

Atomic structure at 2.5 Å resolution of uridine phosphorylase from *E. coli* as refined in the monoclinic crystal lattice

E.Yu. Morgunova^a, A.M. Mikhailov^{a,*}, A.N. Popov^a, E.V. Blagova^a, E.A. Smirnova^a,
B.K. Vainshtein^a, Ch. Mao^b, Sh.R. Armstrong^b, S.E. Ealick^{b,*}, A.A. Komissarov^c,
E.V. Linkova^c, A.A. Burlakova^c, A.S. Mironov^c, V.G. Debabov^c

^aInstitute of Crystallography of Russian Academy of Sciences, Leninsky Prospekt 59, Moscow 117333, Russian Federation

^bSection of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York 14853, USA

^cState Scientific Centre of Russian GNIIGENETIKA, 1st Dorozhny 1, Moscow 113545, Russian Federation

Received 18 april 1995

Abstract Uridine phosphorylase from *E. coli* (Upase) has been crystallized using vapor diffusion technique in a new monoclinic crystal form. The structure was determined by the molecular replacement method at 2.5 Å resolution. The coordinates of the trigonal crystal form were used as a starting model and the refinement by the program XPLOR led to the R-factor of 18.6%. The amino acid fold of the protein was found to be the same as that in the trigonal crystals. The positions of flexible regions were refined. The conclusion about the involvement in the active site is in good agreement with the results of the biochemical experiments.

Key words: Uridine phosphorylase; *E. coli*; X-ray structure

1. Introduction

Uridine phosphorylase (EC 2.4.2.3; Upase) from *E. coli* catalyzes the reversible phosphorolysis of uridine with the formation of riboso-1-phosphate and uracil [1,2]. The protein is involved in degradation of pyrimidine nucleosides and their utilization as carbon and energy sources in *E. coli* cells. Upase along with purine nucleoside phosphorylase (PNP) and thymidine phosphorylase, belongs to the class of nucleoside phosphorylases. Uridine phosphorylase has been identified as the enzyme which is responsible for the cleavage of some pyrimidine nucleoside analogs possessing antitumor activity. Conversion of 5-fluorouridine (FUR) and 5-fluorodeoxyuridine (FdUR) to 5-fluorouracil (5FU) by Upase required the usage of high concentration of 5FU for tumor treatment. This resulted in a strong toxicity by the 5FU and its metabolism products. Thus, the extensive investigation of Upase is connected with the search for new selective inhibitors which might be used as drugs for therapy of some solid tumors [3–5].

In addition to development of specific inhibitors, Upase is of interest as a tool in the laboratory synthesis of nucleoside analogs for certain drugs. Modification of the structure of Upase by site-directed mutagenesis might increase efficiency through broadening of specificity and enhancing stability.

The enzyme molecule is a hexamer which consists of six identical subunits with a total molecular weight 165 kDa [6,7]. The primary structure is known and includes 253 amino acid residues [8]. This amino acid sequence shows a high degree of

homology to the sequence of *E. coli* PNP – the sequence similarity being 49% and the sequence identity 25% [9]. However the sequence homology between Upase and other phosphorylases is very poor. The comparison of the three-dimensional structures of Upase and other phosphorylases has shown that the structure of Upase and human PNP are rather similar but there is no any similarity between Upase and thymidine phosphorylase structures. The study of *E. coli* PNP is in progress and obviously its structure must be similar to that of Upase. The Upase three-dimensional structure has been solved at 3.0 Å resolution using the data from trigonal crystals described previously [9]. This structure contains some disordered regions and little is known about groups involved in catalysis and/or binding at the Upase active site [2,10–12].

In this paper we report results of an X-ray investigation of a monoclinic form at 2.5 Å resolution and the interpretation of the data of the selective chemical modification of the essential amino acid residues.

2. Materials and methods

2.1. Crystallization

The uridine phosphorylase was crystallized in different crystal forms by the vapor diffusion technique at room temperature [13,14]. The monoclinic form was obtained in sitting drops of 0.05 M Tris-HCl buffer at pH 7.3 containing 10–12 mg/ml of the protein and 4–6% PEG (mol.wt. 4,000). The equilibrium solution consisted of 0.1 M Tris-mal/NaOH pH 5.91–5.96, 20–25% PEG (mol.wt. 4,000) and 0.04% sodium azide. The crystals reached the size of 0.7 × 0.3 × 0.5 mm after 4–6 weeks. These crystals belong to monoclinic space group P2₁ with cell dimensions a = 92.6 Å, b = 98.8 Å, c = 93.7 Å, z = 2 and diffract at least to 2.5 Å resolution.

2.2. Data collection

Determination of the space group, cell dimensions and data collection were performed using a diffractometer with a two-dimensional detector KARD-6 (Inst. of Crystallography, Moscow, Russian Federation) based on a five-circle goniometer [15]. The cell dimensions and the crystal orientation were determined from the analysis of the averaged difference vectors with the refinement of the coordinates of the observed reflections. The data set to 2.5 Å resolution was collected. One fourth of the whole diffraction pattern was simultaneously recorded during the crystal rotation around the ω axis with the additional measurement of the blind region after a crystal rotation of 90° around the χ axis. A crystal-detector distance of 586 mm and a scanning rate of 0.14 degrees/min were used to collect data set in 3 days. A summary of the data set collected is shown in Table 1.

2.3. Structure determination

A self-rotation function was calculated by the MERLOT program [16] using data between 10–4 Å. The three-fold non-crystallographic

*Corresponding authors. E-mail: amm@biostr.crystal.msk.ru or see3@cornell.edu

axis was located at an angle of 5.1° to the 'b' axis and 87.5° to the 'a' axis. The Matthew's number calculated for one hexamer per asymmetric unit was 2.24, which corresponds to a solvent content of 47.1%.

The three-dimensional Upside structure was solved at 2.5 Å resolution by molecular replacement method using the AMoRe program package [17] with X-ray diffraction data collected from a single crystal. The previously described structure of the trigonal crystal was used as a starting model at 5.5 Å resolution.

The rotation search resulted in unambiguous solution at $\alpha = 83.5^\circ$; $\beta = 89.5^\circ$ and $\gamma = 18.4^\circ$. The subsequent translation search produced a clear solution at crystallographic coordinates 5.45 Å along x and 32.6 Å along z. Starting from the coordinates calculated from the translation search, rigid body refinements were carried out with reflections in the resolution range 10–5.5 Å that met $F > 2\sigma$ criteria. Rigid-body refinement over the whole hexamer was performed. These refinements produced an R-value of 48.3% that is significantly better than random. The model was then subjected to rigid-body minimization with X-PLOR [18]. After 40 cycles of minimization the model returned to a position with the resulting R-factor = 40.9%.

2.4. Refinement

The crystallographic refinement was carried out without restraint of the non-crystallographic symmetry by the X-PLOR v.3.1 program. The protocol used consisted of a molecular dynamics simulation where the temperature was raised to 3000 K and then slowly cooled to 300 K in steps, followed by energy minimization [19]. The refinement cycles were alternated with the manual rebuilding of the model carried out by the CHAIN (John Sack) v.4.5. The 11,292 nonhydrogen atoms were included in the refinement. After 15 cycles of full refinements and four manual interventions the model was refined to R-factor 18.6% against 42,401 reflections between 6–2.5 Å resolution with overall rms deviations from bond length of 0.012 Å and bond angles 2.095° .

2.5. Water molecules

Water molecules were picked up from the Fo-Fc map and incorporated into the refinement. The water molecules, which had relatively high B-factors and weak density were removed and new water molecules from the Fo-Fc map were added. This procedure was repeated three times until there were no peaks stronger than 3σ in the Fo-Fc map. 150 water molecules were picked in the structure.

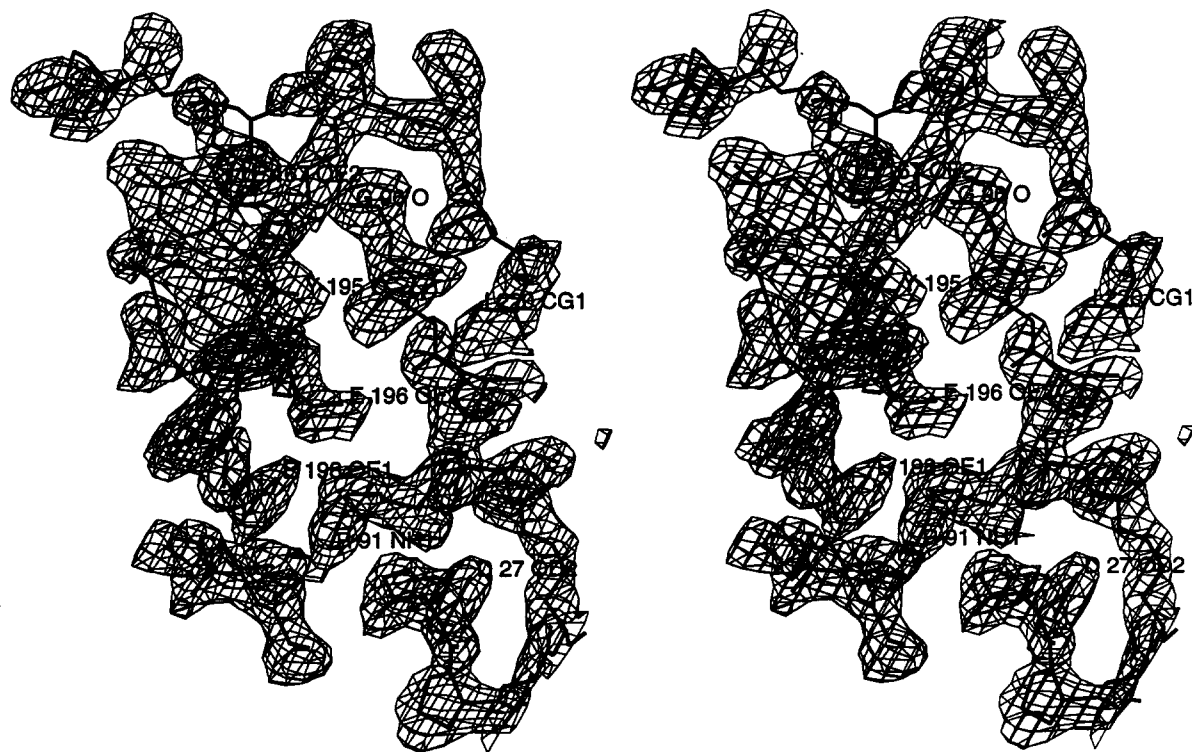
3. Results and discussion

3.1. Electron density map

A 2Fo-Fc electron density map (Fig. 1) for the refined model in monoclinic crystals showed well-defined electron density except the N-terminal region of all subunits and the region of flexible loops (residues 162–180 and 228–234 (Table 2)). The electron density for the five residues at the N-terminus was very similar for all subunits and consists of distinct peaks mostly from the main chain atoms. The main chain of the loop regions was traced at 3.0 Å resolution because the electron density at that point was rather good without any breaks. However the increase in resolution to 2.5 Å showed that these regions were very flexible and we could see several breaks of the electron density. The rebuilding of the side chains in these regions was quite complicated and the individual B-factors were high.

3.2. Packing

The crystallographic asymmetric unit contains a uridine phosphorylase hexamer (Fig. 2) and therefore the structure determination gives independent results for six subunits (Fig. 3a,b). The rms deviations between subunits are presented in Table 3. The hexamers in the cell are packed in layers which have van der Waals contacts between each other. The majority of contacts are of a polar nature with water-mediated hydrogen bonds. Each subunit in the hexamer has a close contact with the identical one related by a 2-fold non-crystallographic axis and three dimers are packed in a hexamer by a 3-fold non-crystallographic axis with is perpendicular to the 2-fold axes. Thus, the hexamer belongs to a 32 point group symmetry. The interface region between two symmetrically (local axis) related subunits involves residues Ala⁴⁸–Glu⁴⁹, Gly⁶⁸–Ile⁶⁹,



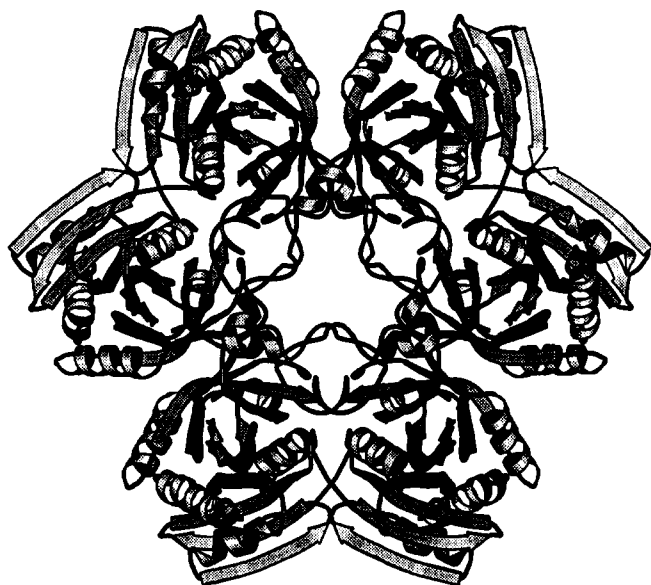


Fig. 2. Ribbon drawing of Upase hexamer in monoclinic crystal form.

Gly¹¹⁸–Ala¹²³, Asp¹⁶⁰–Tyr¹⁶³ and the second α -helix (residues 72–83). The residues in this helix are hydrophobic or ambivalent. Eleven direct hydrogen bonds have been found between symmetrically-related molecules. Moreover, ten water molecules which form the hydrogen bonds with the residues of the interface surface were found. The flexible loop (residues Tyr¹⁶³–Phe¹⁸⁰) extends into the neighboring subunit. This loop is hydrophobic and five hydrogen bonds were found between this loop and neighboring subunit.

3.3. The overall structure

The core of each subunit is essentially the same, as was reported for trigonal crystals [9]. Using more high resolution data permitted us to refine the secondary structure elements (Table 2). It was shown that the two structures show a good superimposition, except the N-terminal end flexible loops. There are no large conformation changes between the two crystal forms as well as between the six subunits in the hexamer.

3.4. Correlation between X-ray data and the results of Upase active site chemical modification

As it was shown for the trigonal crystal form, the active site is located near the interface of two subunits, and this region

includes residues 26–30, 68–70, 91–96, 162–163, 166–168, 195–198 and 220–223 of one subunit and residues 5–8 of another [9].

Recent selective chemical modification studies have indicated that Asp⁵, Tyr¹⁶⁹ and Cys¹³⁶ are probably present near or in the Upase active site. The Asp⁵ residue is located in the N-terminal region of the protein subunit which is close to the putative active site. This residue selectively reacts with Woodward's reagent K (WRK), and modification is accompanied by a complete Upase inactivation. After the reaction of the modified enzyme with hydroxylamine, the voluminous WRK residue was substituted with a hydroxamic group. Despite the addition of only two atoms (N,H) as compared with the native Upase, the resulting enzyme had only 10% of initial activity and demonstrated sufficient increase in

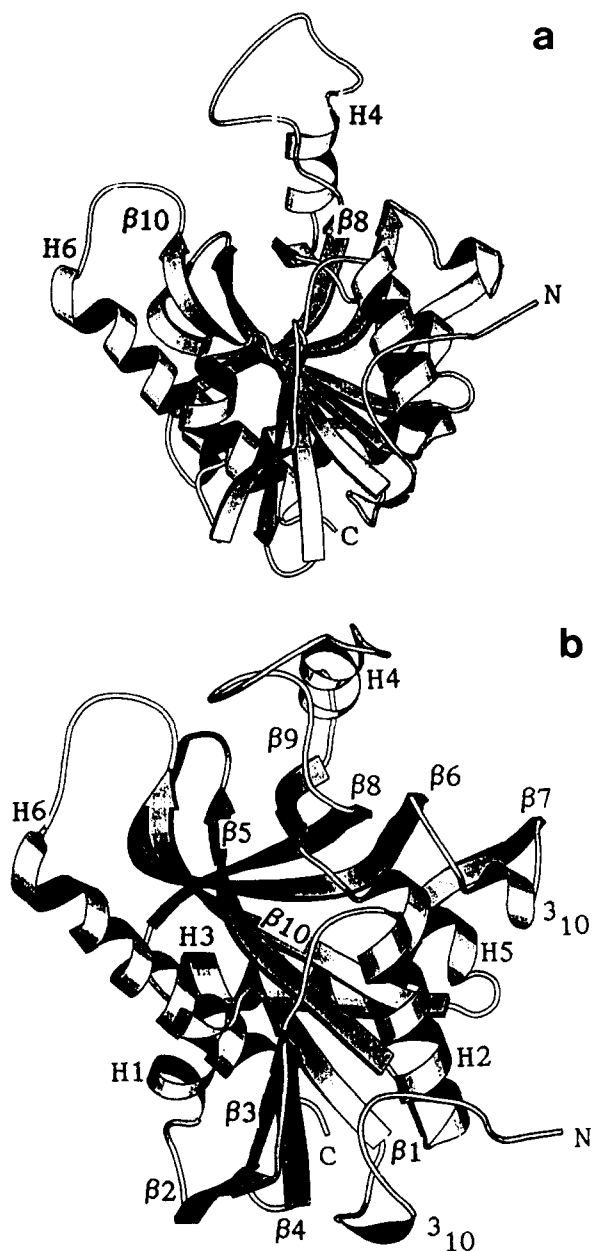


Fig. 3. Ribbon drawing of one Upase subunit in monoclinic crystal form. The secondary structure was calculated by the program Ribbons [23] and drawn by the program MOLSCRIPT [24].

Table 1
Data collection for Upase

Number of crystals	1
Resolution range (Å)	40–2.5
Measured reflections	116,050
Unique reflections	42,403
Completeness (%)	84
R _{merge} (%)	9.7
R _{st} (%)	9.2

$R_{\text{merge}} = \frac{\sum |I_i - \langle I \rangle|}{\sum \langle I \rangle}$, $R_{\text{st}} = \frac{\sum \sigma / \sum I}{\sum I}$, where I_i is the intensity measurement for a reflection and $\langle I \rangle$ is the mean value for this reflection.

thermostability. The investigations of the influence of substrates and added products indicated that only the combined addition of phosphate and uracil binding subsites protected Upase from inactivation. These data permit the suggestion that the Asp⁵ residue is located in the active site between the uridine and phosphate binding subsites and takes part in catalysis or enzyme structure stabilization rather than in substrate or product binding.

The above suggestion is also in agreement with the results of glycosidases and glycosyltransferases studies. It was consistently shown that in these enzymes a carboxyl group mediates glycosyl transfer [20]. For example Asp⁵² residue of lysozyme stabilizes the transition state of the enzymatic reaction. It was demonstrated that substitutions of this residue for homoserine by selective chemical modification [21] or for asparagine by site-directed mutagenesis [22] led to significant enzyme inactivation (90–95%) without sufficient changes in the substrate affinity. The same results were obtained in the experiments on chemical substitution of Asp⁵ carboxylate of Upase for hydroxamate. These data also confirm the suggestion that similar to lysozyme, the loss of Upase activity after such Asp⁵ substitution was due to impaired stabilization of the translation state rather than to a change in the substrate binding ability.

Tyr¹⁶⁹ residue is located in the flexible loop (Fig. 3b), two

parts of which are involved in the active site [9]. Selective nitration of Tyr¹⁶⁹ phenolic ring with tetranitromethane was accompanied by about 90% decrease in the Upase activity and a shift of pH-optimum of activity to the acidic region. Comparison of inactivation rates with diethylpyrocarbonate for native and nitrated enzymes showed a possible interaction of this residue with an essential histidine [10]. It could be supposed, that phenolic group of this residue preserves the orientation and corresponding protonation state of the histidine imidazole ring.

It should be noted that there is an essential SH-group in the Upase molecule [2,11]. Despite the fact that in accordance with X-ray data all three Cys residues are located far from the putative active site, the titration of at least one of them with silver or mercury ions leads to a complete loss of the Upase activity. It was shown that Cys¹³⁶ may be selectively modified with 2,4,6-trinitrobenzene sulfonate, the modification being accompanied by enzyme inactivation. Phosphate addition would accelerate inactivation by a factor of three and shift the pK of the reacting residue to the acidic region. These data confirm the possible location of Cys¹³⁶ near the phosphate-binding subsite.

Thus, it may be concluded that the X-ray data are in a good agreement with the results of Upase selective chemical modification. Nevertheless it is obvious that additional investigations, including high resolution X-ray analysis of the enzyme-sub-

Table 2

The comparison of the secondary structure elements in monoclinic (line down) and trigonal (line up) crystals forms

1	10	20	30
M S K S D V F H L G L T K N D L Q G A T L A I V P G D P D R V E K I A			
	I-3 ₁₀ -I	I-----β1a-----II-----H1----	
	I-3 ₁₀ -I	I----β1----I	I-----H1--I
40	50	60	70
A L M D K P V K L A S H A E F T T W R A E L D G K P V I V C S T G I G			
-I	I-----β2a-----I	I-----β3a-----I	I-----β4a-----I
	I--β2--I	I-----β3-----I	I--β4-----I
80	90	100	
G P S T S I A L E E L A Q L G I A T F L R I G T T G A I Q P H I N V G			
-----H2-----I	I-----β5a-----II--β1b-I	I-----β6a--	
-----H2-----I	I-----β5-----I	I-----β6--	
110	120	130	140
D V L V T T A S V R L D G A S L H F A P L E F P A V A D F E C T T A L			
-----II-----β2b--I	I-3 ₁₀ -I	I--β3b--II-----H3--	
-----I	I-3 ₁₀ -I	I--β7-----II-----H3--	
150	160	170	
V E A A K S I G A T T H V G V T A S S D T F Y P G Q E R Y D T Y S G R			
-----I	I--β7a--II--β4b--I		
-----I	I--β8-----I		
180	190	200	210
V V R H F K G S M E E W Q A M G V M N Y E M E S A T L L T M C A S Q G			
I--H4-----I	I--β5b--I	I-----H5-----I	
I--H4-----I	I--β9--I	I-----H5-----I	
220	230	240	
L R A G M V A G V I V N R T Q Q E I P N A E T M K Q T E S H A V K I V			
I-----β8a--II--β6b-I	I-----H6-----		
I-----β10-----I	I-----H6-----		
250			
V E A A R R L L			
-----I			
-----I			

Table 3
Rms (Å) deviations* for six identical subunits in hexamer

Number of subunit	2	3	4	5	6
1	1.71	1.88	2.11	1.76	1.48
2		1.19	1.31	1.68	1.59
3			1.52	2.10	1.92
4				1.76	1.72
5					1.48

*All calculations are present for C α atoms and were performed by the program X-PLOR at R_f = 19.2% before the addition of water molecules.

strate complexes and site-direct mutagenesis are necessary to reach the molecular level understanding of enzymatic uridine phosphorolysis.

Acknowledgements: We thank Mr. George Soros for financial support. This work was partially supported by Grant MG9000 from the International Science Foundation (ISF). Work at Cornell was supported by NIH Grant GM48874 and the W.M. Keck Laboratory for Molecular Structure.

References

- [1] Leer, J.C., Hammer-Jespersen, K. and Schwartz, M. (1977) *Eur. J. Biochem.* 75, 217–224.
- [2] Vita, A., Amici, A., Cacciamani, T., Lanciotti, M. and Magni, G. (1986) *J. Biochem.* 18, 431–436.
- [3] Iigo, M., Nishikata, K., Nakajima, Y., Szinai, I., Veres, Z., Szabolcs, A. and De Clercq, E. (1990) *Biochem. Pharmacol.* 39, 1247–1253.
- [4] Veres, Z., Neszmelyi, A., Szabolcs, A. and Denes, G. (1988) *Eur. J. Biochem.* 178, 173–181.
- [5] Naguib, F.N.M., Kouni, M.H., Chu, S.H. and Cha, S. (1987) *Biochem. Pharmacol.* 36, 2195–2201.
- [6] Cook, W.J., Koszalka, G.W., Hall, W.W., Narayana, V.L. and Ealick, S.E. (1987) *J. Biol. Chem.* 262, 2852–2853.
- [7] Tsuprun, V.L., Tagunova, I.V., Linkova, E.V. and Mironov, A.S. (1991) *Biokhimiya* 56, 930–934 (Russian).
- [8] Walton, L., Richards, C.A. and Elwell, L.P. (1989) *Nucleic Acids Res.* 17, 6741.
- [9] Zhao, B. (1991) Ph. D. Dissertation, University of Alabama at Birmingham.
- [10] Drabikowska, A.K. and Wozniak, G. (1990) *Biochem. J.* 270, 319–323.
- [11] Bose, R. and Yamada, E.W. (1974) *Biochemistry* 13, 2051–2056.
- [12] Komissarov, A.A., Romanova, D.V., Dmitrieva, N.A., Linkova, E.V., Mironov, A.S. and Debabov, V.G. (1994) *Biochim. Biophys. Acta* 1205, 54–58.
- [13] Mikhailov, A.M., Smirnova, E.A., Tsuprun, V.L., Tagunova, I.V., Vainshtein, B.K., Linkova, E.V., Komissarov, A.A., Siprashvili, Z.Z. and Moironov, A.S. (1992) *Biochem. International* 26, 607–615.
- [14] Cook, W.J., Kolzalka, G.W., Hall, W.W., Narayana, V.L. and Ealick, S.E. (1987) *J. Biol. Chem.* 262, 2852–2853.
- [15] Zanevsky, Yu.V., Donet, D.E., Ivanov, A.B., Movchan, S.A., Ostrovnoy, A.I., Chernenko, S.P., Kheiker, D.M., Andrianova, M.E., Popov, A.N. and Sulyanov, S.N. (1993) *Kristallografiya* 38, 252–260 (Russian).
- [16] Fitzgerald, P.M.D. (1988) *J. Appl. Crystallogr.* 21, 273–278.
- [17] Navaza, J. (1987) *Acta Crystallogr.* A43, 645–653.
- [18] Brunger, A.T., Kuriyan, J. and Karplus, M. (1987) *Science* 235, 458–460.
- [19] Brunger, A.T. and Krukowski, A. (1990) *Acta Crystallogr.* A46, 585–593.
- [20] Mooser, G., Hefta, S.A., Paxton, F.J., Shively, J.E. and Lee, T.D. (1991) *J. Biol. Chem.* 266, 8916–8922.
- [21] Eshdat, Y., Dunn, A. and Sharon, N. (1974) *Proc. Natl. Acad. Sci. USA* 71, 1660–1662.
- [22] Malcolm, B.A., Rosenberg, S., Allen, J.S., De Baetselier, A. and Kirsch, J. (1989) *Proc. Natl. Acad. Sci. USA* 86, 133–137.
- [23] Carson, M. (1987) *J. Mol. Graphics* 5, 103–106.
- [24] Krauklis, P.J. (1991) *J. Appl. Crystallogr.* 24, 946–950.