Structural characterization of extracellular ribonuclease of *Bacillus polymyxa*: amino acid sequence determination and spatial structure prediction

A.A. Lebedev[a], S.V. Shlyapnikov[b], V.N. Pustobaev[b], M.P. Kirpichnikov[c], A.A. Dementiev[c,*]

[a]*Shubnikov Institute of Crystallography, Russian Academy of Sciences, Leninsky pr. 59, Moscow 117333, Russian Federation*
[b]*Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilov str. 32, Moscow 117984, Russian Federation*
[c]*Center of Bioengineering, Russian Academy of Sciences, 60 Letiya Oktyabrya pr. 7/1, Moscow 117312, Russian Federation*

Received 28 May 1996; revised version received 1 July 1996

Abstract The primary structure of extracellular *Bacillus polymyxa* ribonuclease (RNase Bpo) was established by mass spectroscopy analysis and automatic Edman degradation of the individual peptides obtained from protein digestion with Glu-specific protease V8. RNase Bpo consists of 111 amino acid residues, with a relative molecular weight of 12 607. RNase Bpo is a close structural homolog of RNases of *B. amyloliquefaciens* (RNase Ba) and *B. intermedius* (RNase Bi), the similarity of their primary structures being 68%. Molecular modelling of the structure of the complex of RNase Bpo with substrate analog d(CGAC) was performed and a spatial model based on the known crystal structure of RNase Ba complex with the corresponding nucleotide was constructed using the methods of interactive computer graphics and energy minimization. The differences in the primary and tertiary structures of the enzymes were analyzed in order to understand the substrate specificity of *Bacillus* RNases.

Key words: Ribonuclease; Primary structure; Molecular modeling; Protein-nucleic acid interaction; *Bacillus polymyxa*

1. Introduction

Extracellular *Bacillus* RNases belong to the family of structurally related and mutually homologous low molecular weight microbial cyclizing RNases [1,2]. The best-studied representatives of *Bacillus* RNases are the RNases of *B. amyloliquefaciens* (RNase Ba) and *B. intermedius* (RNase Bi). The amino acid sequences of RNases Ba and Bi were found to be different in 18 sites [3–5]. The spatial structures of the enzymes complexed with various specific substrate analogs were solved at high resolution [6–8], and the substrate specificity and mechanism of catalysis were studied in detail [9–13].

Some years ago, three related RNases of *B. thuringiensis* (RNase Bth), *B. coagulans* (RNase Bco) and *B. circulans* (RNase Bci) were obtained in a homogeneous state and their physical-chemical, functional and structural characteristics were studied [5,14,15]. RNases Bth and Bco were shown to have the same primary structure which is identical to that of RNase Bi in all but one residue, while differences in the amino acid sequences of RNases Bci and Ba were found in three positions.

Recently, we have isolated a novel extracellular RNase of *B. polymyxa* (RNase Bpo) and have determined its partial N- and C-terminal amino acid sequences [16]. It is interesting that the substrate specificity of RNase Bpo is markedly different from that of the above-mentioned *Bacillus* RNases and the identity of the N-terminal amino acid sequences (32 residues) of all these RNases is only 59%.

To gain better insight into structural-functional relationships among *Bacillus* RNases, we studied here the complete amino acid sequence of RNase Bpo and carried out molecular modelling of the three-dimensional structure of the enzyme on the basis of its structural similarity to RNase Ba. The predicted spatial structure of the RNase Bpo complex with a substrate analog was characterized and the structure of the enzyme binding site was analyzed to account for the differences in substrate specificity of *Bacillus* RNases.

2. Materials and methods

RNase Bpo was isolated from the filtrate of culture medium of *B. polymyxa* strain 514 as previously described [16].

Proteolysis of RNase Bpo with Glu-specific V8 protease was performed in 10 mM potassium phosphate buffer, pH 7.5, at 50°C for 7 h and at an enzyme-to-substrate ratio of 1:50.

Automatic Edman degradation of RNase peptides was performed on a Knauer model 816 protein sequencer equipped with an on-line Applied Biosystems model 120A PTH analyzer.

Molecular mass analysis of peptides were carried out on a time-of-flight spectrometer using the method of plasma-desorption mass spectroscopy with a ^{252}Cf source as described in [15].

Molecular modelling was performed on a Silicon Graphics INDI work station using the SYBYL program package. Using the structure of RNase Ba from the crystal complex of RNase Ba (M-chain in asymmetric unit) with d(CGAC) as template [7], the structure of RNase Bpo was modelled with the aid of interactive computer graphics [17]. Then, the structure of RNase Bpo with d(CGAC) was obtained from overlay of the RNase Ba-d(CGAC) complex structure with the predicted RNase Bpo structure. The structure of the RNase Bpo-d(CGAC) complex was subjected to energy minimization calculations, employing the TRIPOS force field [17]. The pH of the system was set to 7.0; all residues of Asp, Glu and the C-terminus were taken as negatively charged, all residues of Arg, Lys, His and the N-terminus being considered to be positively charged.

3. Results and discussion

According to structural data obtained previously [16], RNase Bpo was shown to be a close structural homolog of

*Corresponding author.

Abbreviations: RNases Bpo, Ba, Bi, Bth, Bco and Bci, ribonucleases of *Bacillus polymyxa*, *B. amyloliquefaciens*, *B. intermedius*, *B. thuringiensis*, *B. coagulans* and *B. circulans*, respectively (EC 3.1.27.1); V8 protease, endoprotease from *Staphylococcus aureus* strain V8 (EC 3.4.21.19); Barstar, intracellular protein inhibitor of RNase Ba; d(CGAC), 2′-deoxycytosylyl-(3′,5′)-2′-deoxyguanylyl-(3′,5′)-2′-deoxyadenylyl-(3′,5′)-2′-deoxycytidine; rms, root-mean-square.

PII S0014-5793(96)00793-4

extracellular *Bacillus* RNases. Therefore, based on the assumption of a high level of sequence similarity between *Bacillus* RNases, we determined the complete amino acid sequence of RNase Bpo using only one digestion of the protein with Glu-specific V8 protease.

11 individual peptides of RNase Bpo digestion with V8 protease (P-peptides) were isolated by reversed-phase chromatography (Fig. 1). There are six major components (P3, P6–P10 peptides) in the peptide mixture of RNase Bpo. Previously, we identified three glutamic acid residues in the sequence of 32 N-terminal residues of RNase Bpo – Glu^3, Glu^9 and Glu^{30} (Fig. 2) [16]. According to the mass spectroscopy data (Table 1), P3 and P10 peptides are localized in the N-terminal region of RNase Bpo (4–9 and 10–30 fragments, respectively, Fig. 2). The results of automatic sequencing of the P6–P9 peptides and comparison of the data with the sequences of the corresponding fragments of RNases Ba (Fig. 2) make the reconstruction of 94 N-terminal amino acid residues of RNase Bpo possible. The coincidence, within the limits of experimental error, of the theoretical and experimental molecular masses of P6–P9 peptides confirmed the complete amino acid sequences of these peptides as determined by sequence analysis. The positions of other glutamic acid residues, Glu^{45}, Glu^{61} and Glu^{74}, in the RNase Bpo polypeptide chain were determined in the process. Of these, Glu^{61} and Glu^{74} are conserved in all *Bacillus* RNases, being involved in the formation of the active sites [9,13].

Additional information required for deciphering the RNase Bpo complete sequence was obtained from automatic sequencing of the minor components (P1, P2, P4, P5 and P11 peptides) of protein digestion and their mass spectrometric analysis (Table 1, Fig. 2). P11 peptide (residues 4–30) was found to be the product of incomplete cleavage of the peptide bond at Glu^9 in the N-terminal part of the protein. P1, P2, P4 and P5 peptides turned out to be products of the nonspecific cleavage of peptide bonds at the carbonyl groups of Asp^{66}, Asp^{94} and Asp^{102} by V8 protease: P2 peptide (residues 67–74) is part of P7 peptide, while P5 peptide (residues 95–111) is the C-terminal part of the protein and consists of P1 and P4 peptides (Fig. 2). The results obtained were sufficient for the unambiguous establishment of the primary structure of RNase Bpo. Thus, RNase Bpo contains 111 amino acid residues and has a relative molecular mass of 12 607 Da which agrees to within ±0.03% with the experimental molecular mass (12 608 Da) determined by electro-spray mass spectroscopy [16].

Table 1
Mass spectroscopy analysis of the RNase Bpo peptides

Peptide	Relative molecular mass (Da)		Protein chain position
	Experimental	Theoretical	
P1	996±3	994	95–102
P2	960±3	959	67–74
P3	720±2	721	4–9
P4	1239±4	1238	103–111
P5	2218±6	2214	95–111
P6	1616±5	1613	31–45
P7	1525±5	1525	62–74
P8	1685±5	1685	46–61
P9	2243±6	2248	75–94
P10	2361±7	2365	10–30
P11	3071±9	3068	4–30

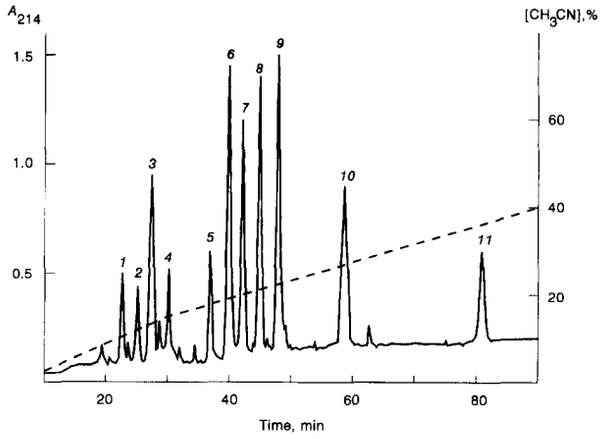

Fig. 1. Fractionation of the V8 protease digest of RNase Bpo (0.5 mg) by reversed-phase HPLC on a Vydac C-18 column (0.46×25 cm, 10 μm). Elution was performed with an acetonitrile gradient (broken line) in 0.1% trifluoroacetic acid at flow rate of 0.7 ml/min. The effluent was monitored at A_{214} (solid line). The numbers of peptides are shown above the peaks.

Sequence alignment (data not shown) indicates 68% similarity between RNase Bpo and other *Bacillus* RNases. This value is below that (83% identity) typical for RNases Ba, Bi, Bth, Bco and Bci. Nevertheless, it should be noted that the residues Lys^{28}, Glu^{74}, Arg^{84}, Arg^{88} and His^{103}, corresponding to catalytically important residues in *Bacillus* RNases, are completely conserved in RNase Bpo. In contrast, there are several substitutions in and near the RNase Bpo fragment (residues 57–63) corresponding to the binding site of the guanine base in *Bacillus* RNases. The relatively high sequence similarity of RNase Bpo to RNase Ba (71% identity, Fig. 2) and complete conservation of their catalytic components enable one to model the tertiary structure of RNase Bpo using the well-characterized three-dimensional structure of RNase Ba as a template. In this work, we used the crystal structure of RNase Ba from the enzyme complex with deoxytetranucleotide d(CGAC) solved at 1.76 Å resolution [7].

As would be expected, the overall polypeptide fold of the molecule in the predicted and energy-minimized structure of RNase Bpo is essentially the same as that of the RNase Ba molecule [7]. The superposition of these structures was performed using all the C_α atoms except for three and four N-terminal residues of RNases Ba and Bpo, respectively, resulting in a mean rms deviation of 0.17 Å. The typical secondary structure elements are conserved with C_α rms deviation of 0.10 Å between the molecules of RNases Bpo and Ba. There are three α-helices (residues 7–19, 27–35 and 42–47) and five-stranded antiparallel β-sheet (residues 51–56, 71–77, 86–92, 95–100 and 107–109) in the RNase Bpo molecule. The hydrophobic core of the protein is formed by the packing of the first α-helix against the β-sheet. The Lys^{28}, Glu^{74}, Arg^{84}, Arg^{88} and His^{103} residues are located on the other side of the β-sheet in the shallow groove. These residues form the active site. The atoms of these catalytic residues of RNases Bpo and Ba can be superimposed to within 0.07 Å rms. Thus, RNases Bpo and Ba share a common catalytic geometry. This conclusion is confirmed by the inhibition of RNase Bpo by the intracellular protein inhibitor of RNase Ba (barstar). The RNase Bpo-bar-

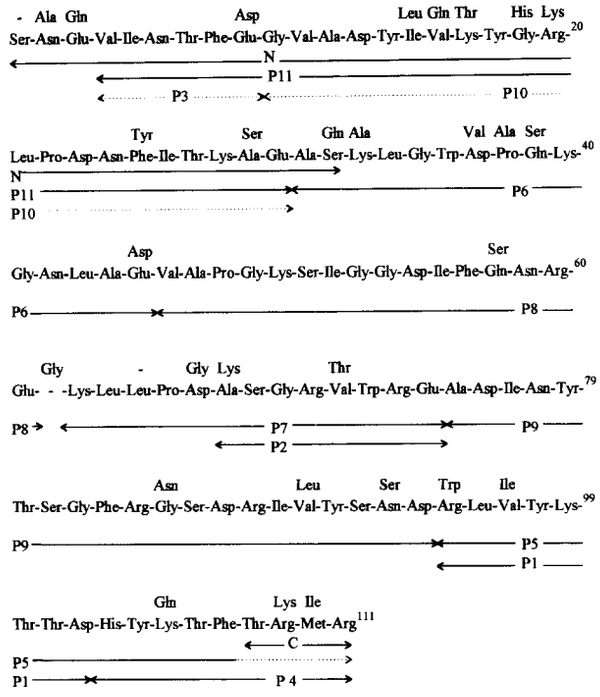

Fig. 2. Amino acid sequence of RNase Bpo. Localization of peptides of RNase digestion with V8 protease (P) and the parts of their structures established by automatic sequencing (solid lines) are shown below the sequence. N and C indicate the data of the N- and C-terminal sequence analysis of the entire protein. The RNase Ba residues different from those of RNase Bpo are shown above the sequence. Deletions are inserted into sequences for optimal alignment.

star 1:1 complex has a dissociation constant of 10^{-12} M, as compared to 10^{-13} M for the RNase Ba-barstar complex [16].

As follows from the analysis of the sequence alignment with RNase Ba, there are 28 amino acid substitutions, two insertions (Ser[1] and Leu[64]) and one deletion (after Glu[61]) in RNase Bpo (Fig. 2) that are scattered throughout the polypeptide chain: nine of those are in the α-helices, four in the β-strands and others in loops, turns, N- and C-terminal regions. 16 amino acid substitutions are not conserved. Essentially all of those occur in the α-helices and loops or in N- and C-terminal parts of the protein, and involve solvent-exposed residues, except for His→Gly[19] and Trp→Arg[95] mutations in the hydrophobic core. It has been shown previously [18] that the His[19] residue interacts with Trp[95] in RNase Ba (RNase Bpo numbering), the distance between the centers of the two rings being 4.0 Å and the angle, 43° (Fig. 3). This histidine-aromatic interaction increases the stability of RNase Ba by 0.8–1.0 kcal/M [18]. The destabilizing effect due to the above-mentioned mutations in RNase Bpo seems to be partly counterbalanced by the putative formation of three hydrogen bonds between the side-chain of Arg[95] and the main-chain CO groups of Val[16] and Lys[17] (Fig. 3). An additional problem is posed by Pro[38] in RNase Bpo which is not present in RNase Ba. We have chosen a *trans*-conformation for Pro[38], since it is because of the proline conformation that the general spatial course of loop region 36–41 in RNases Bpo and Ba is similar.

Extracellular *Bacillus* RNases are guanylyl-specific in the case of nucleoside-2′,3′-cyclophosphates and dinucleoside monophosphates. However, in *trans*-esterification reactions of oligonucleotides more than 4–5 nucleotide long all these enzymes preferentially cleave after guanosine residues, but they also catalyze the cleavage of phosphodiester bonds after other bases [9–11]. The preference for guanine in *Bacillus* RNases is attributed to specific interactions between the guanine base and the enzyme recognition loop through five hydrogen bonds [6,7]. Despite the fact that there are three amino acid substitutions in the recognition loop 57–63 of RNase Bpo, compared to RNase Ba, the network of hydrogen bonds with the guanosine ring in the nucleotide binding site is exactly the same for both enzymes (Fig. 4). In RNase Bpo, there are three hydrogen bonds involving the protein main-chain atoms – Gln[58]NH to N7Gua, Asn[59]NH to O6Gua and Arg[60]NH to O6Gua, and two hydrogen bonds involving the side-chain of Glu[61] – Glu[61]OE1 to NH1Gua and Glu[61]OE2 to NH2Gua. Additionally, the guanine base in the RNase Bpo recognition site is stacked on the aromatic ring of Phe[57] and the side-chain of Arg[60] as in the crystalline structure of RNase Ba [7]. However, the extent of specificity of RNase Bpo towards guanosine and its derivatives in the *trans*-esterification reactions of polyribonucleotides exceeds those for wild-type RNases Ba and Bi by a factor of 18 and 140, respectively

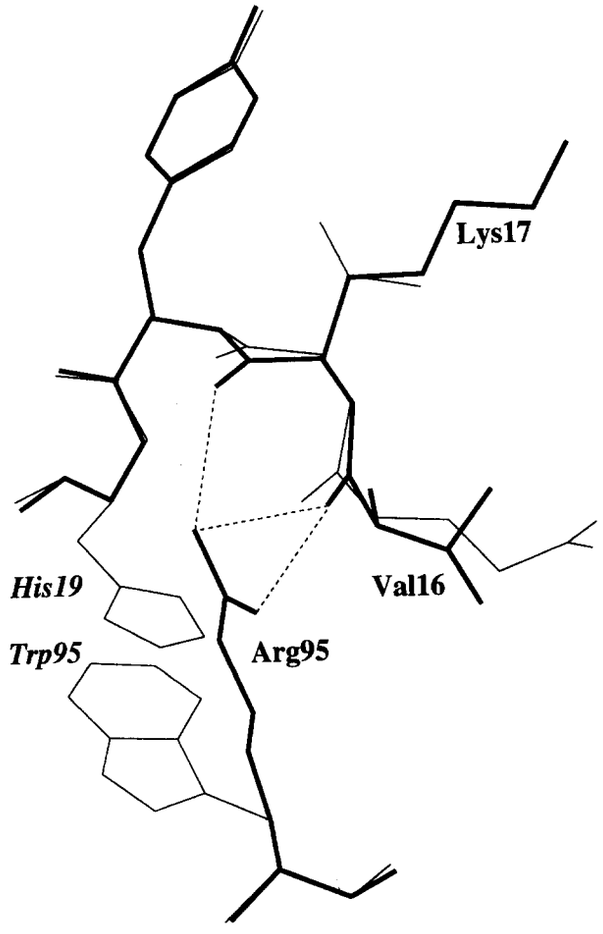

Fig. 3. Superposition of the hydrophobic core part of RNase Bpo (thick lines) with that of RNase Ba (thin lines, italic residue identification). The amino acid residues of RNases Bpo and Ba are labeled by RNase Bpo numbering and potential hydrogen bonds are shown by broken lines.

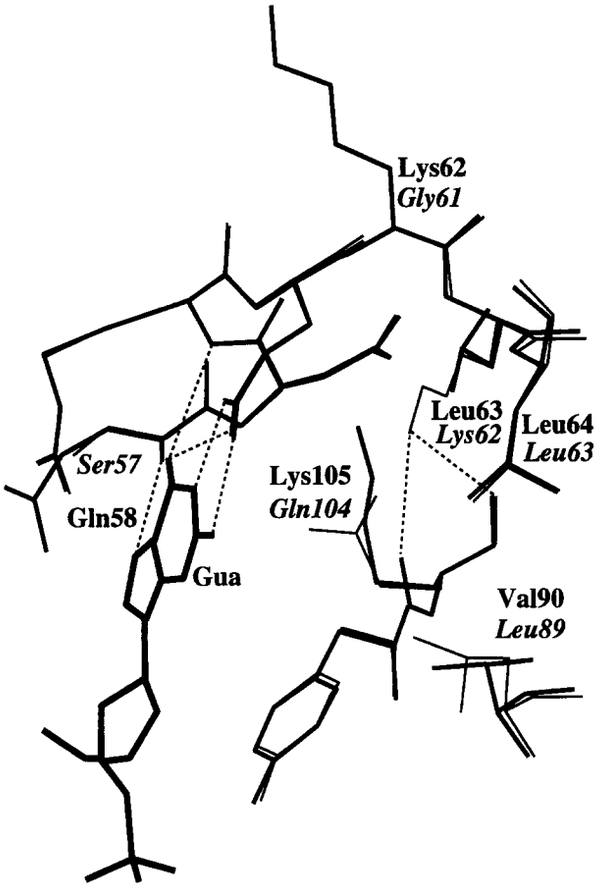

Fig. 4. Superposition of the nucleotide-binding sites of RNases Bpo (thick lines) and Ba (thin lines). The bold lines represent the guanine base of d(CGAC) and potential hydrogen bonds are shown by broken lines. The amino acid residues of RNases Bpo and Ba are labeled by RNase Bpo numbering and italic RNase Ba numbering, respectively.

[16]. The differences in the substrate specificity between RNases Bpo, Ba and Bi may be due to two mutations in the recognition loop that has been confirmed recently by protein engineering experiments [11,12,19]. The mutations of Ser^{57} to Ala and Ser^{57} to Glu in RNase Ba cause a 5- and 35-fold increase, respectively, in the RNase specificity displayed in the cleavage reaction of polynucleotides as compared to that for the wild-type enzyme [11,12]. On the basis of these experiments, it was concluded that the substitution in the RNase Ba recognition site results in elimination of a putative hydrogen bond between the side-chain of Ser^{57} and the N7 atom of purine base of the substrate followed by an increase in the RNase specificity [11]. The substitutions of Ser^{56} and Arg^{61} (structural equivalent of Lys^{62} in RNase Ba) to Gln and Val, respectively, in RNase Bi give rise to a 5-fold increase in the ratio of the cleavage rates of poly(I) and poly(A) in comparison to the wild-type enzyme [19]. Arg^{61} in the RNase Bi molecule participates in the binding of the recognition loop to the rest of the molecule by means of hydrogen bonds with the main-chain CO groups of residues 102 and 103 and due to the side-chain being included in hydrophobic core of the protein [8]. There are similar hydrophobic contacts and hydrogen bonds between the Lys^{62} side-chain and the main-chain of residues 103 and 104 in RNase Ba (Fig. 4) [7]. Thus, the substitution in this site of RNases Ba and Bi eliminates hydrogen bond formation followed by a change in the spatial arrangement of the enzyme recognition loops. This may account for the increased guanylyl preference displayed by the mutant forms of RNases Ba and Bi [19].

According to the spatial alignment between RNases Bpo and Ba, there are three amino acid mutations in the recognition loop of RNases Bpo, namely: Ser^{57}, Gly^{61} and Lys^{62} of RNase Ba are replaced by Gln^{58}, Lys^{62} and Leu^{63} in RNase Bpo, respectively (Fig. 4). The one mutation allows the side-chain of Lys^{62} to be completely exposed to the solvent and, the other two are similar to the above-mentioned substitutions in the recognition loops of RNases Ba and Bi. Leu^{64} of RNase Bpo is structurally equivalent to Leu^{63} of RNase Ba which is completely buried in the hydrophobic core of the protein and faces the side-chain of Leu^{89} (Val^{90} in RNase Bpo, Fig. 4).

Thus, the results of molecular modelling of the RNase Bpo structure supported by the experimental data indicate that the greater guanylyl preference of RNase Bpo towards polyribonucleotides relative to *Bacillus* RNases seems to be the structural consequence of the mutations in the recognition site of the enzyme.

It should be mentioned finally that the precise nature of RNase Bpo-nucleic acid interactions will be understood with the use of the techniques of NMR or X-ray analysis combined with further kinetic studies.

Acknowledgements: This work was supported in part by the Russian Foundation for Basic Research, projects 94-04-12655, 96-04-49850 and 96-04-49852.

References

[1] Hill, C., Dodson, G., Heinemann, V., Mitsuo, Y., Nakamura, K., Borisov, S., Tishenko, G., Polyakov, K. and Pavlovsky, A. (1983) Trends Biochem. Sci. 8, 364–369.
[2] Nomura, H., Inokuchi, N., Kobayasky, H., Koyama, T., Iwama, M., Ohgi, K. and Irie, M. (1994) J. Biochem. 116, 26–33.
[3] Hartley, R.W. and Barker, E.A. (1972) Nat. New Biol. 235, 15–16.
[4] Aphanasenko, G.A., Dudkin, S.M., Kaminir, L.B., Leshchinskaya, I.B. and Severin, E.S. (1979) FEBS Lett. 97, 77–80.
[5] Dementiev, A.A., Ryabchenko, N.F., Protasevich, I.I., Colyshin, P.N., Stepanov, A.I., Orlov, V.M., Pustobaev, V.N., Makarov, A.A., Moiseyev, G.P., Karpeisky, M.Ya., Kirpichnikov, M.P. and Shlyapnikov, S.V. (1992) Mol. Biol. (Russian) 19, 853–861.
[6] Guillet, V., Lapthorn, A. and Mauguen, Y. (1993) FEBS Lett. 330, 137–140.
[7] Buckle, A.M. and Fersht, A.R. (1994) Biochemistry 33, 1644–1653.
[8] Pavlovsky, A.G., Sanishvili, R.G., Borisova, S.N., Strokopytov, B.V., Vagin, A.A., Chepurnova, N.K. and Vainshtein, B.K. (1989) Krystallographiya (Russian) 34, 137–142.
[9] Mossakowska, D.E., Nyberg, K. and Fersht, A.R. (1989) Biochemistry 28, 3843–3850.
[10] Day, A.G., Parsonage, D., Ebel, S., Brown, T. and Fersht, A.R. (1992) Biochemistry 31, 6390–6395.
[11] Yakovlev, G.I., Moiseyev, G.P., Struminskaya, N.K., Romakhina, E.R., Leshchinskaya, I.B. and Hartley, R.W. (1993) Eur. J. Biochem. 215, 167–170.
[12] Kolbanovskaya, E.Yu., Okorokov, A.L., Panov, K.I., Polyakov, K.M., Hartley, R.W. and Karpeisky, M.Ya. (1994) Mol. Biol. (Russian) 28, 602–609.
[13] Yakovlev, G.I., Moiseyev, G.P., Struminskaya, N.K., Borzykh, O.A., Kipenskaya, L.V., Znamenskaya, L.V., Leshchinskaya, I.B., Chernokalskaya, E.B. and Hartley, R.W. (1994) FEBS Lett. 354, 305–306.
[14] Dementiev, A.A., Orlov, V.M. and Shlyapnikov, S.V. (1993) Bioorg. Khim. 19, 853–861.

[15] Dementiev, A.A., Moiseyev, G.P. and Shlyapnikov, S.V. (1993) FEBS Lett. 334, 247–249.
[16] Dementiev, A.A., Mirgorodskaya, O.A., Moiseyev, G.P., Yakovlev, G.I., Shlyapnikov, S.V. and Kirpichnikov, M.P. (1996) Biochem. Mol. Biol. Int. 38, in press.
[17] Clark, M., Cramer III, R.D. and Opdenbosch, N. (1989) J. Comput. Chem. 10, 982–1012.
[18] Loewenthal, R., Sancho, J. and Fersht, A.R. (1992) J. Mol. Biol. 224, 759–770.
[19] Okorokov, A.L., Panov, K.I., Kolbanovskaya, E.Yu., Karpeisky, M.Ya., Polyakov, K.M., Wilkinson, A.J. and Dodson, G.G. (1996) FEBS Lett. 384, 143–146.