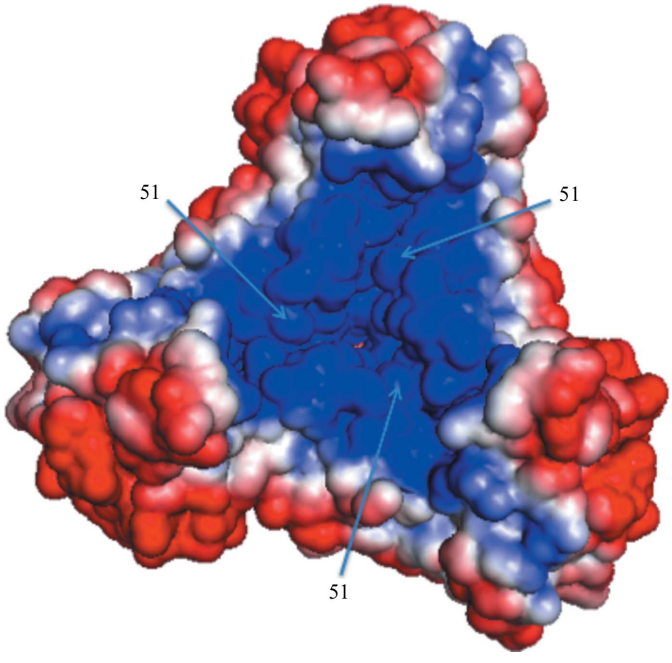


Figure 5
The electrostatic potential as calculated by *APBS* is mapped onto the active sites of the trimer. Solvent-accessible surface-area representation. Red, $-10kT$/e; blue, $+10kT$/e. The calculation was carried out for the trimer in the absence of ligands. Residue Ser51 is indicated by arrows. Sulfate SO4-1 in the central channel has been superimposed on the figure to indicate its position. The other three sulfates (SO4-4) are near the three corners of the aspartate-binding domains behind the 240s loops and are hidden in the figure.

The central channel in the catalytic trimer is formed by the $\alpha2$ helices and $\beta3$ strands from all three chains. The entrance to the channel at the top of the catalytic subunit is formed by the $\alpha1\beta2$ loops. On the other side the channel ends at the 80s loops. The dimension of the cross-section of the channel changes along its length, being wider at the top where it is formed by the $\alpha1\beta2$ loops and towards the active site. The minimal dimension of the channel is at the Glu68 position, where the channel is narrow. The distances of Glu68 OE1 and Glu68 OE2 from the tunnel axis are 2.4 and 2.6 Å, respectively. These distances readily increase to 6.2 and 4.6 Å, respectively, by changing the Glu68 side-chain rotamer. The next

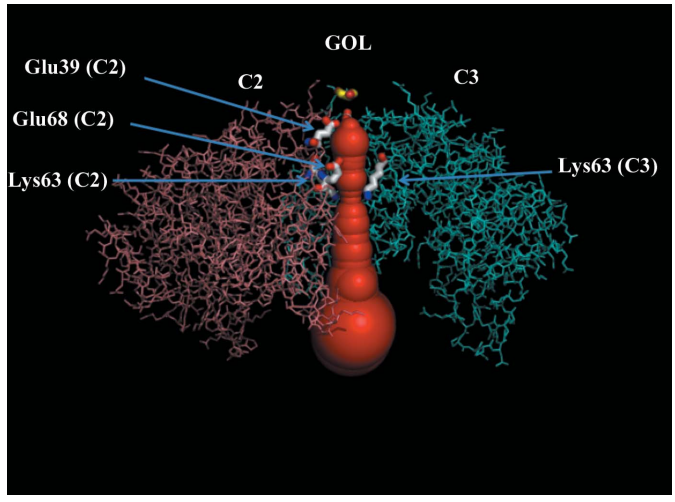


Figure 6
CAVER representation of the channel through the catalytic trimer. The channel is represented as a series of the largest spheres that can be fitted along its length. The starting point of the calculation for this figure was on the threefold axis approximately at the center of the three Glu39 residues. Glu68 is in the alternate orientation that increases the diameter of the channel at this position. Chains C2 (salmon) and C3 (cyan) are shown as lines. Chain C1 was removed from the figure for clarity. The Lys63 residues of C2 and C3 are shown as sticks. Glu68 and Glu39 of C3 are beneath the opaque channel. The glycerol ligand near the entrance to the channel is also shown.

smaller channel dimension corresponds to Lys63 NZ, which is 3.7 Å from the channel axis. Fig. 6 illustrates the central tunnel using the *CAVER* plugin in *PyMOL* with the side chains of Glu68 in the alternate orientation that widens the channel at that position.

It is possible that the central channel may be used for sequestering CP during catalysis. Some support for this idea is provided by the fact that a molecule as small as CP can pass through the channel without steric clashes when Glu68 is in the alternate orientation. In addition, the channel contains two ligands, a K^+ ion and a sulfate, and has a third ligand at its entrance. Finally, a similar suggestion for substrate channeling through the central channel of the catalytic subunit has been made for the DHOase–ATCase complex from *A. aeolicus* (Zhang *et al.*, 2009), with CPSases binding to the outside of the dodecamer and forming an antechamber with threefold symmetry over a shared tunnel through the ATCase trimer.

The second CP site near Arg164 (*M. jannaschii* numbering; Fig. 4) is presumed to be along the path that CP takes to bind to the active site in *E. coli* ATCase (Mendes & Kantrowitz, 2010). *E. coli* is a mesophilic organism and the substrates diffuse to the active sites of ATCase from the surrounding medium. It is possible that a similar diffusion mechanism through this site may partially operate in *M. jannaschii* ATCase at ambient temperatures if the active sites of the enzyme in its functional state are accessible to the solvent as is the case in *E. coli*. The half-life of CP is 5 min at 310 K and its thermal degradation is not a problem. In fact, partial channeling of CP at 310 K has been demonstrated in *P. abyssi* ATCase (Purcarea *et al.*, 1999), but the channeling efficiency increases dramatically at elevated temperatures. In addition, partial channeling of CP has been reported in the pyrimidine-biosynthetic complexes from yeast (Lue & Kaplan, 1970; Belkaïd *et al.*, 1988; Penverne *et al.*, 1994), *Neurospora* (Williams *et al.*, 1970, 1971) and mammals (Coleman *et al.*, 1977; Makoff & Radford, 1978; Mori & Tatibana, 1978; Christopherson & Jones, 1980; Mally *et al.*, 1980; Irvine *et al.*, 1997), and in the mammalian urea-cycle enzymes (Wanders *et al.*, 1984; Cohen *et al.*, 1992). Alternatively, the same path to the active site may be used if the side openings of the ATCase are small pores in the *in vivo* situation of the enzyme. Channeling of the CP could be possible with the CPSases individually aligning their active sites with the pores, as has been suggested for the DHOase–ATCase complex from *A. aeolicus* (Zhang *et al.*, 2009).

The structure of the holoenzyme and additional biochemical studies concerning enzymatic carbamoylation and the molecular organization of the pyrimidine pathway in *M. jannaschii* will provide further insight into these points.

4. Conclusions

We have grown crystals of the catalytic subunit of *M. jannaschii* ATCase in a hexagonal crystal form in the presence of the regulatory subunits. This is the first time that we have obtained crystals of the catalytic subunit that contain only one catalytic chain in the asymmetric unit; all other crystal forms contained multiple chains. The symmetry-related chains form the staggered dimer of trimers observed in other known structures of the catalytic subunit. The structure suggests two possible paths that CP may follow to reach the active site. One path is through the central channel and it is possible that the central channel is involved in channeling CP to the active site. The second path is through a CP-binding site near Arg164 and it is possible that CP may in part diffuse to the active site from the surrounding medium through this site at 310 K. Additional biochemical studies concerning enzymatic carbamoylation and the molecular organization of the pyrimidine pathway in *M. jannaschii* will provide further insight into these points.

This work was supported in part by grant GM071512 (JV) from the National Institutes of Health and by a Faculty Research Development award (JV) from Cleveland State University. Data were measured on beamline X12C of the National Synchrotron Light Source. Financial support comes principally from the Offices of Biological and Environmental Research and of Basic Energy Sciences of the US Department of Energy and from the National Center for Research Resources of the National Institutes of Health (grant No. P41RR012408). The computations were supported in part by an allocation of computing time from the Ohio Supercomputer Center. We thank undergraduate student Nermina Covic (Cleveland State University) for the lysogenization of the ATCase-deficient derivative of *E. coli* C600 cells, Dr E. Kantrowitz (Boston College, Boston, Massachusetts, USA) for providing the EK1911 strain and plasmid pEK407 that were used for this study, Dr R. Cunin (Vrije Universiteit Brussel, Brussels, Belgium) for providing the ATCase-deficient derivative of *E. coli* strain C600 and Dr S. Sandler (University of Massachusetts at Amherst, Amherst, Massachusetts, USA) for the PSJS1240 plasmid. This paper is dedicated to the memory of Dolly Vitali.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Ahuja, A., Purcarea, C., Ebert, R., Sadecki, S., Guy, H. I. & Evans, D. R. (2004). *J. Biol. Chem.* **279**, 53136–53144.
Allen, C. M. & Jones, M. E. (1964). *Biochemistry*, **3**, 1238–1247.
Allewell, N. M. (1989). *Annu. Rev. Biophys. Chem*, **18**, 71–92.
Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.
Beernink, P. T., Endrizzi, J. A., Alber, T. & Schachman, H. K. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 5388–5393.
Belkaïd, M., Penverne, B. & Hervé, G. (1988). *Arch. Biochem. Biophys.* **262**, 171–180.
Brabson, J. S., Maurizi, M. R. & Switzer, R. L. (1985). *Methods Enzymol.* **113**, 627–635.
Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B. & Woods, R. J. (2005). *J. Comput. Chem.* **26**, 1668–1688.
Christopherson, R. I. & Jones, M. E. (1980). *J. Biol. Chem.* **255**, 11381–11395.
Cohen, N. S., Cheung, C.-W., Sijuwade, E. & Raijman, L. (1992). *Biochem. J.* **282**, 173–180.
Coleman, P. F., Suttle, D. P. & Stark, G. R. (1977). *J. Biol. Chem.* **252**, 6379–6385.
Copley, R. R. & Barton, G. J. (1994). *J. Mol. Biol.* **242**, 321–329.
De Vos, D., Van Petegem, F., Remaut, H., Legrain, C., Glansdorff, N. & Van Beeumen, J. J. (2004). *J. Mol. Biol.* **339**, 887–900.
Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. (2004). *Nucleic Acids Res.* **32**, W665–W667.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Endrizzi, J. A., Beernink, P. T., Alber, T. & Schachman, H. K. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 5077–5082.
England, P., Leconte, C., Tauc, P. & Hervé, G. (1994). *Eur. J. Biochem.* **222**, 775–780.
Evans, D. R. & Guy, H. I. (2004). *J. Biol. Chem.* **279**, 33035–33038.
Hack, E. S., Vorobyova, T., Sakash, J. B., West, J. M., Macol, C. P., Hervé, G., Williams, M. K. & Kantrowitz, E. R. (2000). *J. Biol. Chem.* **275**, 15820–15827.
Hervé, G. (1989). *Allosteric Enzymes*, edited by G. Hervé, pp. 61–79. Boca Raton: CRC Press.
Hsin, K., Sheng, Y., Harding, M. M., Taylor, P. & Walkinshaw, M. D. (2008). *J. Appl. Cryst.* **41**, 963–968.
Irvine, H. S., Shaw, S. M., Paton, A. & Carrey, E. A. (1997). *Eur. J. Biochem.* **247**, 1063–1073.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jin, L., Stec, B., Lipscomb, W. N. & Kantrowitz, E. R. (1999). *Proteins*, **37**, 729–742.

Jones, M. E., Spector, L. & Lipmann, F. (1955). *J. Am. Chem. Soc.* **77**, 819–820.
Ke, H., Lipscomb, W. N., Cho, Y. & Honzatko, R. B. (1988). *J. Mol. Biol.* **204**, 725–747.
Kim, R., Sandler, S. J., Goldman, S., Yokota, H., Clark, A. J. & Kim, S.-H. (1998). *Biotechnol. Lett.* **20**, 207–210.
Kleywegt, G. J. (1996). *Acta Cryst.* D**52**, 842–857.
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.
Legrain, C., Demarez, M., Glansdorff, N. & Piérard, A. (1995). *Microbiology*, **141**, 1093–1099.
Lipscomb, W. N. (1992). *Proceedings of the Robert A. Welch Foundation Conference on Chemical Research. XXXVI. Regulation of Proteins by Ligands*, pp. 103–143. Robert A. Welch Foundation, Houston, Texas, USA.
Lipscomb, W. N. (1994). *Adv. Enzymol.* **68**, 67–152.
Lue, P. F. & Kaplan, J. G. (1970). *Biochim. Biophys. Acta*, **220**, 365–372.
Makoff, A. J. & Radford, A. (1978). *Microbiol. Rev.* **42**, 307–328.
Mally, M. I., Grayson, D. R. & Evans, D. R. (1980). *J. Biol. Chem.* **255**, 11372–11380.
Martin, P. D., Purcarea, C., Zhang, P., Vaishnav, A., Sadecki, S., Guy-Evans, H. I., Evans, D. R. & Edwards, B. F. (2005). *J. Mol. Biol.* **348**, 535–547.
Massant, J. & Glansdorff, N. (2005). *Archaea*, **1**, 365–373.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McDonald, I. K. & Thornton, J. M. (1994). *J. Mol. Biol.* **238**, 777–793.
Mendes, K. R. & Kantrowitz, E. R. (2010). *J. Mol. Biol.* **401**, 940–948.
Mori, M. & Tatibana, M. (1978). *Eur. J. Biochem.* **86**, 381–388.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Penverne, B., Belkaïd, M. & Hervé, G. (1994). *Arch. Biochem. Biophys.* **309**, 85–93.
Petrek, M., Otyepka, M., Banás, P., Kosinová, P., Koca, J. & Damborský, J. (2006). *BMC Bioinformatics*, **7**, 316.
Pražnikar, J., Afonine, P. V., Gunčar, G., Adams, P. D. & Turk, D. (2009). *Acta Cryst.* D**65**, 921–931.
Purcarea, C., Ahuja, A., Lu, T., Kovari, L., Guy, H. I. & Evans, D. R. (2003). *J. Biol. Chem.* **278**, 52924–52934.
Purcarea, C., Evans, D. R. & Hervé, G. (1999). *J. Biol. Chem.* **274**, 6122–6129.
Ramón-Maiques, S., Marina, A., Guinot, A., Gil-Ortiz, F., Uriarte, M., Fita, I. & Rubio, V. (2010). *J. Mol. Biol.* **397**, 1261–1275.
Stevens, R. C., Gouaux, J. E. & Lipscomb, W. N. (1990*a*). *Biochemistry*, **29**, 7691–7701.
Stevens, R. C., Gouaux, J. E. & Lipscomb, W. N. (1990*b*). *Biochemistry*, **29**, 11146.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Van Boxstael, S., Cunin, R., Khan, S. & Maes, D. (2003). *J. Mol. Biol.* **326**, 203–216.
Van de Casteele, M., Legrain, C., Desmarez, L., Chen, P. G., Piérard, A. & Glansdorff, N. (1997). *Comp. Biochem. Physiol. A Physiol.* **118**, 463–473.
Vickrey, J. F., Hervé, G. & Evans, D. R. (2002). *J. Biol. Chem.* **277**, 24490–24498.
Vitali, J. & Colaneri, M. J. (2008). *Acta Cryst.* F**64**, 776–780.
Vitali, J., Colaneri, M. J. & Kantrowitz, E. R. (2008). *Proteins*, **71**, 1324–1334.
Wanders, R. J., Van Roermund, C. W. & Meijer, A. J. (1984). *Eur. J. Biochem.* **142**, 247–254.
Wang, J. (2010). *Acta Cryst.* D**66**, 988–1000.
Wang, J., Stieglitz, K. A., Cardia, J. P. & Kantrowitz, E. R. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 8881–8886.
Wang, Q., Xia, J., Guallar, V., Krilov, G. & Kantrowitz, E. R. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 16918–16923.
Wiley, D. C. & Lipscomb, W. N. (1968). *Nature (London)*, **218**, 1119–1121.
Williams, L. G., Bernhardt, S. & Davis, R. H. (1970). *Biochemistry*, **9**, 4329–4335.
Williams, L. G., Bernhardt, S. A. & Davis, R. H. (1971). *J. Biol. Chem.* **246**, 973–978.
Williams, M. K., Stec, B. & Kantrowitz, E. R. (1998). *J. Mol. Biol.* **281**, 121–134.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Zhang, P., Martin, P. D., Purcarea, C., Vaishnav, A., Brunzelle, J. S., Fernando, R., Guy-Evans, H. I., Evans, D. R. & Edwards, B. F. (2009). *Biochemistry*, **48**, 766–778.