Acta Cryst. (1973). B29, 943

Refinement of the Model of a Protein: Rubredoxin at 1.5 Å Resolution

BY K. D. WATENPAUGH, L. C. SIEKER, J. R. HERRIOTT AND L. H. JENSEN

Departments of Biological Structure and Biochemistry, University of Washington, School of Medicine, Seattle, Washington 98195, *U.S.A.*

(*Received* 16 *October* 1972; *accepted* 13 *January* 1973)

Rubredoxin from *Clostridium pasteurianum* is an iron sulfur protein of molecular weight ~6100. Crystals are grown by salting out from aqueous $(NH_4)_2SO_4$ at *p*H 4. They belong to the trigonal system, space group *R*3, hexagonal unit-cell parameters $a=b=64{\cdot}45$, $c=32{\cdot}68$ Å. The structure was solved at 3 Å resolution with phases based on a single HgI_4^{2-} derivative. For the 2 Å resolution electron density map, phases were based on two derivatives, HgI_4^{2-} and UO_2^{2+}, for which both reflections of each Friedel pair were measured. From this map a tentative sequence was determined and coordinates for 401 atoms in the molecule were tabulated. The model for the structure has been refined in the usual crystallographic sense beginning at a conventional *R* of 0·372. Four refinement cycles by ΔF syntheses and four by least-squares have reduced *R* to 0·126 for the 5005 reflections with $2\sigma < I$ and $0{\cdot}05 < \sin\theta/\lambda < 0{\cdot}325$. Important features of the structure which were uncertain or not evident in the 2 Å resolution map have been revealed.

The analysis of protein structure from X-ray diffraction data has been based primarily on electron density maps that have been computed with observed structure amplitudes and multiple isomorphous replacement phases. The ultimate detail that can be revealed in such maps is highly dependent on the accuracy of the phases. In general, the error increases with increasing θ, in part because of the rapid decrease in the structure amplitudes, but also because of lack of isomorphism between native and derivative crystals and other factors which cause errors in the phases. It has not been the practice therefore, to determine phases for reflections with $d<2$ Å ($0{\cdot}25<\sin\theta/\lambda$).

The problem of error in multiple isomorphous phases would be eliminated if a model could be assumed for a protein structure and refined in the ordinary crystallographic sense. Although the efforts to refine the model of myoglobin have met with only limited success (Brändén, Holmes & Kendrew, 1963; Watson *et al.*, 1963; Watson, 1969), there is no reason, in principle, why a protein cannot be refined if the data are sufficiently extensive. Not only would refinement improve the model, but it would provide an estimate of the precision of the coordinates so that an objective assessment of structural features could be made. There are practical difficulties, however, in the sheer size of even the smallest proteins and in collecting accurate data. In addition, half or more of the volume of most protein crystals is fluid which is more or less free to move about in the spaces between the molecules; and parts of the molecules, in particular those R groups on the surface which project into the solution, may suffer disorder of one kind or another and have very large apparent thermal motion.

The above factors all tend to decrease the intensities of the X-ray reflections, and X-ray diffraction from protein crystals, in general, drops off very rapidly with increasing θ. Nevertheless, with presently available diffractometers it is possible, for some of the smaller proteins, to observe most of the reflections with $1{\cdot}5$ Å $< d$ and an appreciable fraction at even smaller *d* spacings with intensities greater than 2σ.

The resolution which can be achieved with three-dimensional X-ray diffraction data is given by the expression (James, 1948):

$$\text{limit of resolution} = 0{\cdot}72\ d_{\min},$$

where $d_{\min}$ is the minimum observed interplanar spacing. This is a theoretical limit which is not realized in practice. In the field of protein crystallography it has become general practice to take the resolution as equal to the minimum observed interplanar spacing, and to speak, for example, of a 2 Å resolution electron density map, meaning one based on reflections with 2 Å $< d$. More simply, such a map is often referred to as a 2 Å map and the data as 2 Å data.

This report describes the course of the refinement of rubredoxin from *Clostridium pasteurianum*, a nonheme iron protein with molecular weight ~6100. A detailed description of the stereochemistry of the rubredoxin molecule and a description of the solvent structure will be published elsewhere.

Experimental

The material used in this work was provided by Dr W. Lovenberg of the National Heart Institute and was extracted from the cells of the bacterium *C. pasteurianum*. Beatiful, deep-red crystals can be grown from aqueous solution at *p*H 4 by salting out with $(NH_4)_2SO_4$. The crystals belong to the trigonal system, space group *R*3, with rhombohedral cell parameters

$$a=38{\cdot}77\ \text{Å},\ \alpha=112{\cdot}37°,$$

and hexagonal cell parameters,

$$a=b=64{\cdot}45(1),\ c=32{\cdot}68(1)\ \text{Å},$$

based on Cu $K\alpha$ ($\lambda=1{\cdot}5418$ Å) (Herriott, Sieker, Jensen & Lovenberg, 1970). The indicated standard deviations for the hexagonal cell parameters are representative of values for a *given* crystal. Crystals from different batches may have cell parameters that differ from those reported here by as much as 20 to 30 times the indicated standard deviations.

The rhomb-shaped crystals grow readily to sizes of a millimeter or more in a few weeks. The maximum crystal dimension is usually normal to the unique axis and the minimum dimensions parallel to it, often in a ratio of about 2:1.

Intensity data were collected on a four-circle diffractometer by the $\omega/2\theta$ scan technique. The crystal was 236 mm from the focal spot, and the scintillation counter 224 mm from the crystal. The Cu target tube was operated at 40 kV constant potential, 14 mA, and set at a 3° take-off angle with a focal spot size $0{\cdot}4\times10$ mm. A 0·009 mm Ni filter was used in the incident beam, and the pulse-height discrimination was set to pass about 95% of the Cu $K\alpha$ photons.

The scan rate was 2° min^{-1} in 2θ and the scan time approximately 40 sec for low 2θ reflections with increasing scan time as 2θ increased to allow for the α_1–α_2 doublet. A background count of 4 sec was collected at each limit of the scan range.

A complete data set with $1{\cdot}54\ \text{Å} < d\ (\sin\theta/\lambda < 0{\cdot}325)$ was collected from one crystal of approximate dimensions $0{\cdot}7\times0{\cdot}7\times0{\cdot}35$ mm. About 9350 reflections were collected of which approximately 2000 were duplicates. Three monitor reflections were measured at intervals of 200 reflections and a large number of standard reflections were measured at longer intervals as a check on deterioration. During the period of data collection, the standard reflections decreased in intensity by 10–15%, and all data were corrected by the appropriate factor. The data were also corrected by experimentally determined absorption factors (North, Phillips & Mathews, 1968).

In the unique data set there were 7345 reflections of which 5033 had peak counts greater than 2σ where σ was calculated from the expression

$$\sigma=[N_{\text{pk scan}}+(T_{\text{scan}}/T_{\text{bkg}})^2N_{\text{bkg}}+(0{\cdot}05N_{\text{pk net}})^2]^{1/2}\ .$$

Under the operating conditions used, approximately 1600–1700 reflections were collected in 24 h. Scan ranges and background count times were kept as short as possible in order to maximize the rate at which data were collected. Since deterioration of protein crystals can seriously impair precision, it is important that data be collected as rapidly as possible. It is advantageous to sacrifice some precision in counting (random error) by increasing the rate of data collection if a disproportionately larger reduction in systematic error can be achieved.

The model

The structure of rubredoxin was originally solved at 3Å resolution with phases determined from the Hg_4^{2-} derivative by using the effects of anomalous scattering to resolve the phase ambiquity (Blow & Rossmann, 1961). The resolution was extended to 2·5 Å, the phases at this resolution being determined from two derivatives, HgI_4^{2-} and UO_2^{2+}, including measurement of both reflections of each Friedel pair for both derivatives (Herriott *et al.*, 1970). Resolution was subsequently extended to 2 Å (Watenpaugh, Herriott, Sieker & Jensen, 1970), again with phases based on two derivatives and measurement of both reflections of each Friedel pair. The 2 Å resolution map was calculated on a grid along **a***, **b**, **c** at intervals of (a cos 30°)/120, b/120, c/60 and with most probable phases and coefficients weighted by figures of merit. It provided the basis for the model used to initiate the refinemcnt (Watenpaugh, Sieker, Herriott & Jensen, 1970). Fig. 1 is a stereogram showing the Fe–S complex and the course of the main protein chain by connecting the α carbon atoms.

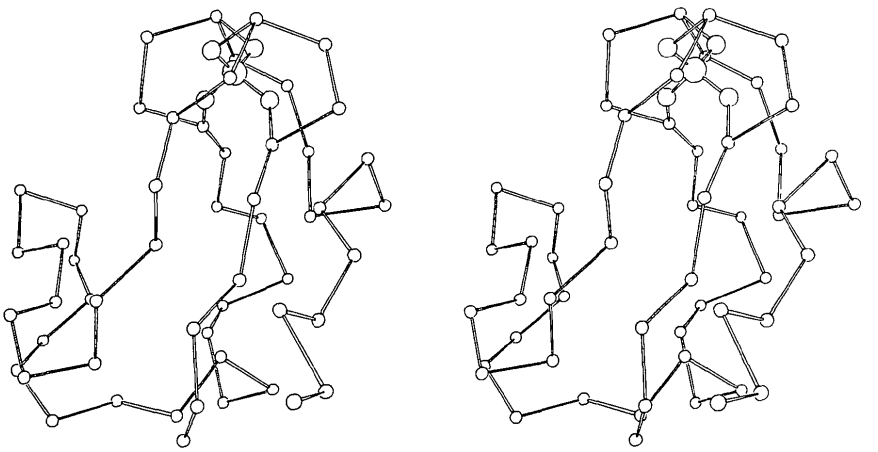


Fig. 1. Stereoview of the C^{α} atoms and iron–sulfur complex of rubredoxin.

At the time the 2 Å map for rubredoxin was calculated the chemical sequence was not known, so that in deriving a model it was necessary to determine a tentative sequence from the appearance of the R groups (Herriott, Watenpaugh, Sieker & Jensen, 1973). About 40 of the 54 amino acids were identified correctly and most of the rest did not differ by more than an atom or two from what they have subsequently turned out to be. Highly doubtful atoms were omitted, *e.g.* one atom in R2, N^{ζ} in what was obviously Lys 3, C^{δ}, C^{ε} and N^{ζ} of R31 which was thought to be Lys, all atoms in R52 and R53 and all atoms in the C terminal amino acid 54. [The numbering of the atoms is that conventionally used for proteins (IUPAC-IUB Commission on Biochemical Nomenclature, 1971).]

Coordinates were derived in the simplest possible way by comparing Kendrew skeleton models of sections of the main chain and R groups with the electron density map contoured on essentially the same scale as the model. The approximate atomic locations were marked on the nearest section of the computer output and the coordinates estimated to 0·1 of a grid interval (0·05 Å). In all, coordinates were derived for 401 atoms in the protein and 23 water molecules, most of which are near the surface of the protein molecule (Watenpaugh, Herriott, Sieker & Jensen, 1970).

Refinement by ΔF syntheses

For structures in which the atoms are either unresolved or only poorly so, corrections in the assumed positions that can be determined from ΔF maps with coefficients $(|F_o|-|F_c|)e^{i\alpha_{calc}}$ lead to better positions for the atoms than those determined directly from an F_o map. Furthermore, difference maps provide an easily interpretable measure of the B values relative to those assumed in calculating the structure factors.

Difference maps are also preferable to the method of least-squares in the initial stages of a refinement because of the easy check on the validity of the indicated corrections in terms of what is known about main-chain and side-chain structure. Moreover, the computing time for a refinement cycle involving a ΔF map is far less (Watenpaugh, Sieker, Herriott & Jensen, 1970, 1972).

The shift in any coordinate x_i from the ΔF map for rubredoxin is given by the equation

$$\delta x_i = -\frac{\partial \Delta\varrho/\partial x_i}{\partial^2\varrho/\partial x_i^2}$$

and similar terms for δy_i and δz_i where the slopes are evaluated at the assumed atomic positions (see, for example, Stout & Jensen, 1968). The cross terms in the full expression have been omitted because the map was calculated on an orthogonal grid. The grid interval was approximately 0·5 Å, sufficiently fine so that slopes could be simply estimated from the grid points bracketing the assumed atomic positions.

The curvature to be used in calculating the shifts for for any atom depends on the resolution of the ΔF map being interpreted. Ideally, it should be taken from an F_o map calculated with the same terms as the ΔF map. It is impractical, however, to evaluate curvatures for each atom because of the lack of resolution in a 1·5 Å electron density map, and it involves the addition effort of calculating another map.

It was sufficient for our purpose to approximate the curvature of a representative atom from the 2 Å electron density map and from it to estimate curvatures to be used for other atom types. We chose the carbonyl oxygen atom of amino acid 39 with a peak density of $\sim$1·5 e Å^{-3}. A Gaussian distribution was assumed according to the equation

$$\varrho = Ae^{-pr^2}$$

(Costain, 1941), where A is the peak density, p is a constant to be determined, and r is the distance from the atomic center. The constant p was evaluated for O(39) by equating the number of electrons associated with an oxygen atom to the integral of $\varrho(r)$ over all space,

$$8 = A(\pi/p)^{2/3}$$
$$p = (1{\cdot}5/8)^{2/3}\pi$$
$$= 1{\cdot}04 \text{ Å}^{-2}.$$

Table 1. *Course of the refinement*

Coordinates	Overall R	R and number of reflections as function of sin θ/λ 0–0·10	0·10–0·167	0·167–0·25	0·25–0·333	Number of reflections	B	Number of H_2O and occupancy
From 2 Å res. F_o map	0·389	0·663 218	0·312 765	0·385 2151	0·397 1899	5033	15	23 1
From 2 Å res. F_o map	0·376*	0·504 212*	0·307 765	0·390 2151	0·392 1899	5027	13·5	23 1
From 2 Å res. F_o map	0·372	0·461 212	0·301 765	0·392 2151	0·394 1899	5027	12	23 1
From 1st ΔF map	0·321	0·389 212	0·271 765	0·328 2151	0·346 1899	5027	12	22 1
From 2nd ΔF map	0·289	0·389 212	0·239 765	0·292 2151	0·312 1899	5027	9–20	22 1
From 3rd ΔF map	0·262	0·389 212	0·209 765	0·261 2151	0·288 1899	5027	7–30	22 1
From 4th ΔF map	0·242	0·373 212	0·192 765	0·239 2151	0·263 1899	5027	6–40	23 1
From 4th ΔF map	0·224	0·306 212	0·170 765	0·225 2151	0·260 1899	5027	6–40	100 0·3–1
1st L. S. cycle	0·179	0·192 212	0·123 765	0·181 2151	0·241 1899	5027	−5–50	106 0·3–1
2nd L. S. cycle	0·148	0·159 212	0·092 765	0·154 2151	0·222 1899	5027	3–50	127 0·3–1
3rd L. S. cycle	0·132†	0·118 190†	0·080 765	0·139 2151	0·188 1899	5005	2–107	130 0·3–1
F_c including 256 H atoms	0·141	0·153 190	0·092 765	0·145 2151	0·192 1899	5005	2–107	130 0·3–1
4th L. S. including 256 H	0·126	0·112 190	0·078 765	0·131 2151	0·183 1899	5005	2–86	130 0·3–1

* Six reflections with sin $\theta/\lambda < 0{\cdot}03$ weighted 0.
† Twenty-eight reflections with sin $\theta/\lambda < 0{\cdot}05$ weighted 0.

Substitution of A and p in the second derivative of $\varrho(r)$ gives

$$\delta^2\varrho/\delta r^2 = -(1{\cdot}5\,(1{\cdot}0)\,(2) = -3{\cdot}0 \text{ e Å}^{-5} \text{ at } r=0\,.$$

Curvatures for all atoms of a given type were assumed to be constant and independent of direction. For oxygen, nitrogen and carbon atoms, curvatures were taken in the ratio 6:5:4 on the basis of values used for a small organic structure (Jensen & Lingafelter, 1961). The curvature of sulfur was taken as twice that of oxygen; that of iron was taken as four times that of oxygen.

It is convenient in practice to evaluate a constant for each atom type by which shifts can be determined by a simple multiplication (or division) of the numerical difference between grid points bracketing the atomic sites. With the difference map on a scale of 100 units/e Å^{-3} for rubredoxin, differences of 10, 12·5 and 15 corresponded to a shift of 0·1 grid spacing for carbon, nitrogen and oxygen respectively. In this way, shifts could be evaluated for approximately 50 atoms per hour. No attempt was made to automate this part of the refinement by adapting the method used for a computer because it was judged essential at the outset to inspect each ΔF map in the vicinity of every atomic site.

For all calculations the data were divided into four groups as a function of sin θ/λ: 0·017–0·10, 0·10–0·167, 0·167–0·25, 0·25–0·325. Scale factors and R values ($R=\sum||F_o|-|F_c||/\sum|F_o|$) were calculated for each group as well as over-all values for the set of data. Information relevant to the course of the refinement is given in Table 1.

The first structure-factor calculation included 401 atoms in the protein molecule and 23 oxygen atoms representing water in the surrounding solution. The over-all R for the 5033 reflections was 0·389 with $B=15$ Å^2 as determined from the Wilson plot. Comparison of the scale factors for the four groups of reflections showed at once that B was too high. It was lowered in two steps to 12 Å^2. Inspection of the very low-angle reflections showed $F_o \ll F_c$. This could be due, in part, to secondary extinction, but it is probably primarily the result of neglecting most of the solvent in calculating the structure factors. Weighting the six reflections with sin $\theta/\lambda < 0{\cdot}03$ zero and with $B=12$ Å^2, R was 0·372. This is to be taken as the value at the beginning of the refinement.

The first ΔF map based on the 1·5 Å data indicated rather large shifts in both positional and thermal parameters. Shifts in the positional parameters were estimated as described above without applying a double shift (Cruickshank, 1950) since the curvatures derived from the 2 Å map are almost certainly underestimates. Ten atoms were added to the model (one of these appeared to be a C atom of a formyl group on the N of the terminal methionine) and one water removed for a total of 433. Although it was evident that an over-all B value of 12 Å^2 was too large for atoms on the inside of the molecule and too small for those on the outside, no change was made since we wished to know the effect on R only of the positional shifts and added atoms. R decreased to 0·321.

In the second ΔF map the large shifts in B values suggested by the first map were also clearly indicated. Accordingly, four different B values were applied: 9, 12, 15 and 20 Å^2, the smaller for atoms inside the molecule, the larger for atoms on the outside. Shifts in the positional parameters were calculated as before, and six more atoms were added to make a total of 439. R decreased to 0·289.

From the third ΔF map, both the number and range range of B values were increased: 7, 9, 12, 15, 20, 25 and 30 Å^2. Positional shifts were again calculated, two atoms were added and one removed to make a total of 440. R decreased to 0·262.

Up to this point, the number of parameters had purposely been restricted to ensure that the decrease in R represented primarily improved parameters rather than addition of parameters. The behavior of the group of the very lowest-angle reflections during the first three refinement cycles suggested, however, that the numerous small peaks in the solution did correspond to structure and should be included in the model. Accordingly, oxygen atoms were added where peaks occurred in *both* the 2 Å F_o map (based on experimentally determined phases) *and* in the third ΔF map. Occupancies were taken proportional to the heights in the electron density map relative to that of O(39) as unity. Only peaks corresponding to occupancies greater than 0·3 were included. By adding 72 water O atoms that fulfilled the above criteria with B values of 25 Å, R decreased only to 0·260.

A fourth ΔF map was calculated and parameter shifts estimated for all atoms as in previous cycles. 21 thermal parameters ranging from 6–40 were applied. 11 water O atoms in the solution were removed and

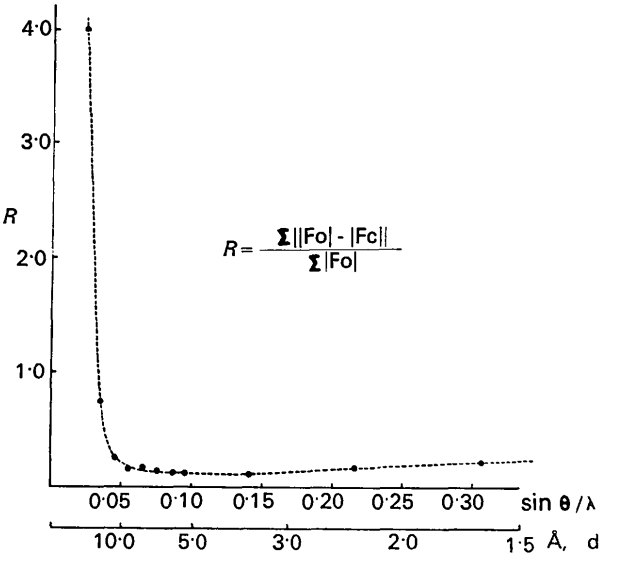


Fig. 2. R as a function of sin θ/λ after second least-squares refinement cycle.

Table 2. *Observed and calculated structure factors*

Each set of five columns contains respectively k, l, $|F_o|$, $|F_c|$, and phase angle (°).

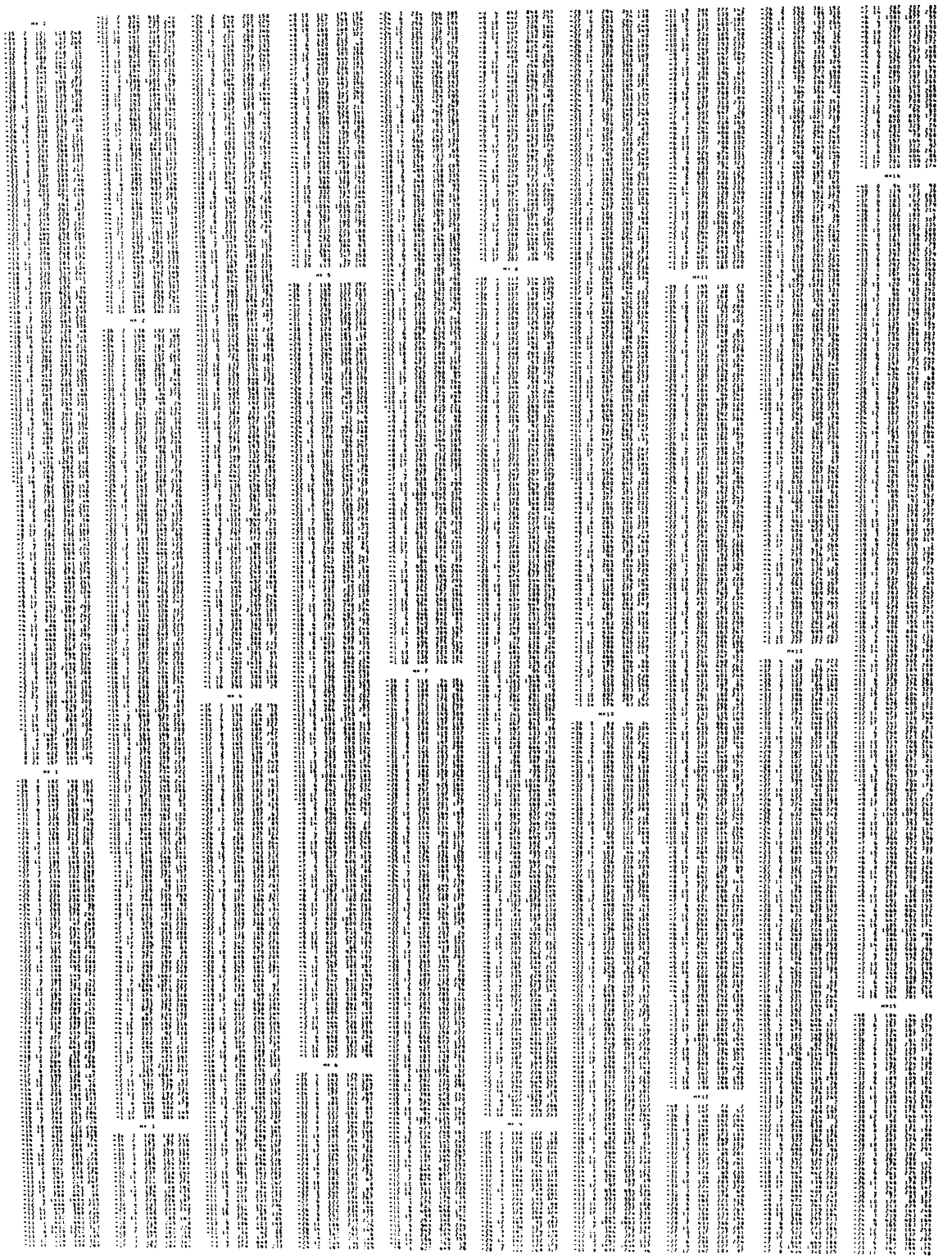
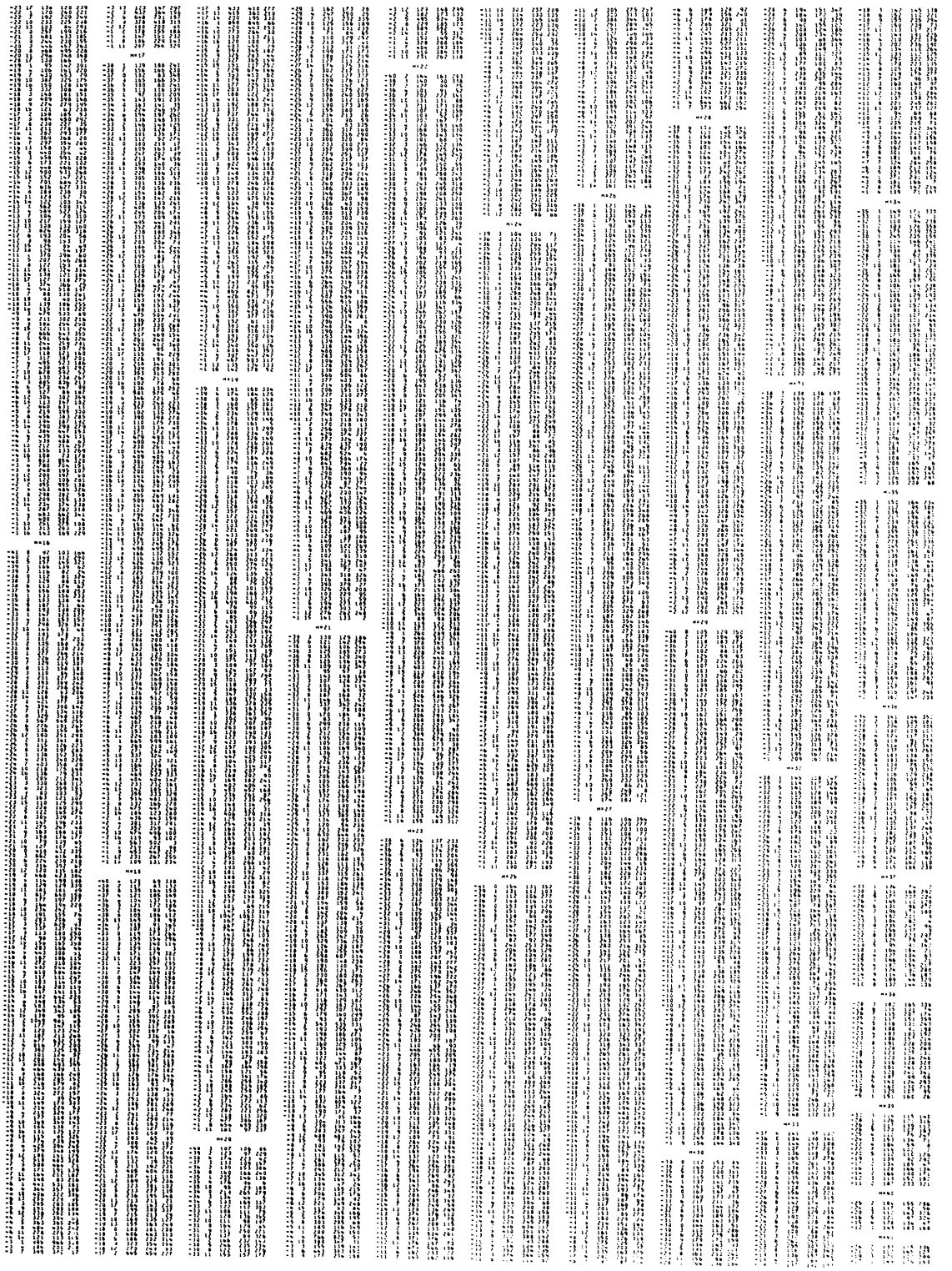


Table 2 (*cont.*)



16 new ones added. Three atoms were added to the molecule (including a trapped water near the C terminus) and one was removed. For comparison with the preceding cycle, R was calculated without the 77 water oxygen atoms from the third and fourth ΔF maps. On the basis of the 442 remaining atoms, R had decreased

to 0·242. Including the 77 additional water oxygen atoms, *i.e.* for the model with 519 atoms, *R* decreased to 0·224.

In the refinement by ΔF syntheses, shifts in parameters were tabulated by hand and new values key-punched. Several copy and key-punch errors were made

Table 3. *Atomic parameters referred to the hexagonal unit cell*

The temperature factor is of the form $\exp(-B\sin^2\theta/\lambda^2)$.

		x	y	z	B
FE		.1225(2)	-.1975(2)	.4764(4)	9.5
*****N-FORMYL*****					
C	0	.161 (7)	-.241 (9)	-.091 (9)	6.4
O	0	.148 (5)	-.264 (5)	-.071 (7)	67.9
*****MET*****					
N	1	.188 (7)	-.209 (8)	-.059 (9)	59.3
C	1A	.194 (6)	-.188 (7)	-.067(10)	68.1
C	1B	.228 (2)	-.177 (2)	-.098 (4)	16.8
C	1G	.229 (4)	-.203 (4)	-.103 (6)	32.5
S	1D	.266 (1)	-.192 (1)	-.072 (2)	28.9
C	1E	.296 (3)	-.156 (3)	-.067 (5)	22.8
C	1	.194 (4)	-.169 (4)	-.074 (5)	19.9
O	1	.207 (2)	-.147 (3)	-.082 (4)	26.4
*****LYS*****					
N	2	.168 (2)	-.187 (3)	-.056 (4)	20.8
C	2A	.167 (3)	-.171 (3)	-.038 (5)	18.1
C	2B	.134 (2)	-.193 (2)	-.045 (4)	13.3
C	2G	.127 (4)	-.207 (4)	-.066 (9)	34.9
C	2D	.127 (3)	-.201 (4)	-.113 (7)	27.2
C	2E	.133 (4)	-.180 (4)	-.157 (7)	30.5
N	2Z	.158 (5)	-.149 (5)	-.129 (8)	52.6
C	2	.176 (2)	-.175 (3)	.008 (4)	11.0
O	2	.178 (1)	-.187 (1)	.025 (2)	13.1
*****LYS*****					
N	3	.167 (1)	-.159 (1)	.021 (3)	10.0
C	3A	.162 (3)	-.159 (4)	.066 (6)	25.6
C	3B	.169 (2)	-.135 (2)	.069 (3)	13.8
C	3G	.193 (2)	-.107 (3)	.059 (3)	15.1
C	3D	.194 (2)	-.086 (3)	.056 (4)	18.1
C	3E	.217 (3)	-.060 (3)	.047 (5)	30.9
N	3Z	.195 (5)	-.042 (5)	.070 (8)	58.5
C	3	.139 (2)	-.179 (2)	.078 (4)	7.9
O	3	.125 (1)	-.187 (1)	.059 (2)	12.0
*****TYR*****					
N	4	.149 (2)	-.178 (2)	.124 (3)	14.0
C	4A	.129 (2)	-.198 (2)	.143 (3)	11.2
C	4B	.128 (2)	-.225 (2)	.151 (3)	4.9
C	4G	.116 (3)	-.242 (3)	.133 (5)	16.3
C	4D1	.098 (2)	-.261 (2)	.118 (3)	9.8
C	4D2	.129 (1)	-.233 (1)	.096 (3)	7.9
C	4E1	.095 (2)	-.276 (2)	.073 (4)	11.7
C	4E2	.129 (1)	-.243 (2)	.046 (2)	6.0
C	4Z	.109 (2)	-.267 (2)	.044 (3)	6.8
O	4H	.099 (1)	-.284 (1)	.011 (3)	15.9
C	4	.131 (2)	-.180 (1)	.174 (3)	6.9
O	4	.145 (1)	-.172 (1)	.210 (2)	8.2
*****THR*****					
N	5	.109 (1)	-.192 (1)	.139 (2)	8.7
C	5A	.102 (1)	-.186 (1)	.242 (3)	6.3
C	5B	.088 (2)	-.169 (2)	.226 (5)	16.4
O	5G1	.091 (2)	-.159 (2)	.271 (3)	18.6
C	5G2	.084 (2)	-.162 (3)	.194 (5)	24.6
C	5	.089 (2)	-.204 (2)	.280 (3)	5.6
O	5	.073 (1)	-.223 (1)	.258 (2)	8.6
*****CYS*****					
N	6	.104 (1)	-.192 (1)	.304 (2)	4.2
C	6A	.091 (1)	-.214 (1)	.343 (2)	4.9
C	6B	.105 (1)	-.204 (1)	.385 (2)	5.4
S	6G	.0939(4)	-.2201(4)	.4261(7)	4.6
C	6	.074 (1)	-.206 (1)	.354 (2)	5.1
O	6	.076 (1)	-.190 (1)	.370 (1)	6.0
*****THR*****					
N	7	.046 (1)	-.229 (1)	.351 (2)	10.2
C	7A	.024 (2)	-.231 (2)	.363 (4)	12.4
C	7B	.008 (2)	-.246 (2)	.330 (4)	12.7
O	7G1	-.004 (1)	-.269 (1)	.363 (2)	9.4
C	7G2	.007 (2)	-.246 (2)	.281 (4)	13.3
C	7	.018 (4)	-.230 (3)	.407 (6)	26.2
O	7	.000 (2)	-.222 (2)	.419 (3)	18.2
*****ILE---VAL*****					
N	8	.035 (2)	-.230 (2)	.429 (3)	14.7
C	8A	.031 (2)	-.227 (2)	.471 (3)	10.1
C	8B	.037 (2)	-.241 (2)	.504 (4)	18.6
C	8G1	.045 (2)	-.236 (2)	.551 (3)	10.4
C	8G2	.018 (2)	-.269 (2)	.486 (4)	17.0
C	8D	.016 (5)	-.295 (5)	.464 (8)	45.8
C	8	.044 (2)	-.200 (2)	.480 (3)	9.7
O	8	.037 (1)	-.195 (1)	.511 (2)	13.6
*****CYS*****					
N	9	.067 (2)	-.187 (2)	.481 (2)	8.8
C	9A	.081 (2)	-.168 (2)	.499 (3)	9.4
C	9B	.103 (2)	-.157 (2)	.517 (3)	15.1
S	9G	.1293(4)	-.1585(4)	.4818(7)	11.4
C	9	.090 (2)	-.146 (2)	.467 (4)	8.6
O	9	.098 (1)	-.123 (1)	.467 (2)	9.0
*****GLY*****					
N	10	.092 (1)	-.144 (1)	.424 (3)	6.6
C	10A	.097 (2)	-.133 (2)	.382 (4)	16.0
C	10	.120 (3)	-.124 (2)	.371 (4)	11.6
O	10	.134 (1)	-.107 (1)	.340 (2)	11.3
*****TYR*****					
N	11	.136 (1)	-.128 (1)	.386 (2)	5.6
C	11A	.160 (2)	-.126 (2)	.365 (4)	9.7
C	11B	.170 (2)	-.132 (1)	.399 (3)	9.0
C	11G	.185 (3)	-.140 (3)	.384 (4)	17.2
C	11D1	.212 (3)	-.124 (2)	.383 (3)	15.8
C	11D2	.191 (3)	-.157 (3)	.356 (5)	21.2
C	11E1	.236 (2)	-.120 (2)	.363 (4)	15.7
C	11E2	.213 (3)	-.155 (2)	.337 (3)	10.0
C	11Z	.232 (3)	-.142 (4)	.340 (5)	17.7
O	11H	.254 (1)	-.137 (1)	.318 (2)	12.4
C	11	.156 (2)	-.132 (2)	.322 (3)	9.5
O	11	.139 (1)	-.131 (1)	.313 (2)	9.8
*****VAL---ILE*****					
N	12	.175 (1)	-.116 (1)	.306 (3)	8.3
C	12A	.177 (2)	-.117 (2)	.263 (3)	8.1
C	12B	.176 (2)	-.098 (2)	.228 (5)	16.5
C	12G1	.180 (2)	-.104 (2)	.185 (3)	9.8
C	12G2	.157 (2)	-.092 (2)	.230 (3)	14.5
C	12	.194 (3)	-.122 (2)	.256 (4)	13.1
O	12	.219 (1)	-.110 (1)	.265 (2)	10.7
*****TYR*****					
N	13	.189 (1)	-.144 (1)	.234 (2)	7.0
C	13A	.202 (2)	-.153 (2)	.214 (4)	11.9
C	13B	.201 (1)	-.178 (1)	.213 (2)	7.3
C	13G	.218 (1)	-.183 (1)	.195 (3)	3.0
C	13D1	.239 (1)	-.178 (1)	.221 (2)	3.2
C	13D2	.218 (1)	-.188 (1)	.154 (3)	5.1
C	13E1	.258 (2)	-.182 (2)	.197 (4)	10.4
C	13E2	.231 (2)	-.194 (2)	.138 (3)	8.5
C	13Z	.249 (2)	-.190 (2)	.155 (3)	8.4
O	13H	.271 (1)	-.190 (1)	.146 (2)	7.6
C	13	.212 (2)	-.143 (2)	.173 (3)	6.8
O	13	.197 (1)	-.147 (1)	.142 (2)	7.3
*****ASX*****					
N	14	.246 (1)	-.122 (1)	.171 (2)	7.6
C	14A	.252 (2)	-.113 (2)	.130 (3)	10.8
C	14B	.260 (2)	-.084 (2)	.142 (4)	16.4
C	14G	.274 (3)	-.070 (4)	.104 (7)	18.9
O	14D1	.274 (1)	-.078 (1)	.068 (3)	16.6
O	14D2	.276 (3)	-.049 (3)	.102 (5)	35.9
C	14	.265 (2)	-.126 (2)	.111 (4)	9.8
O	14	.285 (1)	-.113 (1)	.136 (2)	8.2
*****PRO*****					
N	15	.255 (3)	-.138 (2)	.075 (5)	15.4
C	15A	.275 (3)	-.142 (2)	.063 (3)	12.6
C	15B	.257 (2)	-.161 (2)	.023 (3)	12.7
C	15G	.244 (2)	-.149 (2)	.023 (3)	12.0
C	15D	.233 (2)	-.138 (2)	.056 (3)	15.5
C	15	.308 (4)	-.128 (3)	.057 (5)	7.8
O	15	.313 (3)	-.131 (3)	.054 (5)	28.1
*****GLX*****					
N	16	.302 (1)	-.105 (1)	.046 (2)	7.4
C	16A	.328 (2)	-.086 (2)	.038 (3)	12.0
C	16B	.329 (2)	-.060 (2)	.028 (4)	22.4
C	16G	.325 (3)	-.066 (3)	-.017 (6)	26.9
C	16D	.332(13)	-.044(13)	-.040(24)	105.5
O	16E1	.339 (4)	-.017 (4)	-.023 (7)	48.7
O	16E2	.332 (3)	-.049 (3)	-.079 (5)	43.1
C	16	.350 (2)	-.074 (2)	.069 (3)	8.6
O	16	.370 (2)	-.065 (2)	.073 (3)	19.4
*****ASX*****					
N	17	.333 (1)	-.081 (1)	.099 (3)	8.7
C	17A	.339 (1)	-.074 (2)	.133 (3)	10.2
C	17B	.338 (2)	-.061 (2)	.176 (3)	14.6
C	17G	.346 (5)	-.027 (5)	.150 (8)	19.4
O	17D1	.321 (3)	-.041 (2)	.178 (4)	29.0
O	17D2	.358 (2)	-.026 (2)	.134 (4)	27.8
C	17	.348 (2)	-.092 (2)	.172 (4)	9.8
O	17	.361 (2)	-.090 (1)	.199 (3)	16.4

		x	y	z	B
*****GLY*****					
N	18	.324 (1)	-.113 (2)	.161 (2)	7.4
C	18A	.324 (1)	-.133 (2)	.182 (4)	8.9
C	18	.323 (2)	-.139 (2)	.223 (5)	8.3
O	18	.313 (1)	-.124 (1)	.242 (2)	11.9
*****ASX*****					
N	19	.325 (1)	-.146 (1)	.257 (4)	8.5
C	19A	.326 (2)	-.148 (1)	.294 (4)	6.3
C	19B	.300 (1)	-.164 (1)	.320 (3)	7.9
C	19G	.294 (2)	-.170 (2)	.366 (4)	6.6
O	19D1	.310 (1)	-.148 (1)	.384 (2)	13.8
O	19D2	.286 (1)	-.182 (1)	.390 (3)	14.7
C	19	.340 (2)	-.157 (2)	.316 (3)	7.4
O	19	.342 (1)	-.168 (1)	.331 (2)	8.7
*****PRO*****					
N	20	.366 (2)	-.133 (2)	.303 (2)	10.2
C	20A	.386 (3)	-.136 (2)	.308 (4)	12.3
C	20B	.412 (2)	-.113 (3)	.293 (3)	14.6
C	20G	.399 (3)	-.100 (2)	.288 (3)	12.5
C	20D	.377 (2)	-.107 (2)	.292 (3)	13.4
C	20	.396 (2)	-.145 (2)	.351 (4)	8.6
O	20	.402 (1)	-.156 (1)	.367 (2)	9.7
*****GLX---ASX*****					
N	21	.390 (1)	-.132 (1)	.379 (3)	11.4
C	21A	.389 (2)	-.139 (2)	.427 (4)	13.2
C	21B	.386 (3)	-.122 (3)	.448 (5)	15.7
C	21G	.402 (5)	-.104 (5)	.459 (7)	30.9
C	21D	.428 (5)	-.081 (5)	.421(12)	29.0
C	21E1	.433 (4)	-.061 (4)	.393 (6)	64.4
O	21E2	.450 (4)	-.089 (4)	.389 (6)	64.1
C	21	.374 (2)	-.161 (2)	.448 (4)	10.0
O	21	.381 (1)	-.171 (2)	.465 (3)	17.1
*****ASX*****					
N	22	.356 (2)	-.177 (2)	.418 (3)	14.1
C	22A	.342 (2)	-.204 (2)	.406 (4)	9.6
C	22B	.317 (3)	-.205 (3)	.426 (5)	8.6
C	22G	.307 (5)	-.193 (5)	.448 (9)	33.1
O	22D1	.316 (3)	-.201 (3)	.496 (5)	37.7
O	22D2	.309 (2)	-.179 (2)	.466 (3)	24.0
C	22	.356 (3)	-.205 (2)	.373 (4)	8.7
O	22	.331 (1)	-.231 (2)	.387 (2)	14.2
*****GLY*****					
N	23	.368 (3)	-.199 (3)	.354 (5)	28.7
C	23A	.366 (2)	-.220 (2)	.339 (4)	14.2
C	23	.359 (2)	-.213 (3)	.290 (4)	17.5
O	23	.354 (2)	-.232 (2)	.291 (3)	18.7
*****ILE---VAL*****					
N	24	.345 (1)	-.207 (1)	.281 (2)	7.4
C	24A	.332 (1)	-.211 (1)	.242 (2)	4.4
C	24B	.309 (2)	-.215 (2)	.260 (3)	11.5
C	24G1	.290 (1)	-.235 (1)	.273 (2)	5.9
C	24G2	.295 (2)	-.220 (2)	.216 (4)	14.0
C	24D	.340 (2)	-.204 (2)	.196 (3)	15.0
C	24	.345 (1)	-.187 (2)	.228 (2)	9.3
O	24	.354 (1)	-.165 (1)	.232 (2)	8.8
*****ASX*****					
N	25	.359 (1)	-.189 (1)	.196 (2)	8.8
C	25A	.377 (2)	-.174 (2)	.171 (3)	8.3
C	25B	.394 (1)	-.177 (1)	.143 (3)	7.6
C	25G	.412 (2)	-.185 (2)	.168 (4)	7.6
O	25D1	.426 (2)	-.186 (1)	.149 (3)	18.2
O	25D2	.409 (1)	-.183 (1)	.209 (2)	8.2
C	25	.367 (3)	-.162 (3)	.147 (5)	17.9
O	25	.349 (1)	-.170 (1)	.121 (2)	8.3
*****PRO*****					
N	26	.380 (2)	-.136 (2)	.133 (3)	13.9
C	26A	.374 (1)	-.125 (2)	.093 (3)	8.3
C	26B	.403 (1)	-.096 (1)	.095 (3)	8.9
C	26G	.406 (2)	-.690 (2)	.135 (4)	23.3
C	26D	.405 (2)	-.117 (2)	.147 (3)	14.4
C	26	.372 (2)	-.143 (2)	.058 (3)	7.6
O	26	.382 (2)	-.152 (2)	.040 (3)	18.1
*****GLY*****					
N	27	.353 (1)	-.140 (1)	.037 (3)	10.0
C	27A	.340 (3)	-.149 (3)	.006 (5)	19.9
C	27	.327 (2)	-.176 (2)	-.004 (3)	9.0
O	27	.313 (1)	-.193 (1)	-.027 (2)	14.7
*****THR*****					
N	28	.315 (1)	-.189 (2)	.037 (3)	10.5
C	28A	.298 (2)	-.215 (2)	.051 (3)	9.3
C	28B	.302 (2)	-.222 (2)	.091 (3)	9.8
C	28G1	.322 (1)	-.219 (1)	.106 (2)	7.4
C	28G2	.287 (2)	-.246 (2)	.105 (3)	12.1
C	28	.275 (2)	-.222 (2)	.049 (4)	9.5
O	28	.265 (1)	-.207 (1)	.073 (2)	6.9
*****ASX*****					
N	29	.257 (2)	-.234 (1)	.022 (3)	11.5
C	29A	.234 (2)	-.242 (1)	.011 (3)	10.4
C	29B	.221 (3)	-.265 (3)	-.017 (5)	19.8
C	29G	.195 (8)	-.256 (6)	-.058(12)	24.9
O	29D1	.172 (3)	-.271 (3)	-.023 (5)	36.6
O	29D2	.206 (5)	-.258 (4)	-.052 (7)	58.7
C	29	.217 (2)	-.258 (2)	.051 (3)	9.6
O	29	.226 (1)	-.266 (1)	.070 (2)	8.7
*****PHE*****					
N	30	.200 (2)	-.251 (2)	.061 (3)	15.9
C	30A	.188 (2)	-.264 (2)	.095 (3)	14.3
C	30B	. 63 (2)	-.260 (2)	.107 (3)	12.5
C	30G	.151 (2)	-.270 (2)	.142 (3)	9.9
C	30D1	.129 (2)	-.289 (2)	.136 (3)	6.2
C	30D2	.165 (2)	-.253 (2)	.176 (4)	10.1
C	30E1	.123 (2)	-.292 (2)	.180 (4)	14.0
C	30E2	.158 (2)	-.259 (2)	.214 (4)	14.3
C	30Z	.134 (2)	-.280 (2)	.220 (3)	9.1
C	30	.176 (1)	-.296 (1)	.105 (3)	5.4
O	30	.178 (1)	-.302 (1)	.131 (2)	11.3
*****LYS*****					
N	31	.165 (1)	-.327 (2)	.065 (3)	12.3
C	31A	.158 (2)	-.332 (2)	.062 (3)	9.4
C	31B	.147 (2)	-.342 (2)	.024 (5)	15.3
C	31G	.156 (3)	-.340 (3)	-.014 (6)	26.1
C	31D	.142 (2)	-.361 (2)	-.058 (4)	18.7
C	31E	.157 (4)	-.362 (4)	-.093 (7)	49.4
N	31Z	.134 (4)	-.394 (4)	-.154 (6)	48.6
C	31	.175 (2)	-.344 (2)	.080 (4)	11.5
O	31	.165 (2)	-.360 (2)	.104 (4)	31.3
*****ASX*****					
N	32	.199 (1)	-.322 (1)	.072 (2)	9.7
C	32A	.219 (2)	-.327 (2)	.066 (3)	8.1
C	32B	.244 (2)	-.314 (1)	.044 (2)	5.7
C	32G	.234 (2)	-.313 (3)	-.015 (3)	7.2
O	32D1	.223 (2)	-.338 (2)	-.001 (3)	19.5
O	32D2	.253 (4)	-.299 (4)	-.013 (6)	42.2
C	32	.228 (3)	-.320 (3)	.107 (6)	18.0
O	32	.245 (1)	-.323 (1)	.126 (2)	10.6
*****ILE*****					
N	33	.229 (1)	-.309 (1)	.139 (3)	9.5
C	33A	.227 (2)	-.307 (2)	.181 (5)	14.6
C	33B	.222 (1)	-.287 (1)	.213 (3)	7.7
C	33G1	.231 (1)	-.261 (1)	.186 (2)	7.2
C	33D	.222 (1)	-.246 (1)	.206 (2)	8.6
C	33G2	.225 (1)	-.286 (1)	.261 (2)	5.8
C	33	.221 (2)	-.326 (2)	.211 (3)	7.6
O	33	.198 (1)	-.341 (1)	.203 (2)	9.7
*****PRO*****					
N	34	.234 (2)	-.332 (1)	.224 (3)	6.6
C	34A	.227 (2)	-.351 (2)	.255 (3)	11.9
C	34B	.249 (2)	-.349 (2)	.284 (3)	13.0
C	34G	.268 (1)	-.331 (1)	.264 (2)	8.4
C	34D	.263 (2)	-.313 (2)	.229 (3)	16.4
C	34	.204 (3)	-.363 (3)	.278 (3)	14.3
O	34	.201 (1)	-.346 (1)	.297 (2)	6.6
*****ASX*****					
N	35	.183 (2)	-.381 (1)	.281 (2)	8.7
C	35A	.157 (1)	-.391 (1)	.289 (2)	5.7
C	35B	.143 (1)	-.414 (1)	.278 (2)	2.5
C	35G	.135 (2)	-.412 (2)	.234 (3)	8.8
O	35D1	.125 (1)	-.434 (1)	.214 (2)	11.5
O	35D2	.138 (1)	-.394 (1)	.218 (2)	7.7
C	35	.155 (2)	-.390 (2)	.332 (4)	12.1
O	35	.141 (1)	-.391 (1)	.360 (2)	8.0
*****ASX*****					
N	36	.175 (1)	-.384 (1)	.366 (3)	8.6
C	36A	.181 (2)	-.374 (2)	.407 (4)	15.4
C	36B	.186 (2)	-.391 (2)	.424 (4)	15.6
C	36G	.205 (2)	-.389 (2)	.424 (3)	9.3
O	36D1	.227 (2)	-.387 (1)	.437 (2)	17.1
O	36D2	.228 (1)	-.371 (1)	.394 (2)	13.6
C	36	.190 (2)	-.348 (2)	.438 (4)	10.0
O	36	.193 (1)	-.349 (1)	.467 (3)	14.9

		x	y	z	B
*****TRP*****					
N	37	.199 (1)	-.335 (1)	.386 (2)	8.1
C	37A	.209 (1)	-.310 (1)	.401 (2)	8.3
C	37B	.214 (2)	-.301 (2)	.354 (3)	11.6
C	37G	.229 (2)	-.273 (2)	.355 (2)	9.8
C	37D1	.245 (2)	-.253 (2)	.359 (2)	6.4
C	37D2	.216 (2)	-.254 (2)	.349 (3)	9.7
N	37E1	.254 (1)	-.230 (2)	.357 (2)	9.2
C	37E2	.234 (2)	-.231 (2)	.349 (2)	5.9
C	37E3	.190 (2)	-.268 (2)	.334 (3)	11.9
C	37Z2	.220 (1)	-.216 (1)	.340 (2)	8.7
C	37Z3	.183 (2)	-.252 (2)	.326 (3)	16.4
C	37H2	.194 (2)	-.229 (2)	.330 (2)	7.8
C	37	.187 (2)	-.306 (2)	.411 (3)	8.3
O	37	.167 (1)	-.319 (1)	.399 (2)	11.1
*****VAL*****					
N	38	.196 (1)	-.288 (1)	.438 (2)	6.1
C	38A	.181 (2)	-.281 (1)	.456 (3)	6.2
C	38B	.172 (2)	-.283 (2)	.502 (3)	10.0
C	38G1	.164 (2)	-.317 (2)	.509 (3)	13.8
C	38G2	.190 (2)	-.282 (1)	.528 (2)	11.0
C	38	.169 (2)	-.254 (2)	.448 (2)	7.5
O	38	.21[illegible] (1)	-.239 (1)	.441 (2)	11.0
*****CYS*****					
N	39	.171 (1)	-.245 (2)	.452 (2)	12.1
C	39A	.178 (2)	-.222 (2)	.452 (3)	9.0
C	39B	.154 (1)	-.225 (1)	.463 (2)	6.3
S	39G	.1580(4)	-.1927(4)	.4560(8)	12.9
C	39	.197 (2)	-.207 (2)	.492 (4)	10.0
O	39	.188 (1)	-.219 (1)	.522 (3)	14.0
*****PRO*****					
N	40	.211 (2)	-.192 (2)	.467 (4)	8.3
C	40A	.234 (3)	-.176 (2)	.490 (4)	13.6
C	40B	.247 (2)	-.163 (2)	.472 (4)	13.4
C	40G	.249 (2)	-.154 (2)	.431 (4)	15.3
C	40D	.224 (2)	-.180 (2)	.436 (4)	11.2
C	40	.234 (3)	-.156 (3)	.526 (6)	19.4
O	40	.234 (1)	-.162 (1)	.555 (3)	15.9
*****ILE---LEU*****					
N	41	.211 (2)	-.157 (2)	.523 (3)	15.4
C	41A	.201 (3)	-.142 (4)	.555 (6)	14.3
C	41B	.204 (5)	-.119 (6)	.513 (9)	16.7
C	41G1	.189 (3)	-.144 (3)	.504 (5)	27.9
C	41G2	.202 (4)	-.106 (4)	.526 (7)	47.2
C	41D	.229 (2)	-.104 (2)	.496 (3)	13.8
C	41	.185 (4)	-.156 (4)	.583 (7)	28.0
O	41	.177 (3)	-.168 (3)	.515 (5)	27.0
*****CYS*****					
N	42	.162 (2)	-.184 (2)	.579 (2)	11.7
C	42A	.142 (3)	-.198 (3)	.595 (4)	20.1
C	42B	.119 (2)	-.198 (2)	.580 (3)	11.6
S	42G	.1079(5)	-.2186(5)	.5279(9)	14.5
C	42	.140 (2)	-.224 (2)	.595 (3)	7.9
O	42	.126 (1)	-.237 (1)	.620 (2)	15.0
*****GLY*****					
N	43	.154 (1)	-.232 (2)	.579 (2)	11.5
C	43A	.143 (2)	-.255 (3)	.581 (3)	11.2
C	43	.134 (3)	-.275 (3)	.576 (4)	25.7
O	43	.135 (1)	-.291 (1)	.579 (2)	15.8
*****ILE---VAL*****					
N	44	.122 (1)	-.276 (1)	.549 (2)	11.4
C	44A	.101 (2)	-.292 (2)	.527 (3)	10.0
C	44B	.069 (2)	-.296 (2)	.502 (4)	20.2
C	44G1	.076 (2)	-.277 (2)	.475 (3)	12.2
C	44G2	.064 (2)	-.287 (2)	.547 (5)	19.7
C	44D	.077 (3)	-.277 (3)	.577 (6)	34.7
C	44	.113 (2)	-.301 (2)	.479 (3)	8.7
O	44	.128 (1)	-.286 (1)	.459 (2)	14.0
*****GLY*****					
N	45	.095 (2)	-.322 (2)	.478 (3)	12.2
C	45A	.100 (2)	-.339 (2)	.446 (4)	17.3
C	45	.096 (1)	-.326 (2)	.408 (3)	9.8
O	45	.085 (1)	-.316 (1)	.391 (1)	6.5
*****LYS*****					
N	46	.102 (1)	-.332 (1)	.364 (2)	4.4
C	46A	.099 (1)	-.327 (1)	.320 (2)	3.6
C	46B	.114 (2)	-.337 (1)	.295 (2)	9.3
C	46G	.134 (1)	-.327 (1)	.294 (2)	8.1
C	46D	.151 (1)	-.334 (1)	.262 (3)	8.1
C	46E	.151 (1)	-.329 (1)	.217 (3)	6.8
N	46Z	.156 (1)	-.341 (1)	.186 (2)	4.9
C	46	.072 (1)	-.341 (1)	.308 (2)	4.4
O	46	.070 (1)	-.328 (1)	.277 (2)	11.3
*****ASX*****					
N	47	.055 (1)	-.351 (1)	.329 (2)	5.2
C	47A	.028 (2)	-.371 (2)	.323 (3)	10.9
C	47B	.020 (2)	-.395 (2)	.331 (3)	13.9
C	47G	.015 (2)	-.393 (2)	.373 (5)	11.0
O	47D1	.005 (3)	-.418 (3)	.398 (5)	44.8
O	47D2	.028 (2)	-.380 (2)	.403 (3)	21.5
C	47	.016 (2)	-.358 (2)	.375 (3)	9.4
O	47	-.001 (1)	-.359 (1)	.310 (2)	16.6
*****GLX*****					
N	48	.031 (1)	-.335 (1)	.347 (2)	7.2
C	48A	.025 (1)	-.316 (1)	.357 (3)	3.9
C	48B	.025 (1)	-.310 (1)	.395 (3)	9.3
C	48G	.018 (2)	-.329 (2)	.433 (3)	16.3
C	48D	-.010 (3)	-.335 (3)	.426 (5)	21.3
O	48E1	-.021 (2)	-.366 (2)	.423 (2)	19.5
O	48E2	-.027 (1)	-.339 (1)	.417 (2)	15.2
C	48	.038 (1)	-.298 (1)	.324 (3)	4.1
O	48	.032 (1)	-.279 (1)	.334 (2)	9.0
*****PHE*****					
N	49	.052 (1)	-.292 (1)	.286 (2)	8.6
C	49A	.067 (2)	-.272 (2)	.269 (3)	10.1
C	49B	.087 (2)	-.271 (1)	.257 (3)	8.4
C	49G	.101 (1)	-.259 (1)	.286 (3)	3.0
C	49D1	.105 (1)	-.266 (1)	.323 (2)	3.2
C	49D2	.120 (1)	-.236 (1)	.277 (2)	7.3
C	49E1	.123 (2)	-.255 (2)	.360 (3)	8.3
C	49E2	.142 (1)	-.222 (1)	.321 (3)	9.4
C	49Z	.138 (2)	-.235 (2)	.352 (4)	12.3
C	49	.043 (1)	-.278 (1)	.234 (2)	6.7
O	49	.035 (1)	-.303 (1)	.219 (2)	9.9
*****GLX*****					
N	50	.045 (1)	-.262 (1)	.206 (2)	9.7
C	50A	.033 (1)	-.267 (1)	.168 (2)	6.7
C	50B	.016 (2)	-.260 (2)	.182 (3)	16.2
C	50G	-.005 (2)	-.283 (3)	.206 (4)	22.9
C	50D	-.034 (2)	-.287 (2)	.198 (3)	14.6
O	50E1	-.048 (2)	-.309 (2)	.203 (3)	20.2
O	50E2	-.027 (1)	-.260 (1)	.202 (2)	19.0
C	50	.053 (2)	-.250 (2)	.135 (3)	10.9
O	50	.068 (1)	-.226 (1)	.146 (2)	10.8
*****GLX*****					
N	51	.052 (1)	-.252 (1)	.094 (2)	12.0
C	51A	.063 (2)	-.234 (3)	.067 (4)	10.0
C	51B	.061 (2)	-.249 (2)	.023 (5)	16.4
C	51G	.039 (2)	-.275 (3)	.028 (4)	24.2
C	51D	.069 (5)	-.256 (5)	-.019(10)	47.3
O	51E1	.046 (2)	-.286 (2)	-.033 (3)	29.6
O	51E2	.071 (2)	-.293 (2)	-.053 (4)	30.8
C	51	.060 (3)	-.218 (3)	.066 (3)	16.7
O	51	.035 (2)	-.221 (1)	.070 (2)	22.2
*****VAL*****					
N	52	.074 (2)	-.197 (3)	.067 (3)	20.5
C	52A	.076 (1)	-.173 (2)	.052 (3)	6.1
C	52B	.098 (3)	-.154 (3)	.055 (5)	17.7
C	52G1	.101 (3)	-.137 (3)	.049 (5)	17.2
C	52G2	.100 (2)	-.150 (2)	.096 (5)	18.6
C	52	.079 (4)	-.173 (3)	-.005 (8)	24.3
O	52	.092 (2)	-.170 (2)	-.012 (4)	34.3
*****GLX*****					
N	53	.050 (2)	-.181 (2)	-.010 (3)	17.6
C	53A	.040 (4)	-.184 (5)	-.071 (7)	31.0
C	53B	.023 (5)	-.206 (5)	-.067 (8)	37.5
C	53G	.022 (5)	-.216 (5)	-.094 (8)	42.5
C	53D	.023 (2)	-.240 (3)	-.082 (6)	21.3
O	53E1	.013 (2)	-.252 (2)	-.054 (4)	33.7
O	53E2	.030 (2)	-.241 (2)	-.128 (5)	31.9
C	53	.055 (4)	-.154 (4)	-.087 (7)	31.0
O	53	.075 (3)	-.148 (3)	-.119 (5)	51.9
*****GLX*****					
N	54	.045 (3)	-.147 (3)	-.071 (5)	39.5
C	54A	.047 (4)	-.126 (5)	-.109 (7)	46.6
C	54B	.069(20)	-.107(21)	-.139(34)	15.8
C	54G	.021(14)	-.145(13)	-.152(24)	78.0
C	54D	.027(11)	-.110(11)	-.132(17)	62.4
O	54E1	-.001 (3)	-.124 (3)	-.097 (4)	49.2
O	54E2	.018 (5)	-.097 (4)	-.129 (7)	72.0
C	54	.058 (5)	-.116 (5)	-.029 (9)	29.9
O	54-1	.043 (4)	-.105 (4)	-.018 (7)	63.0
O	54-2	.081 (7)	-.097 (7)	.000(12)	105.9

Table 3 (*cont.*)

	x	y	z	B	M
O 1W	.196 (4)	−.255 (4)	.830 (6)	46.8	.8
O 2W	.284 (5)	−.132 (5)	.832 (8)	53.9	.7
O 3W	.364 (5)	−.054 (5)	.888 (9)	57.2	.7
O 4W	.358 (6)	−.188 (6)	.835(10)	39.9	.6
O 5W	.070 (6)	−.194 (6)	.813(10)	42.6	.6
O 6W	.309 (3)	−.159 (4)	.802 (6)	45.8	.8
O 7W	.296 (3)	−.093 (3)	.815 (5)	26.9	.6
O 8W	.373 (2)	−.113 (2)	.797 (3)	30.8	1.0
O 9W	.128 (3)	−.182 (3)	.728 (6)	41.2	.8
O 10W	.346 (2)	−.061 (2)	.685 (4)	32.4	.9
O 11W	.219 (2)	−.214 (2)	.589 (3)	25.5	1.0
O 12W	.253 (1)	−.221 (1)	.464 (2)	19.3	1.0
O 13W	.129 (1)	−.361 (1)	.348 (2)	11.2	1.0
O 14W	.208 (1)	−.062 (1)	.336 (2)	16.1	1.0
O 15W	.274 (1)	−.111 (1)	.246 (2)	11.5	1.0
O 16W	.006 (2)	−.320 (2)	.908 (3)	30.8	1.0
O 17W	.358 (1)	−.236 (1)	.182 (2)	20.8	1.0
O 18W	.122 (1)	−.321 (1)	.859 (2)	21.9	1.0
O 19W	.075 (2)	−.329 (2)	.726 (4)	27.6	.9
O 20W	.072 (2)	−.312 (2)	.657 (3)	26.5	1.0
O 21W	.142 (1)	−.138 (1)	.966 (2)	19.3	1.0
O 22W	.107 (3)	−.106 (3)	.875 (5)	38.2	.9
O 23W	.189 (1)	−.221 (1)	.019 (2)	19.4	1.0
O 24W	.413 (2)	−.140 (2)	.976 (3)	22.5	.8
O 25W	.402 (2)	−.017 (2)	.961 (4)	27.6	.8
O 26W	.371 (5)	−.248 (5)	.933 (8)	29.6	.4
O 27W	.409 (3)	−.214 (3)	.925 (5)	17.4	.4
O 28W	.371 (4)	−.156 (4)	.922 (6)	14.6	.3
O 29W	.149 (4)	−.296 (4)	.905 (6)	27.6	.4
O 30W	.181 (4)	−.312 (4)	.888 (8)	36.4	.4
O 31W	.443 (5)	−.141 (5)	.877 (9)	49.0	.6
O 32W	.456 (6)	−.238 (6)	.870(10)	38.8	.4
O 33W	.219 (4)	−.169 (4)	.960 (7)	17.2	.3
O 34W	.148 (4)	−.263 (4)	.840 (6)	20.2	.4
O 35W	.230 (9)	−.261 (9)	.862(16)	46.4	.4
O 36W	.110 (3)	−.292 (3)	.783 (4)	38.6	.9
O 37W	.179 (6)	−.316 (6)	.759(10)	36.1	.4
O 38W	.162 (6)	−.367 (7)	.758(11)	38.7	.4
O 39W	.193 (3)	−.248 (3)	.752 (5)	21.2	.5
O 40W	.186(13)	−.171(13)	.720(22)	38.5	.4
O 41W	.349 (3)	−.210 (3)	.732 (6)	26.4	.4
O 42W	.187 (3)	−.218 (3)	.681 (5)	10.3	.3
O 43W	.181 (4)	−.177 (4)	.703 (7)	21.4	.3
O 44W	.143 (4)	−.212 (4)	.730 (7)	46.4	.7
O 45W	.120 (5)	−.355 (5)	.721 (8)	30.5	.5
O 46W	.219 (4)	−.128 (4)	.715 (6)	21.0	.4
O 47W	.336 (4)	−.030 (4)	.737 (8)	21.4	.3
O 48W	.325 (7)	−.150 (7)	.732(11)	21.3	.3
O 49W	.274(10)	−.309(10)	.755(17)	45.5	.4
O 50W	.400 (4)	−.172 (4)	.700 (6)	15.8	.3
O 51W	.244(14)	−.309(14)	.686(26)	86.2	.3
O 52W	.410 (3)	−.116 (3)	.688 (6)	44.0	.7
O 53W	.162 (4)	−.294 (4)	.631 (6)	15.6	.4
O 54W	.262 (4)	−.182 (4)	.618 (7)	23.4	.3
O 55W	.283 (4)	−.107 (4)	.608 (7)	37.2	.6
O 56W	.316(11)	−.272(10)	.552(18)	59.5	.3
O 57W	.424 (4)	−.056 (4)	.595 (7)	24.3	.4
O 58W	.278 (4)	−.300 (4)	.562 (7)	31.6	.5
O 59W	.364 (5)	−.264 (5)	.557 (9)	28.0	.4
O 60W	.005 (8)	−.433 (8)	.495(13)	65.6	.5
O 61W	.235 (4)	−.314 (4)	.490 (7)	22.2	.3
O 62W	.372 (3)	−.219 (3)	.513 (5)	20.7	.4
O 63W	.249 (2)	−.261 (2)	.468 (3)	4.5	.4
O 64W	.303 (9)	−.261 (9)	.481(17)	66.0	.3
O 65W	.330 (7)	−.031 (7)	.506(11)	31.0	.3
O 66W	.117 (2)	−.386 (2)	.464 (4)	14.2	.5
O 67W	.039 (3)	−.372 (3)	.473 (6)	27.1	.6
O 68W	.308 (2)	−.137 (2)	.463 (3)	5.2	.3
O 69W	.337 (4)	−.305 (4)	.401 (7)	25.4	.4
O 70W	.102 (4)	−.433 (4)	.444 (7)	31.3	.6
O 71W	.278 (2)	−.275 (2)	.423 (4)	32.3	.4
O 72W	.363 (2)	−.103 (2)	.389 (3)	23.0	1.0
O 73W	.267 (2)	−.323 (2)	.380 (4)	37.7	.9
O 74W	.294 (1)	−.107 (1)	.354 (2)	19.4	.9
O 75W	.274 (5)	−.055 (5)	.397 (8)	20.8	.3
O 76W	.087 (4)	−.417 (5)	.344 (8)	16.7	.3
O 77W	.335 (2)	−.303 (2)	.303 (3)	22.9	.8
O 78W	.291 (4)	−.060 (4)	.246 (7)	27.4	.4
O 79W	.381 (3)	−.052 (3)	.250 (5)	16.9	.4
O 80W	.289 (3)	−.306 (3)	.126 (4)	12.6	.5
O 81W	.382 (2)	−.252 (2)	.116 (4)	17.1	.5
O 82W	.424 (2)	−.187 (3)	.056 (4)	21.4	.7
O 83W	.349 (4)	−.230 (4)	.049 (6)	26.5	.5
O 84W	.377 (2)	−.194 (2)	.027 (4)	19.4	.6
O 85W	.389 (3)	−.230 (3)	.048 (5)	21.3	.5
O 86W	.413 (3)	−.115 (3)	.181 (5)	21.2	.5
O 87W	.050 (3)	−.341 (3)	.199 (5)	23.6	.6
O 88W	.353 (4)	−.285 (4)	.393 (6)	29.2	.5
O 89W	.372 (3)	−.296 (3)	.450 (5)	22.2	.5
O 90W	.465 (4)	−.126 (4)	.500 (6)	45.9	.6
O 91W	.387 (4)	−.161 (4)	.552 (6)	24.1	.5
O 92W	.034 (3)	−.407 (3)	.576 (5)	28.8	.6
O 93W	.209 (4)	−.294 (5)	.633 (8)	54.1	.6
O 94W	.104 (1)	−.281 (1)	.677 (3)	11.6	.6
O 95W	.153 (5)	−.254 (5)	.713 (8)	58.2	.6
O 96W	.168 (2)	−.230 (2)	.818 (3)	9.6	.5
O 97W	.255 (5)	−.176 (5)	.781 (8)	40.0	.5
O 98W	.114 (4)	−.349 (4)	.803 (8)	36.1	.5
O 99W	.354 (2)	−.107 (2)	.919 (4)	16.5	.6
O 100W	.277 (1)	−.263 (1)	.975 (2)	6.5	.6
O 101W	.219 (4)	−.105 (4)	.117 (7)	20.6	.3
O 102W	.441 (9)	−.163 (9)	.056(14)	17.4	.3
O 103W	.273 (4)	−.349 (4)	.408 (7)	13.6	.4
O 104W	.032 (2)	−.376 (2)	.500 (4)	16.3	.4
O 105W	.091 (5)	−.209 (5)	.932 (9)	37.8	.4
O 106W	.058 (3)	−.227 (3)	.755 (5)	18.9	.4
O 107W	.188 (4)	−.147 (4)	.829 (7)	28.9	.5
O 108W	.415 (6)	−.035 (5)	.164 (9)	52.1	.5
O 109W	.307 (4)	−.276 (4)	.184 (7)	25.5	.4
O 110W	.352 (2)	−.239 (2)	.255 (3)	9.5	.6
O 111W	.043 (4)	−.455 (4)	.316 (6)	21.2	.4
O 112W	.085 (3)	−.401 (3)	.288 (5)	13.4	.4
O 113W	.456 (4)	−.040 (4)	.290 (7)	28.7	.4
O 114W	.478 (4)	−.072 (4)	.306 (6)	21.6	.5
O 115W	.328 (2)	−.079 (2)	.307 (3)	12.8	.5
O 116W	.301 (3)	−.283 (3)	.343 (4)	13.4	.5
O 117W	.411 (4)	−.237 (4)	.376 (6)	30.6	.5
O 118W	.320 (3)	−.118 (3)	.452 (6)	13.3	.4
O 119W	.284 (3)	−.091 (3)	.468 (6)	18.1	.4
O 120W	.212 (3)	−.313 (3)	.605 (6)	22.2	.4
O 121W	.298 (3)	−.037 (3)	.619 (6)	29.5	.5
O 122W	.475 (4)	−.161 (4)	.656 (7)	37.4	.5
O 123W	.138 (4)	−.302 (4)	.667 (6)	20.2	.5
O 124W	.259 (6)	−.253 (6)	.702(11)	51.1	.4
O 125W	.241 (3)	−.237 (3)	.730 (5)	14.8	.4
O 127W	.201 (6)	−.323 (6)	.811(10)	55.0	.5
O 128W	.112 (4)	−.275 (4)	.941 (6)	36.4	.7
O 129W	.148 (8)	−.306 (9)	.939(16)	35.0	.3
O 130W	.142 (3)	−.292 (3)	.002 (6)	26.5	.7

and when detected were used to check the indicated corrections from the ΔF map for the known error. For an error up to about 1 Å, the calculated shift returned the atom roughly to its true position. For an error much greater than 1 Å, no appreciable shift was indicated and the negative peak in the map at the erroneous site was about equal to the peak which appeared at the position where the atom should have been

Least-squares refinement

The refinement by ΔF syntheses suggested that continued refinement by the method of least-squares was warranted (Watenpaugh, Sieker, Herriott & Jensen, 1971*a*, *b*; 1972). It is essential when applying this method in the present instance to use unrestricted B values for each atom so that incorrect or badly misplaced atoms can be removed (Rossmann, Jacobson, Hirshfeld & Lipscomb, 1959). The unrestricted B's increase the number of parameters to over 2000 with only slightly over 5000 observations. Although this is not greatly overdetermined, there is no reason in principle why the refinement should not be meaningful within the limit imposed by the extent and precision of the data.

Because of the large number of parameters, it was necessary to partition the matrix into 10 blocks and to refine one block of approximately 210–240 parameters in each pass. Off-diagonal terms in the matrix correlating atoms bonded together at distances of less than about 1·5 Å are important when data do not extend to smaller spacings. To allow at least in part for such interactions, blocks of parameters for successive passes were overlapped, *i.e.* the parameters of the last main-chain atom included in one block were also included in the next. The function minimized was $w(|F_o|-|F_c|)^2$ with unit weights.

After the first three passes involving refinement of atoms in the Fe–S complex, in 9 residues in the vicinity of the complex, and in the six N terminal residues, a fifth ΔF map was calculated in order to check the least-squares method. The map appeared satisfactory and indicated where a few additional atoms should be placed. Four more atoms were added to the molecule and six more water to the solution, a total of 529 atoms.

Ten passes were required to adjust 2117 parameters in the first least-squares refinement cycle, and R decreased to 0·179.

In refining a structure in space group $R3$ by least squares, the z parameter of one atom should be fixed to determine $z=0$. In rubredoxin we fixed the z coordinate for the Fe atom. This is wrong, however, when refining in passes because the fixed part of the molecule in each pass serves to determine $z=0$. If the z coordinate of one atom is never allowed to vary, it cannot adjust to the origin implied by those atoms which are necessarily fixed in each pass.

The effect of fixing the z coordinate for the Fe atom was clearly evident in the sixth ΔF synthesis calculated after the first least-squares cycle. The Fe coordinate had inadvertently been fixed at the z initially derived from the 2 Å map, and its shifts from the four ΔF refinement cycles had been negated.

Most of the shifts in parameters from the first least-squares cycle were reasonable although there were some exceptions. For example, the B value for $C^{\beta}(39)$ shifted to -5 Å^2, a physically impossible value. The sixth ΔF map clearly indicated that it should be increased to about 10 Å^2, a reasonable value and consistent with its position in the molecule. Inspection of the map raised doubt about the positions of the carbon atom of the formyl group and N(1) to which it is bonded as well as positions of some atoms in R12, R21, R29 and R41. Accordingly, the formyl carbon atom, N(1), and all atoms in the four R groups were removed and a seventh ΔF synthesis calculated. Of the 34 atoms involved, one was removed in the subsequent calculations and positions of the rest adjusted to fit the ΔF map. Three more atoms were added to the molecule and 21 water molecules to the solution to make a total of 552 atoms.

In the second least-squares cycle, 2209 parameters were adjusted in 11 passes and R decreased to 0·148.

An eighth ΔF map was calculated, and from it 4 atoms were added to the molecule and one was removed. Three more water molecules were added to make a total of 558 atoms.

At this point R was plotted as a function of $\sin\theta/\lambda$ where the reflections in the range $0{\cdot}017 < \sin\theta/\lambda < 0{\cdot}10$ were subdivided in groups of 0·01 (Fig. 2). The dramatic increase in R for reflections with $\sin\theta/\lambda < 0{\cdot}05$ is probably primarily due to neglect of the solvent continuum in the model. In order to remove the effect of the 22 reflections with $0{\cdot}03 < \sin\theta/\lambda < 0{\cdot}05$, they were weighted zero in the subsequent refinement. Although there is still some evidence for possible solvent effects for reflections with $\sin\theta/\lambda$ somewhat greater than 0·05, it is clear that giving zero weight to reflections with $\sin\theta/\lambda < 0{\cdot}05$ removes the effects of those most seriously in error.

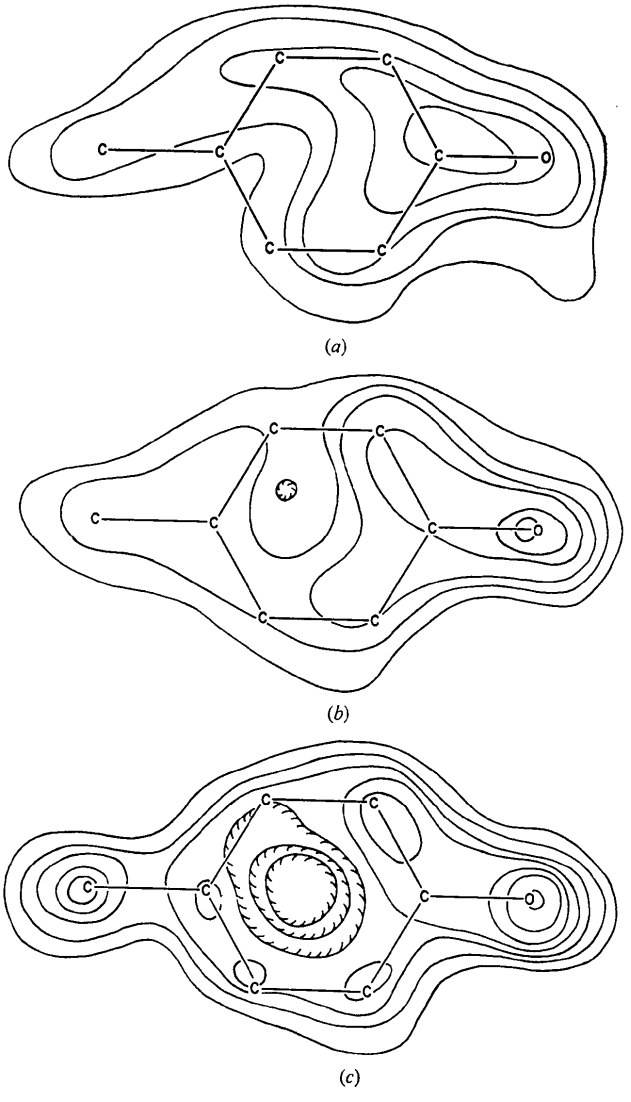


Fig. 3. Section from F_o map in plane of Tyr 11. (*a*) 2 Å data, experimental phases; (*b*) 2 Å data, F_c phases after fourth least-squares refinement cycle; (*c*) 1·5 Å data, F_c phases after fourth least-squares cycle. The 28 reflections with $\sin\theta/\lambda <$ 0·05 were omitted from all three maps because the solvent continuum had not been included in F_c.

In the third least-squares cycle, 2233 parameters were adjusted with 5005 reflections of nonzero weight. The value of R decreased to 0·132 with about 0·002 of the decrease attributable to assigning zero weight to the 22 reflections noted above

The relatively small decrease in R in the third least-squares cycle suggests that the refinement is nearing convergence and no additional large decreases can be expected. One further substantial improvement in the model can be made, however, by including the hydrogen atoms without refining their positions. The positions of 256 hydrogen atoms (of the 385 in the molecule) were calculated from the adjacent nonhydrogen atom positions. (The positions of most of the rest could be calculated by assuming a staggered conformation, and the positions of those involved in hydrogen bonding could be approximated by assuming linear bonds.)

That the effects of hydrogen atoms are by no means negligible can be seen by noting that f_H is approximately $f_C/6$ over most of the range covered by the present data for rubredoxin. Indeed, only for $0{\cdot}23 < \sin\theta/\lambda$ does f_H fall off relatively more rapidly than f_C. Without hydrogen atoms in the model, it is to be expected that the nonhydrogen atoms will tend to be displaced toward the hydrogen positions, and it is important to remove this systematic effect.

In the fourth least-squares cycle, the 256 hydrogen atoms whose positions were determined by adjacent nonhydrogen atoms were included in the model at their calculated positions. *Addition* of the hydrogen atoms *increased* R from 0·132 to 0·141 indicating the substantial effect of hydrogen. *Refining* the nonhydrogen atoms with the 256 hydrogen atoms included in the model *decreased* R in the fourth least-squares cycle to 0·126.

Because the space group for rubredoxin is polar, neglect of the effects of anomalous scattering in the refinement leads to errors in the phases. Although there are effects from the lighter atoms, anomalous scattering from the iron and sulfur atoms is relatively more important and was taken into account after the fourth least-squares refinement cycle by a partial cycle involving only these atoms. R did not change, but the z coordinate for the iron atom shifted by almost 11σ and shifts for the sulfur atoms ranged from $1{\cdot}7$–$2{\cdot}5\sigma$. The average shift in x and y was $0{\cdot}4\sigma$.

Observed and calculated structure factors are listed in Table 2. The parameters in Table 3 are from the fourth least-squares refinement cycle except those for the iron and four cysteine sulfur atoms which are from the partial cycle including anomalous scattering for these atoms.

It should be emphasized that the refinement has not completely converged (see *Discussion*) and some R groups in Table 3 are not yet correct. The chemical sequence indicates the following changes (McCarthy,

1972): R8, Ile → Val; R21, Glu → Asp; R24, Ile → Val; R28; Val → Thr; R41, Ile → Leu; R44, Ile → Val. Ile 8 had been correctly identified as Val from the 2 Å resolution electron density map; the additional atom making it Ile had been added in the fourth least-squares cycle for testing purposes, and the B value suggests that it should be removed. The refinement had indicated difficulty in each of the other groups (as well as some additional ones) and had led to the correct assignment for most of them.

The parameters in Table 3 should only be used with the limitation of incomplete convergence clearly in mind, and for certain uses they should be idealized by imposing constraints. The parameters appear without idealization, however, so that they are representative of the refinement to this point.

Discussion

At the outset of this work, it was questionable whether a protein could be refined with the limited data set that can be observed. For rubredoxin the number of reflections greater than 2σ was only a little more than twice the number of parameters adjusted by least-squares, and in the present data set there was none with $d < 1{\cdot}54$ Å. The number of parameters could have been decreased by use of an idealized model refined by constraining some parameters or by a least-squares group-refinement procedure, but it would have been impossible to derive standard deviations for individual atomic parameters. Accordingly, we attempted to refine the model by techniques applicable to small structures.

Although as noted above the refinement of rubredoxin had not completely converged by the fourth least-squares cycle, it was terminated because it had reached a point of diminishing returns, and a number of questions concerned with refinement of a protein structure had already been answered. First, it is possible to refine rubredoxin in the usual crystallographic sense; second, neglect of the solution continuum in the model appears to be serious for rubredoxin only for reflections with $\sin\theta/\lambda < 0{\cdot}05$; third, the effects of including hydrogen atoms in the model are detectable; fourth, difficulty was experienced with some atoms, primarily in R groups on the surface of the molecule and towards the ends of the chain where B values are very large; fifth, an objective estimate of the standard deviations in the parameters has been derived; sixth, important features of the structure of rubredoxin not evident or uncertain in the 2 Å resolution map based on experimental phases have emerged.

The decrease in R for successive cycles shown in Table 1 indicates that refinement is possible. This suggests that the phases calculated from the model (except for reflections at low $\sin\theta/\lambda$ where the effects of the solution are large) are better than experimental phases. We have used the $hk0$ reflections to monitor the changes in phase that have occurred during refinement. Although these are zonal reflections, their phases are general and independent of any drift of the molecule along z from one refinement cycle to the next. In Table 4 the average differences between the following phase sets for the $hk0$ reflections are given: 1, most probable phases; 2, 'best' phases; 3, phases calculated with coordinates from 2 Å resolution electron density map; 4, phases after four refinement cycles by ΔF syntheses; 5, phases after four ΔF and four least-squares refinement cycles. For the 2 Å data the average change in phase by the refinement by ΔF synthesis was 27·7° and by four least-squares cycles it was 19·3°. The average change of phase over the eight cycles was 33·6°. The average changes from experimental phases (most probable or 'best') to those after refinement are in the range 35°–40° with the difference from most probable phases consistently less than the differences from 'best' phases. The results suggest that errors in the experimental phases are ~40°–45°, a little greater than the 36° indicated by the figure of merit, $m = 0{\cdot}81$. It should be noted that the data also suggest that the calculated phases from the model derived from the 2 Å electron density map are better than experimental phases. For

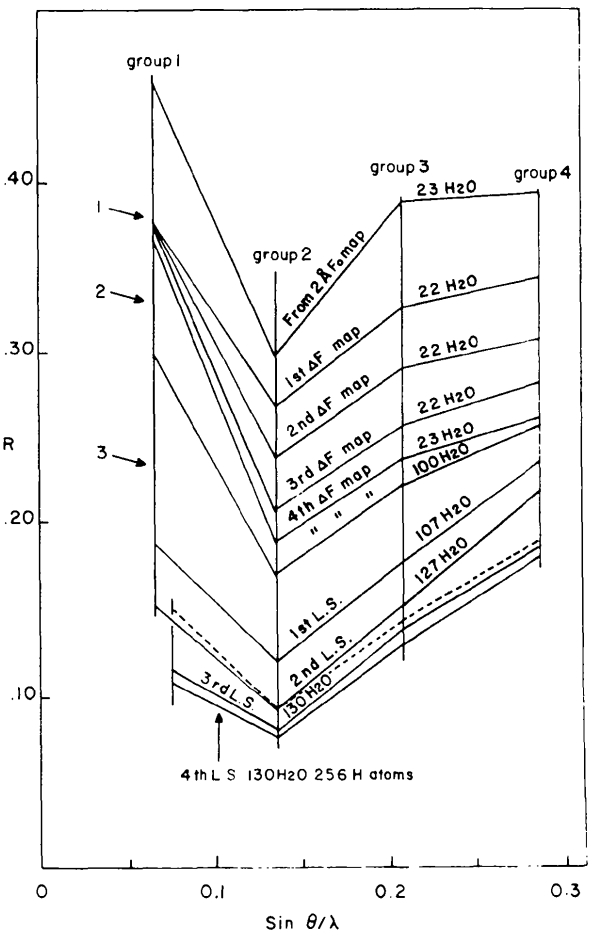


Fig. 4. R *vs.* $\sin\theta/\lambda$ for the refinement.

comparative purposes differences between phase sets 3–5 are also given for the 1·5 Å data.

Table 4. *Comparison of various phase sets for hk*0 *reflections:* 1, *most probable phases;* 2, *'best' phases;* 3, *calculated phases, initial coordinates;* 4, *calculated phases after* 4ΔF *refinement cycles;* 5, *calculated phases after* 4ΔF *and* 4 *least-squares refinement cycles*

(*a*) 2 Å data, 147 phases.

Phase sets	Difference
1–2	10·5°
1–3	41·8
1–4	37·7
1–5	35·2
2–3	43·8
2–4	41·6
2–5	39·2
3–4	27·7
3–5	33·6
4–5	19·3

(*b*) 1·5 Å data, 213 phases.

3–4	29·4°
3–5	35·2
4–5	20·3

Fig. 3 shows the effect of improved phases for Tyr 11, the poorest of the six ring R groups in the 2 Å electron density map. In Fig. 3(*a*) the electron density in a plane through the ring is shown at 2 Å resolution with experimental phases, and in Fig. 3(*b*) the same with phases from the refinement. In Fig. 3(*c*) the electron density is greatly improved with the 1·5 Å data and phases from the refinement. In these maps the 28 reflections with $\sin\theta/\lambda < 0{\cdot}05$ were not included because the solution continuum was omitted in the model.

Certain features of the refinement are most clearly shown by a plot of the residual, R, for each group of reflections in Table 1 *vs.* $\sin\theta/\lambda$ (Fig. 4). For groups 2–4, R increases with increasing $\sin\theta/\lambda$ as is normally found. Contrary to expectation, however, R for group 1 is greater than that for group 2; in fact, in the initial stages of the refinement it is greater than the R values for any of the other groups. Although this may be attributable in part to secondary extinction, the extremely rapid decrease of the effect with increasing $\sin\theta/\lambda$ as shown in Fig. 2 suggests that neglect of the solvent continuum in the model is a more likely explanation. Indeed, three features of the plot in Fig. 4 support this view. (1) In the first three ΔF refinement cycles when there was little change in the model for the solution and in the fourth cycle before most of the water oxygen atoms were added, there was little decrease in R for group 1 although the other three groups were refining satisfactorily (see 1 in Fig. 4). (2) When most of the water was added after the third ΔF cycle and refined in the fourth, there was a disproportionately large decrease in R for group 1 as compared with the other three groups (see 2 in Fig. 4). (3) In the first least-squares cycle when B values for individual atoms were allowed to vary, there was again a disproportionately large decrease in R for group 1 (see 3 in Fig. 4). This can be accounted for by the great variability in B values for the water oxygen atoms. Thus improving the model for the solution has produced the greatest decrease in R for group 1 reflections where the effects of neglect of solvent are largest.

For group 2 reflections, $R = 0{\cdot}078$ when the refinement was terminated. Although approximately $\frac{3}{4}$ of the electrons associated with the solution are unaccounted for in the model, the low R value for this group indicates little effect from the solution even for planes with spacings as great as 5 Å, Fig. 2.

On including the 256 hydrogen atoms after the third least-squares cycle, R for each group increased (broken line in Fig. 4). This can be explained by noting that neglect of hydrogen atoms in the model causes errors both in the positions and B values of the nonhydrogen atoms. Including the hydrogen atoms impairs the model as shown by the increased R, but subsequent refinement of the nonhydrogen atoms shifts them to more nearly approximate their true positions and R decreases.

Although the refinement progressed steadily as indicated by the decrease in R from one cycle to the next, some atoms refined to positions giving unreasonable bond lengths or angles. These were primarily atoms with high B values. Such values could represent excessive thermal motion, but they may also cover disorder, both static and dynamic, or indicate either that the model is wrong for the atoms in question or convergence is incomplete. It is likely, therefore, that both the B values and the standard deviations in the positions of the atom involved are not realistic.

The lowest standard deviations (for a given atom type) are for atoms with the lowest B values. These are primarily main-chain atoms or atoms in side chains tightly packed together within the molecule. Atoms toward the ends of the chain and those in R groups on the outside of the molecule tend to have high B values.

In Fig. 5 the average B values for N_n, C_n^z and C_n, *i.e.* the main chain atoms for the nth residue, have been plotted against n. It will be seen that B values near the chain termini are much larger than in any other region of the molecule. From residues 4–50, the average B's range from 4–13 Å^2 except at residues 7–8, 23, and in the region 40–45. Residues 7–8 and 40–41 would be expected to have large B's because they project from the surface of the molecule. In fact, the whole region from residues 40–45 is not tightly bound to the rest of the molecule except through the Fe–S complex, and even these heavy atoms have B's greater than many main chain C and N atoms. The dramatic difference of the average B for residue 23 from surrounding values probably stems primarily from error in the atomic positions (Rossmann *et al.*, 1959). Indeed, error in atomic positions may explain in part the large B values near the chain terminii.

A good example of the variation of B with atomic position is provided by Lys 3 which is fully extended, projecting from the molecular surface straight into the solution. B values and standard deviations for $C^{\beta}(3)$ through $N^{\zeta}(3)$ are given in Table 5(*a*). It can be seen that B values increase with increasing distance from the main chain and B for $N^{\zeta}(3)$ is unrealistically large. The difficulty with this atom is readily apparent by inspecting the bond lengths in Fig. 6(*a*). Although the C–C bond lengths are consistent with the standard deviations of the atoms involved, the $C^{\varepsilon}(3)$–$N^{\zeta}(3)$ bond is almost one Å longer than the accepted value. This can be explained by a shift of $N^{\zeta}(3)$ in the model out toward hydrogen-bonded oxygen atoms of the water with an accompanying increase in the B value of the N atom to cover the several atoms in the vicinity of the terminus of Lys 3 and possible disorder of $N^{\zeta}(3)$. (The same phenomon was also observed for a number of carboxyl oxygen atoms in some of the Asp and Glu groups on the surface of the molecule.) In contrast to N^{ζ} of Lys 3, N^{ζ} of Lys 46 is firmly bound by H bonds to O(30) and O(33). For comparative purposes, B values and standard deviations for Lys 46 are listed in Table 5(*b*) and bond lengths are given in Fig. 6(*b*).

Table 5. *B Values and σ position for atoms in* (*a*) *Lys* 3, (*b*) *Lys* 46

	(*a*)			(*b*)	
	B	σ position		B	σ position
$C^{\beta}(3)$	14 Å²;	0·12	$C^{\beta}(46)$	9 Å²;	0·09
$C^{\gamma}(3)$	15	0·14	$C^{\gamma}(46)$	8	0·07
$C^{\delta}(3)$	18	0·15	$C^{\delta}(46)$	8	0·08
$C^{\varepsilon}(3)$	31	0·18	$C^{\varepsilon}(46)$	7	0·08
$N\zeta(3)$	58	0·30	$N\zeta(46)$	5	0·06

The estimated standard deviations from the inverse matrix can be checked by comparison with values derived from measurements of the length of bonds of a given type. The C^{α}–C^{β} bonds were selected for this purpose. Of the 48 in the molecule, those in the N-terminal methionine and in the two C-terminal amino acids were excluded because the model is unsatisfactory in these regions. The remaining 45 have a range of 1·13 Å, a mean value of 1·57 Å and a σ of 0·20 Å in essential agreement with the r.m.s. σ of 0·19 Å derived from the inverse matrices of the last refinement cycle.

For model building it has been assumed that the peptide groups are planar (Pauling, Corey & Branson, 1951). Recent evidence suggests, however, that some peptide groups may be significantly nonplanar (Ramachandran, 1968; Winkler & Dunitz, 1971). We have calculated least-squares planes for the peptide groups in rubredoxin and tabulated the deviations of the atoms from each. Again neglecting the N-terminal methionine and the two C-terminal residues, we find that 4 of the remaining 51 peptides have χ^2 distributions greater than 9·2, *i.e.* they are nonplanar at the 0·99 confidence level.

The Fe–S complex is an important feature of the structure and the refinement has provided bond lengths and angles with greatly improved precision. Bond lengths and angles for the Fe–S complex from the last three refinement cycles and the partial cycle including anomalous scattering are listed in Table 6.

All four Fe–S bonds are chemically equivalent in the sense that all four atoms coordinated to Fe are cysteine S. Three of the Fe–S bonds do not differ significantly in

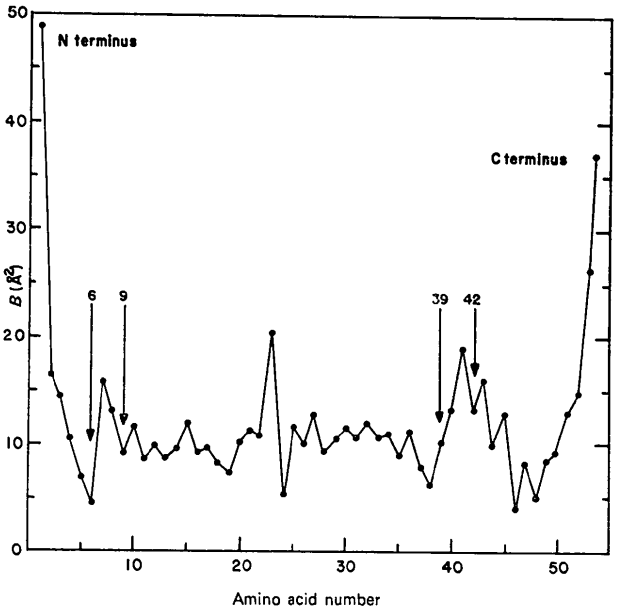


Fig. 5. Plot of $\bar{B}$ *vs.* amino acid number.

Table 6. *Bond lengths and angles for the* Fe–S *complex for the last three least-squares refinement cycles and for the partial cycle including anomalous scattering*

Bond or angle	From 2nd l.s.	From 3rd l.s.	From 4th l.s.	Incl. anom. scat.
Fe–$S^{\gamma}(6)$	2·39 (3) Å	2·40 (3) Å	2·40 (3) Å	2·34 (2) Å
Fe–$S^{\gamma}(9)$	2·33 (4)	2·33 (4)	2·32 (4)	2·32 (3)
Fe–$S^{\gamma}(39)$	2·31 (4)	2·27 (4)	2·27 (4)	2·24 (3)
Fe–$S^{\gamma}(42)$	1·97 (4)	1·98 (4)	1·98 (3)	2·05 (3)
$S^{\gamma}(6)$—Fe–$S^{\gamma}(9)$	112 (2)°	111 (1)°	111 (1)°	113 (1)°
$S^{\gamma}(6)$—Fe–$S^{\gamma}(39)$	105 (1)	106 (1)	105 (1)	108 (1)
$S^{\gamma}(6)$—Fe–$S^{\gamma}(42)$	101 (1)	102 (1)	102 (1)	101 (1)
$S^{\gamma}(9)$—Fe–$S^{\gamma}(39)$	102 (1)	103 (1)	104 (1)	104 (1)
$S^{\gamma}(9)$—Fe–$S^{\gamma}(42)$	118 (2)	117 (2)	117 (1)	115 (1)
$S^{\gamma}(39)$–Fe–$S^{\gamma}(42)$	118 (2)	117 (2)	117 (2)	115 (1)

length from an average value of 2·27–2·28 Å for 70 such bonds reported in the literature, but the fourth one is shorter, and two of the four are significantly less than the 2·339–2·380 Å lengths in the tetrahedral FeS_4 core of bis(imidotetramethyldithiodiphosphino-*S,S*)-iron-(II) (Churchill & Wormald, 1971). The S–Fe–S angles in the latter compound range from 100·5–114·9°, a range similar to that in rubredoxin where five of the six angles appear to differ from the tetrahedral value.

Refinement by constraints

The extent to which refinement of any protein structure is warranted will depend on the quality and extent of the X-ray data. For most proteins it will be necessary with presently available data to restrict the number of parameters by use of an idealized model in which parameters for structural features such as peptide residues, ring R groups, and bond lengths and angles are fixed at accepted values. Diamond (1966, 1971) has developed an excellent pair of programs, one to fit a model to a set of coordinates, the other to refine the model by fitting an electron density map. In both programs, various torsion and bond angles may be constrained or varied to give the best fit. In general, if parameters are known with better precision than they can be determined with the available data, they should be constrained in the refinement.

Improved model

The model for rubredoxin can be improved in several ways. Correcting it as indicated by the chemical sequence would aid convergence. Including more of the off-diagonal terms in the least-squares matrix by use of a computer with mass storage would also improve the convergence. In any one refinement cycle in the present work only about 10% of the off-diagonal terms were included.

It is now known that considerable data can be observed beyond the present 1·54 Å limit. A survey of reciprocal space in the range $0{\cdot}325 < \sin\theta/\lambda < 0{\cdot}40$ suggests that as much as 40% of the data can be observed with intensities greater than 2σ. The model can be expected to improve substantially by including additional data, but it should be noted that little improvement can be expected for atoms with high B values, because such atoms contribute little to the high-angle data. It may still be necessary to fit regions of the molecule with high B values by an idealized model unless the B's can be reduced experimentally *e.g.* by collecting data at low temperatures.

Two completely different refinement methods may provide an improved model. One of these involves improving the phases of the X-ray reflections by a phase refinement technique (Sayre, 1972). The other involves minimizing the energy of interaction between various main chain and side chain components of the protein molecule (Scheraga, 1968; Levitt & Lifson, 1969; Hermans, 1972).

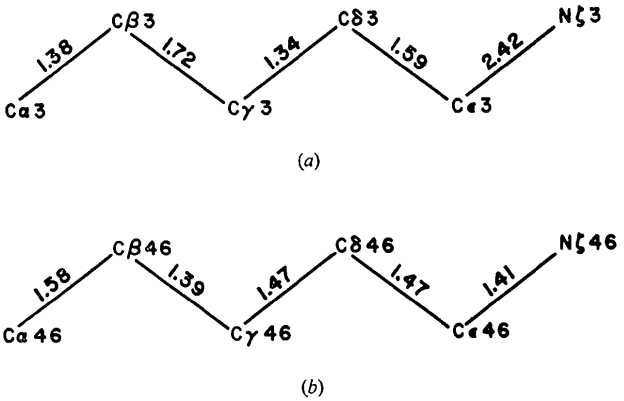


Fig. 6. Bond lengths (*a*) Lys 3, (*b*) Lys 46.

This work was supported under USPHS Grant GM-13366 from the National Institutes of Health, and a grant of computer time, Computer Center, University of Washington. Programs used were those in the X-ray System (Stewart, Kundall & Baldwin, 1970) and ORTEP (Johnson, 1965). An F_o synthesis based on the phases from the 4th least-squares refinement cycle was calculated for us by Dr D. Sayre.

References

IUPAC–IUB Commission on Biochemical Nomenclature (1971). *Biochem. J.* **121**, 577–585.
Blow, D. M. & Rossmann, M. G. (1961). *Acta Cryst.* **14**, 1195–1202.
Brändén, C. -I., Holmes, K. C. & Kendrew, J. C. (1963). *Acta Cryst.* **16**, A175.
Churchill, M. R. & Wormald, J. (1971). *Inorg. Chem.* **10**, 1778–1782.
Costain, W. (1941). Ph. D. Thesis, Univ. of Birmingham.
Cruickshank, D. W. J. (1950). *Acta Cryst.* **3**, 10–13.
Diamond, R. (1966). *Acta Cryst.* **21**, 253–266.
Diamond, R. 1971). *Acta Cryst.* A**27**, 436–452.
Hermans, J. (1972). Personal communication.
Herriott, J. R., Sieker, L. C., Jensen, L. H. & Lovenberg, W. (1970). *J. Mol. Biol.* **50**, 391–406.
Herriott, J. R., Watenpaugh, K. D., Sieker, L. C. & Jensen, L. H. (1973). To be published.
James, R. W. (1948). *Acta Cryst.* **1**, 132–134.
Jensen, L. H. & Lingafelter, E. C. (1961). *Acta Cryst.* **14**, 507–520.
Johnson, K. C. (1965). Oak Ridge National Laboratory Report ORNL-3794, revised.
Levitt, M. & Lifson, S. (1969). *J. Mol. Biol.* **46**, 269–279
McCarthy, K. (1972). Ph. D. Thesis, George Washington Univ.
North, A. C. T, Phillips, D. C. & Mathews, F. S. (1968). *Acta Cryst.* A**24**, 351–359.
Pauling, L., Corey, R. B. & Branson, H. R. (1951). *Proc. Natl. Acad. Sci. U. S.* **37**, 205–211.
Ramachandran, G. N. (1968). *Biopolymers*, **6**, 1494–1496.

ROSSMANN, M. G., JACOBSON, R. A., HIRSHFELD, F. L. & LIPSCOMB, W. N. (1959). *Acta Cryst.* **12**, 530–535.
SAYRE, D. (1972). *Acta Cryst.* A**28**, 210–212.
SCHERAGA, H. A. (1968). *Advanc. Phys. Org. Chem.* **6**, 103–184.
STEWART, J. M., KUNDALL, F. A. & BALDWIN, J. C. (1970). Computer Science Center Report, Univ. of Maryland.
STOUT, G. H. & JENSEN, L. H. (1968). *X-ray Structure Determination: A Practical Guide.* pp. 376–381. New York: Macmillan.
WATENPAUGH, K. D., HERRIOTT, J. R., SIEKER, L. C. & JENSEN, L. H. (1970). *Amer. Cryst. Assn. Abstr.* p. 78, Tulane Univ., New Orleans.
WATENPAUGH, K. D., SIEKER, L. C., HERRIOTT, J. R. & JENSEN, L. H. (1970). *Amer. Cryst. Assn. Abstr.* p. 44, Carleton Univ., Ottawa.
WATENPAUGH, K. D., SIEKER, L. C., HERRIOTT, J. R. & JENSEN, L. H. (1971*a*). *Amer. Cryst. Assn. Abstr.*, p. 68, Univ. of South Carolina, Columbia.
WATENPAUGH, K. D., SIEKER, L. C., HERRIOTT, J. R. & JENSEN, L. H. (1971*b*). *Amer. Cryst. Assn. Abstr.*, p. 52, Iowa State Univ., Ames.
WATENPAUGH, K. D , SIEKER, L. C., HERRIOTT, J. R. & JENSEN, L. H. (1972). In *Cold Spring Harbor Symposia on Quantitative Biology*, XXXVI, pp. 359–367.
WATSON, H. C., KENDREW, J. C., COULTER, C. L., BRÄNDÉN, C. -I., PHILLIPS, D. C. & BLAKE, C. F. (1963). *Acta Cryst.* **16**, A81.
WATSON, H. C. (1969). In *Progress in Stereochemistry*, Vol. 4, pp. 299–333. London: Butterworth.
WINKLER, F. K. & DUNITZ, J. D. (1971). *J. Mol. Biol.* **59**, 169–182.

Acta Cryst. (1973). B**29**, 956

The Crystal Structure of a Tellurium(IV,VI) Oxyhydroxide, $H_2Te_2O_6$

BY OLIVER LINDQVIST

Department of Inorganic Chemistry, Chalmers' University of Technology and the University of Göteborg, P.O. Box, S-402 20 Göteborg 5, Sweden

AND JACQUES MORET

Laboratoire de Chimie Minérale C, Université des Sciences et Techniques du Languedoc, Place Eugène-Bataillon, 34000 Montpellier, France

(*Received* 30 *December* 1972; *accepted* 2 *January* 1973)

Single crystals of $H_2Te_2O_6$ have been prepared by hydrothermal synthesis. The space group is $Pbn2_1$ and the cell dimensions are $a=8{\cdot}037$, $b=12{\cdot}070$, $c=4{\cdot}735$ Å, and $Z=4$. The structure was solved from three-dimensional Patterson and electron density calculations and the structure parameters, excluding those of the hydrogen atoms, were refined to an R value of 0·036 using 2480 independent reflexions. The structure contains $Te(VI)O_6$ octahedra and four-coordinated Te(IV) units with Te(VI)–O and Te(IV)–O bond distances in the ranges 1·864–1·952 Å and 1·861–2·107 Å, respectively. The Te(VI) octahedra are linked through corners to form chains, which are connected *via* –Te(IV)–O–Te(IV)–O– chains to form infinite sheets. These sheets are held together by hydrogen bonds and van der Waals forces only, resulting in cleavage planes in the crystals. The $H_2Te_2O_6$ structure is closely related to that of Te_2O_5, the three-dimensional structure of which can be regarded as a condensation of $H_2Te_2O_6$ layers.

Introduction

Phases of composition Te_2O_5 have been known for some years to exist as microcrystalline powders (Rosicky, Loub & Pavel, 1965; Moret & Maurin, 1968). Recently, single crystals of Te_2O_5 have been prepared and the crystal structure determined by X-ray methods (Lindqvist & Moret, 1973). The Te_2O_5 single crystals were obtained in a hydrothermal investigation of the TeO_2–TeO_3–H_2O system. In the same series of experiments it was also possible to isolate a new compound of composition $H_2Te_2O_6$ and to determine the conditions required for the growth of single crystals. Formally, $H_2Te_2O_6$ ought to be obtained by the reaction of Te_2O_5 with water, and the present investigation was undertaken to examine the structural relationship between Te_2O_5 and $H_2Te_2O_6$.

Experimental

The hydrothermal synthesis of crystals of $H_2Te_2O_6$ and the analysis of the compound have been described elsewhere (Moret, 1972). Cell dimensions, as obtained from measurements with a calibrated Siemens powder diffractometer, and possible space groups have been given previously (Moret & Lindqvist, 1972) and a summary of the crystallographic data is given in Table 1.