Acta Cryst. (1993). D**49**, 257–271

Complex of Ribonuclease from *Streptomyces aureofaciens* with 2′-GMP at 1.7 Å Resolution

BY JOZEF SEVCIK

Institute of Molecular Biology, Slovak Academy of Sciences, Dúbravská cesta, 84251 *Bratislava, Czechoslovakia*

CHRISTOPHER P. HILL*

Molecular Biology Institute and Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90024-1569, *USA*

AND ZBIGNIEW DAUTER AND KEITH S. WILSON

European Molecular Biology Laboratory, Hamburg Outstation, c/o DESY, Notkestrasse 85, 2000 *Hamburg, Germany*

(*Received* 9 *April* 1992; *accepted* 2 *July* 1992)

Abstract

The crystal structure of a complex of ribonuclease from *Streptomyces aureofaciens* (RNase Sa) with guanosine-2′-monophosphate (2′-GMP) has been refined against synchrotron data recorded from a single crystal using radiation from beamline X31 at EMBL, Hamburg, and an imaging plate scanner. The crystals are in space group $P2_12_12_1$ with cell dimensions $a = 64.7$, $b = 78.8$ and $c = 39.1$ Å. The structure has two enzyme molecules in the asymmetric unit, complexed with 2′-GMP inhibitor with occupancies of 1 and $\frac{2}{3}$ (different to the 3′-GMP complex crystal structure where only one of the two independent RNase Sa molecules binds nucleotide), 492 associated water molecules and one sulfate ion, and was refined using all data between 10.0 and 1.7 Å to a final crystallographic R factor of 13.25%. Binding of the base to the enzyme confirms the basis for the guanine specificity but the structural results still do not provide direct evidence of the identity and role of the particular residues involved in the catalytic process. New native RNase Sa data to 1.8 Å were recorded to provide a reference set measured under comparable experimental conditions. The crystals are in the same space group and have the same lattice as those of the 2′-GMP complex. The native structure with 423 water molecules was refined in a similar manner to the complex to a final R factor of 13.87%. 1.77 Å resolution data were independently measured on a 2′-GMP complex crystal at UCLA using an R-AXIS II image plate scanner mounted on a conventional source. The cell dimensions were essentially the same as above. 2′-GMP was bound more fully to molecule *A* than to molecule *B* of the RNase Sa. The structure was refined to an R factor of 14.64% with 388 water molecules. This work follows on from the structure determination of native RNase Sa and its complex with 3′-GMP [Sevcik, Dodson & Dodson (1991). *Acta Cryst.* B**47**, 240–253].

Introduction

Ribonucleases have been widely studied and the results obtained have made a major contribution to the understanding of the structure, function and stability of enzymes. The pancreatic and microbial families of ribonucleases have very different structures in spite of the fact that they both catalyse the hydrolysis of single-stranded RNA with the same mechanism suggested by Findlay, Herries, Mathias, Ranbin & Ross (1961), Takahashi (1970) and Takahashi & Moore (1982). The reaction occurs in two steps, the first is a transesterification to produce a terminal cyclic 2′,3′-phosphate and the second a hydrolysis to a 3′-phosphate. The most studied member of the pancreatic ribonuclease group is RNase A, while that of the microbial group is RNase T_1 from the fungus *Aspergillus oryzae* (Pace, Heinemann, Hahn & Saenger, 1991).

The family of microbial ribonucleases was originally proposed by Hartley (1980) and later more generally discussed by Hill *et al.* (1983). The family is divided into fungal (eukaryotic) and bacterial (prokaryotic) subgroups. Members of the fungal subgroup are more closely related to one another than

* Present address: Department of Biochemistry, University of Utah School of Medicine, 50 North Medical Drive, Salt Lake City, UT 84132, USA.

0907-4449/93/020257-15$06.00

are members of the bacterial subgroup. Both subgroups have five amino-acid residues strictly conserved, namely Asn 39, Glu 41, Glu 54, Arg 69 and His 85 (numbered according to RNase Sa) (Hill *et al.*, 1987; Hill, 1989). Their tertiary structures, while showing considerable differences, all have a common central structural motif consisting of a three-stranded antiparallel β-sheet and a single strand positioned above the sheet in a direction perpendicular to it (Sevcik, Sanishvili, Pavlovsky & Polyakov, 1990). This conserved structural motif is functionally important as it includes the substrate binding site and the catalytic site with all the totally conserved amino-acid residues. Some of the microbial ribonucleases possess properties which are becoming important in practical applications, *e.g.* in plant engineering (Mariani, Beuckeleer, Truettner, Leemans & Goldberg, 1990).

The ribonuclease from *Streptomyces aureofaciens*, Rnase Sa (E.C. 3.1.4.8), is a member of the bacterial subgroup. It is a small enzyme composed of 96 amino-acid residues. Its relative molecular weight determined from the known amino-acid sequence (Shlyapnikov *et al.*, 1986) is 10544. RNase Sa, a guanylate endoribonuclease, highly specifically hydrolyses the phosphodiester bond of RNA at the 3′-side of guanosine nucleotides (Zelinkova, Bacova & Zelinka, 1971). Because of its properties (*e.g.* small size, stability, specificity, possible practical applications) the enzyme is the subject of intensive biochemical and structural study. In addition RNase Sa shows some antiviral activity (Both, Gallo & Sevcik, 1991). In spite of the very low sequence and structural homology between RNase Sa and T_1, they have very similar enzymatic properties with respect to specificity, catalytic activity and kinetic parameters (Both, Moiseyev & Sevcik, 1991).

The crystal structure of RNase Sa was previously determined by multiple isomorphous replacement. This structure and that of the complex of RNase Sa with guanosine-3′-phosphate (3′-GMP) were refined against 1.8 Å synchrotron data measured on film at the SRS Daresbury Laboratory to crystallographic R factors of 17.2 and 17.5%, respectively (Sevcik, Dodson & Dodson, 1991). There are two enzyme molecules in the asymmetric unit, referred to as *A* and *B*. 3′-GMP was only found at the active site of molecule *A*. Binding of 3′-GMP to the active site of molecule *B* was proposed to be prevented by steric hindrance in the crystal. Atomic coordinates and structure factors of these two structures are in the Brookhaven Protein Data Bank (Bernstein *et al.*, 1977) as coordinate sets 1SAR and 2SAR respectively. We refer to these as the 'York' structures throughout.

In this paper we present the refinement of the 2′-GMP complex and the re-refinement of native RNase Sa with data recorded under comparable conditions from crystals grown recently in Bratislava. We also report the structure of a second 2′-GMP complex with data recently recorded at UCLA.

Data collection

Crystals of the 2′-GMP complex were prepared by diffusion of 2′-GMP into native crystals stored in mother liquor. The concentration of 2′-GMP was approximately 50 mg ml^{-1}, which was close to saturation. Binding of 2′-GMP to the enzyme was monitored by precession photographs and was judged to be complete after about one week. The space group is $P2_12_12_1$ with cell dimensions $a = 64.7$, $b = 78.8$ and $c = 39.1$ Å.

Data were collected from a single crystal (0.8 × 0.6 × 0.4 mm) on the EMBL X31 beamline at the DORIS storage ring, DESY, Hamburg, with an imaging plate scanner built at EMBL (Hendrix & Lentfer, unpublished work). The c axis of the crystal was sufficiently offset from the rotation axis to effectively reduce the fraction of unique non-measurable reflections in the blind region close to zero. 1.7 Å data were measured with an oscillation range of 1.0° per image, and subsequently 2.8 Å data with 2.5° per image. For both sets about 90° rotation was covered and the total data collection time was about 10 h. An analogous regime was used for recording the native data. The system and the software used for data reduction were as given in Dauter, Terry, Witzel & Wilson (1990).

X-ray data collection is summarized in Table 1 for the 2′-GMP and native crystals. The reflections (as a percentage of the unique set) measured with intensities greater than three standard deviations are given in Fig. 1 as a function of resolution. The merging $R(I)$ factor for symmetry-related reflections is shown as a function of resolution in Fig. 2. The Wilson plot (Wilson, 1942) for the complex data (Fig. 3) gives a satisfactory fit to the theoretical line; the substantial deviations in the lowest resolution ranges are typical for protein structures.

Crystals of a second 2′-GMP complex were prepared as above but with a crystal from the same batch used in the original work on the native enzyme and the 3′-GMP complex. This crystal was grown about five years ago but, nevertheless, diffracted reasonably well to high resolution. There were no modifications to the complex preparation protocol. One of these crystals was used at UCLA to record data to 1.8 Å resolution using a Rigaku R-AXIS II imaging plate scanner mounted on a rotating anode RU-200 with Cu $K\alpha$ radiation. The data collection statistics are summarized in Table 1. The completeness is close to 100% as was that for the EMBL data

Table 1. *Data-collection parameters*

	2′-GMP	Native	2′-GMP
Site	EMBL	EMBL	UCLA
Source	DORIS	DORIS	Cu $K\alpha$
Wavelength (Å)	0.95	0.95	1.54
Max. resolution (Å)	1.7	1.8	1.77
No. of measurements	77200	71868	56410
Fully recorded	60437	53026	39128
Partially recorded	16763	18842	14589
$R(I)_{merge}$ (%)	3.3	3.3	4.2
Independent reflections	22455	18996	19476
Completeness (%)	98.9	99.3	96.3

sets. Data collection took about 15 h, for a set with substantially less redundancy of measurements than the EMBL data.

Scaling of the data sets

Here and in the following sections the temperature factors are given as B values in units of Å^2 as used in the expression $\exp(-B\sin^2\theta/\lambda^2)$ for the effects of isotropic thermal vibration. 2′-GMP data were scaled to native. The relative B value for EMBL 2′-GMP was −2.6 and for the UCLA complex +23.5, *i.e.* for the EMBL complex the data fell off with resolution slightly less than the native and the UCLA complex substantially more. The mean fractional isomorphous difference between the EMBL and native data was 27.4% and for the UCLA 23.5%, Fig. 4. The rise in the difference with resolution suggests a small degree of non-isomorphism between the native and 2′-GMP complex crystals. The cell dimensions are almost identical, and there is an almost negligible rotation and translation required to superimpose the refined coordinates of

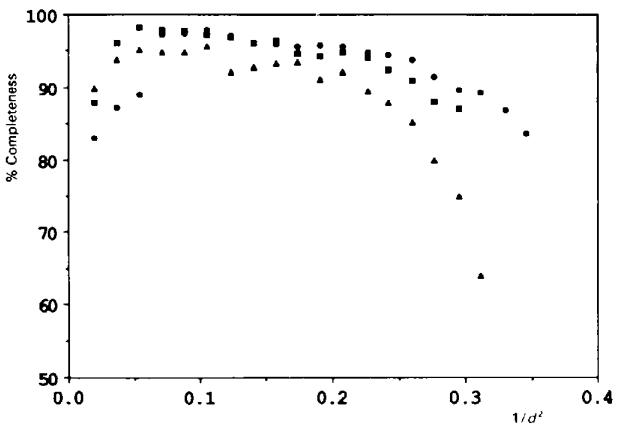


Fig. 1. Reflections with $I > 3\sigma(I)$ as a percentage of the total theoretical number of reflections as a function of resolution. (●) EMBL 2′-GMP, (■) EMBL native and (▲) UCLA 2′-GMP.

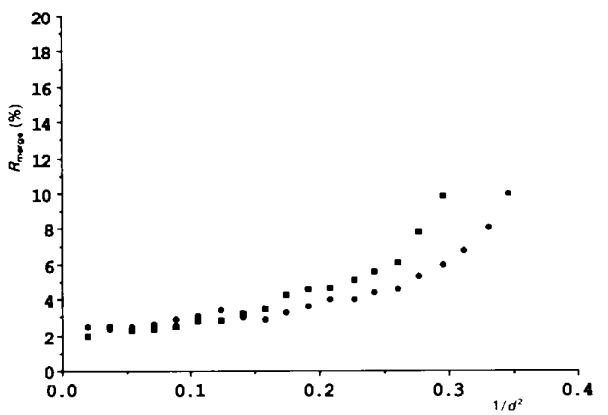


Fig. 2. The merging R factor, $R(I) = \Sigma|I - \langle I\rangle|/\Sigma I$, as a function of resolution for the EMBL native (■) and 2′-GMP complex (●).

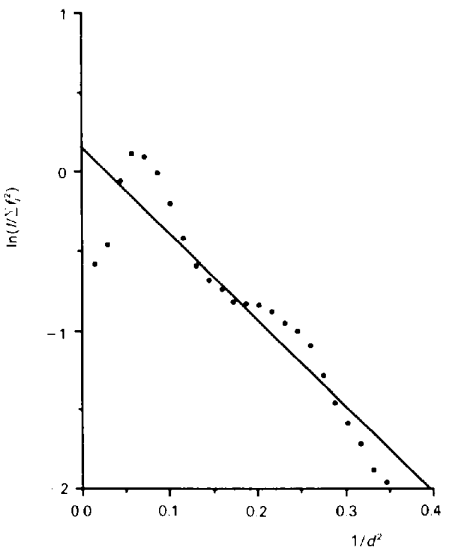


Fig. 3. The Wilson plot for the 10.0–1.7 Å resolution EMBL 2′-GMP complex data. The line was fitted using the full resolution range.

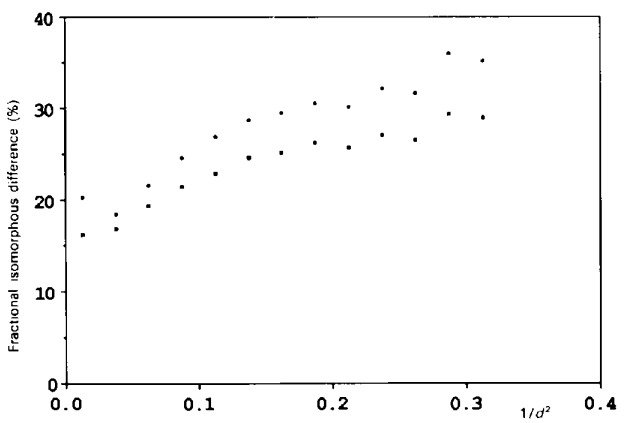


Fig. 4. The mean fractional isomorphous difference for the EMBL (●) and UCLA (■) 2′-GMP data compared to the EMBL native data as a function of resolution.

the protein atoms in the crystals. The mean fractional isomorphous difference between the two 2′-GMP data sets was 12.8%.

The 1SAR native and 2SAR 3′-GMP data were scaled to the EMBL native. The 2SAR 3′-GMP data had been previously scaled to the 1SAR, and hence the results for both are very similar. The data fell off much more rapidly with resolution than the EMBL native, having relative *B* factors of +10.8 and +11.1 for York native and 3′-GMP respectively. The mean fractional isomorphous difference was 8.2% for York native data and 25.1% for the 3′-GMP. The value for the two native sets gives a base line for two theoretically identical sets. The difference would be much less for 'identical' data sets collected on an EMBL imaging plate. The difference between all three complex sets and the EMBL native after scaling is in the range 23–28%.

Refinement

Refinement was carried out by restrained least-squares minimization (Konnert & Hendrickson, 1980; Baker & Dodson, 1980). The stereochemical restraints applied and the parameters of the final model are given in Table 2. No σ cutoff was applied to the amplitudes. Rebuilding of the model was carried out on an Evans and Sutherland PS330 using *FRODO* (Jones, 1978).

EMBL 2′-GMP complex

The crystals used for the preparation of the EMBL 2′-GMP complex were crystallographically identical to those previously used for the structure determination of native RNase Sa (Sevcik *et al.*, 1991) and the corresponding atomic coordinates (PDB set 1SAR) were used as a starting model. The initial *R* factor was 35.94%. The course of refinement is not presented in detail. The occupancy of 2′-GMP on molecule *B* was set to $\frac{2}{3}$ on the basis of low electron density and high *B* factors for the inhibitor. Water molecules at $\frac{1}{3}$ occupancy were included at this site at positions from the native structure. There were 239 water molecules in the initial model. Some were absent from the 2′-GMP complex, but many new ones were added using *FRODO*. In the last steps an automatic water search program (V. Lamzin, unpublished work) was used to locate new waters and check the current ones. In the final model there were 492 water molecules and the *R* factor was 13.25%.

Some side-chain atoms for five residues were assigned two conformations. In molecules *A* and *B* there were two alternative positions for the sulfur atom of Cys 72, which were accorded occupancies of 0.4 and 0.6 on the basis of the electron density and *B* factors. The three other examples of two conformations all lay in molecule *A*. Val 6*A* had two clearly defined conformations for the C*G* atoms representing a rotation of about 120° around the C*A*—C*B* bond. In contrast Val 6*B* had a single well defined orientation. The five atoms beyond C*G* of the side chain of Arg 40*A* and the O*E*1 and N*E*2 atoms of Gln 77*A* were assigned two alternative positions. The equivalent atoms in molecule *B* had high temperature factors, but in neither case were two conformations identifiable.

Table 2. *Refinement parameters, restraint weighting scheme and standard deviations after the last cycle of refinement*

The weight of restraints is $1/\sigma^2$.

	σ	EMBL 2′-GMP	EMBL native	UCLA 2′-GMP
Resolution limits (Å)		10.0–1.70	10.0–1.80	10.0–1.77
No. of reflections		22372	18895	19319
Protein atoms in mol. *A*; mol. *B*		745; 745	745; 745	745; 745
Protein sites in mol. *A*; mol. *B*		755; 746	752; 746	748; 746
No. of waters		492	423	388
No. of sulfate ions		1	1	1
R factor (%)		13.25	13.87	14.64
σ_A error estimate (Å)		0.013	0.015	0.020
Stereochemistry				
Bond length (1—2) (Å)	0.020	0.015	0.018	0.020
Bond angles (1—3) (Å)	0.040	0.038	0.041	0.047
Dihedral angles (1–4) (Å)	0.050	0.048	0.051	0.052
Planar groups	0.020	0.014	0.015	0.017
Chiral volumes (Å^3)	0.200	0.216	0.228	0.252
Peptide angle ω (°)	5.0	3.3	3.5	3.7

Native structure

To compare properly the EMBL 2′-GMP complex with the native enzyme, a refinement of the native structure was carried out using the EMBL native data to 1.8 Å (Table 1). 1SAR coordinates were again used as the starting model. Double conformations were only used for the sulfur atom of Cys 72 in both molecules *A* and *B*, the O*E*1 and O*E*2 atoms of Asp 1*A* and the four atoms after C*B* for Glu 74*A*. The final *R* factor was 13.87%. The estimated errors in the coordinates (Table 2) are a little higher than those for the 2′-GMP complex, reflecting the slightly lower resolution of the analysis.

UCLA 2′-GMP complex

The inhibitor was first introduced at an occupancy of $\frac{2}{3}$ as in the EMBL complex. After inspection of the difference map, occupancy of molecule *B* 2′-GMP was reduced to $\frac{1}{2}$ and the active site waters correspondingly increased to $\frac{1}{2}$. The final *R* factor was 14.64%.

Water molecules

After use of *FRODO* in the refinement of the EMBL 2′-GMP complex there were a total of 298

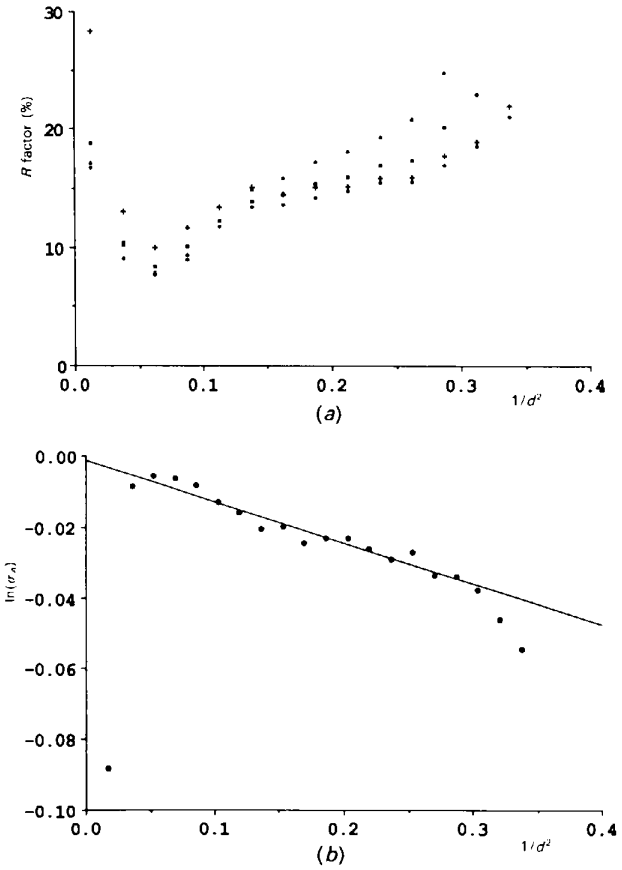


Fig. 5. (*a*) Plot of final *R* factor against resolution for the EMBL native (■) and the EMBL (●) and UCLA (▲) 2′-GMP data. The *R* factor is also shown for the EMBL 2′-GMP model (+) after refinement step 9 before the increase in the number of waters using the automatic search program. (*b*) σ_A plot for the EMBL 2′-GMP data.

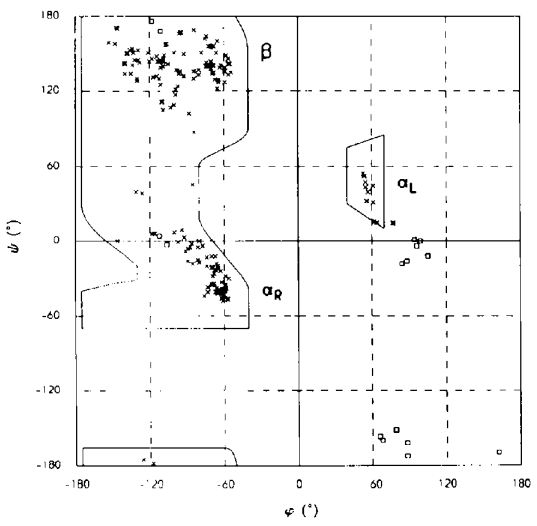


Fig. 6. The Ramachandran plot for the EMBL 2′-GMP complex. Glycine residues are represented by □.

solvent water molecules in the model. This was subsequently increased to 492 by the automatic water search. It is reasonable to ask whether these water positions are significant. Firstly they all obeyed the criteria for hydrogen-bond formation, lying within 2.5–3.5 Å of a hydrogen-bond donor or acceptor. Secondly it was checked that the density for all of them in the final $(2F_o - F_c)$ synthesis was greater than 0.4 e Å $^{-3}$.

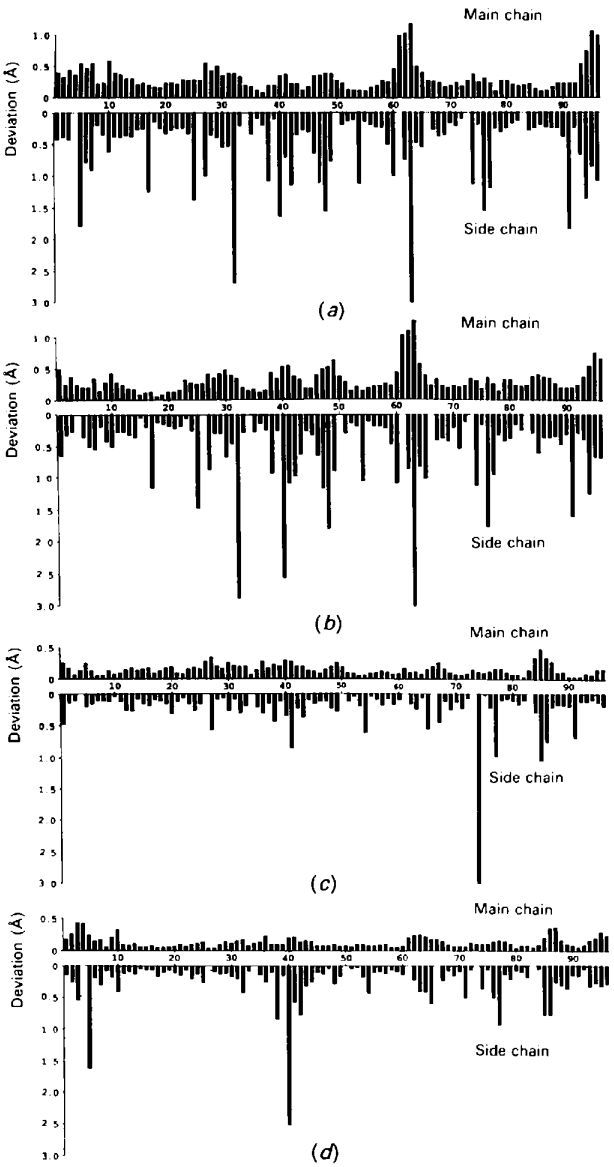
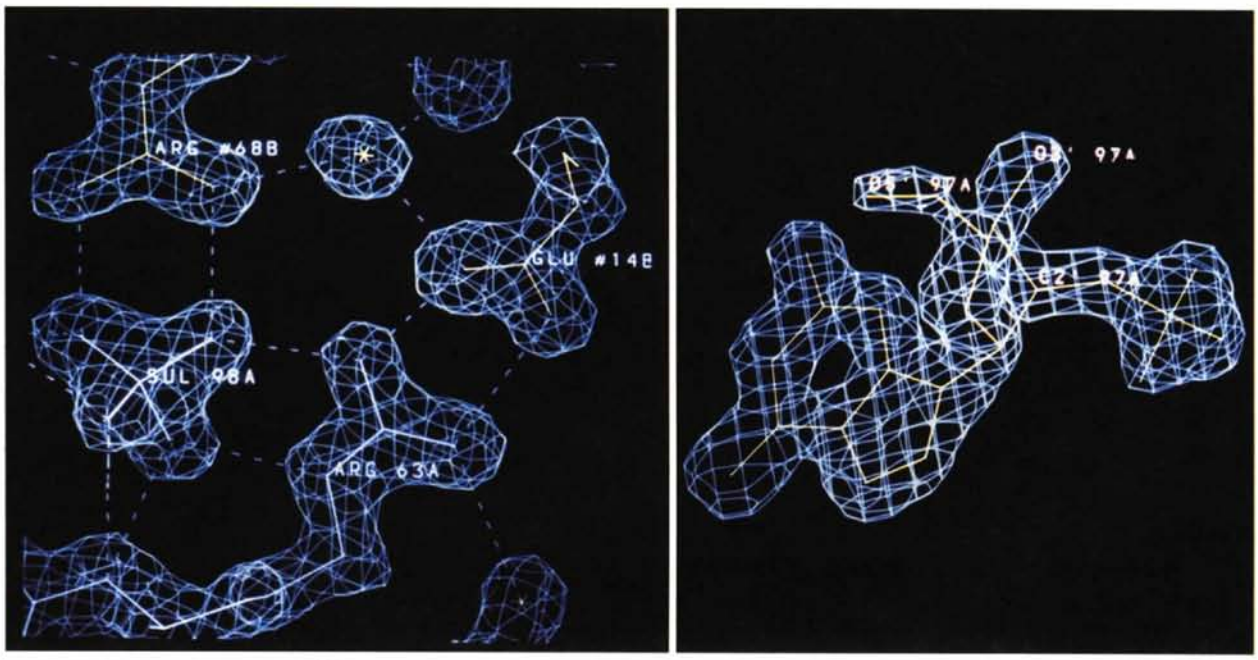
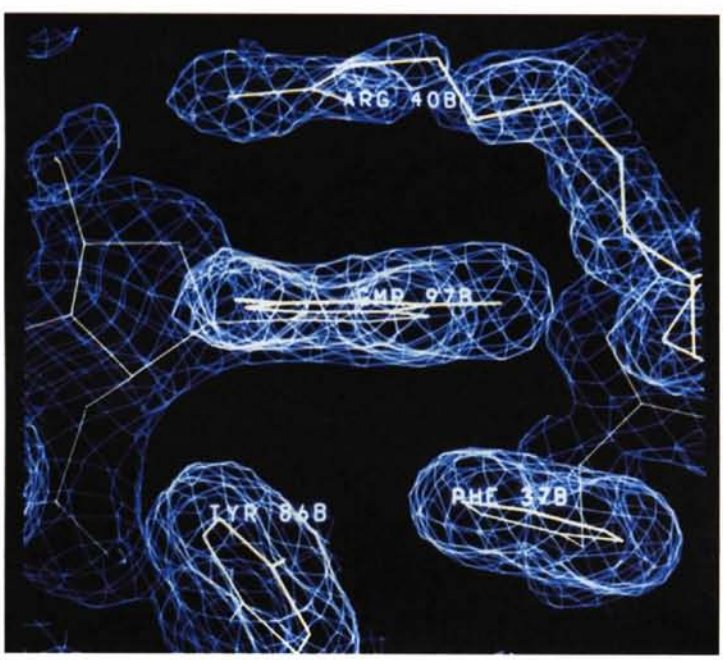


Fig. 7. Average root-mean-square deviations of main-chain and side-chain atoms as a function of residue number. In each case the structures were superimposed on each other by least-squares overlap of all the C*A* atoms. (*a*) Molecule *A* of the 2′-GMP on molecule *B* of the 2′-GMP. (*b*) Molecule *A* of the native on molecule *B* of the native. (*c*) Molecule *A* of the 2′-GMP on molecule *A* of the native. (*d*) Molecule *B* of the 2′-GMP on molecule *B* of the native.

In Fig. 5(*a*) the *R* factor is given as a function of resolution both for the 298 water model and for the final model with 492 waters. Other differences between these two models are small: some minor adjustments of side-chain conformation and the introduction of the multiple conformations as above.



(*a*) (*b*)



(*c*)

Fig. 8. The electron density for the EMBL 2′-GMP complex in three regions of the final $(2F_o - F_c)$ synthesis, Table 4. (*a*) The intermolecular contact region between the two enzyme molecules in the crystal which involves the sulfate ion. The contour level is 1.2 e Å^{-3}. (*b*) The 2′-GMP in molecule *A*. (*c*) The aromatic residues Phe 37*B* and Tyr 86*B* forming one side of the guanine binding cavity. In (*b*) and (*c*) the contour level is 0.6 e Å^{-3}.

The major effect of the extra waters is easily seen. The R factor in the lower resolution ranges drops substantially (from 28.3 to 16.8% in the first range) and in the high ranges falls by about 1%. Inspection of the $(2F_o - F_c)$ and $(F_o - F_c)$ maps and the stereochemistry suggest that the whole model has improved and that this has been achieved by improving the poorly ordered solvent. The introduction of the extra water molecules improved the protein electron-density maps and caused a small decrease of about 0.7 Å^2 in the thermal parameters of the protein. Clearly not all the extra waters lie in well defined spherical density. Higher resolution data would be required to validate with certainty the less well ordered water structure.

The waters are tabulated in ranges of their B factor in Table 5. For the EMBL complex three broad peaks can be identified. The first at about 20–30 Å^2, the second at about 40–50 Å^2 and the third between 70 and 80 Å^2. In the EMBL native the situation is analogous. For the UCLA 2′-GMP complex the distribution is shifted to higher values and the third peak is either missing or, possibly, the upturn only starts to show in the last range. We believe that the EMBL models are more accurate, and reflect real distribution of waters into successive shells: this point will be analysed in more detail at higher resolution.

Accuracy of the models

The refinement, stereochemical and thermal parameters for the final restrained models for the EMBL 2′-GMP and native and the UCLA 2′-GMP are given in Tables 2 and 3. The σ_A plot (Read, 1986) was used to estimate the overall mean coordinate error for all three models (Table 2) and is shown for the EMBL 2′-GMP in Fig. 5(*b*). This is likely to overestimate the true coordinate error for highly refined structures as stated by Read. The plot of the R factor *versus* resolution for all three models is shown in Fig. 5(*a*). The electron densities were calculated on an approximately absolute level and the statistics are given for the final $(2F_o - F_c)$ and $(F_o - F_c)$ maps in Table 4. The highest residual feature in the final difference maps was less than 7σ of the density.

The Ramachandran plot (Ramakrishnan & Ramachandran, 1965) for both molecules of the EMBL complex in the asymmetric unit is shown in Fig. 6. The dihedral angles (φ, ψ) of two residues Tyr 86*A* and Tyr 86*B*, which overlap in the figures, lie only sightly outside the allowed region, *i.e.* the region in which no unacceptable van der Waals interatomic contacts occur. Tyr 86 is localized in a β turn in the nucleotide binding site. Almost identical results are found for the other two models (data not shown).

Table 3. *Statistics of the average isotropic atomic temperature factors* (Å^2)

	EMBL 2′-GMP	EMBL native	UCLA 2′-GMP
Wilson plot estimate	12.2	12.6	20.0
Molecule *A*			
Main chain	9.3	11.2	18.3
Side chain	13.7	16.7	22.5
Protein	11.5	13.8	20.4
2′-GMP	15.3	—	26.7
Molecule *B*			
Main chain	11.0	13.2	20.4
Side chain	15.9	19.1	25.5
Protein	13.4	16.0	22.9
2′-GMP*	19.1	—	32.7
Sulfate	13.8	15.9	23.8
Waters	48.7	48.0	51.7
Total	21.3	22.2	27.9

* The occupancy of the EMBL 2′-GMP on molecule *B* is 2/3 and on the UCLA 1/2.

Table 4. *The electron density* (e Å^{-3}) *in the final* $(2F_o - F_c)$ *and difference* $(F_o - F_c)$ *maps for the EMBL native and the two 2′-GMP complexes*

The electron densities were calculated on an approximately absolute level using the unit-cell volume (200 000 Å^3) and the number of protein electrons in the unit cell (50 704) for the $F(000)$ term in the $(2F_o - F_c)$ map. No $F(000)$ term was included in the difference Fourier synthesis.

	EMBL 2′-GMP	EMBL native	UCLA 2′-GMP
$(2F_o - F_c)$ synthesis			
Maximum	6.53	4.71	3.61
Minimum	−1.40	−1.10	−0.86
Mean	0.25	0.25	0.25
R.m.s. from mean	0.54	0.43	0.36
$(F_o - F_c)$ synthesis			
Maximum	0.48	0.33	0.25
Minimum	−0.37	−0.29	−0.23
R.m.s. from mean	0.07	0.06	0.05

Table 5. *The number of the water molecules in ranges of B-factor values* (Å^2)

	EMBL 2′-GMP	EMBL native	UCLA 2′-GMP
0–10	12	9	0
10–20	47	46	20
20–30	63	52	49
30–40	62	51	57
40–50	86	69	57
50–60	51	57	82
60–70	55	48	43
70–80	68	56	37
80–90	39	33	18
90–100	9	2	25

The EMBL 2′-GMP data are clearly of the highest quality and at the highest resolution. For these data only, five trial cycles of refinement were carried out after step 14 without stereochemical restraints to check further the stability and the effects of the restraints on the final model. The R factor dropped to 12.09%. For the main-chain atoms of the two molecules the root-mean-square shift was 0.08 and the average 0.06 Å. The maximum displacement was

0.44 Å for the C atom of the rather poorly ordered Ser 3*A*, roughly twice the next highest shift of 0.23 Å, about 3σ of the distribution of the displacements. For all protein atoms the root-mean square difference between the restrained and unrestrained model was 0.32, the average 0.10 Å, with a maximum shift of 6.48 Å for the NH2 atom of the highly disordered side chain of Arg 40*A*. These values give some indication of the errors in the final model. The main chain and well ordered side chains behaved well in the unrestrained refinement even at 1.7 Å resolution. The badly defined side chains simply exploded as was observed for rubredoxin at 1.0 Å (Dauter, Sieker & Wilson, 1992), *i.e.* they moved large distances (up to several ångströms) to stereochemically unrealistic positions.

To avoid the latter problem the five cycles of unrestrained refinement were repeated keeping the disordered atoms fixed as a constant contribution with positions and *B* factors as obtained after step 14. 63 atomic sites were thus fixed. These were composed of the 22 positions representing the double conformations and in addition the 41 sites with *B* factors greater than 34 Å^2. The *R* factor dropped to 12.31%. For all protein atoms the root-mean-square difference between the restrained and unrestrained protein model was now 0.073 Å, with maximum shift 0.320 Å for the C*G* atom of Glu 74*A*. The average discrepancy in position was only 0.060 Å. This suggests that the mean positional error in the ordered part (>95%) of the protein is indeed close to that indicated by the σ_A plot.

The refinement of the UCLA 2′-GMP led to a structure which is closely similar to the EMBL complex. There are, nevertheless, significant differences between the two data sets and resulting models, which are now described. The EMBL data nominally extend to 1.7 Å and the UCLA to 1.77 Å. The overall merging $R(I)$ factors are 3.3 and 4.2% and the *R* factors for the refinements are 13.25 and 14.64%, respectively (Tables 1 and 2). The root-mean-square densities in the final difference Fourier syntheses are 0.07 and 0.05 e Å^{-3}. At first sight these criteria indicate that the two models are of comparable quality.

However, a closer inspection reveals some irksome problems with the UCLA data which can be explained by the age (five years) of crystals used at UCLA, compared to relatively fresh crystals at EMBL and York. Firstly, the merging $R(I)$ factor as a function of resolution rises to a considerably higher value for the UCLA than for the EMBL data (not shown) and this is reflected in the much lower proportion of significant reflections in the high-resolution ranges, Fig. 1. Consequently, the *B* factors for the UCLA data estimated from the Wilson plot, and indeed those for the protein in the refined model, are higher by about 8 Å^2 than the values for the EMBL complex. This is also seen in the scaling of the two 2′-GMP data sets to the native. This demonstrates directly that the UCLA data fall off much more rapidly with resolution than the EMBL data. The *B* factors for the EMBL complex are discussed in more detail below, and possess physically meaningful values in the range 7–10 Å^2 for most of the main-chain and ordered side-chain atoms: such values are not higher than those found in small-molecule analyses. As a conseqence, while the overall refinement *R* factors only differ by 1.5%, Fig. 5(*a*) shows a much sharper rise for the UCLA complex as a function of resolution, with the value in the 1.8 Å range being 29.5% for the UCLA complex and that for the EMBL complex only 18.6%. This gives rise to a much smaller error estimated from the σ_A plot in the EMBL model of 0.13 as opposed to 0.20 Å in the UCLA model (Table 2). A final indicator is the density in the $(2F_o - F_c)$ maps (Table 4). The root-mean-square and maximum features of the map are diminished by more than 30% in going from the EMBL to the UCLA map. This is much greater than the difference between the EMBL 2′-GMP and native syntheses and confirms that the effective resolution of the UCLA data is actually lower than 1.8 Å. The results emphasize the weakness of criteria such as the overall *R* factor alone in evaluating the quality of refined structures. Given this we have not used tests such as that proposed by Hamilton (1965) for the significance of differences in *R* factor for different numbers of parameters, *e.g.* number of water sites.

In conclusion the upper limit of the errors in atomic coordinates is estimated at 0.13 Å from the σ_A plot. This is in keeping with the root-mean-square difference of 0.07 Å in atomic position after refinement with and without restraints for the well defined 95% of the model.

Comparison with previous refinements

The data recorded in the present study are more accurate than those reported for the 2SAR and 1SAR data sets. The merging $R(I)$ factors for the latter were 11.7% (3′-GMP complex) and 5.6% (native), respectively, to a resolution of 1.8 Å. For the EMBL 2′-GMP complex $R(I)$ is 3.3% to 1.7 Å and for the EMBL native data also 3.3% to 1.8 Å (Table 1). $R(I)$ is 4.2% for the UCLA 2′-GMP complex to 1.77 Å. The drop in $R(I)$ compared to the York data is significant, and certainly reflects more accurate data at high resolution, though unfortunately the variation of $R(I)$ was not presented by Sevcik *et al.* (1991). The EMBL and UCLA image-plate data are in addition more complete. The higher $R(I)$ of the York data probably results from use of

film for data collection and the much higher $R(I)$ for the York 3′-GMP complex probably reflects an enforced hiatus in data collection due to temporary loss of beam. EMBL and UCLA data were collected on image plate detectors, the higher quality of the EMBL is probably due to the use of synchrotron radiation and fresher crystals. The advantages of synchrotron radiation in terms of high beam intensity, low divergence, fine collimation and shorter data-collection time are unlikely to lead to such large differences in data quality as a function of resolution. Unfortunately, it is impossible to repeat data collection on the aged crystals as they are not longer available.

The overall R factors for the York structures 2SAR and 1SAR for data between 10.0 and 1.8 Å were 17.5 and 17.2% respectively. The variation of R with resolution was not reported by Sevcik *et al.* (1991) but the values in the highest (1.8 Å) range were 25.7 and 29.7% respectively. This is reflected in the substantial relative B factors in scaling the York to the EMBL data (see above). The more rapid fall off of intensity with resolution also shows in the higher mean B values for the 2SAR (16.8 and 19.4 Å^2 for molecules A and B) and 1SAR (17.6 and 19.2 Å^2) structures. These should be compared with the values in Table 3.

Discussion

Overall structure

The overall fold of the protein is the same in all the crystals. It has been described previously and readers are referred to Sevcik *et al.* (1991) for a detailed analysis. In Fig. 7 the results of superimposing the refined protein molecules for the EMBL native and 2′-GMP complex are shown. Deviations are plotted as a function of residue number for main- and side-chain atoms independently. They are discussed in the following sections.

The strength and the beauty of the intermolecular contacts between the two enzyme molecules in the crystal are depicted in Fig. 8(*a*). A sulfate anion and water molecules involved in these contacts saturate the hydrogen-bonding potential of the residues involved. Not all hydrogen bonds are shown for the sake of clarity. Arg 63*A* is part of a loop protruding outside the surface of the molecule. Its binding to a neighbouring molecule causes shifts in the main-chain atoms of more than 1 Å compared to the analogous region of molecule *B*, where the crystal packing is completely different, Figs. 7(*a*) and 7(*b*).

B factors and related properties

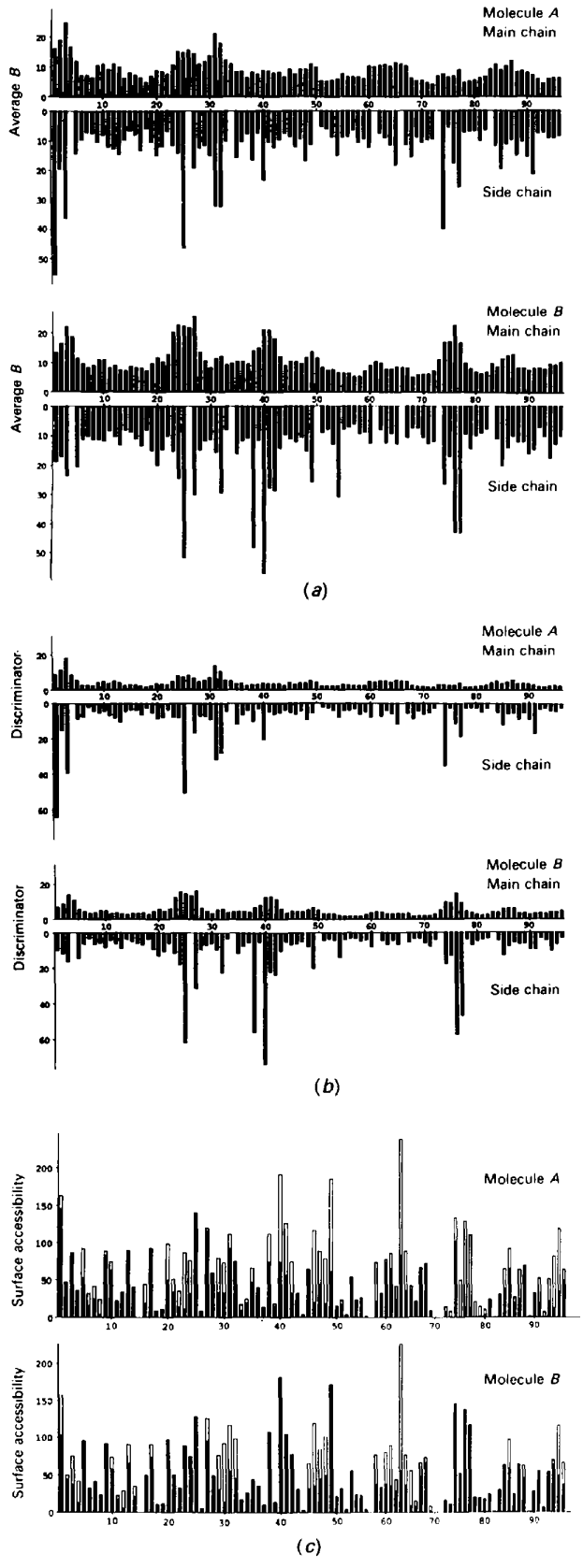
The B factors for the EMBL native and the two 2′-GMP complexes are summarized in Table 3. The values deduced from the Wilson plot using data from the full resolution range correspond quite well with the average of those for the protein atoms in the final models rather than with those of the solvent. The B factors for the EMBL native structure are slightly higher than those of the EMBL 2′-GMP, but both are considerably lower than those for the UCLA complex. This is not surprising. Firstly the UCLA data were collected on crystals which were roughly five years old, while the EMBL crystals had been relatively freshly grown in Bratislava: the advantage of fresh protein is clear. Secondly the EMBL data were collected with synchrotron radiation of wavelength 0.95 Å: this gives an additional advantage in recording high-resolution data. The significance of this is relevant to the accuracy of the analyses discussed above.

For the EMBL 2′-GMP the mean B factors are plotted for the main- and side-chain atoms separately as a function of residue number in Fig. 9(*a*). Similar results (not shown) are found for the native and the UCLA complex. The data are plotted separately for molecules A and B. For those residues where the side chain had two conformations only the data for the one with lower thermal parameters were included in the averaging. In Fig. 9(*b*) a parameter we refer to below as the 'discriminator' is plotted as a function of residue number.

Fig. 9(*c*) shows the surface area accessible to solvent calculated according to Kabsch & Sander (1983). The area is shown both for the molecules as packed in the crystal and when removed from the crystal and 2′-GMP environment. The total areas accessible to solvent for the A and B molecules in the crystal are 3456 and 4247 Å^2, respectively, and for A and B removed from the crystal 5587 and 5577 Å^2. Molecule A has considerably less area exposed to solvent region in the crystal than does molecule B. This is probably the cause of the consistently lower average B factors for molecule A compared to molecule B (about 2 Å^2 lower in all crystals; Table 3).

Flexible and disordered residues

Figs. 7(*a*) and 7(*b*) identify those residues which, through different crystal environments in molecules A and B, show substantial deviations in the positions of the side chains. For a number of these (including residues 5, 17, 25, 32, 38, 40, 41, 42, 47, 48, 49, 60, 63, 74, 76, 77, 91 and 94) the deviations can be directly correlated with the exposed surface area in the crystal (Fig. 9*c*). Four of these, residues 17, 27, 32 and 60, are exposed in A and almost completely buried by intermolecular contacts in B. Ten residues, 5, 38, 41, 42, 49, 74, 76, 77, 91 and 94, are, conversely, buried more in A. Three residues, 47, 48 and 63, are buried in both molecules A and B, but in differ-



ent ways. Residue 25 is highly exposed in both molecules but, nevertheless, takes up different conformations. This emphasizes the importance of considering intermolecular crystal packing in evaluating the significance of the conformations of residues on the surface of the protein. All of these residues clearly have a degree of flexibility in the enzyme to accommodate such changes.

A few residues were modelled with two conformations in the three structures. These are: (1) EMBL 2′-GMP complex: Val 6*A*, Arg 40*A*, Cys 72*A*, Gln 77*A* and Cys 72*B*. (2) EMBL native: Asp 1*A*, Cys 72*A*, Glu 74*A* and Cys 72*B*. (3) UCLA 2′-GMP complex: Val 6*A*, Cys 72*A* and Cys 72*B*. The two positions of the relatively heavy cysteine sulfur atoms can be easily seen in all six molecules in the structures. From the present study the positions of the other double conformations are most clearly defined in the 1.7 Å EMBL complex. However, the density for these other double conformations lies close to the error level of the electron-density maps, and would require improved resolution for their validation. The significance of mean *B* factors *etc.*, Fig. 9, is not meaningful for these side chains.

In addition to these residues with multiple conformations there are several others with *B* factors substantially higher than the average. The residues with mean side-chain *B* factors (Fig. 9*a*) more than 2σ above the average for the EMBL 2′-GMP complex are Asp 1*A*, Ser 3*A*, Asp 25*A*, Ser 31*A*, Gln 32*A*, Glu 74, and Asp 25*B*, Gln 38*B*, Arg 40*B*, Thr 76*B* and Gln 77*B*. From experience with the structure of rubredoxin (Dauter *et al.*, 1992) it is to be expected that several, if not all, of these residues would be assigned more than one conformation if sufficiently accurate and high-resolution data were available.

In Fig. 9(*b*) the mean *B* factor divided by the mean electron density in the final $(2F_o - F_c)$ map is plotted as a function of residue number, which we propose as a useful 'discriminator' for identifying poorly defined residues. The units of the discriminator (e^{-1} Å^5) do not have a simple physical meaning and use of the discriminator is purely empirical. From the figure it is very easy to pick out the most poorly defined residues. In the EMBL 2′-GMP these are Asp 1*A*, Ser 3*A*, Asp 25*A*, Ser 31*A*, Gln 32*A*, Arg 40*A*, Ile 70*A*, Glu 74*A*, Gln 77*A*, Leu 91*A*, and

Fig. 9. Mean values of three parameters for main-chain and side-chain atoms in molecules *A* and *B*, as a function of residue number. (*a*) The atomic *B* factors (Å^2). (*b*) The 'discriminator' parameter, *i.e.* the average *B* factor (Å^2) divided by the average density (e Å^{-3}) in the final $(2F_o - F_c)$ map (see text). (*c*) The solvent accessibility. In (*c*) the solid bars represent the accessibility in the crystal and the unfilled bars the area which becomes accessible on removing the protein from its crystalline and 2′-GMP environment.

Table 6. *Contact distances* (Å) *less than* 3.5 Å *between* 2′*-GMP and the enzyme*

For comparison, bonds found in the complex with 3′-GMP are given.

Molecule *A*				Molecule *B*				3′-GMP		
		EMBL	UCLA			EMBL	UCLA			York
O21—NH2	65	2.75	2.83	O21—*OE*1	54	2.77	2.71	O31—NH2	65	2.78
				—NH1	65	3.09	3.24	*NE*2	32	3.15
				—NH2	65	3.15	2.88			
O22—NH1	69	2.74	2.93	O22—NH2	65	3.46	3.50	O32—*NE*2	85	2.54
—*NE*	69	3.50		—NH1	69	3.04	2.98			
—*NE*2	85	2.76	2.78	—*NE*2	85	3.30	3.24			
O23—OE1	54	2.48	2.42	O23—*OE*1	54	3.13	2.84	O33—NH2	65	2.80
—OH	86	2.34	2.40	—*NE*	69	3.00	3.02	—*NE*	69	3.05
				—*NE*2	85	3.42	3.48			
				—OH	86	2.24	2.18			
				O3′—*NE*2	85	3.29	3.00			
O5′—NH1	40	3.06		O5′—NH2	40	3.06				
N7—N	38	2.96	2.95	N7—N	38	2.90	3.01	N7—N	38	2.88
O6—N	39	2.71	2.68	O6—N	39	2.73	2.72	O6—N	39	2.83
—N	40	2.87	2.83	—N	40	2.80	2.70	—N	40	2.66
N1—*OE*1	41	2.67	2.59	N1—*OE*1	41	2.70	2.42	N1—*OE*1	41	2.63
N2—*OE*1	41	3.07		N2—*OE*2	41	2.85	2.76			

Asp 25*B*, Pro 27*B*, Gln 32*B*, Gln 38*B*, Arg 40*B*, Glu 41*B*, Ser 42*B*, Tyr 49*B*, Glu 74*B*, Thr 76*B* and Gln 77*B*. Each of them certainly needs careful inspection on the graphics system before accepting the final model.

The nucleotide complexes

For the EMBL 2′-GMP complex the final coordinate set contains all 745 non-hydrogen protein and 24 inhibitor atoms for both molecules, 492 water molecules, and the five atoms of a sulfate anion lying close to Arg 63*A*. There are two conformations for the sulfur atoms of Cys 72 in both the *A* and *B* molecule, and some side-chain atoms of Val 6*A*, Arg 40*A* and Gln 77*A*.

Superposition of main-chain atoms of EMBL and UCLA 2′-GMP complexes gives a root-mean-square difference of 0.098 Å, with a maximum of 0.8 Å for the poorly ordered carbonyl oxygen of Ser 3*B*. The two complexes are essentially identical and further discussion is restricted to the EMBL 2′-GMP complex, with the exception of the pucker of the ribose ring.

Comparison of the protein in the 2′-GMP complex with the native structures does not show any significant differences in the overall fold as indicated by the root-mean-square deviations in main- and side-chain atomic positions after superimposing all C*A* atoms by least squares, Figs. 7(*c*) and 7(*d*). These should be compared with Figs. 11(*c*) and 11(*d*) of Sevcik *et al.* (1991). The deviations in the main-chain positions are all less than 0.5 Å. The largest movements occur around residues His 85 and Tyr 86 which are involved in nucleotide binding. Large side-chain deviations are restricted to residues involved in binding of nucleotide or those showing considerable disorder. These are somewhat larger in the *B* molecule than was the case for 3′-GMP where molecule *B* was unliganded. The protein is in a similar conformation in the 3′-GMP complex (data now shown).

Conformational differences between protein molecules *A* and *B* for 2′-GMP and for native models, are substantially larger than between analogous molecules (see Figs. 7*a*–7*d*) and seem to arise mainly from crystal-packing interactions. They are caused to a somewhat lesser extent by different modelling of disordered residues and hardly at all by binding of the inhibitor. This is discussed more fully in Sevcik *et al.* (1991).

The electron density for both 2′-GMP molecules is very clear for all atoms, Fig. 8(*b*). The 2′-GMP molecules form ten potential hydrogen bonds with molecule *A* and twelve with molecule *B*, with lengths within 2.5–3.5 Å (Table 6). The different number of nucleotide-to-protein hydrogen bonds for molecules *A* and *B* is due to different positions of the phosphate groups in the two molecules and slight differences in conformation of residues making contacts with the phosphate groups. Comparison of the 2′-GMP and associated amino-acid residues in molecules *A* and *B* is shown in Fig. 10(*a*). Slight differences in the position of the phosphate group and some of the amino-acid residues indicate flexibility in the structure, and reflect the different crystalline environments for the two sites. The lower occupancy of the 2′-GMP on molecule *B* presumably reflects more restricted access of inhibitors *etc.* by diffusion to this site in the crystal.

In the 3′-GMP complex, density for the inhibitor was found at the active site of molecule *A* only. It was proposed that the phosphate of the 3′-GMP could not fit into the active site of molecule *B* because of the close proximity of the neighbouring

molecule *A*, as the *OD*1 atom of Asn 20*A* formed a hydrogen bond with the NH2 atom of Arg 64*B*. In constrast the neighbouring *A* molecule presents no such obstacle to the binding of 2′-GMP as the distance between the base and the phosphate is about 1 Å shorter than in 3′-GMP and its phosphate group binds in a somewhat different position, see below.

Comparison of positions of 2′-GMP and 3′-GMP based on a least-squares overlap of all C*A* atoms, Fig. 10(*b*), shows only very small differences in the positioning of the base due to the essentially identical network of hydrogen bonds formed with the enzyme. It proves that there is a set of specific interactions between the enzyme and the base which unambiguously confers the guanine specificity on the enzyme.

In the *B* molecule the *OE*1 and *OE*2 atoms of the Glu 41*B* side chain form hydrogen bonds with the N1 and N2 atoms of the base. The hydrogen-bonding network formed between the enzyme and the base is shown in Fig. 11. This is in constrast to the situation in molecule *A* where, due to steric hindrance, the carboxyl group of Glu 41*A* is constrained to a position approximately perpendicular to the base, allowing only the *OE*1 atom to bind, Fig. 10(*a*). Therefore, the situation in molecule *B* is probably most like that for a complex in free solution.

Arg 40 is well placed to form a 'lid' covering the guanine base bound to the active site, analogous to the situation in RNase T_1 where the guanosine base is sandwiched between two tyrosine residues (Martinez-Oyanedel, Hui-Woog, Heinemann & Saenger, 1991). Arg 40 has high *B* factors in both molecules and for Arg 40*A* two conformations were refined. The first (open) conformation is identical to that found in the native structure and in the 3′-GMP complex (Fig. 10*b*), where it is bound to the neighbouring protein molecule and points away from the active site. The second (closed) conformation is like that adopted by Arg 40*B* in the *B* molecule, Fig. 10(*a*), and the arginine side chain lies parallel to the 2′-GMP base at a separation of 3.5 Å which is the same as that between the base and the aromatic residues Phe 37*A* and Tyr 86*A* which form the bottom of the binding side cavity, Fig. 8(*c*). This conformation was not observed in the 3′-GMP complex, possibly because of the lower accuracy of the analysis. The closed conformation of Arg 40*A* is not visible in the electron density for the UCLA 2′-GMP complex, probably reflecting the lower overall occupancy of the inhibitor. A rigid conformation of Arg 40 would prevent the substrate base from entering the binding site, and its flexibility is of functional importance. In molecule *A* the Arg 40 side chain fluctuates between two apparently energetically equivalent positions, therefore its disorder is much higher than that of the base. In the *B* molecule the occupancy of 2′-GMP is lower than in the *A* molecule, which should result in greater disorder of the Arg 40*B* side chain and indeed the thermal param-

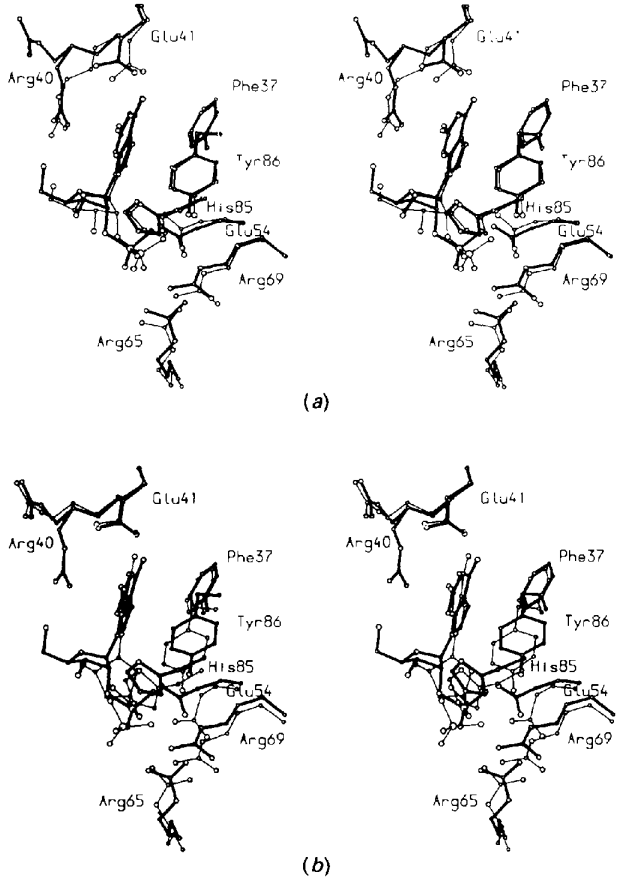


Fig. 10. Stereoviews of superimposed nucleotides bound to RNase Sa. The protein environment of the nucleotides is also shown. (*a*) The 2′-GMP in molecule *A* (thick line) and in molecule *B* (thin line). (*b*) The EMBL molecule *A* 2′-GMP (thick line) overlapped with the 3′-GMP (thin line, coordinate set 2SAR). The superpositions were based on the least-squares fit of all the C*A* atoms in the relevant molecules.

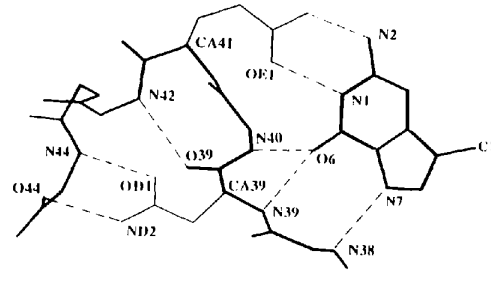


Fig. 11. The hydrogen-bonding network formed between the enzyme and guanine base of the EMBL 2′-GMP inhibitor in the *B* molecule.

eters of Arg 40*B* are higher than those of the two individual Arg 40*A* conformations.

Thus the bases in the 2′- and 3′-GMP complexes lie in extremely similar positions. However, the torsion angles around the glycosyl bonds are distinctly different, with 2′- and 3′-GMP in *syn* and *anti* conformations respectively. In spite of this the enzyme is able to bind 2′-GMP effectively. The ribose ring of the 3′-GMP is buried in the enzme causing conformational changes for residues Ser 31*A*, Gln 32*A* and the side chain of Glu 54*A* in relation to the native enzyme. These changes are not observed in the 2′-GMP complex as the ribose ring of the inhibitor is directed towards the outside of the enzyme molecule and less structural change is necessary for 2′-GMP to bind to the enzyme. The largest changes caused by 2′-GMP binding are observed in the positions of residues His 85 and Tyr 86 which move towards the phosphate group in order to bind to it. These two residues are very flexible and show a degree of disorder in the native structure also.

The puckers adopted by the ribose moiety in the *A* and *B* molecules in EMBL 2′-GMP and the *A* molecule in the UCLA 2′-GMP are essentially identical. They are not ideal C(3′)-*endo* conformations, but closely similar, with the C2′ atom lying slightly above the plane of the C1′, O4′ and C4′ atoms rather than somewhat below as in the ideal conformation (Fig. 12*a*). The torsion angle around the glycosyl link is in the *syn* conformation, in contrast to the 3′-GMP complex where a C(2′)-*endo* pucker and the *anti* conformation was adopted (Fig. 12*b*). Indeed the riboses lie in distinctly different positions in the 2′- and 3′-GMP complexes (Fig. 10*b*). The pucker for the 2′-GMP in the *B* molecule of the UCLA model refines much closer to the ideal C(3′)-*endo* conformation (Fig. 12*c*). This is almost certainly a result of the low occupancy ($\frac{1}{2}$) and overall higher *B* factor for this structure, rather than a real structural difference. The occupancy of 2′-GMP is 1 in both *A* molecules, and $\frac{2}{3}$ in the EMBL *B* molecule. This again emphasizes the need of accurate high-resolution data for such detailed analyses.

The binding of the phosphate group is not so specific as the binding of the base. The phosphate binding ligands are the side chains of Glu 54, Arg 65, Arg 69, His 85 and Tyr 86 (Fig. 13), which lie in a relatively flexible part of the structure equipped with a number of potential binding sites capable of

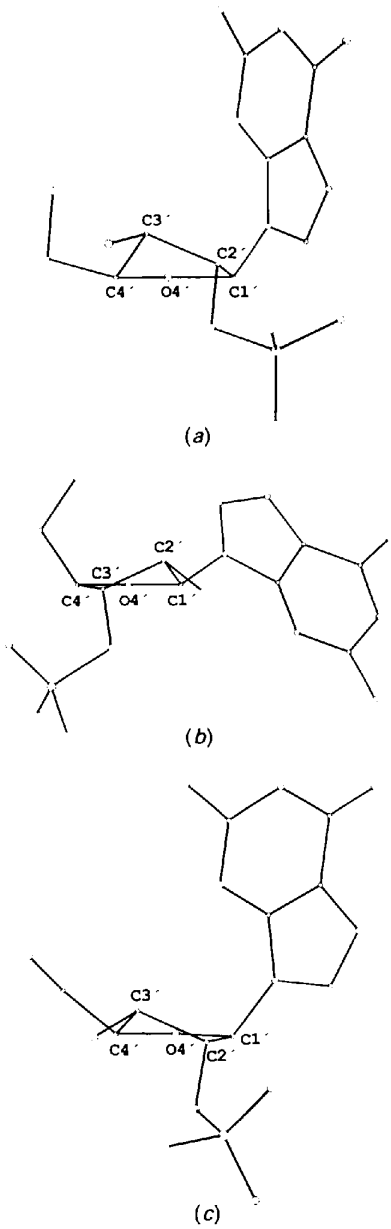


Fig. 12. The nucleotide molecules oriented so as to show the pucker of the ribose rings: (*a*) EMBL 2′-GMP molecule *A*, (*b*) York 3′-GMP and (*c*) UCLA 2′-GMP molecule *B*.

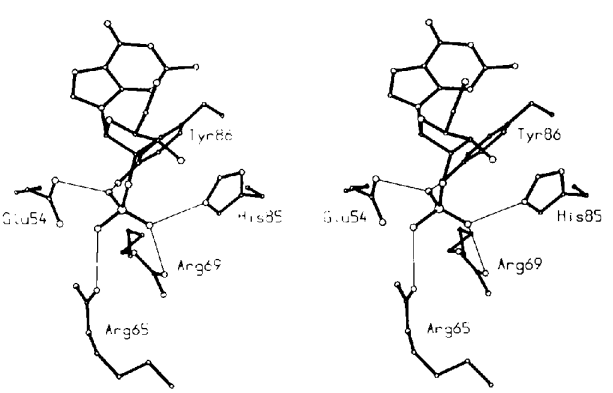


Fig. 13. Binding of the 2′-GMP phosphate group to molecule *A*. The hydrogen bonds to the phosphate oxygens are shown as thin lines.

forming contacts with differently positioned phosphate groups. Indeed the phosphate groups of 3′-GMP and 2′-GMP are bound at locations more than 1 Å apart. In spite of this, both form a number of hydrogen bonds to the enzyme. Glu 54, which is thought to abstract a proton from the 2′-hydroxyl during the first step in catalysis, does not form any hydrogen bonds with the ribose ring, a phenomenon surprisingly observed also in the structure of the 3′-GMP complex. Perhaps it is to be expected that neither the product of the reaction (3′-GMP) nor the inhibitor (2′-GMP) show a catalytically productive mode of binding in the crystal structure.

The inorganic anion

Throughout the text the inorganic anion bound between two molecules is described as sulfate (Fig. 8*a*). It is present in the native and in the 2′-GMP and 3′-GMP complexes. However, there is no absolute chemical evidence that the central atom of the anion is sulfur, as the buffer in which the crystallization was performed was 0.1 *M* phosphate, in a solution 25% saturated with ammonium sulfate. The mean S/P—O distances in the five structures are almost identical with a mean of 1.50 Å. The ion could therefore be sulfate or phosphate and at the resolution of the present analysis it is not possible to discriminate between the two anions.

Concluding remarks

The present work forms part of a wider study of ribonucleases from closely related strains of *S. aureofaciens*. The ultimate goal is a deeper understanding of the mechanism of action and the structure–function relationship of these enzymes. The chemically synthesized gene of RNase Sa has been obtained and expressed. In addition the ribonuclease genes from two different strains of *S. aureofaciens* coding for RNase Sa2 (Nazarov *et al.*, manuscript in preparation) and RNase Sa3 (Homerova *et al.*, manuscript in preparation) have been isolated and sequenced. RNases Sa, Sa2 and Sa3 represent an *S. aureofaciens* subgroup within the microbial ribonucleases and are the subjects of a comparative study. Comparison of the deduced amino-acid sequences of RNases Sa2 and Sa3 to the amino-acid sequence of RNase Sa shows that there are 53 (55%) and 65 (68%) identical amino-acid residues, respectively. All 17 amino-acid residues conserved in bacterial ribonucleases are also present in these two sequences. Amino-acid residues which differ from those of RNase Sa are generally localized on the surface of the molecule. It is expected that the enzymatic properties and tertiary structures of these three enzymes will be similar. The Sa2 gene has been expressed in *Escherichia coli* with reasonably high yield and the properties of the recombinant material are currently under investigation. Expression of the Sa3 gene is underway. This should shed light on why these closely related bacterial strains produce such different ribonucleases. The crystallization and three-dimensional study of Sa2 and Sa3 are planned.

The structures of RNase Sa complexes with 2′- and 3′-GMP explain the basis of guanine specificity but do not provide direct evidence for the expected role of Glu 54 in catalysis as this residue is not in direct contact with the ribose. However, given the flexibility of this enzyme and its substrate, it seems probable that the Glu 54 side chain may move to form a hydrogen bond with the 2′-hydroxyl in the active complex. The flexibility in the phosphate binding site observed between 2′- and 3′-GMP complexes is consistent with the expected movement of the phosphate group during catalysis. Preparation of complexes of RNase Sa with other ligands, including transition-state analogues, is underway and it is hoped that this type of structural work combined with selective point mutations of the enzyme will clarify details of the enzymatic mechanism.*

JS thanks EMBL for a fellowship which facilitated this work.

* Atomic coordinates and structure factors of the EMBL 2′-GMP complex (Reference: 1GMP, R1GMPSF), the native enzyme (Reference: 1GMQ, R1GMQSF) and the UCLA 2′-GMP complex (Reference: 1GMR, R1GMRSF) have been deposited with the Protein Data Bank, Brookhaven National Laboratory, and are available in machine-readable form from the Protein Data Bank at Brookhaven. The data have also been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 37070 (as microfiche). Free copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

References

BAKER, E. N. & DODSON, E. J. (1980). *Acta Cryst.* A**36**, 559–572.

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MAYER, E. F. JR, BRYCE, M. D., RODGERS, J. R., KENNARD, O., SIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

BOTH, V., GALLO, J. & SEVCIK, J. (1991). *Biologia (Bratislava)*, **46**(3), 273–275.

BOTH, V., MOISEYEV, G. P. & SEVCIK, J. (1991). *Biochem. Biophys. Res. Commun.* **177**, 630–635.

DAUTER, Z., SIEKER, L. C. & WILSON, K. S. (1992). *Acta Cryst.* B**48**, 42–59.

DAUTER, Z., TERRY, H., WITZEL, H. & WILSON, K. S. (1990). *Acta Cryst.* B**46**, 833–841.

FINDLAY, D., HERRIES, D. G., MATHIAS, A. P., RANBIN, B. R. & ROSS, C. A. (1961). *Nature (London)*, **190**, 781–784.

HAMILTON, W. C. (1965). *Acta Cryst.* **18**, 502–510.

HARTLEY, R. W. (1980). *J. Mol. Biol.* **15**, 355–358.

HILL, C. P. (1989). *Structure and Chemistry of Ribonucleases*, edited by A. PAVLOVSKY & K. POLYAKOV, pp. 171–179. Moscow: Viniti.

HILL, C. P., DODSON, G. G., BORISOVA, S. N., PAVLOVSKY, A. G., POLYAKOV, K. M., STROKOPYTOV, B. V. & SEVCIK, J. (1987). *Metabolism and Enzymology of Nucleic Acids*, Vol. 6, edited by J. ZELINKA & J. BALAN, pp. 317–322. Bratislava: Veda.

HILL, C. P., DODSON, G. G., HEINEMANN, U., SAENGER, W. MITSUI, Y., NAKAMURA, K., BORISOV, S., TISCHENKO, G., POLYAKOV, K. & PAVLOVSKY, S. (1983). *Trends Biochem. Sci.* **8**, 364–369.

JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.

KABSCH, W. & SANDER, C. (1983). *Biopolymers*, **22**, 2577–2637.

KONNERT, J. H. & HENDRICKSON, W. A. (1980). *Acta Cryst.* A**36**, 344–350.

MARIANI, C., BEUCKELEER, M. D., TRUETTNER, J., LEEMANS, J. & GOLDBERG, R. B. (1990). *Nature (London)*, **347**, 737–741.

MARTINEZ-OYANEDEL, J., HUI-WOOG, C., HEINEMANN, U. & SAENGER, W. (1991). *J. Mol. Biol.* **222**, 335–352.

PACE, C. N., HEINEMANN, U., HAHN, U. & SAENGER, W. (1991). *Angew. Chem. Int. Ed. Engl.* **30**, 343–360.

RAMAKRISHNAN, C. & RAMACHANDRAN, G. N. (1965). *Biophys. J.* **5**, 909–933.

READ, R. J. (1986). *Acta Cryst.* A**42**, 140–149.

SEVCIK, J., DODSON, E. J. & DODSON, G. G. (1991). *Acta Cryst.* B**47**, 240–253.

SEVCIK, J., SANISHVILI, R. G., PAVLOVSKY, A. G. & POLYAKOV, K. N. (1990). *Trends Biochem. Sci.* **15**, 158–162.

SHLYAPNIKOV, S. U., BOTH, V., KULIKOV, V. A., DEMENTIEV, A. A., SEVCIK, J. & ZELINKA, J. (1986). *FEBS Lett.* **209**, 335–339.

TAKAHASHI, K. (1970). *J. Biochem. (Tokyo)*, **67**, 833–839.

TAKAHASHI, K. & MOORE, S. (1982). *The Enzymes*, 3rd. ed., Vol. 15, *Nucleic Acids*, Part B, pp. 435–467. New York: Academic Press.

WILSON, A. J. C. (1942). *Nature (London)*, **150**, 151–152.

ZELINKOVA, E., BACOVA, M. & ZELINKA, J. (1971). *Biochim. Biophys. Acta*, **235**, 342–344.