Acta Cryst. (1995). D**51**, 767–771

Structure of the Crystalline Complex of Deoxycytidylyl-3′,5′-guanosine (3′,5′-dCpdG) Cocrystallized with Ribonuclease at 1.9 Å Resolution

BY J. N. LISGARTEN, D. MAES AND L. WYNS

Department of Ultrastructure, Instituut voor Moleculaire Biologie, Vrije Universiteit Brussel, B-1640 Sint-Genesius-Ride, Belgium

AND C. F. AGUILAR AND R. A. PALMER*

Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, England

(*Received* 29 *July* 1994; *accepted* 1 *February* 1995)

Abstract

The X-ray structure of bovine ribonuclease A cocrystallized with the dinucleotide deoxycytidylyl-3′,5′-guanosine has been determined at 1.9 Å resolution and refined by restrained least squares to $R = 0.218$ for 7807 reflections. The structure established that the recently observed retrobound mode of attachment of substrate analogues cytidylyl-2′,5′-guanosine and deoxycytidylyl-3′,5′-guanosine found in soaked RNase A crystals is also present in the cocrystallized complex. Retrobinding is thus unlikely to be the result of restrictions imposed by the crystalline environment as the ligands soak into the lattice but rather a phenomenon specific to small nucleotides containing guanine.

Introduction

Bovine pancreatic ribonuclease A (RNase A; E.C. 3.1.27.5) is one of the most intensively studied enzymes (see for example the review of Blackburn & Moore, 1982). Its two-step catalytic cleavage of the phosphodiester bond of ribonucleotides is understood in some detail (Richards *et al.*, 1971; Richards & Wyckoff, 1973; Wodak, Liu & Wyckoff, 1977; Pavlovsky, Borisova, Borisov, Antonov & Karpeisky, 1978).

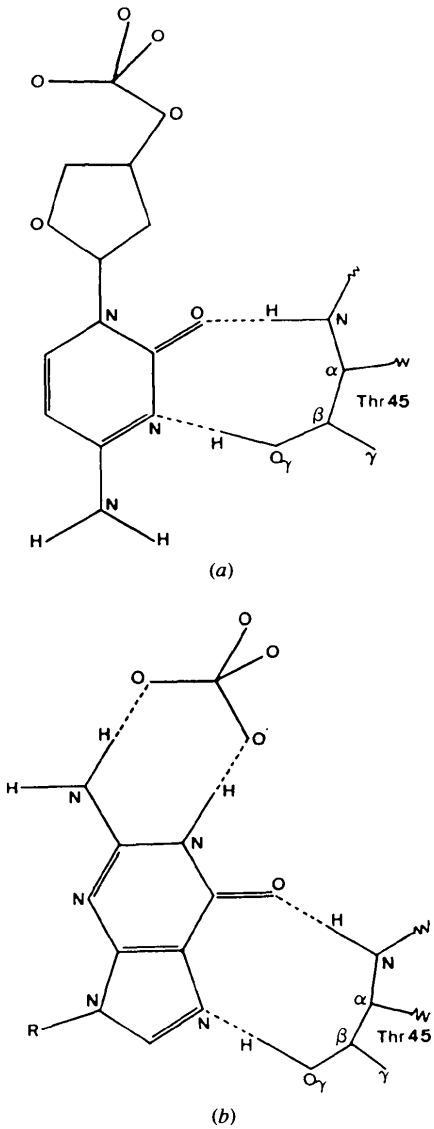
Recently, a novel form of nucleotide binding to RNase has been observed in high-resolution studies of complexes formed with the inhibitors cytidylyl-2′,5′-guanosine (2′-5′-CpG) and deoxycytidylyl-3′,5′-guanosine (3′,5′-dCpdG), respectively (Aguilar, Thomas, Moss, Mills & Palmer, 1991; Aguilar, Thomas, Mills, Moss & Palmer, 1992). When these dinucleotide substrate analogues were soaked into RNase crystals they were found to bind in a non-productive spatial disposition with the guanine moiety occupying the site which had previously been identified (Richards, Wyckoff & Allewell, 1970) as involved in the recognition of pyrimidine bases of the substrate (see Fig. 1). In this retrobinding mode the phosphate group of the inhibitor molecule does not interact with any of the residues associated with the active site of the enzyme: His12, His119, Lys41 or Glu11 (see for example, Lisgarten *et al.*, 1993). Instead, the phosphate protrudes into the intermolecular solvent region, pointing directly away from both the active site and the pyrimidine base binding site occupied by Thr45. The guanine base of the retrobound inhibitors 2′,5′-CpG and 3′,5′-dCpdG forms a perfect hydrogen-bonded fit with Thr45, the side chain normally involved in the productive pyrimidine base-recognition mechanism of RNaseA, a bound solvent SO_4^{2-} anion and a bound water molecule, and thus forms a very stable complex occupying the pyrimidine base binding site. From the guanine base onwards the dinucleotide moieties form fewer interactions with the protein and consequently become progressively disordered. Mechanically, RNase cleaves single-stranded RNA at the 3′ end of pyrimidine nucleosides, highly specifically, Thr45 being responsible for anchoring the base in position during this process. It is difficult to reconcile the perfect fit of the guanine base in the Thr45 pyrimidine-base binding site which is observed in retrobinding with the notion that it is fortuitous. If there is any biological significance for retrobinding it is possibly a mechanism whereby a distant part of a substrate RNA molecule is constrained while undergoing enzymatic cleavage by a neighbouring RNase molecule. However, since the two dinucleotide complexes studied previously by Aguilar *et al.* (1992) were formed by soaking pre-grown native monoclinic crystals in the inhibitor solutions, it is conceivable that retrobinding is brought about by serendipitous juxtapositioning of the guanine base in the Thr45 pocket as the inhibitor molecules infiltrate through the crystal solvent channels. As shown recently by Lisgarten *et al.* (1993) the smaller mononucleotide inhibitor 2′-CMP does bind productively with RNase in soaked crystal complexes. It is known however that di- and trinucleotides damage the crystals during soaking experiments, and this effect may be caused by disruption of existing solvent-channel structures (Perry, 1987).

* Author for correspondence.

In order to determine whether the retrobinding mode is an artifact resulting from the technique of soaking the ligand into the crystal we have undertaken a crystallographic analysis of the cocrystallized dCpdG–RNase A complex, the results of which are reported here.

Experimental

Large crystals, 0.5–1.0 mm in linear dimensions, using RNase A (ex-Sigma) were grown in a similar manner to that previously used to produced native RNase A crystals



(*a*)

(*b*)

Fig. 1. (*a*) Hydrogen-bonding arrangement at Thr45 in the *B*1 site of RNaseA for cytosine in the postulated (Richards *et al.*, 1971) productive binding mode, and (*b*) guanine in retrobinding mode, as observed in the present study and by Aguilar *et al.* (1991) in studies on dinucleotide substrate analogues (2′,5′-CpG and 3′,5′-dCpdG) soaked into RNase crystals.

Table 1. *X-ray data and refinement statistics for the RNaseA–3′,5′-dCpdG cocrystallized complex*

(*a*) X-ray data

	Native RNase A	RNase A–3′,5′-dCpdG
Space group	$P2_1^*$	$P2_1$
Cell constants		
a (Å)	30.3	30.5
b (Å)	38.3	38.7
c (Å)	52.9	53.7
β (°)	105.9	106.3
Wavelength λ(Cu *Kα*) (Å)	1.542	1.542
Total number of measured reflections		17427
Number of unique reflections†		7807
Merging *R* factor‡ (intensities)		

R_{merge}	Resolution (Å)
0.059	All data
0.041	6.71
0.038	4.74
0.044	3.87
0.045	3.35
0.057	3.00
0.068	2.74
0.065	2.54
0.107	2.37
0.081	2.24
0.097	2.12
0.110	2.02
0.114	1.94

Completeness of data to 1.94 Å (%)	89
Completeness of data 2.54–1.94 (%)	85
Completeness of data 2.12–1.94 (%)	76

(*b*) Refinement

R.m.s. deviation from target bond lengths (Å)	0.023 (distances <2.12 Å)
R.m.s. deviation from target angle distances (Å)	0.058 (distances in range 2.12–2.62 Å)
R.m.s. deviation from non-bonded target distances	0.062 (distances >2.62 Å)
R.m.s. distance from least-squares main-chain planes (Å)	0.084
R.m.s. deviation from least-squares side-chain planes (Å)	0.024
Mean coordinate error § (Å)	0.15
Resolution range (Å)	20.0–1.9
R factor (all data)	0.218
Correlation coefficient (final)	0.949
Number of cycles	110
Mean U_{iso}'s ¶	
Enzyme (Å)	0.288 (951)
Inhibitor (Å)	0.512 (28)
Waters (Å)	0.468 (59)

* *International Tables for X-ray Crystallography* (1974, Vol. IV).

† All unique data used throughout the analysis with no σ cutoff.

‡ $R_{merge} = \sum_h \sum_i |I_{hi} - \langle I_h \rangle| / \sum_h N_h \langle I_h \rangle$, where h = reflection number, i = observation number and N_h = number of observations of reflection h.

§ Estimated from a Luzzati plot (Luzzati, 1952).

¶ Number of atoms is shown in parentheses.

(Carlisle, Palmer, Mazumdar, Gorinsky & Yeates, 1974) but with an approximately fourfold molar excess of the dinucleotide dCpdG present. Full details of the crystallization procedure have been published elsewhere (Mills *et al.*, 1992). The cocrystals are monoclinic with unit-cell dimensions $a = 30.47(7)$, $b = 38.72(7)$, $c = 53.68(8)$ Å, $\beta = 106.3(2)°$, $P2_1$, $Z = 2$ and are isomorphous with the native $P2_1$ form. A suitable crystal for diffractometry was mounted in a capillary tube with its monoclinic *b* axis parallel to the length of the tube. Using only one crystal of the complex, X-ray intensity data

were collected by FAST diffractometery (Arndt, 1985) to 1.9 Å resolution using the *MADNES* software (Messerschmidt & Pflugrath, 1987). The crystal orientation was refined prior to and during data collection, allowing reflections to be integrated and Lp corrected online. Data-collection details are given in Table 1(*a*). Merging of equivalents was performed with the Fox & Holmes (1966) algorithm by dividing the data into 5° batches on the scan angle (ω), analogous to using separate films in the photographic method. The resulting measured structure-factor amplitudes $|F_{PI}|$, were combined with $|F_c|$ values derived from native ribonuclease A coordinates (Borkakoti, Moss & Palmer, 1982; Borkakoti, Palmer, Haneef & Moss, 1983).

Luzzati-weighted coefficients $(2m|F_{PI}| - D|F_c|)$ (Read, 1986) were prepared by the program *SIMWT* (Tickle, 1988), which also performs the derivative-to-native scaling. Prior to calculating difference electron-density maps, all solvent water molecules and the protein side chains Lys7, Lys41, His12, His119, Phe120 and Thr45, in the vicinity of the active site and known to be highly labile, were removed from the native atom coordinate data set. The solvent and active-site residues were subsequently remodelled and refined. Difference Fouriers were calculated by the program *FFT* (Ten Eyck, 1973) and displayed using *FRODO* (Jones, 1978) on an Evans & Sutherland PS390 raster graphics terminal. The initial maps allowed the inhibitor structutre to be built in prior to least-squares refinement, which was performed with the program *RESTRAIN* (Haneef, Moss, Stanford & Borkakoti, 1985) using Vax and Convex computers. The structure of the RNase–inhibitor complex was subjected to several rounds of rebuilding and restrained least squares. A total of 59 solvent water molecules were modelled from the electron density and successfully included in the refinement. The results of the refinement procedure are given in Table 1(*b*). The geometrical restraints libraries used by *FRODO* and *RESTRAIN* were extended to accommodate the geometrical features of the inhibitor molecule.

Results and discussion

Fig. 2 is an electron-density map showing the position of the 3′,5′-dCpdG molecule as it occurs bound to the RNase molecule in the cocrystallized intermolecular complex. Fig. 3 is a stereoview of the cocrystallized RNase 3′,5′-dCpdG complex superimposed onto the soaked RNase 3′,5′-dCpdG complex.

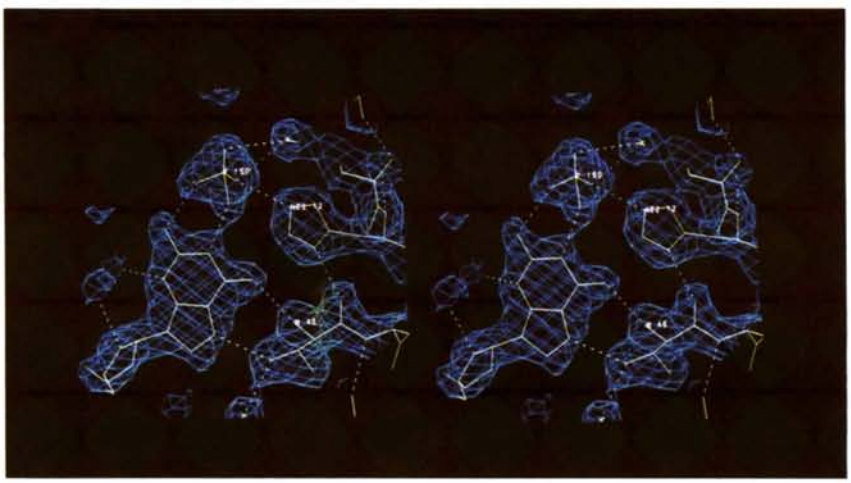


Fig. 2. Stereoview of the electron density ($2|F_o| - |F_c|$ map contoured at 0.6 e Å^{-3}) at the active-site region in the RNaseA–3′,5′-dCpdG cocrystallized complex.

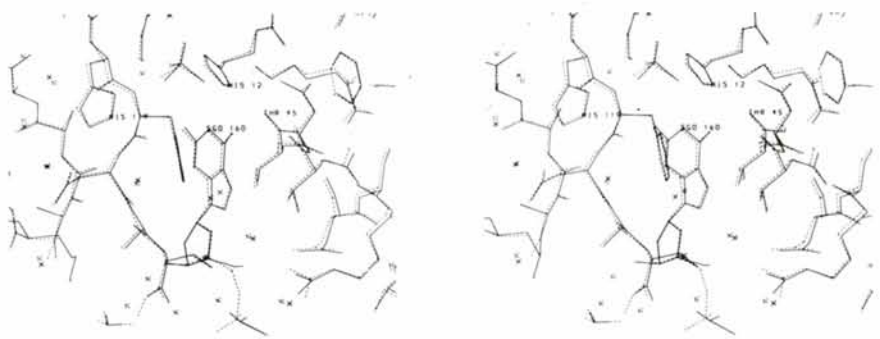


Fig. 3. Stereoview of the RNaseA–3′,5′-dCpdG retrobound complexes superimposed showing the active-site region. Dotted lines, represent the 'soaked' dCpdG complex.

Table 2. *Close contacts and torsion angles for the active-site residues in the RNaseA–3′,5′-dCpdG complexes*

(*a*) Close contacts (< 3.5 Å) between atoms of the active-site residues and the dCpdG inhibitor with equivalent distances in 'soaked' complex shown in parentheses.

			Distance (Å)
O1P	150 (anion)	OE1 Gln11	3.09 (2.98)
O2P	150 (anion)	O500	3.35 (2.96)*
O2P	150 (anion)	NE2 His12	3.03 (2.71)*
O2P	150 (anion)	N Phe120	2.82 (3.17)*
O2′	150 (anion)	NZ Lys41	2.83 (2.90)
O2′	150 (anion)	N1 160 (guanine)	2.72 (2.76)
O5′	150 (anion)	ND1 His119	2.57 (2.69)
O5′	150 (anion)	N2 160 (guanine)	3.27 (3.16)
O500		OE1 Gln11	3.01 (3.12)
NE2	Gln11	NZ Lys41	2.96 (3.06)
O6	160 (guanine)	N Thr45	3.04 (2.91)
N7	160 (guanine)	OG1 Thr45	2.76 (2.77)
NE2	His119	OD1 Asp121	2.65 (2.70)
ND1	His12	O Thr45	2.74 (2.72)
NZ	Lys41	ODI Asn44	2.94 (2.71)
N3	160 (gua)	O537	2.99 (3.19)

(*b*) Torsion angles (°) for the active-site residues in the RNase A–3′,5′-dCpdG cocrystallized (I) and soaked (II) structures

	$-\chi_1-$†		$-\chi_2-$†		$-\chi_2^1-$†		$-\chi_3-$†		$-\chi_4-$†	
	(I)	(II)	(I)	(II)	(I)	(II)	(I)	(II)	(I)	(II)
His12	−65	−57*	−67	−70	114	110				
His48	−70	−69	−72	−81*	106	99*				
His105	−57	−52	−99	−96	82	82				
His119	142	160*	84	72*	−95	−106*				
Lys41	−161	−162			177	−168*	158	164*	171	171

* Indicates a significant difference.

† $\chi_1 = N-C_\alpha-C_\beta-C_\gamma$, $\chi_2 = C_\alpha-C_\beta-C_\gamma-N_\delta$, $\chi_2^1 = C_\alpha-C_\beta-C_\gamma-C_\delta$, $\chi_3 = C_\beta-C_\lambda-C_\delta-C_\varepsilon$ and $\chi_4 = C_\lambda-C_\delta-C_\varepsilon-N_\zeta$.

Inspection of Fig. 2 clearly shows the guanine base retrobound at the *B*1 site, hydrogen bonded to Thr45 and to the sulfate/phosphate anion which has remained undisplaced between His12 and His119, the remainder of the dinucleotide displaying increasing disorder as the cytidylyl moiety extends across the solvent channel to an adjacent protein molecule. This was also found to occur in pre-soaked native RNaseA crystals containing guanine substrate analogues 2′,5′-CpG; 3′,5′-dCpdG (Aguilar *et al.*, 1991; Aguilar *et al.*, 1992) and 8-oxo-guanosine-2′-phosphate (O8-2′GMP) (Thomas, 1992), but not in similarly treated crystals containing the modified cytidylyl monophosphate bases cytidine N(3)-oxide-2-phosphate (N3O-CMP) (Palmer, Moss, Haneef & Borkakoti, 1984) and cytidine-2′-monophosphate (2′-CMP) (Howlin, Harris, Moss & Palmer, 1987; Lisgarten *et al.*, 1993).

Inspection of Fig. 3 reveals a marked similarity between the active-site regions of the cocrystallized RNaseA–3′,5′-dCpdG complex and the soaked RNaseA–3′,5′-dCpdG complex. Table 2 parts (*a*) and (*b*) giving the active-site inhibitor–protein close contacts and torsion angles, for both complexes also serve to establish the close geometrical similarity between soaked and cocrystallized structures. The only significant difference between the two structures appears to be a repositioning of the catalytically important residue His119. Considering the well known lability of this residue, this result is not too surprising. For example, in the native structure His119 is disordered in two sites, *A* (80%) and *B* (20%) (Borkakoti *et al.*, 1982, 1983), whereas in the RNaseA–3′,5′-dCpdG complex it exists in a conformation close to native His119*A* with no occupancy for His119*B*.

Comparison of the native monoclinic RNaseA structure with the soaked 3′,5′-dCpdG inhibitor structure and the cocrystallized 3′,5′-dCpdG

A detailed comparison of the refined monoclinic RNase structures, native (Brookhaven Protein Data Bank citation p3rn3.brk, Bernstein *et al.*, 1977), soaked 3′,5′-dCpdG (p3rn3.brk) and cocrystallized 3′,5′-dCpdG (dcpdg-jl.brk) was performed using the program *PROCHECK-NMR* (Laskowski, 1994). This revealed no substantial changes in the structures other than discussed above. Comparison of the main-chain conformations summarized in the three-way ensemble Ramachandran plot given in Fig. 4 (with statistics in Table 4) shows the close correspondence of the three structures. R.m.s. deviations of the atomic coordinates of the three structures are included in Table 3. Details of the water structures of the native and soaked 3′,5′-dCpdG complex structures have been discussed by Aguilar *et al.* (1991) and Aguilar (1991). No significant differences in the water structures of the soaked and cocrystallized 3′,5′-dCpdG complex structures were observed.

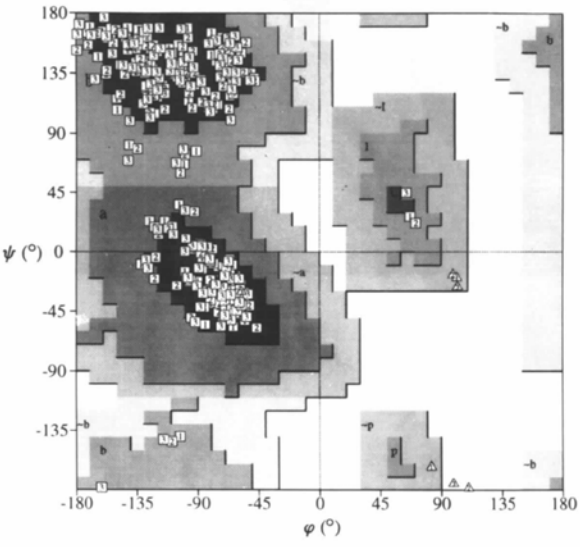


Fig. 4. Ensemble Ramachandran plot for (1) native refined RNaseA structure (Howlin, Harris, Moss & Palmer, 1987) (monoclinic form, p3rn3.brk), (2) soaked 3′,5′-dCpdG complex (Aguilar *et al.*, 1992) (dcpdg.brk) and (3) cocrystallized 3′,5′-dCpdG complex (this structure) (dcpdg-jl.brk). For plot statistics see Table 4.

Table 3. *R.m.s. deviations* (Å) *of atomic coordinates for the structures* (1) *native* (*p3rn3.brk*), (2) *soaked complex with* 3′,5′-*dCpdG* (*dcpdg.brk*) *and* (3) *cocrystallized complex with* 3′,5′-*dCpdG* (*dcpdg-jl.brk*)

	Main chain/main chain	Side chain/side chain	All
(1)–(2)	0.091	0.428	0.305
(1)–(3)	0.127	0.539	0.387
(2)–(3)	0.106	0.345	0.253

Table 4. *Plot statistics for Ramachandran plot*

Based on an analysis of 118 structures of resolution of at least 2.0 Å and *R* factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Residues in most favoured regions [*A*, *B*, *L*] (%)	303	87.8
Residues in additional allowed regions [*a*, *b*, *l*, *p*] (%)	42	12.2
Residues in generously allowed regions [~ *a*, ~ *b*, ~ *l*, ~ *p*] (%)	0	0.0
Residues in disallowed regions (%)	0	0.0
Number of non-glycine and non-proline residues	345	100.0
Number of end-residues (excl. Gly and Pro)	6	
Number of glycine residues (shown as triangles)	9	
Number of proline residues	12	
Total number of residues	372	

Concluding remarks

The observations reported here suggest that retrobinding to RNase is a phenomenon that occurs in complexes with mono- and dinucleotides containing guanine. Further, the results of the present investigation establish unequivocally that retrobinding can occur in complexes obtained by cocrystallization. In conclusion, this study has demonstrated that this mode of binding is not a phenomenon associated exclusively with the process of soaking RNase crystals with inhibitor solutions but is probably a function of some recognition mechanism as yet to be fully characterized.*

Dr D. Maes acknowledges support by the National Fund for Scientific Research (NFWO), Belgium. The authors acknowledge the receipt of NATO grant No. 900 270, which greatly facilitated completion of this study.

* Atomic coordinates have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1RCA). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: LI0191).

References

AGUILAR, C. F. (1991). PhD thesis, Univ. of London, England.

AGUILAR, C. F., THOMAS, P. J., MILLS, A., MOSS, D. S. & PALMER, R. A. (1992). *J. Mol. Biol.* **224**, 265–267.

AGUILAR, C. F., THOMAS, P. J., MOSS, D. S., MILLS, A. & PALMER, R. A. (1991). *Biochim. Biophys. Acta*, **1118**, 6–20.

ARNDT, U. W. (1985). *Methods Enzymol.* **114**, 472–485.

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F. JR, BRICE, M. D., ROGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

BLACKBURN, P. & MOORE, S. (1982). *The Enzymes*, 3rd ed. Vol XV, edited by P. D. BOYER, pp. 317–433. New York: Academic Press.

BORKAKOTI, N., MOSS, D. S. & PALMER, R. A. (1982). *Acta Cryst.* B**38**, 2210–2217.

BORKAKOTI, N., PALMER, R. A., HANEEF, N. & MOSS, D. S. (1983). *J. Mol. Biol.* **169**, 743–755.

CARLISLE, C. H., PALMER, R. A., MAZUMDAR, S. I., GORINSKY, B. A. & YEATES, G. R. (1974). *J. Mol. Biol.* **85**, 1–18.

FOX, G. C. & HOLMES, K. C. (1966). *Acta Cryst.* **20**, 886–891.

HANEEF, I., MOSS, D. S., STANFORD, M. J. & BORKAKOTI, N. (1985). *Acta Cryst.* A**41**, 426–433.

HOWLIN, B., HARRIS, G. W., MOSS, D. S. & PALMER, R. A. (1987). *J. Mol. Biol.* **196**, 159–164.

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

LASKOWSKI, R. (1994). *PROCHECK-NMR. Protein geometry checking program*. Personal communication.

LISGARTEN, J. N., GUPTA, V., MAES, D., WYNS, L., ZEGERS, I., PALMER, R. A., DEALWIS, C. G., AGUILAR, C. F. & HEMMINGS, A. M. (1993). *Acta Cryst.* D**49**, 541–547.

LUZZATI, V. (1952). *Acta Cryst.* **5**, 802–810.

MESSERSCHMIDT, A. & PFLUGRATH, J. W. (1987). *J. Appl. Cryst.* **20**, 306–315.

MILLS, A., GUPTA, V., SPINK, N., LISGARTEN, J., PALMER, R. A. & WYNS, L. (1992). *Acta Cryst.* B**48**, 549–550.

PALMER, R. A., MOSS, D. S., HANEEF, I. & BORKAKOTI, N. (1984). *Biochim. Biophys. Acta*, **785**, 81–88.

PAVLOVSKY, A. G., BORISOVA, S. N., BORISOV, V. V., ANTONOV, I. V. & KARPEISKY, M. Y. (1978). *FEBS Lett.* **92**, 258–262.

PERRY, A. (1987). PhD thesis, Univ. of London, England.

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

RICHARDS, F. M. & WYCKOFF, H. W. (1973). In *Atlas of Molecular Structures in Biology*, Vol 1, *Ribonuclease S*, edited by D. C. PHILLIPS & F. M. RICHARDS. Oxford: Clarendon Press.

RICHARDS, F. M., WYCKOFF, H. W. & ALLEWELL, N. (1970). *Neurosciences Second Study Program*, edited by T. O. SCHMIDT, pp. 901–912. New York: Rockerfeller Univ. Press.

RICHARDS, F. M., WYCKOFF, H. W., CARLSON, W. D., ALLWELL, N. M., LEE, B. & MITSUI, Y. (1971). *Cold Spring Harbour Symp. Quant Biol.* **36**, 35–43.

TEN EYCK, L. F. (1973). *Acta Cryst.* A**29**, 183–191.

THOMAS, P. J. (1992). Personal communication.

TICKLE, I. J. (1988). *Improving Protein Phases*, edited by S. BAILEY, E. DODSON & S. PHILLIPS, pp. 130–137. Warrington, England: SERC Daresbury Laboratory.

WODAK, S. Y., LIU, M. Y. & WYCKOFF, H. W. (1977). *J. Mol. Biol.* **116**, 855–875.