Crystal and Molecular Structure of Yeast Phenylalanyl Transfer RNA. Structure Determination, Difference Fourier Refinement, Molecular Conformation, Metal and Solvent Binding

By C. D. Stout,* H. Mizuno, S. T. Rao, P. Swaminathan, J. Rubin,† T. Brennan‡
AND M. Sundaralingam§

Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin–Madison, Madison, Wisconsin 53706, USA

Our model of the structure of monoclinic yeast tRNA^{Phe} from the 3·0 Å MIR map has been rebuilt through a series of cycles of $2F_o - F_c$ maps omitting oligonucleotide segments of 6–10 residues at a time. The improved model coordinates have been subjected to several rounds of difference synthesis refinement using data at 2·5 Å resolution that have reduced the R value from 0·470 to 0·312 (constrained model) and 0·276 (unconstrained model). The binding sites for the heavy metals Os, Au, Pt and the lanthanides (Sm^{3+},Gd^{3+}) used in the derivative studies and possible sites for some of the cations (Mg^{2+}, Na$^+$ and spermine) and solvents have been identified from the difference electron density maps. The lanthanides show a preference for the anionic phosphate O atoms while the transition metals prefer the base ring atoms. The conformation of the tRNA molecule has now been established more precisely. All the 42 nucleotides of the double-helical sections, including those of the 'wobble' pair G4–U69, display the preferred conformational combination: C(3')-*endo*, *anti* (χ), *gauche*$^+$ (ψ); the backbone phosphodiesters are all within the preferred *gauche*$^-$, *gauche*$^-$ domain. Over half of the remaining 34 nucleotides of the 'single-stranded' regions also display the above preferred conformations. The less-preferred conformations are confined to about 15 residues of the loops and bends which exhibit the C(2')-*endo* sugar pucker (6), the *trans* or *gauche*$^-$ conformation (9) around the C(4')–C(5') bond, and the non-helical phosphodiester conformations (12). The glycosyl conformation of all 76 nucleotide residues is in the *anti* range.

Introduction

The molecular structure of phenylalanine specific tRNA from yeast has been solved independently in three different laboratories. The groups at MIT and Duke have reported their results on the orthorhombic crystal form from an analysis at 3·0 Å and 2·5 Å resolution (Suddath, Quigley, McPherson, Sneden, Kim & Rich, 1974; Quigley *et al.*, 1975; Quigley & Rich, 1976). The MRC group have reported their work on the monoclinic crystal form at 3·0 Å and at 2·5 Å from MIR (multiple isomorphous replacement) maps and also recently a real-space refinement of their model at 2·5 Å resolution (Robertus *et al.*, 1974; Ladner *et al.*, 1975; Jack, Ladner & Klug, 1976). We have been studying the structure of the monoclinic form since 1971 (Ichikawa & Sundaralingam, 1972) and have reported the results of our MIR studies at 5·5 Å (Rubin

et al., 1975), the chain tracing at 4·0 Å (Sundaralingam, Stout, Mizuno, Rubin, Rao & Brennan, 1975a) and the complete structure at 3·0 Å (Sundaralingam *et al.*, 1975b). The atomic coordinates for the orthorhombic (Quigley *et al.*, 1975; Sussman & Kim, 1976) and monoclinic (Ladner *et al.*, 1975; Stout, Mizuno, Rubin, Brennan, Rao & Sundaralingam, 1976) forms have been published. We have also published a conformational analysis of the molecule based on our preliminary coordinates (Sundaralingam, Mizuno, Stout, Rao, Liebman & Yathindra, 1976). Since our experimental procedure and the MIR work have not been previously published in detail, we present this and the results of our refinement here.

Since the publication of our 3·0 Å work, we have rebuilt the entire model using electron density maps with coefficients $2F_o - F_c$, initially at 3·0 Å and later at 2·5 Å resolution (Mizuno, Rao, Stout, Swaminathan & Sundaralingam, 1976), following a novel procedure where we omitted oligonucleotide segments 6–10 residues in length from the phasing calculations and fitted these residues to the resulting density. This procedure biases the phases minimally and, at the same time, enables conformational changes to be made in these residues more easily than in a map calculated

---

* Present address: Department of Crystallography, University of Pittsburgh, Pittsburgh, PA, USA.

† Present address: Medical Research Council Laboratory, Cambridge, England.

‡ Present address: Department of Chemistry, Northeastern University, Boston, Massachusetts, USA.

§ To whom correspondence should be addressed.

including the residues. Using this approach, we went through three rounds of adjustment of the entire model with the optical comparator. This model has now been

Fig. 1. (*a*) Cloverleaf arrangement of the nucleotide sequence of yeast tRNA^{Phe}. The abbreviations are DH = dihydrouridine, AC = anticodon, VL = variable loop and Tψ = pseudouridine. Solid circles indicate invariant bases; dashed circles indicate semi-invariant bases – either purines or pyrimidines. The Watson-Crick base pairs are indicated by dots while the G4–U69 'wobble' base pair in the amino acid stem is indicated by a circle. The seven tertiary base interactions having two or three hydrogen bonds between bases are joined by solid lines and the three having a single hydrogen bond between bases are joined by broken lines. (*b*) A trace of the polynucleotide fold showing the secondary structural base-pairing interactions.

subjected to crystallographic refinement by the difference Fourier technique using a carefully recollected set of data out to 2·5 Å resolution. At the current stage of our refinement, several potential solvent and cation molecules have been identified. Here we report the details of this analysis.

Yeast phenylalanine tRNA consists of 76 nucleotide residues and 1652 atoms, excluding H atoms. A cloverleaf representation of the molecule is given in Fig. 1(*a*) and a backbone trace in Fig. 1(*b*), where the different regions of the molecule are identified.
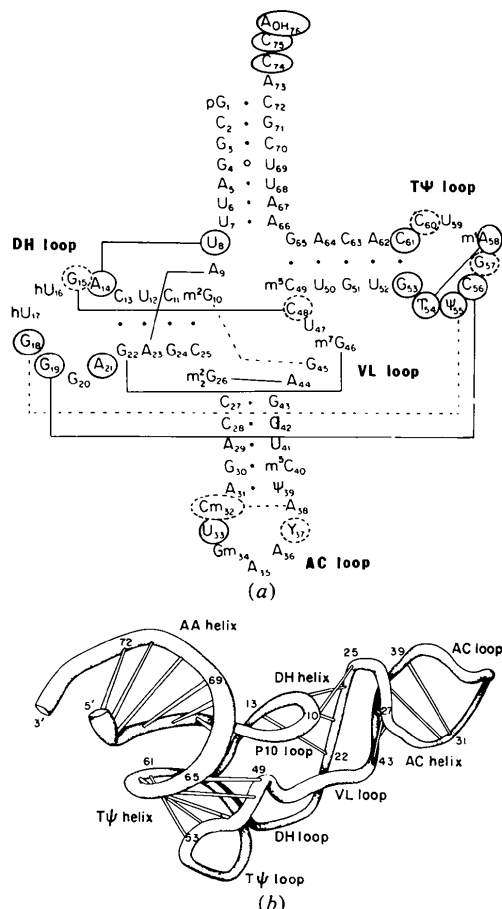
Experimental

Samples of brewer's yeast phenylalanine tRNA were purchased from Boehringer–Mannheim Corporation. Crystallization was carried out in the cold room (8–10 °C) in a buffer of 10 mM MgCl$_2$, 1 mM spermine hydrochloride and 10 mM Na cacodylate at pH 6·0. The details of the conditions used in the preparation of the monoclinic tRNA^{Phe} crystals have already been presented (Rubin *et al.*, 1975). Crystals used in the analysis were typically 0·3 × 0·8 × 0·08 mm. The unit-cell parameters are $a = 55\cdot8$, $b = 33\cdot3$, $c = 63\cdot0$ Å and $\beta = 90\cdot4°$; the space group is $P2_1$ with one molecule in the asymmetric unit.

The derivative crystals were prepared by addition of solutions of the heavy-metal compounds to droplets of the native crystals in the microscope slides. Among the numerous reagents explored (Rubin, 1974), five gave useful derivatives, *viz* Sm, Gd, Pt, Os and Au. A double derivative was also prepared by using the combination of Sm and Os reagents. The Os derivative (Rubin *et al.*, 1975) was made by soaking the crystals in a 25 mol excess of the reagent [OsO$_3$(pyridine)$_2$] for two weeks. The Sm derivative was prepared by gradual addition of samarium acetate over a period of a few days to a final mole ratio of about 2:1. The Gd derivative was prepared similarly by adding gadolinium sulfate. It was found that higher concentrations of lanthanide tended to fracture the crystals. The Sm/Os double derivative was prepared by addition of the Os reagent to the Sm derivative using the same procedure. The Pt derivative (Robertus *et al.*, 1974) was prepared by adding *trans*-Pt(NH$_3$)$_2$Cl$_2$ (supplied by Dr J. Hoeschele, Engelhard

Table 1. *Cell dimensions and other data for native and derivative crystals*

Crystal	a	b	c	β	Number of crystals	Resolution	Number of reflections	r.m.s. $\Delta F/F$
Native	55·8 Å	33·3 Å	63·0 Å	90·4°	3	3·0 Å	3837	–
Sm	56·0	33·2	62·8	90·1	3	3·0	3502	0·187
Gd	56·0	33·3	62·9	90·1	3	3·0	3418	0·223
Pt	55·8	33·3	62·8	89·9	4	3·0	3436	0·213
Os	55·7	33·5	62·9	90·4	2	4·0	1609	0·146
Au	55·7	33·5	63·0	90·0	1	6·0	549	0·189
Sm/Os	55·9	33·4	63·0	90·3	2	4·0	1781	0·151

Industries, Edison, New Jersey) over a three-day period to a final mole ratio of 2·5:1. The Au derivative was prepared by soaking the crystals in a 10 mol excess of NaAuCl$_2$ for about 7 weeks.

The unit-cell dimensions of the derivatives were within 0·2 Å of those of the native crystal (Table 1). The most sensitive parameter was the β angle, which decreased to 90·1° in the lanthanide derivatives and to 89·9° for the Pt derivative. The standard deviations in the cell edges were typically 0·1 Å, and about 0·2° in the angle.

Data collection, processing and scaling

Crystals were mounted for data collection on the *b* axis in thin-walled capillaries plugged with a stabilizing solution consisting of the crystallization buffer and 20% by volume of 2-methyl-2,4-pentanediol (MPD). Intensity data for both the native and derivative crystals were collected on a Picker FACS-I diffractometer equipped with a graphite monochromator and a high-brilliance Cu X-ray tube operated at 40 kV and 30 mA. The programs used to control the diffractometer were those of Lenhert (1975). Reflections were measured with θ–2θ continuous scans of 1·8 to 2·0° with scan rates of 2° min^{-1} for the 12·6 to 4·0 Å data and 1° min^{-1} for the 4·0–3·0 Å data. Background counts were measured on both sides of the scan range for 10 s.

In the X-ray beam, the native crystals were generally stable for 3 to 4 d at 16°C and during this period standard reflections dropped 20–25% of their initial intensity values. The Os- and Pt-containing derivative crystals were somewhat less stable in the X-ray beam, whereas the lanthanides appeared to have no effect on crystal decay. To correct for differential decay during data collection, crystal deterioration was monitored with three to six standard reflections distributed over the 2θ range being scanned and measured every 3 to 4 h. The variation of the intensities of the standard reflections with time were used to correct for decay. Decay was linear with time and essentially uniform throughout the reciprocal lattice. The validity of this method was checked by measuring 55 strong reflections, distributed over the reciprocal space, at the beginning and end of each data-collection session. After application of the decay correction, structure amplitudes for the 55 reflections agreed consistently to within 2–3%.

An empirical absorption correction was made during data collection by measuring the intensity of the 060 reflection at $\chi = 90°$ at 10° intervals in φ (North, Phillips & Mathews, 1968). The curves showed a change in absorption of about 40% from the broad dimension to the thin dimension of the crystal. After application of the absorption and decay corrections and data reduction with Lorentz and polarization factors,

equivalent reflections were found to agree to within 5% ($100\Sigma|F_i - \bar{F}|/\Sigma\bar{F}$).

Data from different crystals of the native or a given derivative were scaled together with the 55 selected strong reflections, which typically agreed to within 2%. The number of crystals used for each data set and the number of reflections with $F > 3·0 \sigma$ for each are given in Table 1. The derivative data sets were scaled to that of the native with a scale factor and isotropic shape parameter followed by anisotropic fitting with seven parameters:

$$|F_N| = |F_{NH}|K \exp(\beta_{11}h^2 + \ldots + \beta_{12}hk + \ldots).$$

The r.m.s. fractional differences on F for the different derivative data with the native data set after scaling are given in Table 1.

For the high-resolution analysis and further refinement of the structure, a complete set of native data out to 2·5 Å were recollected without using a monochromator. The data were collected using the θ–2θ scan mode and a scan rate of $\frac{1}{2}°$ min^{-1} for 12·6–3·4 Å and $\frac{1}{4}°$ min^{-1} for 3·4–3·0 Å. The data between 3·0 and 2·5 Å were collected by a count-15–drop-6 procedure. The data sets between the crystals had an R value of 0·06 based on F's. Of the 8300 reflections available, 6542 were considered observed at a level of $F > 3·0$. Thus, about 80% of all the reflections out to 2·5 Å resolution were contained in this data set.

Location of heavy-atom sites and native phase calculation

Low resolution analysis. The 5·5 Å diffractometer data were used for the initial interpretation of the major Os and Sm sites (Rubin *et al.*, 1975). Analysis for the minor heavy-atom sites and refinement of the heavy-atom parameters was then carried out in two stages, first at 4·0 Å and then at 3·0 Å. The heavy-atom positions at 4·0 Å were deduced from the difference Patterson maps with $(|F_{NH}| - |F_N|)^2$ coefficients. Representative difference Patterson maps for the Os, Sm, Gd and Pt derivatives are shown in Fig. 2. The Harker section at $V = 0·5$ for the Os derivative (Fig. 2a) revealed the Os(1) site. Since this was our first interpreted derivative, the arbitrary y coordinate was initially assigned as zero for Os(1). A second minor site, Os(2), was also deduced from these maps. The Sm and Gd derivatives are similar and the Harker section at $V = 0·5$ for the Gd derivative (Fig. 2c) and the section $V = 0·0$ for the Sm derivative (Fig. 2d) are shown. Using these difference Patterson maps, two major sites with nearly the same y coordinate in each lanthanide derivative were characterized. The Sm/Os double derivative was analyzed in a similar way. The major heavy-atom sites in different derivatives were referred to the same origin and relative y coordinate by the use of correlation maps (Steinrauf, 1963) and

checked against the double-derivative maps. These sites were then used for preliminary native phase calculation and the minor sites in each derivative were picked out using difference electron density maps. In this manner, the minor site Os(2) was confirmed and two minor sites (3 and 4) were assigned for the lanthanides. It may be mentioned that the Harker vector corresponding to site 4 is present in the difference Patterson map, while that corresponding to site 3 is not (Fig. 2c). A Bijvoet difference Fourier map (Kraut, 1968) established the hand of the heavy-atom coordinates.

The final rounds of calculation included Os, Sm, Gd, Au and Sm/Os derivative data out to 4·0 Å and the heavy-atom parameters were refined by the standard procedure (Dickerson, Kendrew & Strandberg, 1961). Phases were calculated for 1518 native reflections which had a mean figure of merit of 0·66.

3·0 Å *Analysis*. The 3·0 Å analysis was carried out using the three best derivatives: Pt, Gd and Sm. Alternative cycles of native phase calculations and refinement of the heavy-atom parameters led to a mean figure of merit of 0·63 for 3126 reflections after 12 cycles. The parameters for these three derivatives at the end of the refinement are given in Table 2. A plot of the
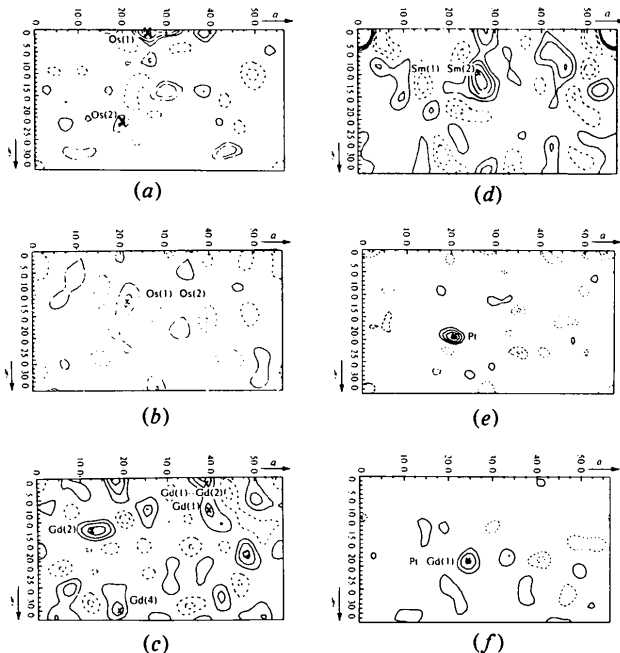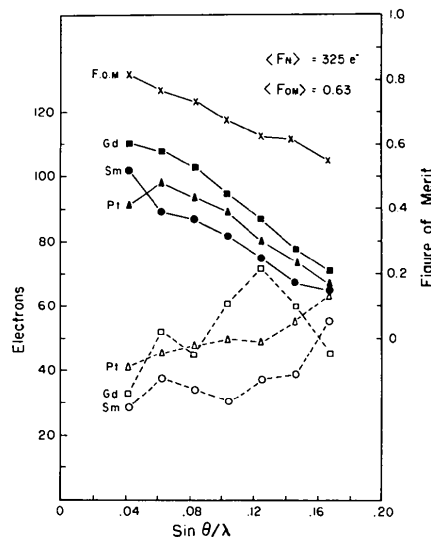

Fig. 3. Plot of the figure of merit, heavy-atom structure amplitudes (solid symbols) and lack-of-closure errors (open symbols) *versus* sin θ/λ for the three derivatives used in the 3·0 Å phase determination. Errors in the Gd data probably account for the hump in the lack-of-closure curve around 4·0 Å.

heavy-atom structure factors, the lack-of-closure errors, and the figure of merit with resolution is shown in Fig. 3.

The electron density maps and tracing the polynucleotide chain

The 4·0 Å electron density map was used to trace the polynucleotide backbone (Sundaralingam *et al.*, 1975a, b). In the map, the phosphate–ribose backbone was clearly seen as a continuous density with strong peaks spaced at approximately 6 Å intervals representing the phosphate groups. Coordinates of 69 such peaks representing phosphates were measured. The ribose groups were generally weaker. The ribose–phosphate backbone stood out clearly from the much weaker surrounding densities, and was sufficiently clear to allow the tracing of the polynucleotide chain.

The 3·0 Å map was calculated at grid spacings of 0·8 × 0·75 × 0·75 Å and sectioned along the *a* axis. The phosphate groups were clearer here and they were often resolved from adjacent ribose and base densities. Overall, the polynucleotide chain was seen to possess a distinct zigzag for many of the phosphate–ribose– phosphate groups. Generally, the density was strong in the core region of the molecule and progressively weaker as one approached the ends of the molecule. Since almost all of the bases were seen either edge-on or tilted in the *a*-axis view, sections of the map correspond- ing to the base-pair planes were also calculated and used in conjunction with the original map for interpre- tation (see also Stout, 1976).

Fig. 2. Difference Patterson maps for tRNA^{Phe} derivatives. (a) Harker section at $V = 0\cdot5$ for Os, 5·0 Å resolution. (b) Section at $V = 0\cdot15$ for Os, 5·0 Å, showing the cross-vector Os(1)–Os(2). (c) Harker section at $V = 0\cdot5$ for Gd, 5·5 Å. Difference electron density maps do not indicate sites corresponding to some of the larger peaks. The peak corresponding to the minor site Gd(4) is indicated. (d) Section at $V = 0\cdot0$ for Sm, 5·5 Å, showing Sm(1)–Sm(2) and Sm(3)–Sm(4) cross vectors. (e) Harker section at $V = 0\cdot5$ for Pt, 4·0 Å. (f) Correlation Patterson function $\Delta F_{Pt} . \Delta F_{Gd}$. Section at $V = 0\cdot15, 5\cdot0$ Å.

Table 2. *Heavy-atom parameters and refinement statistics*

Derivative	Resolution (Å)	Site occupancy (e)		x	y	z	B (Å²)	r.m.s. $|f_H|$	r.m.s. ε	R_M	R_w
Sm(OAc)₃	3·0	Sm(1)	43	0·6261	0·7717	0·4501	21	72	42	0·45	0·19
		Sm(2)	48	0·1009	0·7269	0·5914	19				
		Sm(3)	14	0·1918	0·0543	0·4411	21				
		Sm(4)	20	0·8301	0·8991	0·2458	17				
Gd₂(SO₄)₃	3·0	Gd(1)	59	0·6250	0·7816	0·4499	18	86	57	0·57	0·26
		Gd(2)	48	0·1036	0·7217	0·5913	27				
		Gd(3)	19	0·1664	0·9904	0·4403	22				
		Gd(4)	25	0·8238	0·8967	0·2508	27				
trans-Pt(NH₃)₂Cl₂	3·0	Pt	75	0·1776	0·6207	0·1475	*	82	58	0·54	0·28
OsO₃(pyridine)₂	4·0	Os(1)	33	0·2170	0·0674	0·9896	21	43	51	1·05	1·22
		Os(2)	10	0·1525	0·6807	0·1539	26				
NaAuCl₂	6·0	Au(1)	25	0·2013	0·0155	0·9857	20	86	59	0·79	0·55
		Au(2)	42	0·1801	0·6217	0·1441	20				
		Au(3)	19	0·7125	0·0480	0·3856	16				
		Au(4)	17	0·5420	0·4694	0·0948	12				
		Au(5)	22	0·9707	0·0574	0·4719	29				
		Au(6)	17	0·3894	0·4865	0·2560	20				
Sm/Os	4·0	Os(1)	25	0·2215	0·0674	0·9814	21	98	91	0·79	0·84
		Os(2)	10	0·1464	0·6753	0·1540	24				
		Os(3)	9	0·7033	0·0416	0·4106	23				
		Sm(1)	13	0·6446	0·7892	0·4607	18				
		Sm(2)	23	0·0944	0·7290	0·5893	26				
		Sm(3)	7	0·1984	0·0672	0·4460	20				

r.m.s. $|f_H| = [\Sigma|f_H|^2/n]^{1/2}$, $|f_H|$ = heavy-atom structure amplitude; r.m.s. ε (lack of closure) = $[\Sigma|F_{NH} - (F_N + f_H)|^2/n]^{1/2}$; $R_M = \Sigma||F_{NH}| - |F_N + f_H||/\Sigma|f_H|$; $R_w = \Sigma w(|F_{NH}| - |F_N + f_H|)^2/\Sigma w|f_H|^2$, $w = 1/\varepsilon^2$.

* Anisotropic temperature factors: $\exp[-(\beta_{11}h^2 + ... + \beta_{12}hk + ...)]$ where $\beta_{11} = 0.001490$, $\beta_{22} = 0.004766$, $\beta_{33} = 0.001375$, $\beta_{12} = 0.000017$, $\beta_{23} = -0.000686$, $\beta_{13} = 0.000141$.

The complete model was built on a scale of 2 cm Å⁻¹ with Kendrew brass components for the nucleotides. In a few regions of the map, the density could be equally well explained by more than one model conformation. In these cases, the most preferred conformation was used which also avoided close contacts. Thus, residues 16, 17, 37, 39, 75 and 76 were built in this way and could be changed during the subsequent work.

The coordinates of the atoms in the Kendrew scale model were measured with two scopes mounted on two faces of the model cage. The coordinates of the atoms measured this way were idealized using a program written by Isaacs (Dodson, Isaacs & Rollett, 1976) to restore the bond lengths and bond angles to standard values while at the same time constraining the atoms of the base to remain planar. The standard bond lengths and bond angles used for the backbone are listed in Table 3. The positional coordinates were assigned a standard deviation of 0·4 Å to enable them to be flexible enough to restore proper geometry. The co-ordinate set obtained by this procedure has been published (Stout *et al.*, 1976). An extended region of the molecule superposed on the MIR density map is shown in Fig. 4.

The inherent quality of the MIR map was such that

Table 3. *Standard values for bond lengths (Å) and bond angles (°) for the backbone chain (Sundaralingam, 1972)*

The geometry used for the common bases was from Voet & Rich (1970) and for modified bases from Sundaralingam (1973).

P—|O(3′),O(5′)|	1·60	|O(1P),O(2P)|—P—|O(3′),O(5′)|	109
P—|O(1P),O(2P)|	1·49	O(1P)—P—O(2P)	118
O(5′)—C(5′)	1·44	O(3′)—P—O(5′)	102
C(5′)—C(4′)	1·51	P—O(5′)—C(5′)	120
C(4′)—O(1′)	1·45	O(5′)—C(5′)—C(4′)	111
O(1′)—C(1′)	1·41	C(5′)—C(4′)—O(1′)	109
C(1′)—C(2′)	1·52	C(4′)—O(1′)—C(1′)	109
C(2′)—C(3′)	1·52	O(1′)—C(1′)—C(2′)	105
C(4′)—C(3′)	1·52	C(1′)—C(2′)—C(3′)	101
C(2′)—O(2′)	1·41	C(2′)—C(3′)—C(4′)	101
C(3′)—O(3′)	1·44	C(3′)—C(4′)—C(5′)	116
		C(3′)—C(4′)—O(1′)	105
		C(4′)—C(3′)—O(3′)	112
		C(2′)—C(3′)—O(3′)	112
		C(3′)—O(3′)—P	120

an unequivocal interpretation was difficult in some regions; particularly in the region of the terminal residues 73–76 and the DH and AC loops. It is known that the MIR phases, while being sufficiently precise to

enable general chain tracing, can be improved by using phases calculated from the model. However, such a procedure biases the corresponding maps to the input model. A normal procedure has been to compute maps with coefficients $2F_o - F_c$ and calculated phases. This map is a superposition of the F_o map and the difference map and has the property that the incorrect regions in the input model will have density that tends to show the corrections to be made. Nevertheless, there still tends to be a bias from the input model. We decided, therefore, always to omit the regions to be rebuilt from the phasing model which results in minimum bias from the input model. The density corresponding to the omitted regions was generally lower than that for the regions included in phasing.

Rebuilding the model using $2F_o - F_c$ maps

The entire model was refitted in the optical comparator using $2F_o - F_c$ maps at 3·0 Å, always omitting the oligonucleotide segment of 5–10 residues from the structure factor calculation. The maps generally looked much better than the corresponding regions of the MIR map and therefore allowed the omitted segments to be better fitted into the resulting density. Several adjustments were made, particularly in the loop regions and two base pairs (29–41 and 6–67) in the AC and AA stems respectively.

Further work was carried out with a new set of native data collected out to 2·5 Å resolution comprising 6542 reflections. Again, a complete refitting of the model was done in the optical comparator using the new data. The maps were more well defined than the earlier maps due to the more extensive and improved quality of the data, enabling better fitting of the polynucleotide chain. Several changes were indicated and were made in the model. This involved the ACCA end of the molecule, bases 16, 17, 20 in the DH loop and the exposed base 47 in the variable loop. A comparison of the old model and the refitted model for the DH loop

region is shown in Fig. 5. The atomic coordinates were measured and idealized as before. The R value was 0·470 and the mean deviation in the atomic coordinates from the MIR model was 1·12 Å. These coordinates formed the starting point for the refinement of the structure described in the next section.

Difference Fourier refinement at 2·5 Å

The structure was refined using difference Fourier maps at 2·5 Å resolution to obtain shifts in individual atomic positions. One round of refinement consisted of (i) a structure factor calculation using the current coordinates, (ii) calculation of a difference electron density map, (iii) estimation of shifts in the atomic coordinates using the relation: shift $= -K$ (slope/curvature), where K is a damping factor, usually set to 0·5. The slope was evaluated by considering the density in the difference map at two points bracketing the atomic
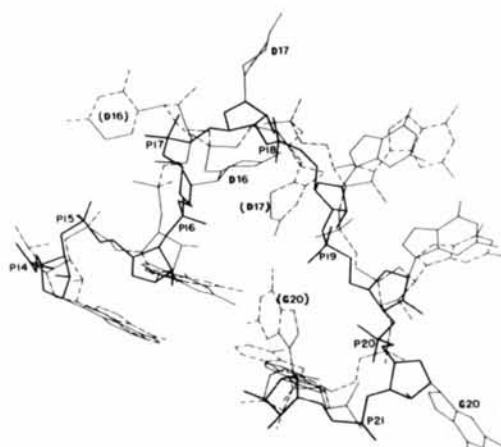

Fig. 5. The superposition of the DH loop region of the current model (solid lines) on the earlier MIR model (broken lines). Note the new conformations and positions of residues hU16, hU17 and G20. Notice that these residues are exposed (see also Fig. 7).
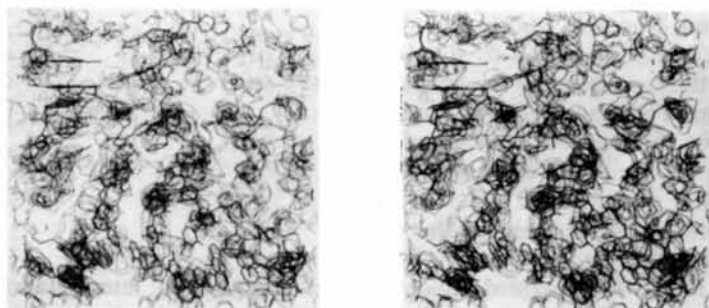

Fig. 4. Stereoscopic view of a section of our earlier model superposed on the 3·0 Å MIR map (16·8 Å deep). The origin is at the upper left with y horizontal (−0·04 to 0·95) and z vertical (0·25 to 0·85).

position. The curvature of each atom type was assumed to be $-0.2Z$, where Z is the atomic number. This empirical value proved to be adequate, particularly since the process of refinement is iterative in nature. Using the new coordinates, the three steps were repeated. The scattering factors used for the atoms were those of Cromer & Waber (1965).

The progress of the refinement is shown in Fig. 6. Four rounds of refinement resulted in the R value dropping from 0·470 to 0·382. The average shift in the atomic coordinates in a typical cycle was about 0·2 Å. At this stage, the coordinates were idealized using a relatively large standard deviation for the coordinates (0·4 Å). This provides a good deal of flexibility for the movement of the atoms and the resulting coordinate set usually has very good geometry, the average deviations in bond lengths and angles being 0·01 Å and 1·2°. This relatively large movement of the atoms tended to offset the shifts produced by the refinement process and the R value jumped to 0·448. In the initial stages we felt that such a procedure is not unreasonable but in later rounds the standard deviations were assigned lower
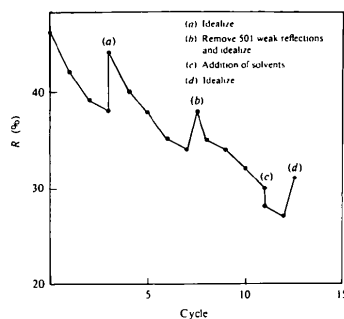
Fig. 6. The plot of R values for the different cycles during the difference Fourier refinement.

values, *viz* 0·05 Å, to retain the gains made in the refinement while at the same time substantially improving the geometry of the molecule. Four rounds of refinement reduced the R value to 0·345. At this stage 501 reflections were removed from the list for which the individual R value was more than 200% and further refinement was carried out only with the remaining 6041 reflections. The coordinates were idealized and the current R value was 0·384. Three further cycles of refinement brought the R value down to 0·307. At this stage we adjusted the relative occupancies of the atoms of the individual residues. The average residual density over all atoms belonging to one residue was computed and used to empirically assign a residue occupancy between 0·95 and 1·05 to all residues. The contributions from the atoms of each residue were modified by these factors during the structure factor calculation. The difference map showed indications of solvent molecules. They appeared as significant positive density close to the molecular boundary. Fifty peaks with the highest positive density were chosen and were analyzed for contacts with the set of solvent peaks and the atoms of the tRNA molecule. Thirty sites which survived such a screening process were assigned scattering factors of oxygen and included in the structure factor calculation. Even though many of these peaks probably represent cations like Mg^{2+}, Na^+ *etc.* they could not be identified as such from the maps. The R value dropped to 0·284. An additional round of refinement followed by inclusion of more solvent sites resulted in the current R value of 0·276. At this stage, a restrained idealization produced a set of coordinates with an R value of 0·314. In this model the average deviation in bond lengths and angles was 0·05 Å and 6·8° respectively. The average deviation of all the atoms in the current model from the starting model was 0·64 Å. Further refinement using
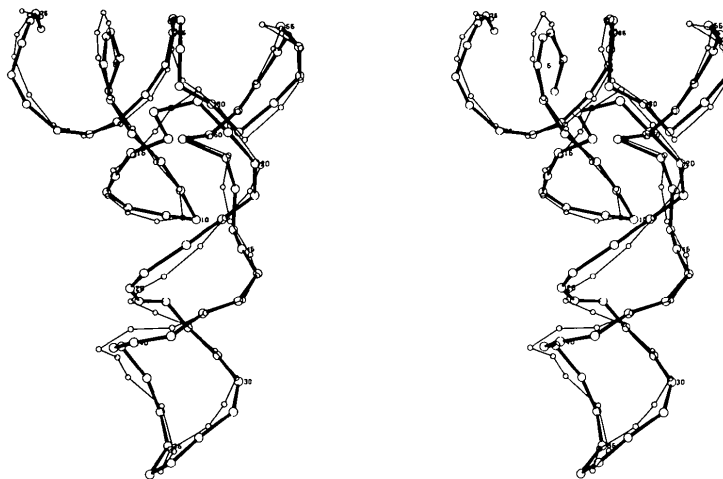

Fig. 7. Stereoview of the current phosphorus coordinates (larger circles connected by thick lines) superposed on the coordinates of our MIR model (smaller circles and thin lines) (Stout *et al.*, 1976).

individual thermal parameters for the atoms is in progress. A stereodiagram comparing the phosphate backbone of the refined model with the MIR model is shown in Fig. 7.

Discussion

The maps

In Fig. 8 is shown a $2F_o - F_c$ map around the region where the DH and Tψ loops meet to engage in the tertiary Watson–Crick base-pairing interaction G19–

C56, and the G18–ψ55 one hydrogen-bonded base–base interaction. The residues omitted in the phasing calculation were 16 to 20 of the DH loop and 55 to 57 of the Tψ loop. All of the omitted residues could be fairly well fitted into the densities with the possible exception of 16 and 17. The densities for the latter residues are weaker; nevertheless, this map suggested an alternative conformation for this segment of the DH loop in contrast to our earlier model (Fig. 5). The map also indicated a markedly different position for G20 which is now exposed to solvent.
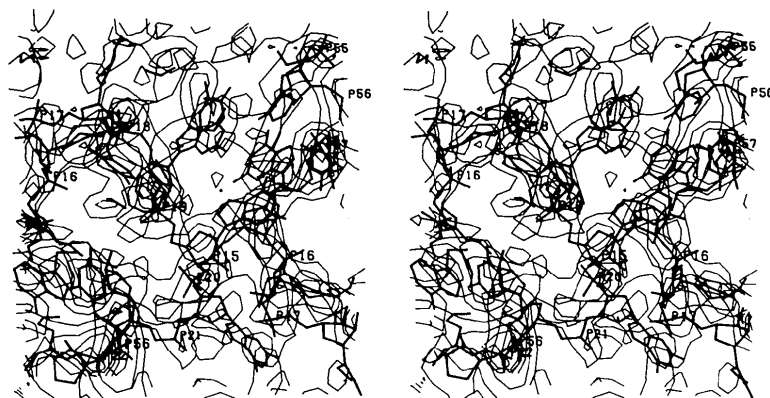
Fig. 8. A stereoview down the a axis of the $2F_o - F_c$ map showing the DH loop and the Tψ loop region. The density box is 10·8 Å deep but the superposed model has a greater depth for the sake of clarity. Residues 16 to 20 and 55 to 57 were omitted from the phasing calculation. Interactions between the two loops are seen at the top right. Both the Tψ loop at the bottom left and the DH loop at the bottom right belong to neighboring molecules.
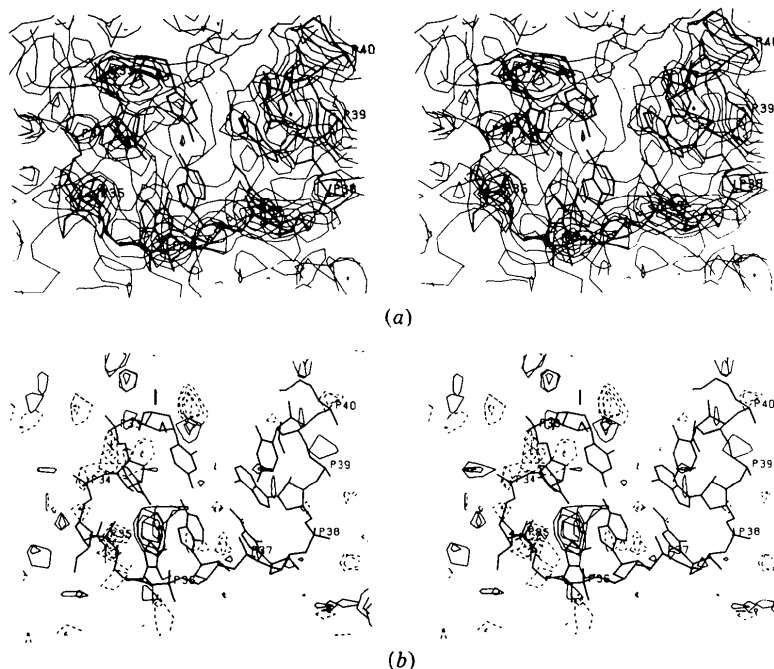

(a)

(b)

Fig. 9. (a) The $2F_o - F_c$ map and (b) the $F_o - F_c$ map showing identical sections in the region of the AC loop viewed down the b axis. The density boxes are 8·2 Å deep but the superposed model has a greater depth for the sake of continuity. Residues from Cm32 to A38 and the cations and solvents were omitted from the phasing calculations of both maps.

The AC loop has posed some difficulty in defining accurately. The $2F_o - F_c$ map omitting residues 32 to 38 shows impressive densities for many of the residues, although the bases A35 and A36 appear to be somewhat weak. Thus, these residues could not be defined with the same confidence as the other residues in the loop. The corresponding $F_o - F_c$ density map (Fig. 9) in addition to showing a significant positive peak also shows some areas of negative density which suggests possible shifts in the positions of the residues in its vicinity. Because of the negative residuals, we cautiously assign the positive peak to a cation (Mg^{2+} or Na^+) which is also at good bonding contacts to the sites O(2) and N(7) of the surrounding bases U33 and A36 respectively and the phosphate P35. It is also conceivable that the positive peak is a solvent water. Other potential sites for solvent molecules are also indicated in the map and some of these have been included in the refinement studies.

An extremely convincing electron density segment is that comprising residues 52 to 56 of the Tψ stem and loop (Fig. 10). It is seen that the backbone and the side-chain bases of both the secondary and tertiary paired residues of the Tψ stem and loop appear with remarkable clarity.

In Fig. 11(a) through (d) are shown both the MIR and difference density maps of the terminal ACCA residues 73 to 76. The MIR map (a) shows the ambiguity we encountered in fitting residue 76 because there were peaks on both the right and left of P76. The phosphates 74 and 75 are displaced from the peak centers, partly from the errors that arose in transforming the measured model coordinates to the electron density. The $2F_o - F_c$ map calculated by omitting the residues 73–76 appears to suggest unambiguously the location of A76 and the rest of the aminoacyl segment. In Fig. 11(b), the current model is overlayed on the MIR map. The main difference between our earlier model and the current one is in the rotation around the P–O(3') (ω') bond from g^+ to the extended t

Fig. 10. A view down the c axis of the $2F_o - F_c$ map, 7·5 Å deep, showing a portion of the Tψ stem and loop. The five residues from U52 to C56 were omitted from the phasing calculation.

conformation. Including the residues 73–76 in the phasing model does not seem to enhance the peak densities significantly in the $2F_o - F_c$ map, *cf.* Fig. 11(c) *vs* (d). Several residual peaks bordering the ACCA sugar–phosphate backbone perhaps indicate the solvent shell and cation coordination.

Representative sections illustrating the $2F_o - F_c$ densities of the G42–C28 Watson–Crick base pair within the AC stem are shown in Fig. 12. In Fig. 12(a), where the two residues have been omitted from phasing, the pyrimidine and purine bases show a good fit to the density. The corresponding $2F_o - F_c$ map with the paired residues included in the phasing calculation is shown in Fig. 12(b). The MIR density for the same region is also shown in Fig. 12(c) for comparison with the $2F_o - F_c$ maps. The MIR model fit is certainly not as convincing as the $2F_o - F_c$.

Molecular conformation

Previous data on the stereochemistry of nucleic acids were mainly derived from X-ray diffraction studies on single crystals of the constituents of nucleic acids. Although considerable information was gleaned from these studies on the stereochemical rules of the nucleotide building blocks and the internucleotide phosphodiesters, it is of great interest to test these theories on a biological macromolecular nucleic acid comprised of both secondary and tertiary structures. The crystallization of transfer RNA's and subsequent structure determination and refinement at 2·5 Å resolution has provided an opportunity to test the applicability of the stereochemical principles that have been developed, as well as obtain additional insights into the mechanics of polynucleotide folding and nucleic acid architecture.

The trace of the polynucleotide backbone fold is shown in Figs. 1(b) and 7.

The double helices. The four double-helical segments, *viz* AA, DH, AC and Tψ, exhibit overall conformations similar to those of fibrous RNA duplexes (Arnott, Hukins, Dover, Fuller & Hodgson, 1973). There are, in tRNA, a number of intrachain hydrogen bonds involving the O(2') hydroxyl and the ring oxygen O(1') of contiguous residues as indicated by the O(2')–H···O(1') distances. Besides these hydrogen bonds in the helices, there are a number of hydrogen bonds involving O(2') in the tertiary structure (Table 4).

As we had earlier pointed out (Sundaralingam *et al.*, 1976) and confirmed by our refinement studies, the sugar–phosphate chains of the helices are all within the normal ranges of conformation despite the occurrence of (a) a 'wobble' G–U base pair within the AA stem helix and (b) tertiary structural constraints in the spatial organization of the helical domains.

The loops and kinks. The 'single-stranded' stretches in tRNA vary in length from 2–8 residues and they display some of the most interesting conformational
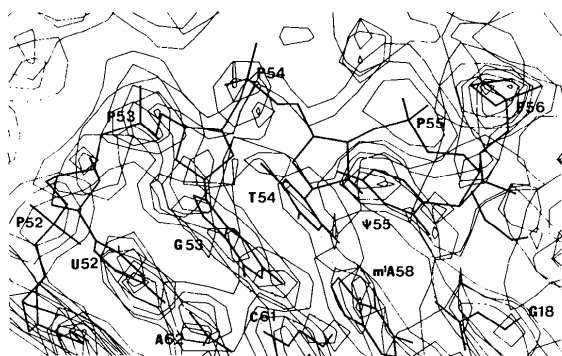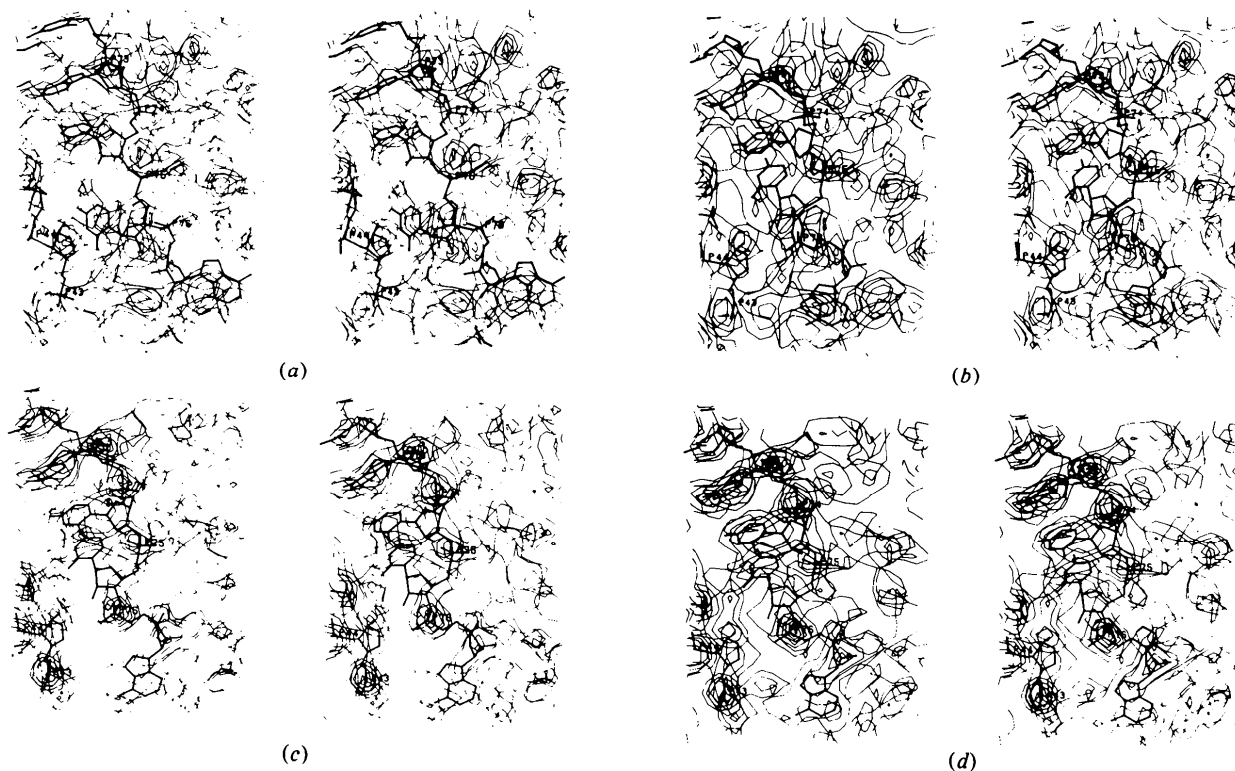
(a)

(b)

(c)

(d)

Fig. 11. Identical stereoviews down the b axis of the MIR and $2F_o - F_c$ maps, 7·2 Å deep, around the ACCA terminal region. (a) MIR map with our earlier model (Stout et al., 1976), (b) MIR map with the current model, (c) $2F_o - F_c$ map omitting residues 73 to 76 in the phasing and (d) same as (c) but including residues 73 to 76 in the phasing.

features in folding the tRNA molecule into its unique tertiary structure. The consequence of the three-dimensional folding of the tRNA molecule is the formation of a large variety of tertiary structural base–base, base–sugar, base–phosphate and sugar–phosphate hydrogen-bonding interactions in addition to metal-ion binding sites and dye (Liebman, Rubin & Sundaralingam, 1977) and amino acid (Stout & Sundaralingam, 1977) binding pockets. In spite of the influence of all of these interactions, the conformations around the sugar–phosphate and the sugar–base bonds are remarkably within the limits of the values previously established (Sundaralingam, 1969, 1973, 1975). The nucleotide residues and the internucleotide phosphodiesters within the helices all exhibit the preferred conformation, viz the C(3')-endo, anti (χ), gauche$^+$(ψ) combination and the gauche$^-$, gauche$^-$ combination respectively; the alternative conformations are also exhibited, as expected, by the residues in the loops and kinks (Sundaralingam, 1969, 1973; Yathindra & Sundaralingam, 1973, 1974). Most striking among these is the recurrent occurrence of the $g^-,t/g^+$ phosphodiester in the sharp turns of the AC loop (P34), Tψ loop (P56) and P10 loop (P10) (Sundaralingam et al., 1976; Sussman & Kim, 1976; Jack et al., 1976).

Similarly, the backbone, particularly the phosphodiesters in the loops and bends, is distorted from the helical g^-,g^- conformation and this is best seen in the (ω',ω) map (Fig. 13a). We find that there are only about a dozen phosphodiesters distorted from the helical g^-,g^- conformation in contrast to the considerably larger number indicated in our earlier unrefined structure (Stout et al., 1976). The terminal ACCA fragment displays an extended (t,g^-) phosphodiester conformation at P76 which results in destacking of the terminal and penultimate bases as observed in the crystal structure of dinucleotide UpA (Rubin, Brennan & Sundaralingam, 1972; Sussman, Seeman, Kim & Berman, 1972). The possible significance of such an extended conformation for aminoacylated tRNA has been pointed out (Sundaralingam & Arora, 1972). The double gauche$^+$ (g^+,g^+) conformation which is commonly encountered in dinucleoside monophosphates is not observed in the polynucleotide in agreement with our predictions (Yathindra & Sundaralingam, 1974).

The riboses in tRNA exhibit the preferred C(3')-endo pucker (Fig. 14) with the exception of six in the C(2')-endo pucker, viz 7, 9, 19, 48, 58 and 60. Similarly, the conformation around the C(4')–C(5') bond is in the preferred gauche$^+$ domain in the majority of the
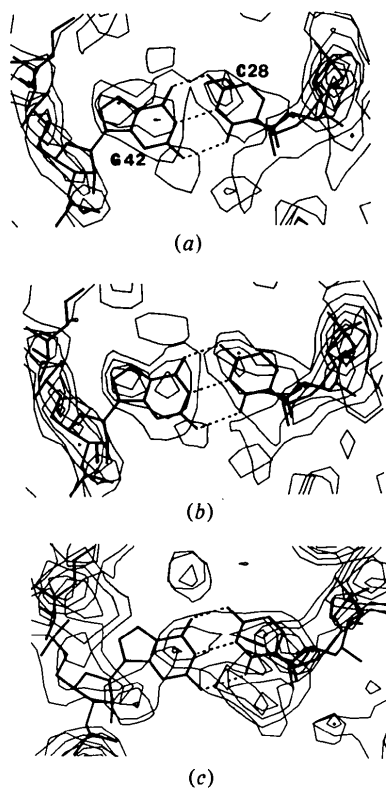
(a)

(b)

(c)

Fig. 12. Densities for the Watson–Crick G42–C28 base pair. (a) $2F_o - F_c$ map omitting residues involving the base pair in the phasing calculation. (b) $2F_o - F_c$ map including the residues in the phasing calculation. (c) 3 Å MIR map.

residues with the exception of about eight in the less preferred *trans* and *gauche⁻* conformations (Fig. 14) which are distributed in the loops and bends of the molecule. The bases are all *anti* and exhibit the greatest reluctance to deviate from their preferred conformation irrespective of their location (helix or loops) in the molecule. The phosphodiesters exhibit a greater tendency for conformational change than the sugar ring pucker or the exocyclic C(4')—C(5') bond in tRNA, in conformity with earlier deductions from the 'rigid' nucleotide theory (Sundaralingam, 1973).

Conformational maps. The remaining conformation maps correlating the adjacent backbone bond rotations pairwise are given in Fig. 13(b) through (g). Fig. 13(h) shows the map correlating the glycosyl torsion χ with the sugar pucker (pseudorotation phase angle). These maps are in striking agreement with the results obtained from studies of the nucleic acid constituents (Sundaralingam, 1969, 1973, 1975).

Conformation wheels. The six backbone torsion angles (Sundaralingam, 1969), the ribofuranose ring pucker (Sundaralingam, 1965, 1969; Altona & Sundaralingam, 1972) and the base–sugar glycosyl torsions (Sundaralingam, 1969; Donohue & Trueblood, 1960) are depicted in Fig. 14. The restricted distribution of the conformations around the various single bonds of the polynucleotide chain is reminiscent of the distribution found for the consituent nucleotide building blocks themselves.

Thus, the refined structure of tRNA lends further support for the conformational principles of nucleic acids derived from X-ray studies of their constituents. Further details of the molecular conformation will be presented elsewhere.

Metal-binding sites

Up until the structural study of tRNA, direct information on the mode of binding of metals to nucleic acids had been mainly derived from studies of crystal structures of metal complexes of nucleotides. A recent survey of these crystal structures and a comparison with the mode of binding of the metals used in the heavy-atom derivatives of tRNA have been covered elsewhere (Vijayalakshmi & Sundaralingam, 1977). Therefore, here we present briefly the binding sites of the transition metals Os, Pt and Au, and the lanthanides Sm and Gd to tRNA. The shorter contacts between the metal ion and the surrounding tRNA atoms are listed in Table 5. Among the heavy-atom derivatives used, there are at least eight different unique sites in the tRNA molecule. What is striking is that many of the transition-metal sites are also found to be common to more than one metal and none of the lanthanide sites are shared by the transition metals (Table 2). Further, it is interesting that as a class, the lanthanides bind in regions of the molecule different

Table 4. *Tertiary hydrogen bonding of the ribose 2'-hydroxyl group*

For nomenclature of O(1P) and O(2P) see Stout *et al.* (1976).

Backbone–base				Backbone–backbone			
Residue	Atom	Residue	Atom	Residue	Atom	Residue	Atom
U8	O(2')	m⁷G46	N(7)	U7	O(2')	m⁵C49	O(2P)
A9	O(2')	m²G10	N(7)	U7	O(2')	m⁵C49	O(5')
G18	O(2')	C(6)	N(4)	G45	O(2')	A9	O(3')
C48	O(2')	U59	O(2)	C48	O(2')	U59	O(2')
ψ55	O(2')	G57	O(6)	m¹A58	O(2')	C60	O(2P)
ψ55	O(2')	G57	N(7)	C60	O(2')	D17	O(1')

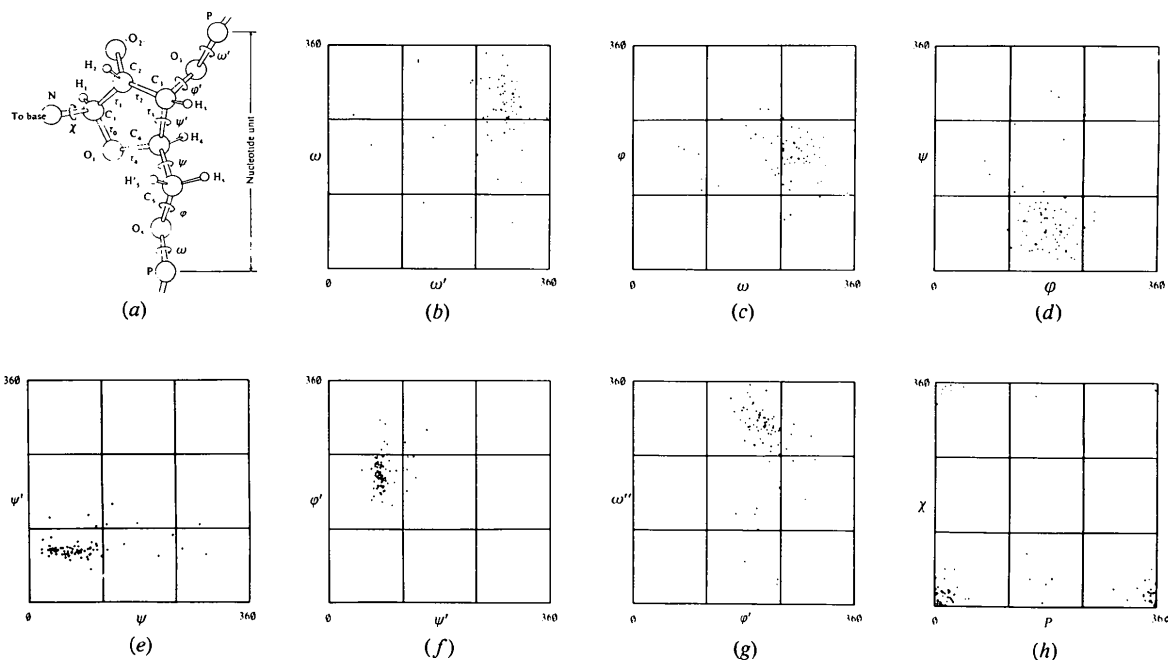Fig. 13. (a) Notations and definitions of the torsion angles in the nucleotide units: P—O(5′) = ω, O(5′)—C(5′) = φ, C(5′)—C(4′) = ψ, C(4′)—C(3′) = ψ′, C(3′)—O(3′) = φ′, O(3′)—P = ω′. (b)–(g) Conformation maps showing pairwise correlation of the adjacent sugar–phosphate backbone torsion angles for the 76 nucleotide residues of tRNA. (b) ω′,ω, (c) ω,φ, (d) φ,ψ, (e) ψ,ψ′, (f) ψ′,φ′ and (g) φ′,ω′ maps. (h) A map correlating the pseudorotation phase angle (P) of the ribofuranose ring and the glycosyl torsion angle χ.
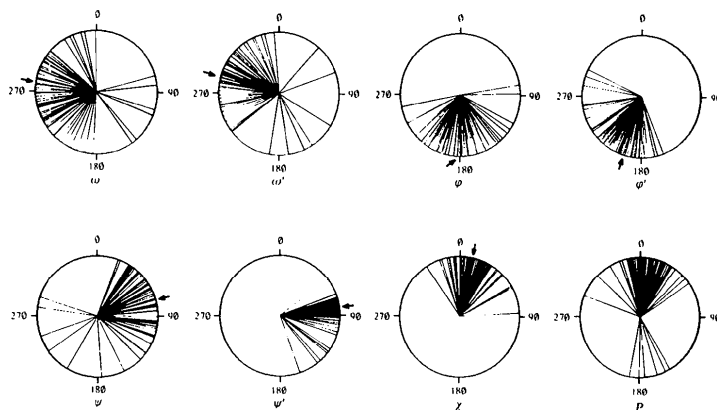

Fig. 14. Conformation wheels for the six sugar–phosphate backbone single bonds (ω,ω′,φ,φ′,ψ,ψ′), the pseudorotation phase angle (P) circle for the ribofuranose ring and the glycosyl bond (χ) for the 76 nucleotide residues of tRNA. The mean (indicated by an arrow) and r.m.s. deviations for the most populous conformation for the backbone bonds are: ω = 280 ± 30°, ω′ = 292 ± 24°, φ = 181 ± 27°, φ′ = 206 ± 19°, ψ = 68 ± 26° and ψ′ = 84 ± 9°. The corresponding values for the glycosyl bond χ are 18 ± 24°.

from those in which transition metals bind. While the former metals show a characteristic preference for the folded regions of the molecule where polynucleotide chains from different parts of the molecule or from different molecules come close together, the transition metals bind to both single- and double-helical regions of the molecule.

All the metal sites appear to be intramolecular with the exception of Sm(3) and Au(4) which are found in

positions on the molecule that also place them quite close to a neighboring molecule in the crystal. In general, the lanthanides are preferentially coordinated to the anionic O atoms of the phosphates, consistent with the 'hardness' of the rare-earth ions. A further important feature of the lanthanide binding sites is that they are associated with the phosphate backbone in single-stranded or loop regions where there are close intrachain phosphate and intermolecular contacts.

Sm(1) and Sm(4) are embedded in the molecule near the DH stem region where tertiary base interactions also occur, and they bind to the U8–A9 link and variable loop respectively. In this unique region, which forms the core of the tRNA molecule, four segments of the polynucleotide chains are in close proximity and these two lanthanide ions find the necessary ligands to be chelated. In a similar manner, the other two lanthanide ions [Sm(2), Sm(3)] are lodged in the lattice on opposite sides of the DH loop where there is close intermolecular interaction between two symmetry-related DH loops. The sites 1 and 2 are at least 70% occupied by the lanthanides and represent the major sites of attack in tRNA^{Phe}. Since site 1 is associated with the adjacent phosphates of residues 8 and 9, it could well be a common binding site for lanthanides (or Mg) in all tRNA's.

The OsO_3(pyridine)$_2$ reagent reacts with tRNA^{Phe} to irreversibly stain the crystals dark purple (Schevitz *et al.*, 1972). The major Os(1) binding site is located in the deep groove of the AC stem, near the bases A29, G30 and A31. A possible coordination scheme by these bases to the Os reagent has already been advanced by the MRC group (Jack, Ladner, Rhodes, Brown & Klug, 1977). The other site, Os(2), is similar to the Pt site (Table 2).

Au reacts less strongly and binds in multiple sites with a specificity somewhat similar to that of Os. For instance, the major Au(1) site is similar to Os(1) (Table 2).

Square-planar Pt^{II} complexes have a unique reactivity towards nucleic acids, and the *cis* isomers have been shown to have antitumor properties (Rosenberg, Vancamp, Trosko & Mansour, 1969). However, several attempts in our laboratory to use *cis*-Pt(NH$_3$)$_2$Cl$_2$ and Pt(ethylenediamine)Cl$_2$ for making derivatives of tRNA were unsuccessful. Meanwhile, the MRC group have shown that *trans*-Pt(NH$_3$)$_2$Cl$_2$ reacts strongly with tRNA^{Phe} both in solution (Rhodes, Piper & Clark, 1974) and in the crystal to form a single-site isomorphous derivative (Robertus *et al.*, 1974). We have employed this compound in our structure analysis. The Pt is bonded to the residue Gm34 as found by Jack *et al.* (1977).
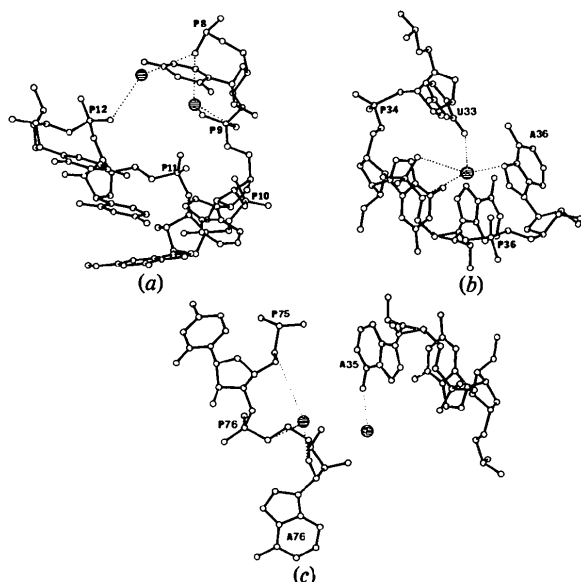
Fig. 15. Some examples of interactions between sections of the polynucleotide chain and solvent and cations. (a) P10 loop region, (b) AC loop region and (c) intermolecular region involving the CCA end and the AC loop.

Table 5. *Binding sites of heavy-atom reagents*

Heavy atom	Atom on tRNA	Distance	Heavy atom	Atom on tRNA	Distance	Heavy atom	Atom on tRNA	Distance
Os(1)	N(6)(A29)	4·3 Å	Au(2)	O(1P)(Gm34)	3·5 Å	Sm(2)	O(2')(D16)	2·6 Å
[OsO$_3$+	N(7)(A29)	3·4		O(6)(Gm34)	4·5		O(5')(G20)	3·0
(pyridine)$_2$]	N(7)(G30)	4·5		N(7)(Gm34)	3·1		O(1')(G20)	1·2
	N(7)(A31)	4·2	Au(3)	N(6)(A21)	2·2	Sm(3)	O(1P)(A14)	2·8
Os(2)	O(3')(U33)	2·4		O(3')(m^{7}G46)	3·3		O(2P)(A14)	4·3
	O(1P)(Gm34)	1·8		O(2P)(U47)	3·0		O(5')(A14)	3·6
				O(5')(U47)	1·9			
Pt	O(1P)(Gm34)	3·4		O(1')(U47)	2·3	Sm(4)	O(2P)(A44)	4·4
[trans-	O(6)(Gm34)	4·7	Au(5)	O(5')(G19)	2·8		O(6)(G45)	3·3
Pt(NH$_3$)$_2$Cl$_2$]	N(7)(Gm34)	3·3		O(1')(G19)	3·1		N(7)(G45)	4·8
				O(3')(G18)	3·3			
Au(1)	N(7)(A29)	4·8		O(1P)(G20)	3·0	*Gd(3)	O(1P)(G57)	2·8
(NaAuCl$_2$)	O(6)(G30)	3·9				[Gd$_2$(SO$_4$)$_3$]	O(5')(G57)	3·2
	N(6)(A31)	2·9	Sm(1)	O(2P)(U8)	3·4		O(1P)(A14)	3·7
	N(7)(A31)	3·5	[Sm(OAc)$_3$]	O(2P)(A9)	3·5			
				O(1P)(A9)	3·5			

* Gd(1), Gd(2) and Gd(4) bind in similar positions as Sm(1), Sm(2) and Sm(4) respectively.

In the Sm/Os double derivative, the occupancies of sites Sm(1) and Sm(2) were markedly reduced (Stout, Mizuno, Rubin, Mallikarjunan, Rao & Sundaralingam, 1975) in contrast to the Sm derivative itself (Table 2), even though the soaking conditions were the same. This observation suggests that the initial binding of the Os in the anticodon probably induces a conformation change in the 'core' region of the tRNA which significantly reduces the binding affinity for the lanthanides. Further studies on the Sm, Os and Sm/Os complexes with tRNA in the crystal and in solution should provide more insight into the conformational changes which occur in tRNA upon complexation with heavy atoms.

Possible binding sites for spermine, indicated by our MIR maps, occur in the deep grooves of the AA stem (residues 1–4 and 64–67), Tψ stem (residues 51–54 and 57–59) and AC stem (residues 37–43 and 24–28). These sites are being carefully investigated in our difference density maps.

Solvent binding sites

Most of the solvent peaks occur close to either phosphate or ribose O atoms. A number also occur, as expected, close to the ring N(3) and N(7) sites and the potential exocyclic sites of the bases. Some examples of solvent binding sites in tRNA are illustrated in Fig. 15. These modes of interaction of water with the polynucleotide chain resemble the variety of hydrogen-bonding schemes that have been observed in crystal structure studies of the nucleic acid constituents.
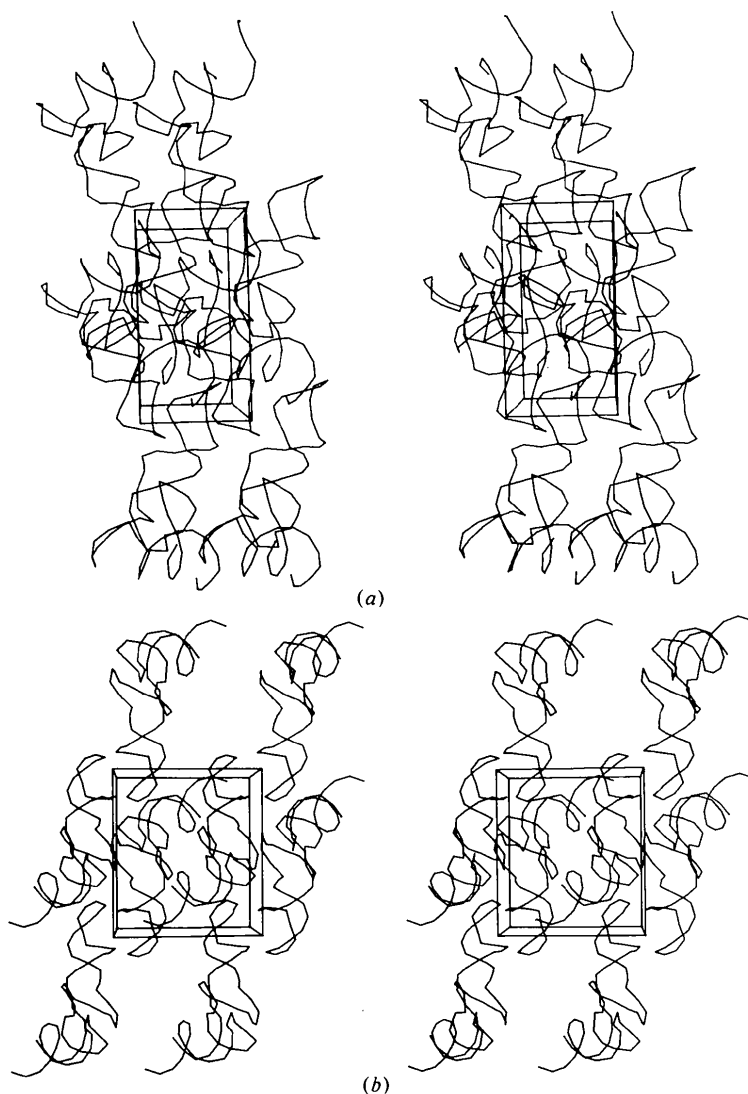

(a)

(b)

Fig. 16. Packing diagrams viewed down (a) the a axis and (b) the b axis. For clarity, each molecule is represented by lines joining the contiguous phosphorus atoms.
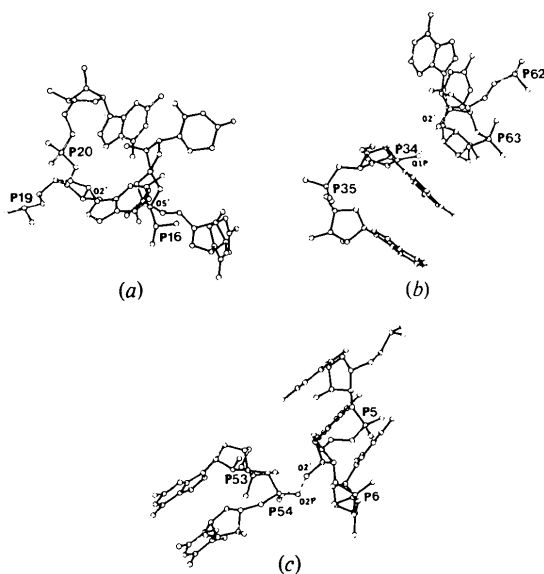
Fig. 17. Intermolecular contacts: (a) close contacts of DH loop backbones of screw-related molecules, (b) head-to-tail contacts involving the sugar–phosphate backbone of the AC loop residue Gm34 and the Tψ stem residue A62 (ribose) of a c-translated molecule, (c) backbone–backbone contacts involving the AA stem residue A5 and Tψ stem and loop residues G53 and T54 of a b-axis translated molecule.

Molecular packing and nucleic acid–nucleic acid interaction

The packing and the intermolecular contacts are illustrated in Fig. 16 in both the *a* and *b* axis views. The crystal structure of tRNA does not afford much information on the possible modes of interaction between nucleic acids. There are only three regions where neighboring molecules approach each other. These regions are shown in atomic detail in Fig. 17(*a*) through (*c*). Fig. 17(*a*) shows the interactions between DH loop residues 16 and 17 with DH residues 19 and 20 of a screw-related molecule. In Fig. 17(*b*) is shown the interaction between the 2'-hydroxyl group of residue 62 and the backbone of the AC loop of a translation-related molecule. In Fig. 17(*c*) is shown the interaction between the AA stem residues 5 and 6 and the residues 53 and 54 of the Tψ region. All of these close contacts involve interactions between sugar–phosphate backbone atoms.

Note added in proof: We have carried out further refinement of the structure using isotropic thermal parameters for the individual atoms instead of relative occupancies. The current *R* value is 0·275 including solvent molecules. The atomic coordinates of the model are being deposited in the Brookhaven Protein Data Bank.

References

ALTONA, C. & SUNDARALINGAM, M. (1972). *J. Am. Chem. Soc.* **94**, 8205–8212.

ARNOTT, S., HUKINS, D. W. L., DOVER, S. D., FULLER, W. & HODGSON, A. R. (1973). *J. Mol. Biol.* **81**, 107–122.

CROMER, D. T. & WABER, J. T. (1965). *Acta Cryst.* **18**, 104–109.

DICKERSON, R. E., KENDREW, J. C. & STRANDBERG, B. E. (1961). *Acta Cryst.* **14**, 1188–1196.

DODSON, E. J., ISAACS, N. W. & ROLLETT, J. S. (1976). *Acta Cryst.* A**32**, 311–315.

DONOHUE, J. & TRUEBLOOD, K. N. (1960). *J. Mol. Biol.* **2**, 363–371.

ICHIKAWA, T. & SUNDARALINGAM, M. (1972). *Nature New Biol.* **236**, 174–175.

JACK, A., LADNER, J. E. & KLUG, A. (1976). *J. Mol. Biol.* **108**, 619–649.

JACK, A., LADNER, J. E., RHODES, D., BROWN, R. S. & KLUG, A. (1977). *J. Mol. Biol.* **111**, 315–328.

KRAUT, J. (1968). *J. Mol. Biol.* **35**, 511–512.

LADNER, J. E., JACK, A., ROBERTUS, J. D., BROWN, R. S., RHODES, D., CLARK, B. F. C. & KLUG, A. (1975). *Proc. Natl Acad. Sci. USA*, **72**, 4414–4418.

LENHERT, P. G. (1975). *J. Appl. Cryst.* **8**, 568–570.

LIEBMAN, M., RUBIN, J. & SUNDARALINGAM, M. (1977). *Proc. Natl Acad. Sci. USA*, **72**, 4821–4825.

MIZUNO, H., RAO, S. T., STOUT, C. D., SWAMINATHAN, P. & SUNDARALINGAM, M. (1976). Abstracts, Am. Crystallogr. Assoc., Evanston, Illinois, August 9–12.

NORTH, A. C. T., PHILLIPS, D. C. & MATHEWS, F. S. (1968). *Acta Cryst.* A**24**, 351–359.

QUIGLEY, G. J. & RICH, A. (1976). *Science*, **194**, 796–806.

QUIGLEY, G. J., WANG, A. H. J., SEEMAN, N. C., SUDDATH, F. L., RICH, A., SUSSMAN, J. L. & KIM, S. H. (1975). *Proc. Natl Acad. Sci. USA*, **72**, 4866–4870.

RHODES, D., PIPER, P. W. & CLARK, B. F. C. (1974). *J. Mol. Biol.* **89**, 469–475.

ROBERTUS, J. D., LADNER, J. E., FINCH, J. T., RHODES, D., BROWN, R. S., CLARK, B. F. C. & KLUG, A. (1974). *Nature (London)*, **250**, 546–551.

ROSENBERG, B., VANCAMP, L., TROSKO, J. E. & MANSOUR, V. H. (1969). *Nature (London)*, **222**, 385–386.

RUBIN, J. (1974). PhD Thesis, Univ. of Wisconsin–Madison.

RUBIN, J., BRENNAN, T., STOUT, C. D., MIZUNO, H., MALLIKARJUNAN, M., MCMULLAN, R. K., ICHIKAWA, T., RAO, S. T. & SUNDARALINGAM, M. (1975). *Structure and Conformation of Nucleic Acids and Protein–Nucleic Acid Interactions* (*Proceedings of the Fourth Annual Harry Steenbock Symposium*), edited by M. SUNDARALINGAM & S. T. RAO, pp. 25–37. Baltimore: Univ. Park Press.

RUBIN, J., BRENNAN, T. & SUNDARALINGAM, M. (1972). *Biochemistry*, **11**, 3112–3128.

SCHEVITZ, R. W., NAVIA, M. A., BANTZ, D. A., CORNICK, G., ROSA, J. J., ROSA, M. D. H. & SIGLER, P. B. (1972). *Science*, **177**, 429–431.

STEINRAUF, L. K. (1963). *Acta Cryst.* **16**, 317–318.

STOUT, C. D. (1976). PhD Thesis, Univ. of Wisconsin–Madison.

STOUT, C. D., MIZUNO, H., RUBIN, J., BRENNAN, T., RAO, S. T. & SUNDARALINGAM, M. (1976). Nucleic Acids Res. 3, 1111–1123.

STOUT, C. D., MIZUNO, H., RUBIN, J., MALLIKARJUNAN, M., RAO, S. T. & SUNDARALINGAM, M. (1975). Acta Cryst. A31, S39.

STOUT, C. D. & SUNDARALINGAM, M. (1977). Abstract F3, Am. Crystallogr. Assoc., East Lansing, Michigan, August 7–12.

SUDDATH, F. L., QUIGLEY, G. J., McPHERSON, A., SNEDEN, D., KIM, S. H. & RICH, A. (1974). Nature (London), 248, 20–24.

SUNDARALINGAM, M. (1965). J. Am. Chem. Soc. 87, 599–606.

SUNDARALINGAM, M. (1969). Biopolymers, 7, 821–860.

SUNDARALINGAM, M. (1972). Jerusalem Symp. Quantum Chem. Biochem. 4, 73–101.

SUNDARALINGAM, M. (1973). Jerusalem Symp. Quantum Chem. Biochem. 5, 417–456.

SUNDARALINGAM, M. (1975). Structure and Conformation of Nucleic Acids and Protein–Nucleic Acid Interactions (Proceedings of the Fourth Annual Harry Steenbock Symposium), edited by M. SUNDARALINGAM & S. T. RAO, pp. 487–524. Baltimore: Univ. Park Press.

SUNDARALINGAM, M. & ARORA, S. K. (1972). J. Mol. Biol. 71, 49–70.

SUNDARALINGAM, M., MIZUNO, H., STOUT, C. D., RAO, S. T., LIEBMAN, M. & YATHINDRA, N. (1976). Nucleic Acids Res. 3, 2471–2484.

SUNDARALINGAM, M., STOUT, C. D., MIZUNO, H., RUBIN, J., RAO, S. T. & BRENNAN, T. (1975a). Abstracts, Am. Crystallogr. Assoc., Charlottesville, Virginia, March 9–13.

SUNDARALINGAM, M., STOUT, C. D., MIZUNO, H., RUBIN, J., RAO, S. T., MALLIKARJUNAN, M. & BRENNAN, T. (1975b). Acta Cryst. A31, S39.

SUSSMAN, J. L. & KIM, S. H. (1976). Biochem. Biophys. Res. Commun. 68, 89–96.

SUSSMAN, J. L., SEEMAN, N. C., KIM, S. H. & BERMAN, H. B. (1972). J. Mol. Biol. 66, 403–421.

VIJAYALAKSHMI, S. & SUNDARALINGAM, M. (1977). In preparation.

VOET, D. & RICH, A. (1970). Prog. Nucleic Acid Res. Mol. Biol. 10, 183–265.

YATHINDRA, N. & SUNDARALINGAM, M. (1973). Biopolymers, 12, 297–314.

YATHINDRA, N. & SUNDARALINGAM, M. (1974). Proc. Natl Acad. Sci. USA, 71, 3325–3328.

[1]Benzoxepino[2,3-b]quinoxaline and [3]Benzoxepino[1,2-b]quinoxaline: Two Previously Unknown Photoproducts from Benzo[a]phenazine 7-Oxide

BY ROBERTA OBERTI AND ALESSANDRO CODA

Centro di Studio per la Cristallografia Strutturale del CNR, Via Bassi 4, 27100 Pavia, Italy

AND LUCIA INCOCCIA AND FABIO COMIN

Istituto di Cristallografia dell'Università di Pavia, Via Bassi 4, 27100 Pavia, Italy

Two new products from the photoisomerization of benzo[a]phenazine 7-oxide have been isolated and crystallized: $C_{16}H_{10}N_2O$, m.p. 145–147°C, orthorhombic, Fdd2, $a = 39.931$ (3), $b = 26.864$ (2), $c = 4.471$ (2) Å, $Z = 16$; and $C_{16}H_{10}N_2O$, m.p. 129–130°C, orthorhombic, Pna2$_1$, $a = 12.201$ (3), $b = 7.443$ (2), $c = 13.328$ (3) Å, $Z = 4$. Intensity data were collected at room temperature on a Philips PW 1100 computer-controlled diffractometer (Cu $K\alpha$ graphite-monochromatized radiation). Both crystal structures were solved by direct methods and refined anisotropically to final conventional unweighted R indices of 0.042 and 0.052 respectively for the reflections with $I > 2\sigma(I)$. The molecular conformations of both compounds are discussed.

Introduction

Irradiation of benzo[a]phenazine 7-oxide (I) yields the annulene (II) via an oxaziridine or 'quasi-oxaziridine' intermediate (Albini, Barinotti, Bettinetti & Pietra, 1976).

Compound (II) is photostable under irradiation at 366 nm (maximum of the first absorption band: 364 nm), while irradiation at 254 nm (maximum of the second absorption band: 261 nm) leads to isomerization, although with a low quantum yield.

The structures of the two principal products obtained