Crystallization and structure determination of the catalytic trimer of *Methanococcus jannaschii* aspartate transcarbamoylase

Jacqueline Vitali,[a] Tatyana Vorobyova,[a] Gordon Webster[b] and Evan R. Kantrowitz[a]*

[a]Department of Chemistry, Merkert Chemistry Center, Boston College, Chestnut Hill, MA 02467, USA, and [b]Division of Experimental Medicine, Beth Israel Deaconess Medical Center, Harvard Institutes of Medicine, 4 Blackfan Circle, Boston, MA 02115, USA

Correspondence e-mail: evan.kantrowitz@bc.edu

Aspartate transcarbamoylase (ATCase) catalyzes the first step in the pyrimidine biosynthetic pathway, the reaction between carbamoyl phosphate and L-aspartate to form *N*-carbamoyl-L-aspartate and phosphate. The structural analysis of the ATCase catalytic trimer from *Methanococcus jannaschii*, a unicellular thermophilic archaea-bacterium, has been undertaken in order to gain insight into the structural features that are responsible for the thermostability of the enzyme. As a first step, the catalytic trimer was crystallized in space group *R*32, with unit-cell parameters $a = b = 265.3$, $c = 195.5$ Å and two trimers in the asymmetric unit. Its structure was determined using molecular replacement and Patterson methods. In general, structures containing multiple copies of molecules in the asymmetric unit are difficult to determine. In this case, the two trimers in the asymmetric unit are parallel to each other and use of the Patterson function greatly simplified the structure solution.

1. Introduction

The enzyme aspartate transcarbamoylase (ATCase) catalyzes the first step in the pyrimidine biosynthetic pathway, the reaction between carbamoyl phosphate and L-aspartate to form *N*-carbamoyl-L-aspartate and phosphate (Jones *et al.*, 1955; Reichard & Hanshoff, 1956). In many organisms, this enzyme is involved in the control of this pathway by feedback inhibition.

The aspartate transcarbamoylase from *Escherichia coli* is a dodecameric enzyme and is composed of two types of subunits (Gerhart & Schachman, 1965). The two larger or catalytic trimers are each composed of three identical 33 kDa polypeptide chains, while the three smaller or regulatory dimers are each composed of two identical 17 kDa polypeptide chains. ATCase exists in different quaternary structures in different organisms. For example, in *Bacillus subtilis* the enzyme exists as a trimer of catalytic chains without any regulatory chains (Lerner & Switzer, 1986).

The X-ray structure of the dodecameric enzyme from *E. coli* has been determined in both the absence and presence of substrates/inhibitors (Kim *et al.*, 1987; Ke *et al.*, 1988; Gouaux & Lipscomb, 1988, 1990; Gouaux *et al.*, 1990; Jin *et al.*, 1999). Furthermore, the structure of the catalytic trimer from *E. coli* in the absence of the regulatory subunits has been determined (Beerbink *et al.*, 1999) and a low-resolution α-carbon trace of the *B. subtilis* enzyme is also available (Stevens *et al.*, 1991).

No X-ray structural data are yet available for a thermophilic ATCase. X-ray structural data on such an enzyme would be important because they would provide insight into which structural features must be changed to provide an enzyme with enhanced thermostability. These data may also provide an understanding of how a thermophilic organism can survive when one of the substrates for ATCase (carbamoyl phosphate) is thermally unstable.

The DNA sequence of the unicellular thermophilic archaeabacterium *M. jannaschii* (Bult *et al.*, 1996) revealed genes that are homologous to the *E. coli pyrB* and *pyrI* genes coding for the catalytic and regulatory chains of ATCase. These genes were inserted into expression systems that yielded substantial amounts of both gene products (Hack *et al.*, 2000) and a purification scheme for the catalytic trimer of ATCase was devised. Studies with the *M. jannaschii* cell-free extract showed that the ATCase has a similar molecular weight to that of the *E. coli* holoenzyme (Hack *et al.*, 2000). Furthermore, the *M. jannaschii pyrB* gene product exists as a trimer, has catalytic activity and exhibits hyperbolic kinetics. The *M. jannaschii pyrI* gene product exists as a dimer with no catalytic activity. The *M. jannaschii pyrB* and *pyrI* gene products have 47 and 35% sequence identity with the catalytic and regulatory chains of *E. coli* ATCase, respectively.

We have crystallized the *M. jannaschii* catalytic trimer in space group *R*32 with two trimers in the asymmetric unit. The crystals

crystallization papers

diffract to 3.5 Å resolution. The orientations and positions of the two trimers in the cell were obtained using molecular-replacement and Patterson methods. In general, structures containing multiple copies of molecules in the asymmetric unit are difficult to determine. This was a special case with two parallel molecules in the asymmetric unit and use of the Patterson function simplified the structure solution. We are currently testing cryo-cooling conditions in order to obtain higher resolution data.

2. Protein preparation

The expression system and purification scheme described by Hack *et al.* (2000) were used. The expression system consisted of *E. coli* strain EK1594 (Sakash & Kantrowitz, 1998), which has an inducible gene for T7 RNA polymerase and a deletion in the *pyrBI* region on the chromosome, and the plasmid pEK406, which was derived from pET23a and contained the *M. jannaschii pyrB* gene under the control of the T7 promoter. In some preparations, the plasmid pSJS1240 (Kim *et al.*, 1998) was also co-transformed, as the tRNAs that it codes for improve the expression of the protein. Typically, 8 l of M9 media (Maniatis *et al.*, 1982) supplemented with 0.5% cas-amino acid, 100 mg ml^{-1} ampicillin and 100 mg ml^{-1} spectinomycin (when the pSJS1240 plasmid was used) were inoculated with the transformed strain. The cells were grown at 310 K for 4 h after induction with 0.4 mM isopropyl-β-D-thiogalactoside (IPTG) and were subsequently pelleted by centrifugation. The cell pellet was resuspended in 100 mM Tris–HCl buffer pH 9.2 and the cells were lysed by sonication. After centrifugation of the lysed cells, the cell-free supernatant was brought to 30% saturation in ammonium sulfate and heated at 363 K for 15 min. After heating and subsequent centrifugation to remove heat-denatured *E. coli* proteins, the supernatant was further purified with a Q-Sepharose Fast Flow

anion-exchange column and a phenyl Sepharose column. The quality of the purification was evaluated with SDS–PAGE. Aspartate transcarbamoylase activity was measured by the calorimetric method (Pastra-Landis *et al.*, 1981) and by the pH-stat method (Wu & Hammes, 1973).

3. Crystallization

The purified protein was dialyzed against 40 mM KH$_2$PO$_4$, 0.2 mM EDTA, 2 mM 2-mercaptoethanol pH 7.0 and concentrated to 15 mg ml^{-1} using a Centricon-50 concentrator (Amicon). Crystals were grown at 295 K by the sitting-drop method from reservoirs containing 2.0 M ammonium sulfate and 5% 2-propanol. The drops consisted of equal volumes of reservoir and protein solution. The crystal used for this study was cut from a larger cluster. It was shown to be rhombohedral, space group $R32$, with unit-cell parameters $a = b = 265.3$, $c = 195.5$ Å and two catalytic trimers in the asymmetric unit, $V_M = 3.2$ Å^3 Da^{-1} (Matthews, 1968).

4. Data measurement

X-ray diffraction data were measured at the Macromolecular Crystallography Laboratory of the Division of Experimental Medicine at Beth Israel Deaconess Medical Center using an R-AXIS IV image plate installed on a Rigaku RU300H rotating-anode X-ray generator operated at 50 kV and 100 mA with Molecular Structure Corporation mirrors. Oscillations were measured for 30 min each in 1.0° intervals of φ. The crystal-to-detector distance was 210.0 mm. The data were processed with the *HKL* software package (Otwinowski & Minor, 1997). A summary of the X-ray data-collection statistics is presented in Table 1. The crystal decayed by 24% after 16.5 h exposure to X-rays and by 37% over the course of the measurements.

5. Structure solution

The structure was solved using molecular-replacement and Patterson methods. The molecular-replacement calculations were carried out with *X-PLOR* 3.1 (Brünger, 1988, 1990, 1993a), data with $F > 4\sigma$ in the resolution range 4.0–15.0 Å and a shell of integration in the rotation function of 5.0–24.0 Å. A native Patterson function was computed with the package *PHASES* (Furey & Swaminathan, 1997) using data to 3.5 Å resolution with $F > 4\sigma$.

Homology modeling (Hack *et al.*, 2000) of the *M. jannaschii* ATCase *pyrB* gene product based on X-ray structural data of the catalytic chain of the *E. coli* enzyme (Jin *et al.*, 1999) using *Swiss Model* (Peitsch, 1995) suggested similar overall structures for the catalytic subunits of *E. coli* and *M. jannaschii*, with major discrepancies occurring on the surface loops in the form of insertions and deletions. The amino-acid sequences were 47% identical and 67% similar for the two catalytic chains. Accordingly, the search model used for molecular replacement consisted of the isolated ATCase catalytic trimer from *E. coli* whose structure has been determined and refined to 1.9 Å resolution (Beerbink *et al.*, 1999). In this model, all surface loops which were expected to differ between the two were deleted. Additionally, all residues were replaced with Ala unless they were common to the two enzymes.

The self-rotation function did not contain peaks other than the origin peak and those corresponding to the crystallographic symmetry operations, suggesting that the two trimers should be parallel in the crystal and that the non-crystallographic threefold axes relating monomers in each trimer should be parallel to the crystallographic threefold.

Both cross-rotation function and PC refinement strongly indicated an orientation for the catalytic trimers in the crystal. It corresponded to the highest peak in both searches; the signal-to-noise ratio, taken as the ratio of the top peak to the next highest peak, was 1.08 for the rotation function and 2.19 for the PC refinement. In this orientation, the trimers had their non-crystallographic threefold axes parallel to the crystallographic threefold, in agreement with the results of the self-rotation function.

The top peak of the translational search was taken to correspond to one of the trimers at (0.17, 0.33, 0.19). It had a correlation coefficient of 0.40 and a signal-to-noise ratio of 1.03. If the two trimers were parallel in the crystal, as suggested by the rotational searches, then the native Patterson would contain a very strong peak corresponding to the trimer–trimer vector. It turned out that the Patterson function contained two strong peaks, one at (0.67, 0.33, 0.81) with height approximately one-fourth of the height of the origin peak and a second at (0.33, 0.17, 0.14) with height approximately one-sixth of the height of the origin peak. The possible positions of the second trimer would then be at (0.83, 0.67, 1.00) or at (0.50, 0.50, 0.33). Position (0.83, 0.67, 1.00) is equivalent to position (0.50,

0.50, 0.33) but with the trimer rotated around its local non-crystallographic threefold axis. Accordingly, trimer 2 was placed at (0.50, 0.50, 0.33), keeping in mind that the trimer residues deviating from the non-crystallographic threefold may need to be refitted at later stages of the analysis. The position also corresponds to the third independent peak of the translational search; the correlation coefficient is 0.36 and the signal-to-noise ratio is 0.94. The correlation coefficient when both trimers were included in the model was 0.52.

An initial refinement of the model was carried out using *X-PLOR* 3.1. The data was divided in two sets: a working set composed of 90% of the data and a test set composed of the remaining 10% for cross-validation purposes (Brünger, 1993*b*). Of the residues included in the molecular-replacement calculations, those that deviated from the non-crystallographic threefold symmetry were not included in this refinement. A total of 246 of the 306 residues per monomer were refined and, of them, 113 were treated as Ala. Firstly, a rigid-body refinement was carried out. This started with the two trimers, was followed by the six monomers and finally by the carbamoyl phosphate and aspartate domains of each monomer. The data used in the rigid-body refinement had $F > 2\sigma$ and the resolution ranges were 20.0–9.0, 10.0–8.0, 10.0–6.0, 10.0–4.0 and 7.0–3.5 Å, in sequence. The lower resolution refinements were carried out to ensure
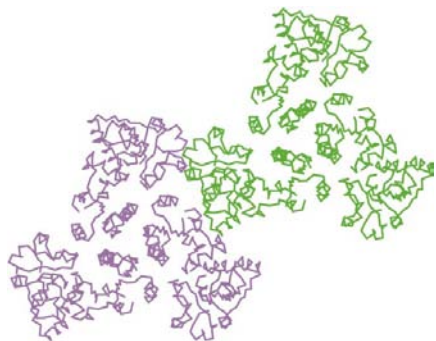

Figure 1
A view of the C^α backbone of the two catalytic trimers in the asymmetric unit looking down the non-crystallographic threefold axes. These are at the center of each trimer and are parallel to the crystallographic threefold axis. As discussed in the text, several loops and residues deviating from the non-crystallographic threefold symmetry in the trimers are not included in the model and is the reason that there are several discontinuities in the polypeptide chains.

convergence to the true solution. Subsequently, the refinement employed molecular dynamics. The slow-cool annealing protocol was used for simulating annealing with a dynamics temperature starting at 500 K. Furthermore, because of the low resolution of the analysis only an overall temperature factor was refined. The refinement was carried out with data in the resolution range 3.5–7.0 Å, $F > 2\sigma$, weights set to half the weights recommended by the check stage and the protein parameters of Engh & Huber (1991). During the simulated-annealing refinement, the atomic coordinates of the six monomers were restrained by the non-crystallographic symmetry using a force constant of 300 kcal mol^{-1} Å^{-2} (1 kcal = 4.184 kJ). The final R and $R_{\rm free}$ values after rigid-body refinement were 0.48 and 0.50, respectively. The corresponding figures after simulated annealing were 0.43 and 0.48, respectively. A view of the C^α backbone of the two catalytic trimers at the present stage of the analysis is given in Fig. 1.

We are confident about the correctness of the solution. Firstly, the parallel orientation of the two trimers in the crystal that was suggested by the rotational searches was supported by the presence of a very strong peak in the native Patterson function. Secondly, the correlation coefficients in the molecular-replacement calculations were high. Additionally, even though the signal-to-noise ratios in the translation function were somewhat low, the solutions were consistent with the trimer–trimer vector obtained from the Patterson function. Thirdly, the R and $R_{\rm free}$ values are reasonable for this stage of the analysis. Fourthly, the packing is reasonable with no short contacts between the various trimers in the unit cell. Furthermore, when only one trimer was included in the model, the $2F_o - F_c$ and $F_o - F_c$ electron-density maps had density for the second trimer. Finally, there is electron density for parts missing in the current model such as several side chains and loops. We are currently testing cryo-cooling conditions in order to obtain higher resolution data.

References

Beerbink, P. T., Endrizzi, J. A., Alber, T. & Schachman, H. K. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 5388–5393.
Brünger, A. T. (1988). *J. Mol. Biol.* **203**, 803–816.
Brünger, A. T. (1990). *Acta Cryst.* A**46**, 46–57.
Brünger, A. T. (1993*a*). *X-PLOR Manual, Version 3.1*. Yale University, New Haven, CT, USA.
Brünger, A. T. (1993*b*). *Acta Cryst.* D**49**, 24–36.
Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J.-F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F. & Weinstock, K. (1996). *Science*, **273**, 1058–1073.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Furey, W. & Swaminathan, S. (1997). *Methods Enzymol.* **277**, 590–620.
Gerhart, J. C. & Schachman, H. K. (1965). *Biochemistry*, **4**, 1054–1062.
Gouaux, J. E. & Lipscomb, W. N. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 4205–4208.
Gouaux, J. E. & Lipscomb, W. N. (1990). *Biochemistry*, **29**, 389–402.
Gouaux, J. E., Stevens, R. C. & Lipscomb, W. N. (1990). *Biochemistry*, **29**, 7702–7715.
Hack, E. S., Sakash, J. B., Vorobyova, T., West, J. M., Macol, C. P., Herve, G., Williams, M. K. & Kantrowitz, E. R. (2000). *J. Biol. Chem.* **275**, 15820–15827.
Jin, L., Stec, B., Lipscomb, W. N. & Kantrowitz, E. R. (1999). *Proteins Struct. Funct. Genet.* **37**, 729–742.
Jones, M. E., Spector, L. & Lipmann, F. (1955). *J. Am. Chem. Soc.* **77**, 819–820.
Ke, H.-K., Lipscomb, W. N., Cho, Y. & Honzatko, R. B. (1988). *J. Mol. Biol.* **204**, 725-747.
Kim, K. H., Pan, Z., Honzatko, R. B., Ke, H.-M. & Lipscomb, W. N. (1987). *J. Mol. Biol.* **196**, 853–875.
Kim, R., Sandler, S. J., Goldman, S., Yokota, H., Clark, A. J. & Kim, S.-H. (1998). *Biotech. Lett.* **20**, 207–210.
Lerner, C. G. & Switzer, R. L. (1986). *J. Biol. Chem.* **261**, 11156–11165.
Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual*. New York: Cold Spring Harbor Laboratory.
Matthews, B. W. (1968). *J. Mol. Biol.* **133**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pastra-Landis, S. C., Foote, J. & Kantrowitz, E. R. (1981). *Anal. Biochem.* **118**, 358–363.
Peitsch, M. C. (1995). *Biotechnology*, **13**, 658–660.
Reichard, P. & Hanshoff, G. (1956). *Acta Chem. Scand.* **10**, 548–560.
Sakash, J. B. & Kantrowitz, E. R. (1998). *Biochemistry*, **37**, 281–288.
Stevens, R. C., Reinisch, K. M. & Lipscomb, W. N. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 6087–6091.
Wu, C. W. & Hammes, G. G. (1973). *Biochemistry*, **12**, 1400–1408.