Albion College ILL

ILLiad TN: 35633

Borrower: WLD

Lending String: *EXA,EZC,EYE,GZP,WRF

Patron:

Journal Title: Progress in stereochemistry.

Volume: 4 **Issue:**
Month/Year: 1969 **Pages:** 299-333

Article Author: Klyne, William; De la Mare, Peter Bernard David; Aylett, B. J.; Harris, Margaret M H.C. Watson

Article Title: unknown

Imprint: London : Butterworths Scientific Publications ; [New York] : [Academic Press], 1954-

ILL Number: 169962276

Call #: QD481 .P85 v.4

Location: Stockwell 110 AVAILABLE

We send our documents electronically using Odyssey. ILLiad libraries can enable Odyssey for receiving documents seamlessly. Non-ILLiad libraries can download Odyssey stand alone for FREE. Check it out at: http://www.atlas-sys.com/Odyssey.html

Charge
Maxcost: 0.00IFM

Shipping Address:
ILL
Walter Schroeder Library Milwaukee School of Engineering
1025 N Broadway
Milwaukee, Wisconsin 53202-3109
United States

Fax: 414-277-7186
Ariel:
Email: interlibraryloan@msoe.edu

7

THE STEREOCHEMISTRY OF THE PROTEIN MYOGLOBIN

H. C. Watson

PAGE
1 Introduction . . . 299
2 Basic Chemical and Crystal Data . . . 300
3 The Structure Determination . . . 301
A. The application of the isomorphous replacement method to a resolution of 2 Å 301
B. Detailed interpretation of the electron-density map . . . 302
C. Recording the structure's atomic co-ordinates . . . 304
D. The Fourier refinement of the structure . . . 304
4 Conformation of the Molecule . . . 305
A. The general nature of the structure . . . 305
B. The helical regions of the main chain . . . 305
C. The non-helical regions of the main chain . . . 309
D. Side-chain conformation . . . 316
5 Forces Stabilizing the Structure . . . 318
A. Specific interactions . . . 318
B. Hydrophobic interactions . . . 319
C. Non-polar bonding . . . 320
APPENDIX: The myoglobin co-ordinate data . . . 321

1 INTRODUCTION

Myoglobin is a relatively small protein whose principal biological function appears to be that of undergoing reversible combination with molecular oxygen. It is present in appreciable amounts in vertebrate muscle tissue from which it can be isolated and crystallized.

Crystals of sperm whale (*Physeter catodon*) myoglobin have proved to be particularly suitable for x-ray diffraction studies (Kendrew and Parrish, 1956). The use of the method of multiple isomorphous replacement, developed in the study of haemoglobin by Perutz and his co-workers (Green *et al.*, 1954), led first to the structure of myoglobin as seen at a resolution of 6 Å (Bodo *et al.*, 1959), then to one of 2 Å (Kendrew *et al.*, 1960). More recent work has extended the resolution to 1·4 Å and, with the help of the chemical sequence data (Edmundson, 1965), the co-ordinates of almost all the non-hydrogen atoms in the molecule have been determined (see Frontispiece).

In this Chapter, the structure of the myoglobin molecule will be described both in a qualitative and quantitative manner. Emphasis will be placed, where possible, on the general stereochemical principles which appear to govern the configuration of the folded polypeptide

chain. A brief factual description of the structure determination is also included, so that the reader can evaluate for himself the conclusions drawn from the co-ordinate data given in the Appendix.

2 BASIC CHEMICAL AND CRYSTAL DATA

Sperm whale myoglobin is made up from 153 amino-acids linked to form one peptide chain and a single prosthetic or haem group, protoporphyrin IX. Of the 20 commonly occurring amino-acids, only cysteine is not represented, as shown in TABLE 1. From the chemical

TABLE 1. THE AMINO-ACID COMPOSITION OF SPERM WHALE MYOGLOBIN

Amino-acid	*Number of residues*
Glycine	11
Alanine	17
Valine	8
Leucine	18
Isoleucine	9
Serine	6
Threonine	5
Methionine	2
Proline	4
Phenylalanine	6
Tyrosine	3
Aspartic acid	6
Asparagine	2
Glutamic acid	14
Glutamine	5
Histidine	12
Lysine	19
Arginine	4
Tryptophan	2
Total	153

composition and sequence data (Edmundson, 1965) the protein's molecular weight has been found to be 17,816.

The two physiologically important forms of the protein are those in which the molecule has reacted, or can react, with molecular oxygen. In both the oxygenated and deoxygenated forms, the haem iron is in the ferrous state. When the protein is removed from its normal biological environment and exposed to the atmosphere, it oxidizes to the ferric form. In this state the ligand attached to the haem iron is a water molecule. Except for the change in the haem ligand, no significant

structural differences have been observed between these three forms of the protein (Nobbs *et al.*, 1966; Watson and Nobbs, 1968).

Crystals of the stable ferric form of sperm whale myoglobin (metmyoglobin) grown from solutions containing approximately 3 M ammonium sulphate and a small amount of strong phosphate buffer (pH 6·8–7·2) are monoclinic with $a = 64{\cdot}5$ Å, $b = 30{\cdot}9$ Å, $c = 34{\cdot}7$ Å, $\beta = 106°$, space group $P2_1$. The unit cell's two asymmetric units each contain one myoglobin molecule together with bound ions and water molecules and mother liquor. The two protein molecules alone account for only $\frac{2}{3}$ of the unit cell volume.

3 THE STRUCTURE DETERMINATION

A. *The application of the isomorphous replacement method to a resolution of* 2 Å

Three-dimensional x-ray diffraction data were obtained from 22 sets of precession photographs of the native, and of 4 heavy-atom derivatives of myoglobin (see Kendrew *et al.*, 1960). For each derivative some 50,000 intensities were measured and reduced after scaling to a unique set of 9,600 structure amplitudes. No attempt was made to utilize the effects of anomalous dispersion.

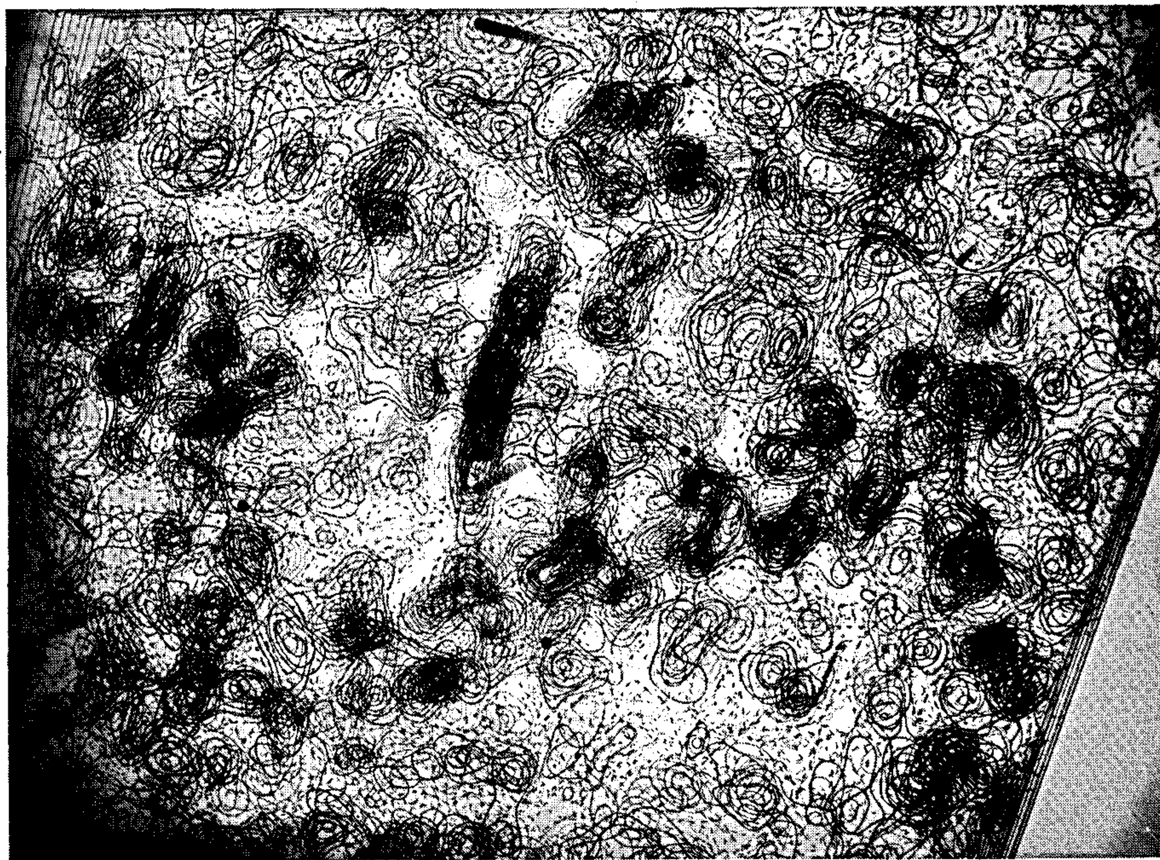

FIGURE 7.1. Section of the myoglobin 2 Å resolution electron-density map showing the haem group edge-on and part of the helix running from right to left below it

Reproduced by kind permission of Dr. J. C. Kendrew, F.R.S.

The positions of the added heavy-atom groups were first determined by difference Patterson procedures and the position, occupancy and isotropic temperature parameters refined by least-squares methods (Dickerson *et al.*, 1961). The phase angles used in the Fourier synthesis were calculated by the centroid method of Blow and Crick (1959). The electron-density map was calculated at intervals of $x/96$, $y/48$, $z/48$ and the results presented, first, in the form of equal electron-density contours on pieces of perspex sheet. FIGURE 7.1 shows a stack of several adjacent sections cut through the unit cell in a direction perpendicular to the *b* axis. In these sheets it is possible to recognize predominant stereochemical features of the molecule such as the planar haem group with its central electron-dense region representing the ferric iron. The regions of the peptide chain which fold in helical arrays (Pauling and Corey, 1951) appear as hollow cylinders of high density. Since covalently bonded atoms are not, in general, represented by separate peaks, at 2 Å resolution the detailed structure had to be determined from the known shapes of the various chemical groupings (see Kendrew *et al.*, 1962).

B. *Detailed interpretation of the electron-density map*

It has recently been demonstrated that a 2 Å resolution electron-density map of a protein molecule can be interpreted if the sequence of amino-acids is known (Blake *et al.*, 1965). When the myoglobin electron-density map was first calculated, the chemical sequence had not been determined and the task of identifying the various electron-density features presented a formidable problem. The difficulty was partly overcome by representing the electron-density levels by coloured clips placed at appropriate points on a matrix of steel rods (FIGURE 7.2). Skeletal models of the correct scale (5 cm = 1 Å) were then placed into the density, as represented by the clouds of coloured clips, thus allowing the various chemical groups to be compared in shape and size with the region under consideration.

By first fitting the peptide groups to the helical main-chain density and by identifying the more obvious amino-acid side-chain groups (FIGURE 7.3) such as histidine, phenylalanine, tyrosine and tryptophan, the course of the complete polypeptide chain backbone could be defined.

In general, the absolute identification of the side-chain groups proved less easy. The longer side chains, for example, tend to protrude from the surface of the molecule and to take up ill-defined configurations. This effect led to erroneous side-chain identifications, although the atomic positions actually observed were invariably correct. A further complication arose with the side chains of valine and threonine, glutamic acid and glutamine, and with aspartic acid, asparagine and leucine which are so closely related in a structural sense that it is impossible to distinguish between these respective groups at 2 Å resolution,

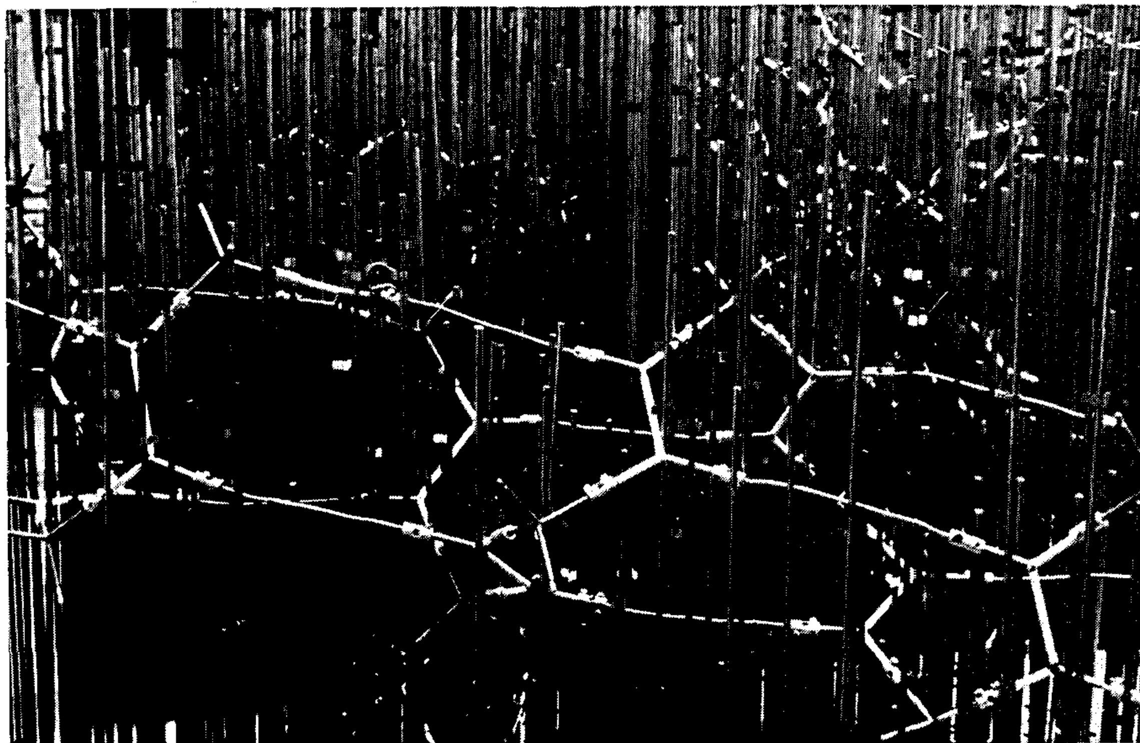

FIGURE 7.2. Model of part of the myoglobin molecule near the haem group; the vertical rods support coloured clips indicating height of electron density at each grid point

Reproduced by kind permission of Dr. J. C. Kendrew, F.R.S.

FIGURE 7.3. Small section of the 'forest' model, showing side chain of the amino acid leucine built into the electron density as defined by the cloud of coloured clips

Reproduced by kind permission of Dr. J. C. Kendrew, F.R.S.

unless special hydrogen-bonding properties are involved which rule out all but one of the alternatives. The subsequent amino-acid sequence determination by Edmundson (1965) has now allowed such anomalies to be resolved and has provided strong independent evidence concerning the correctness of the overall structure.

C. *Recording the structure's atomic co-ordinates*

The positions of the model components were altered in a systematic manner until the best overall compromise was obtained between the structure as indicated by the electron density and that built from the correct linkage of the various components. The co-ordinates of the individual atoms were then recorded from the complete model to the nearest $\frac{1}{4}$ grid interval (1 grid interval is approximately $\frac{2}{3}$ Å).

Subsequent analysis of the co-ordinate set showed that the known bond distances had standard deviations of 0·15 Å (see TABLE 2) indicating a mean error in recording the atomic co-ordinates of approximately

TABLE 2. ANALYSIS OF THE RECORDED ATOMIC CO-ORDINATE DATA IN TERMS OF KNOWN BOND DISTANCES OF POLYPEPTIDE CHAIN (150 OBSERVATIONS OF EACH TYPE OF BOND)

Bond type	*Mean distance,* Å	*Standard deviation,* Å
N–C	1·50	0·16
Cα–C	1·55	0·14
C=N	1·38	0·15
C=O	1·28	0·14
Cα–Cβ	1·55	0·15

0·1 Å. It should be emphasized, however, that the co-ordinates also contain errors associated with the positioning of the standard-model components within the density, with diffraction errors which affect the position of the electron density peaks and with errors due to deviations, either locally or in general, of interatomic distances and angles from those used in the standard models.

D. *The Fourier refinement of the structure*

Several attempts have been made to improve the myoglobin structure determination using conventional Fourier refinement techniques (for details of the method, see Lipson and Cochran, 1966). This has involved calculating a series of Fourier syntheses at the maximum experimentally permissible resolution (1·4 Å) using phase information calculated from the positions of the atoms of the structure which are known (Watson *et al.*, 1964). Although these procedures have led to the assignment of

co-ordinates for many 'new' atoms, they have not produced any significant improvement in the positioning of the 'known' atoms. The co-ordinates listed in the Appendix originate, therefore, essentially from those derived from the experimentally phased 2 Å electron-density maps as described in the previous Section.

4 CONFORMATION OF THE MOLECULE

A. *The general nature of the structure*

The overall shape of the myoglobin molecule is that of a flattened triangular prism which approaches an oblate ellipsoid of dimensions $a = 22$ Å, $b = 12{\cdot}5$ Å.

The polypeptide chain has all its peptides in the *trans* conformation and is arranged in 8 major (containing two or more regular hydrogen bonds) α-helical segments. The latter are joined together by two sharp corners (containing no non-helical residues) and 5 segments composed of from 1 to 8 residues. In addition, there are two non-helical end segments of 2 and 4 residues*. The whole is folded in a complex and unsymmetrical manner to produce an extremely compact structure in which the haem group lies buried except for one of its edges (that containing the propionic acid groups) which forms part of the molecular surface.

B. *The helical regions of the main chain*

The 8 major helical sections are all right-handed. For convenience of describing the structure, the helices have been labelled alphabetically, starting at the amino end of the chain. Residues are defined as being helical if either the amino or carbonyl group is involved in the system of helical hydrogen bonds*. The residues between helices are numbered from the amino end of the segment and identified by the two letters denoting the adjoining helices (see Table 5, p. 307). The number of residues in each helical section is listed in Table 3, together with the direction cosines of the respective helices.

The total average value for the axial translation per residue and the number of residues per helix turn are very close to the values of 1·5 Å and 3·6 predicted by Pauling and Corey for the 'classical' α-helix. When the helices are examined individually, however, they show considerable deviations from these values (cf. Table 4). With the possible exception of the longest helix, (H), they all exhibit some distortion. The A and E helices, for example, are not straight (the E helix has a bend of some 7° at or near residue E10) and the C and G helices have the regular hydrogen-bond pattern interrupted by side chain/main chain interactions (see Table 5). In many cases, particularly when the main

* See note to legend of Table 5.

TABLE 3. NUMBER OF RESIDUES IN EACH HELICAL SECTION AND DIRECTION COSINES l, m AND n OF HELICES AND HAEM NORMAL

Orthogonal co-ordinate system: a, b, c^* where a, b and c = axes of the crystal

Helix	*Number of residues*	l	m	n
A	16	−0·951	−0·302	0·063
B	16	0·241	0·181	0·955
C	7			
D	7	0·078	−0·599	−0·797
E_1	10	0·872	0·075	−0·485
E_2	10	0·843	0·212	−0·494
F	10*	−0·361	0·004	0·933
G	19	−0·490	0·603	−0·629
H	26*	0·339	−0·823	0·457
Haem normal		0·890	−0·360	0·280

* cf. note to TABLE 5 on assignment of helical residues.

TABLE 4. PARAMETERS OF THE HELICES (KENDREW, 1962)

n = number of residues per turn
ϕ = angular rotation per residue, rad
h = axial translation per residue, Å

Helix	n	ϕ	h
A	3·63	1·73	1·50
B	3·72	1·69	1·47
C*	—	—	—
D	3·63	1·73	1·45
E_1	3·61	1·74	1·52
E_2	3·67	1·71	1·49
F	3·70	1·70	1·46
G	3·59	1·75	1·53
H	3·63	1·73	1·49
α-helix	3·61	1·74	1·50

* The C helix is too short and irregular for the parameter refinement to be meaningful.

chain is in contact with the solvent, the carbonyl residues appear to point slightly away from the helix axis. This effect is much more pronounced in six of the eight carbonyl terminal residues (the only real exception is the F helix) where the nitrogen appears to point towards the carbonyl group three residues back along the chain. The resulting hydrogen-bonding arrangement for the last turn of the helix, therefore, more closely resembles that of the $3{\cdot}0_{10}$ helix (Huggins, 1943) than that of the classical α-helix (for examples see FIGURES 7.6 and 7.10).

TABLE 5. THE SPECIFIC NON-COVALENT BONDING ARRANGEMENT

Residues are identified in three ways: (*a*) by the position within the polypeptide chains; (*b*) by the standard abbreviation for the amino-acid name; (*c*) by a convenient structural code in which a single letter refers to one of the helical segments (A–H) and a double letter to an irregular region between helices. Connecting lines and pointing arrows indicate the interacting groups (atoms are identified according to the convention proposed by Edsall *et al.*, 1966). Squared symbols indicate evidence for the interaction with ions (■ definite; □ possible; ⊠ sterically restricted). Bound sulphate ions [1] and [2] are related by the crystal symmetry, as also are [3] and [4]. Enclosed polar groups have not been located in the electron-density distribution.

Note—The assignment of helical residues adopted is that proposed by Kendrew *et al.* (1962). The bonding pattern follows the original table devised by Dr. J. C. Kendrew (unpublished) and indicates that residues FG1, GH6 and HC1 should also be considered helical.

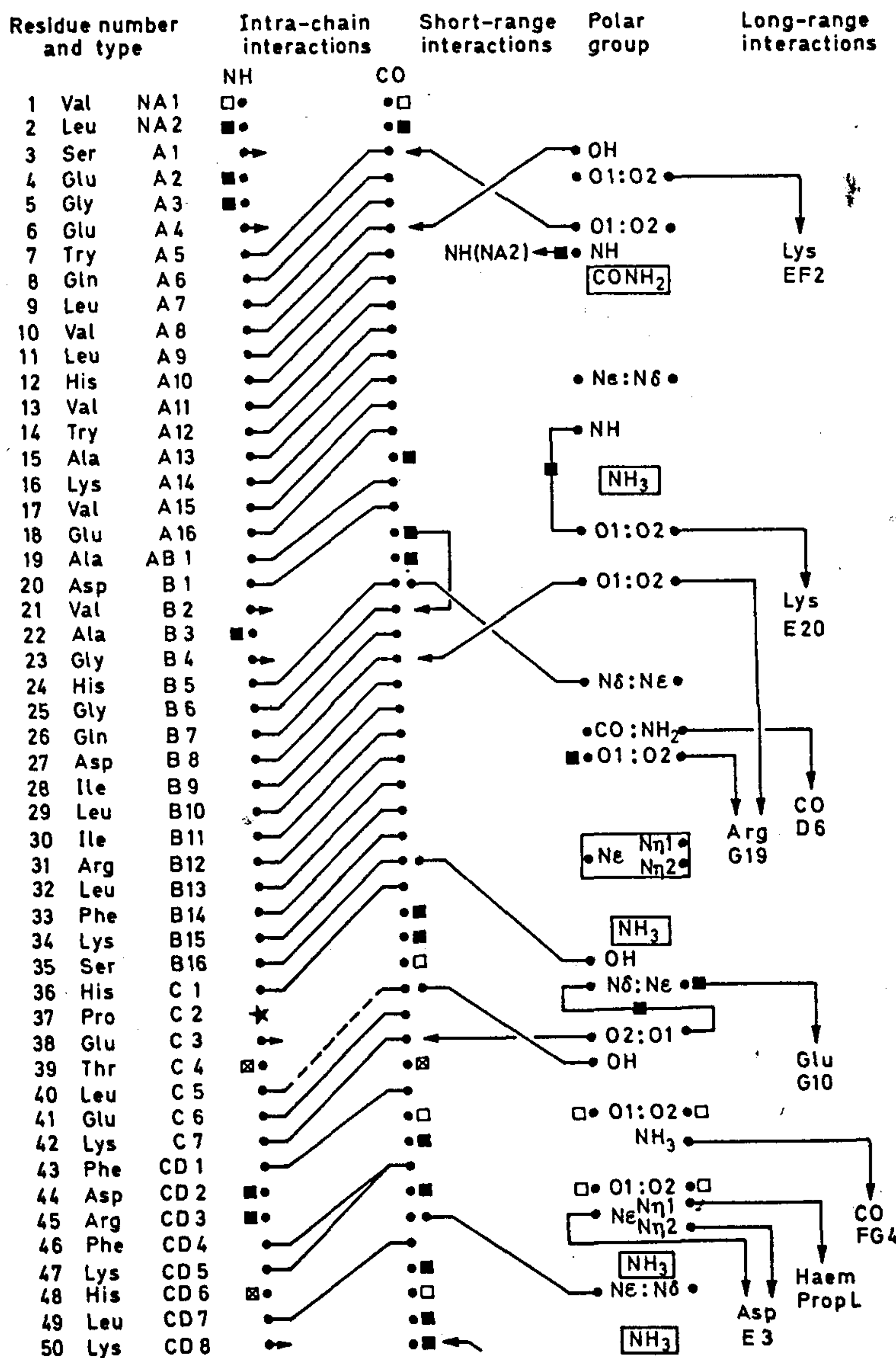

TABLE 5 (continued)

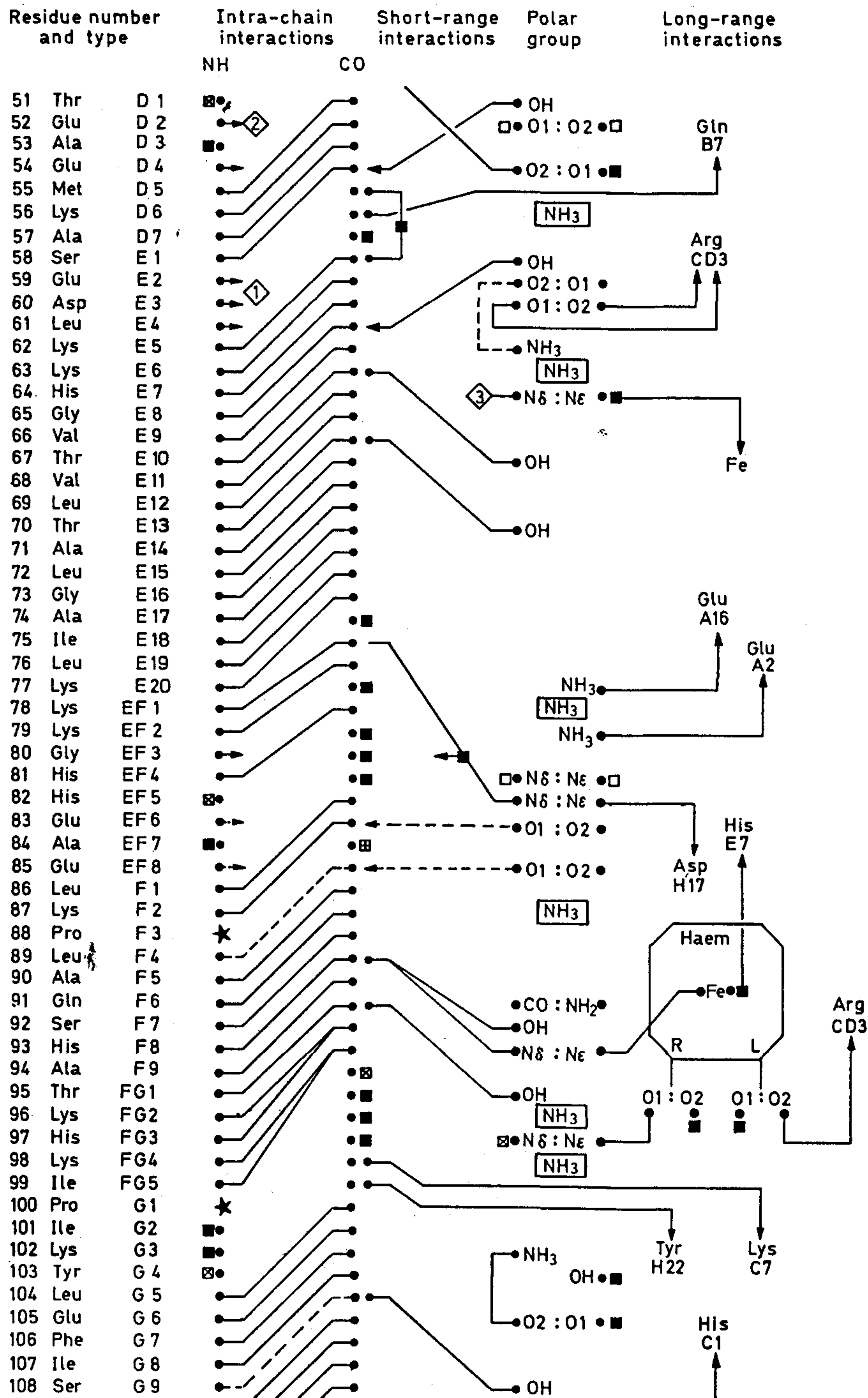

TABLE 5 (continued)

Residue number and type			Intra-chain interactions NH	CO / Short-range interactions	Polar group	Long-range interactions
109	Glu	G10	•	•	•O1 : O2•■	—
110	Ala	G11	•	•		
111	Ile	G12	•	•		
112	Ile	G13	•	•		
113	His	G14	•	••	■•Nδ : Nε •■	
114	Val	G15	•	•		
115	Leu	G16	•	•⊠		
116	His	G17	•		•Nδ : Nε •	◇(4) → Asp B8
117	Ser	G18	•	•□	•OH	
118	Arg	G19	•	•■	•Nε Nη1• Nη2•	Nη1 → Asp B8; Nη2 → Asp B1
119	His	GH1	•	•	•Nδ : Nε.•	
120	Pro	GH2	★	•■		
121	Gly	GH3	⊠•	•□		
122	Asn	GH4	■•	•■	CONH2	
123	Phe	GH5	•	•■		
124	Gly	GH6	•►	•◄		
125	Ala	H1	■•	•		
126	Asp	H2	•►	•◄	•O2 : O1•■	
127	Ala	H3	⊠•	•		
128	Gln	H4	•	•	CONH2	
129	Gly	H5	•	•		
130	Ala	H6	•	•		
131	Met	H7	•	•		
132	Asn	H8	•	•	•CO : NH_2•	
133	Lys	H9	•	•	NH_3	
134	Ala	H10	•	•		
135	Leu	H11	•	•		
136	Glu	H12	•	•	■•O1 : O2 •■	
137	Leu	H13	•	•		
138	Phe	H14	•	•		
139	Arg	H15	•	•	•Nε Nη1• Nη2•	
140	Lys	H16	•	•	NH_3•	
141	Asp	H17	•	•	■•O1 : O2•	His EF5
142	Ile	H18	•	•		
143	Ala	H19	•	•		
144	Ala	H20	•	•		
145	Lys	H21	•	•	NH_3	
146	Tyr	H22	•	•	OH•	CO FG4
147	Lys	H23	•	•	NH_3	
148	Glu	H24	•	•⊠	■•O1 : O2 •■	
149	Leu	HC1	•	•■		
150	Gly	HC2	•	•■		
151	Tyr	HC3	•	•■	OH•	
152	Gln	HC4	■•	•□	CONH2	
153	Gly	HC5	□•	•□		

C. *The non-helical regions of the main chain*

Chain conformation may be described in terms of the dihedral angles ϕ, ψ which relate the planes of successive peptide groups to the α-carbon atom through which they are linked together (FIGURE 7.4); the complete myoglobin main-chain dihedral angles are listed in the Appendix. In FIGURE 7.5, the values of these angles are plotted for the non-helical residues together with the first residue in each of the major helices.

The broken lines enclose regions outside which, according to Ramachandran *et al.* (1963), unacceptable van der Waals contacts are formed between the atoms of neighbouring units. Recent work by the same authors (1965), by Dunnill (1965) and by others has shown that the accessible areas should be enlarged somewhat: in particular, there is a continuous region of permissible values running from the right-handed α-helix region to the extended chain area. Much of the normally forbidden area is accessible when the included residue is glycine.

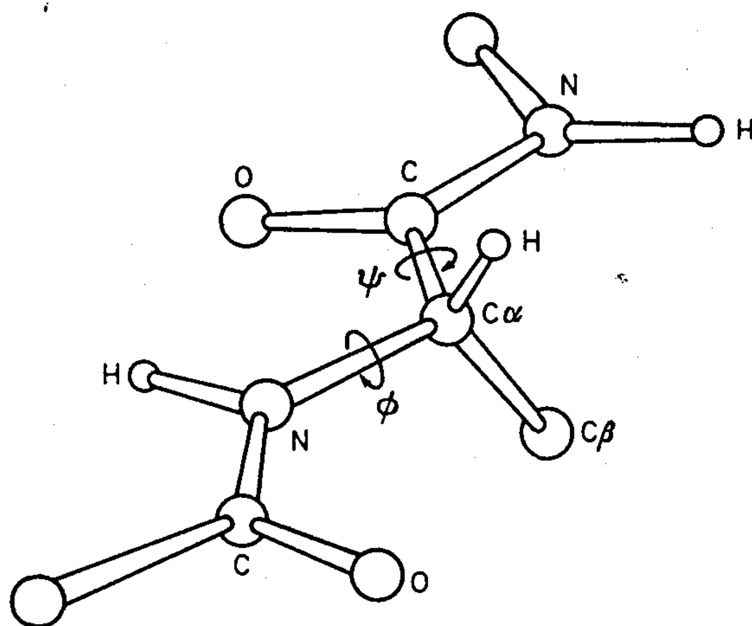

FIGURE 7.4. Two planar peptide groups: the variable dihedral angles ϕ and ψ define the configuration about the linking carbon atom

With the exception of the residues EF3 and HC2, which are both glycine, the main-chain dihedral angles fall, within the experimental error, in the permissible regions (it is difficult to assess the true error involved, but the scatter of α-helical values would indicate $\pm 15°$). The densest cluster of points lies near the right-handed α-helix position but with the centre of gravity displaced slightly towards the $3{\cdot}0_{10}$ position.

A point near the α-helix position does not necessarily imply that the residue is part of a recognizable helix. To generate a helix, a succession of residues must have the same conformational relationship. This situation exists between residues EF5 to F1 and GH2 to GH4 which each form one turn of a regular α-helix. The region EF5 to F1, in terms of the main-chain dihedral angles, appears to form part of the F helix (see FIGURE 7.8), although the proline residue at F3 restricts the formation of at least one regular hydrogen bond (for the bonding arrangement see TABLE 5). In addition, the corner residues CD2 to CD6 and FG1 to FG3 also form continuous sequences in which the main-chain dihedral angles have values in the region of either the α-helix or $3{\cdot}0_{10}$

positions. Although these regions do not show any regular bonding pattern, they nevertheless form spiral structures which have the same right-handed sense as the regular α-helices.

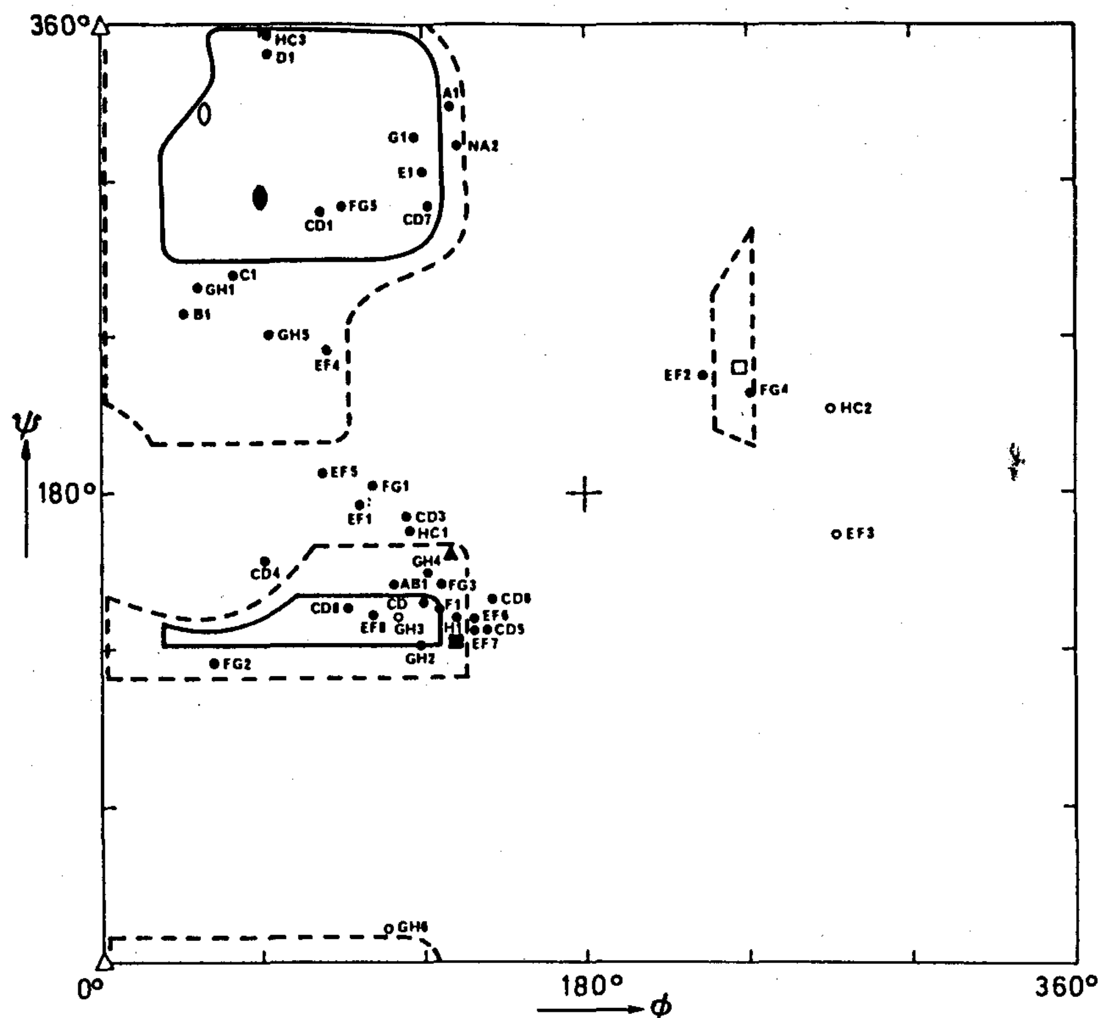

FIGURE 7.5. Plot of the dihedral angles ϕ and ψ for the non-helical regions of the myoglobin polypeptide chain and the first residue in each helix

—, - - - allowed regions calculated by Ramachandran *et al.* (1963) using two criteria for permissible interatomic contact distances
glycine residues, subject to less stringent limits
position of ■ □ right- and left-handed α-helix
▲ 3·0_{10} helix
△ fully extended chain
● ○ parallel and anti-parallel pleated sheet (see Schellman and Schellman, 1964)

From the helix definition given in Section 4B it follows that the first and last-plus-one residues of each helical section have main-chain dihedral angles which differ significantly from those of the regular helical residues. Allowing for these residues and for the three at the *N*- or *C*-terminal ends of the chain which have not been resolved, there are only 9 residues which have main-chain dihedral angles dissimilar to those of the 3·0_{10} and α-helix. The apparent affinity for the 3·0_{10}/α-helix conformation is such that there are only 5 isolated chain regions where non-α-helical configurations are used, within a corner, to form the necessary interhelical sections.

The four longer corners are all quite different in conformation (see FIGURES* 7.7, 7.8 and 7.10). The two sharp corners have a superficial similarity but are different in detail, as is indicated by the main-chain dihedral angle values for the residues C1 and E1 (see also FIGURES 7.7 and 7.9). The BC corner is essentially the same as that between the end of the G helix and the one turn of α-helix in the GH corner (see FIGURES 7.7 and 7.10). It is perhaps more than coincidence that in both these corners the first and second residues of the carboxyl helix are histidine and proline, respectively.

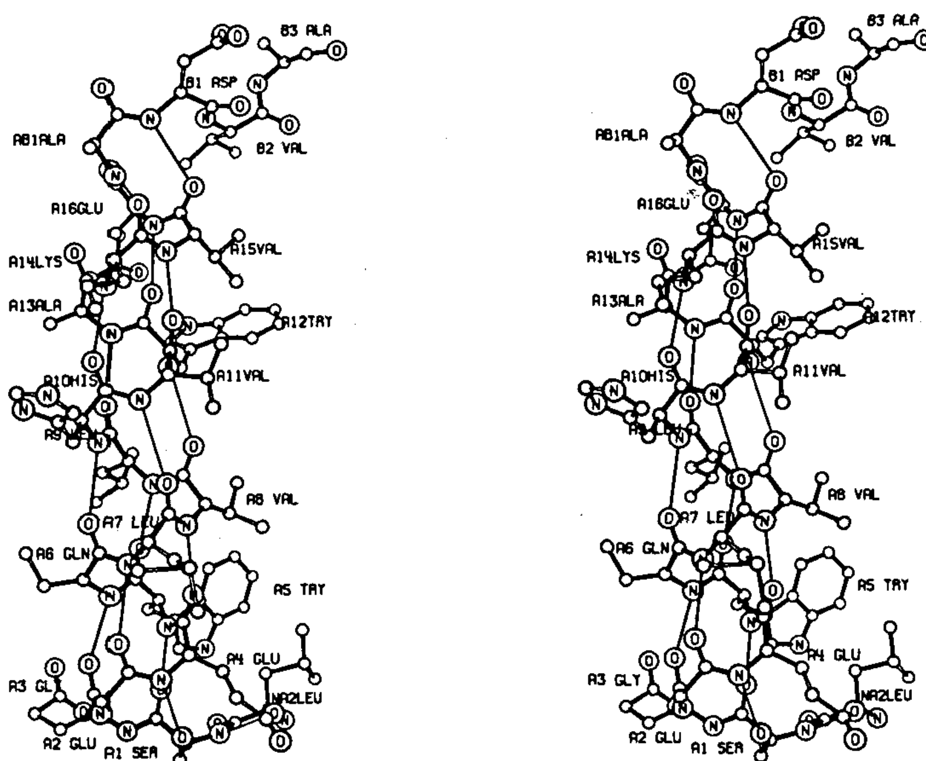

FIGURE 7.6. Amino end of the molecule including the first helix, A, and AB corner (NA2–B3)

* FIGURES 7.6–7.13 were prepared by Dr. C. K. Johnson using a modified version of his OR.TEP computer program. The covalent bonds of the myoglobin main chain are shown as wide solid lines, those of the side chains as parallel lines, and hydrogen bonds indicated by narrow solid lines. Carbon-atom positions are designated by unlettered small circles, those of nitrogen, oxygen, sulphur and iron atom positions by larger circles enclosing their chemical symbols; hydrogen atoms are not shown. The identifying symbols contain an abbreviation for the residue name together with the position within the helical or non-helical segment. Atoms which have not been located with certainty (for details see Appendix) have not been included; the terminal ends of the polypeptide chain and several side chains thus appear incomplete.

The stereo-photograph of the frontispiece and these stereo-drawings are best seen with a stereo pocket viewer having lenses set approximately 6·5 cm apart. Best results are obtained in soft but even illumination. It is recommended that the viewer should direct his initial attention to some small portion of the more complicated stereo-drawings before attempting to look at the whole field of view.

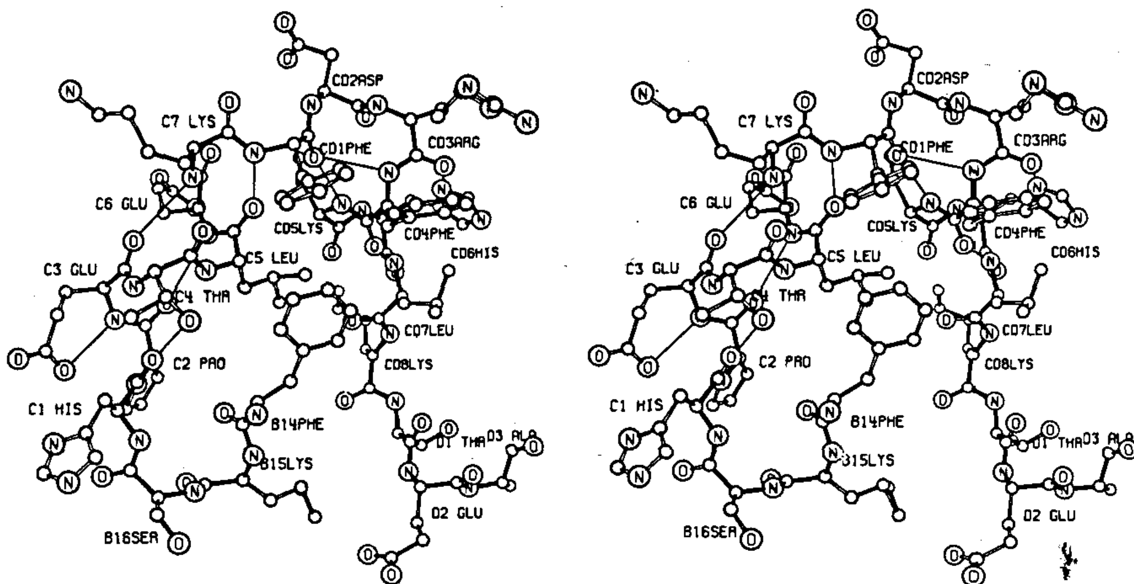

FIGURE 7.7. CD corner region (B14–D3)

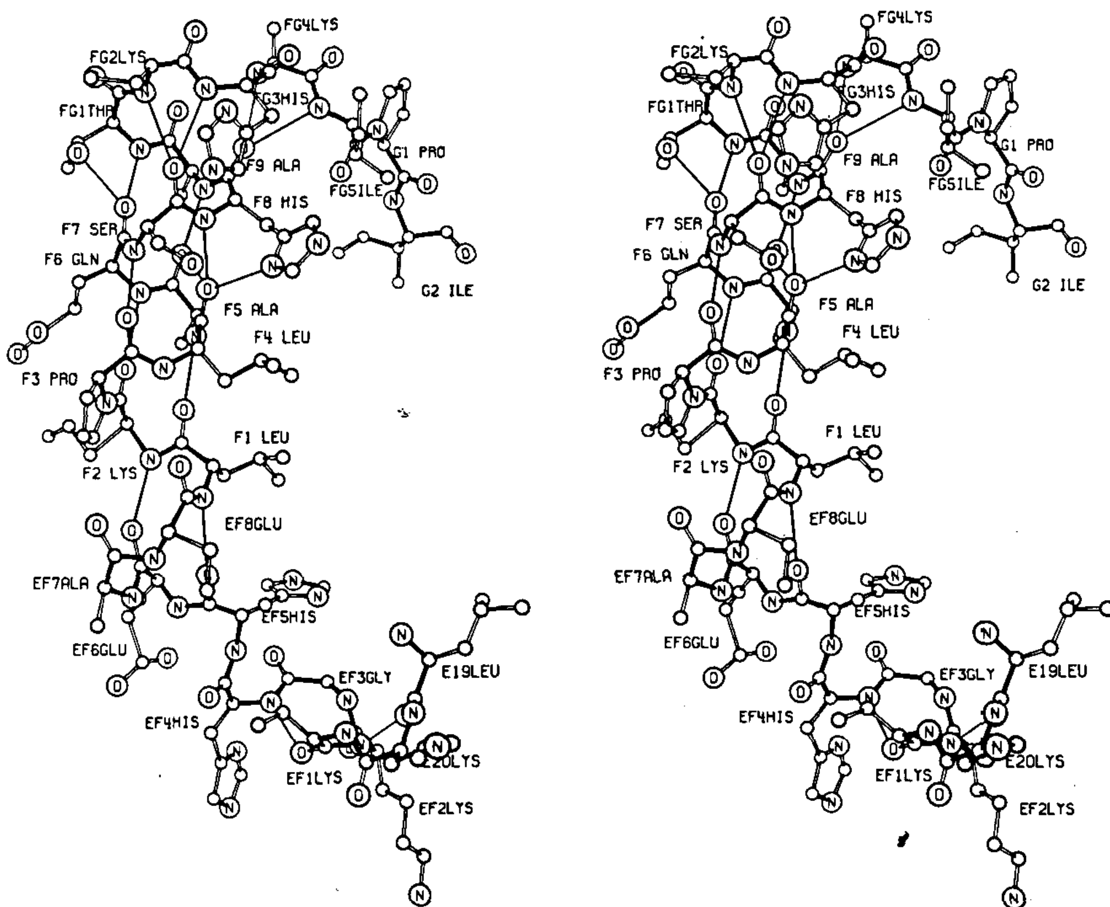

FIGURE 7.8. F helix including EF and FG corner regions (E19–G2)

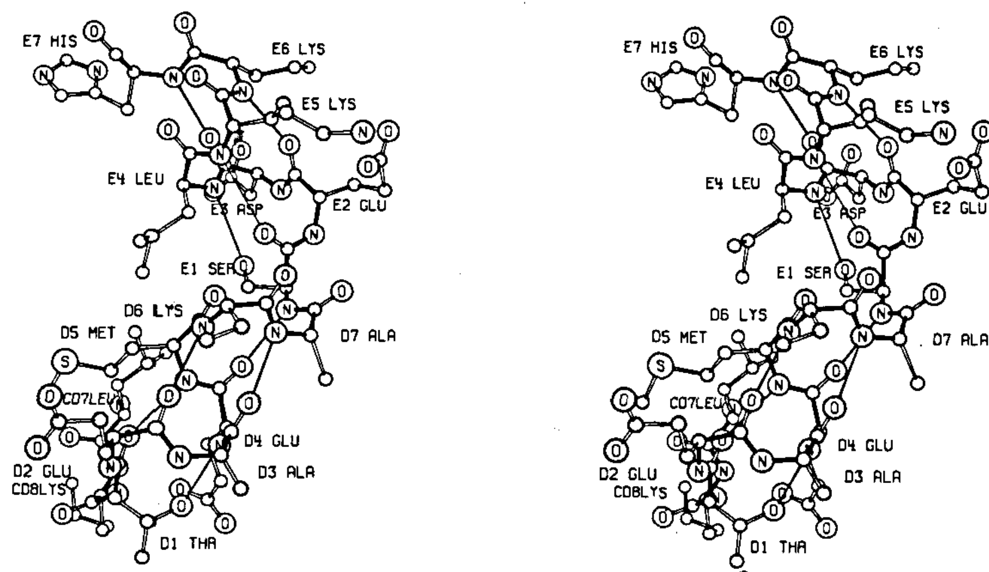

FIGURE 7.9. D helix and DE corner region (CD7–E7)

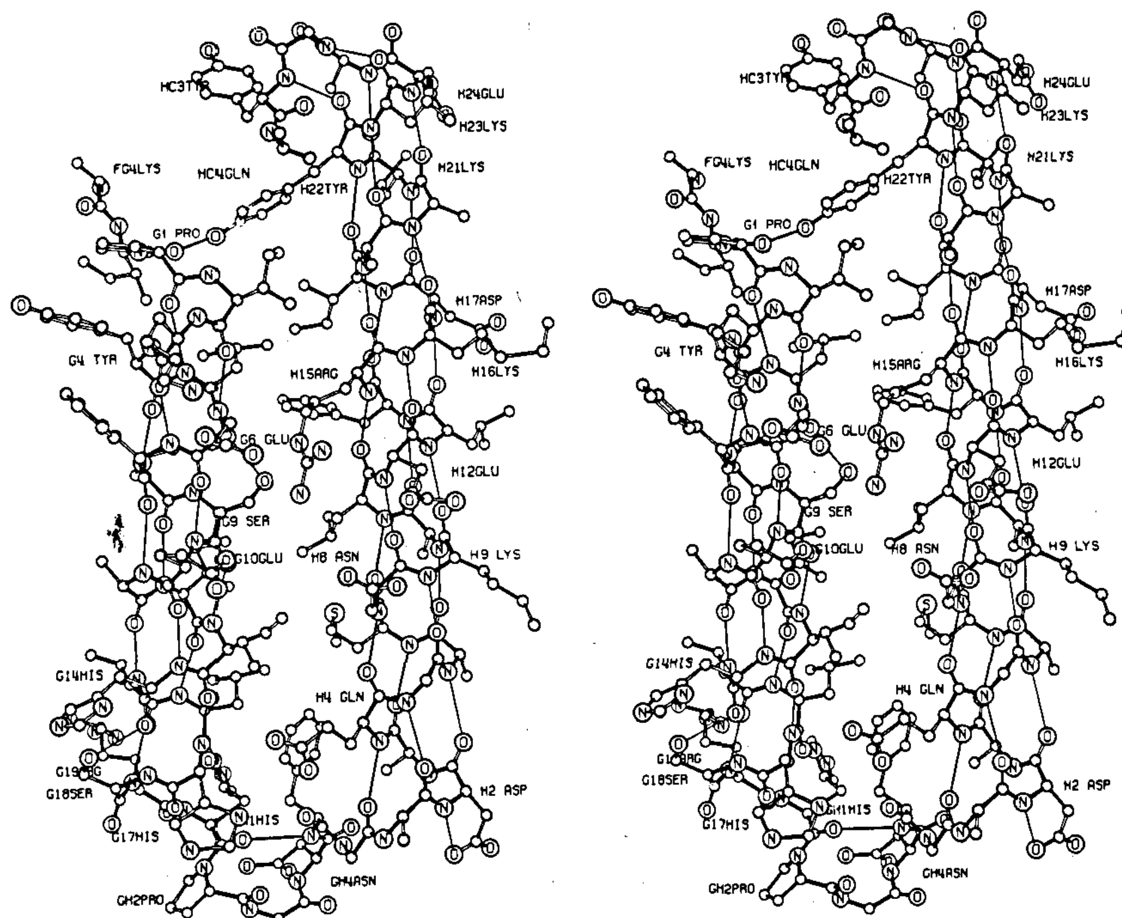

FIGURE 7.10. Carboxyl end of the molecule including G and H helices; note hydrogen bond between tyrosine hydroxyl (H22) and main-chain carbonyl of residue FG5

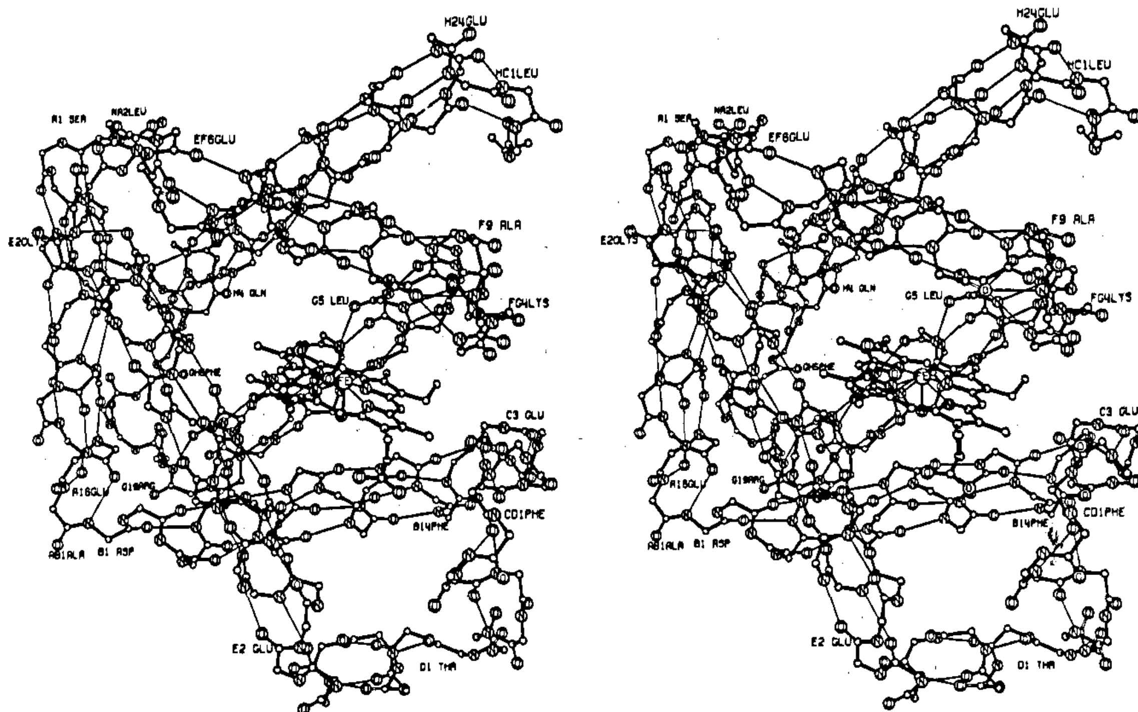

FIGURE 7.11. Main chain and haem group without side chains

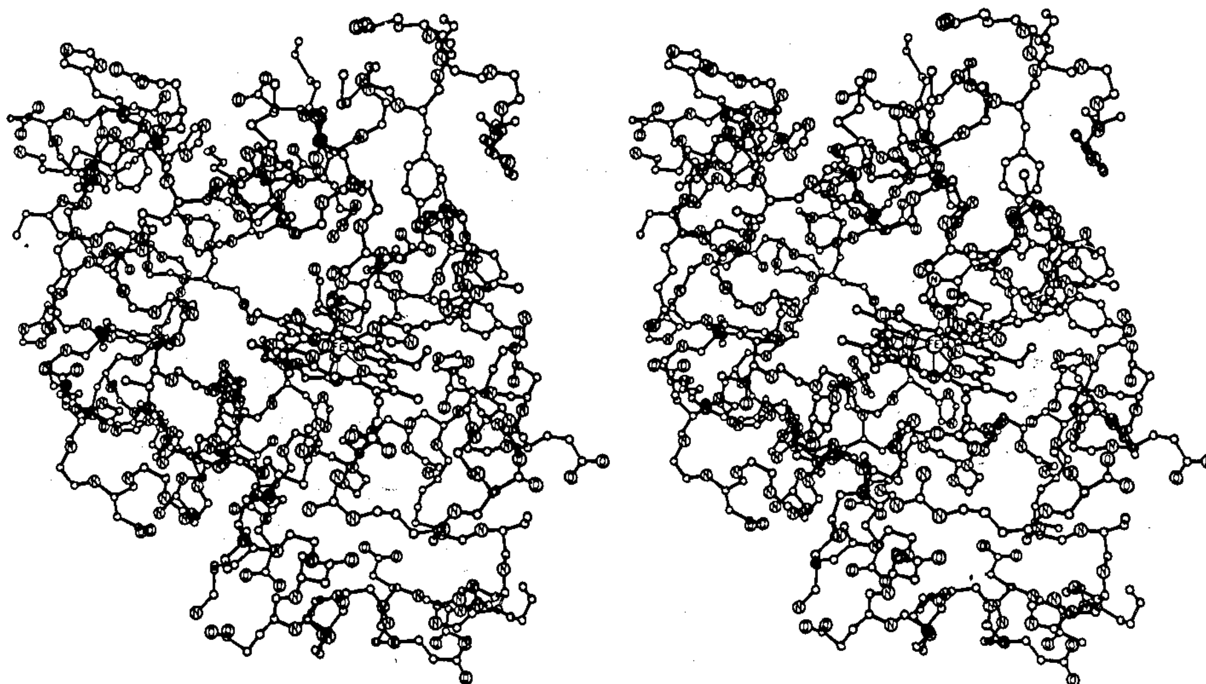

FIGURE 7.12. Main chain, haem and polar side chains only (main-chain carbonyl oxygen atoms not included); note polar side-chain groups distributed over the molecular surface

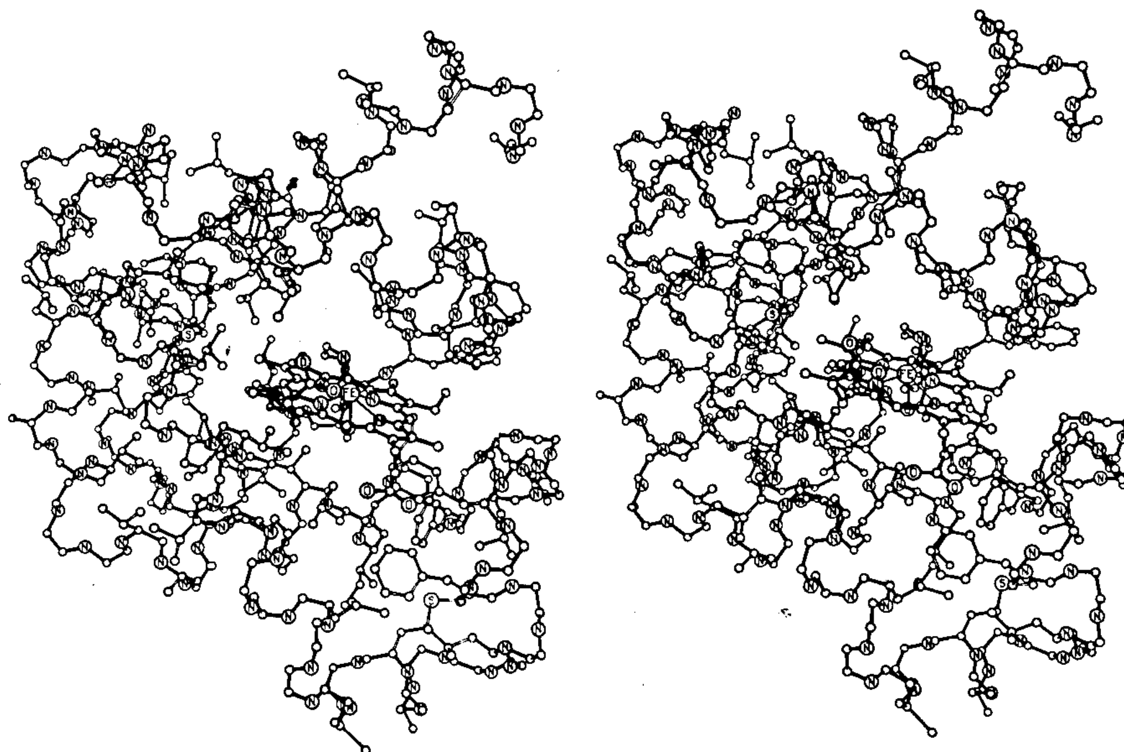

FIGURE 7.13. Main chain, haem and non-polar side chains only, emphasizing the hydrophobic nature of the centre of the molecule (main-chain carbonyl oxygen atoms not included)

D. *Side-chain conformation*

There are three preferred or staggered directions for a carbon–carbon bond in an aliphatic side chain. According to the convention proposed by Edsall *et al.* (1966), these positions have dihedral angles χ of 60°, 180° and 300° (see FIGURE 7.14). FIGURE 7.15 shows the distribution of

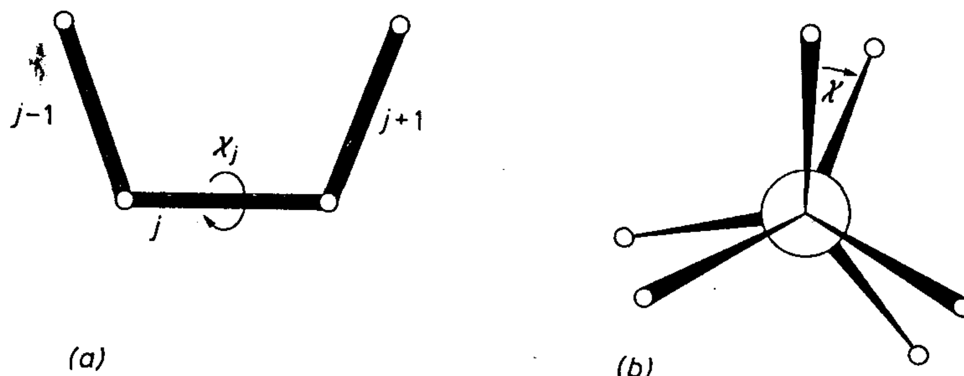

FIGURE 7.14. Definition of the angle of rotation, χ (as defined by Edsall *et al.*, 1966) for rotation about single bonds in a side chain

the dihedral angles about the Cα–Cβ bond, χ_1, for residues which have an atom in the γ position. The experimental error involved in determining χ_n values varies according to the type of residue and to its

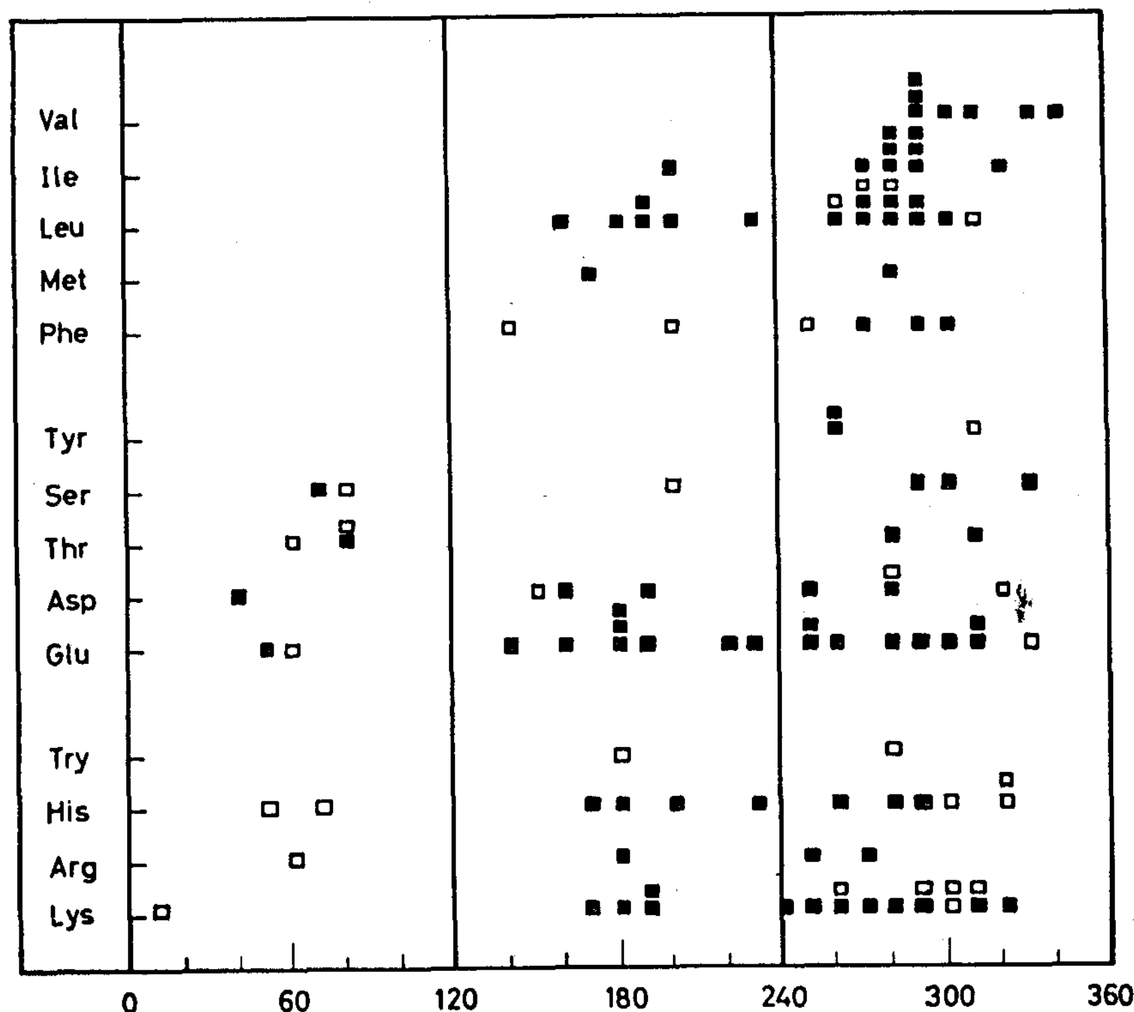

FIGURE 7.15. Rotation about Cα–Cβ bond

□ ■ points for helical and non-helical residues; staggered positions are at 60, 180 and 300°

position within the molecule. Of the 118 χ_1 values, 84 fall within $\pm 20°$ of the staggered positions. A further 22 residues, however, have χ_1 values between 250° and 280°, which suggests that the 300° stagger position is displaced some 15° towards the fully-extended chain configuration ($\chi = 180°$). For residues within an α-helix, this corresponds to a movement of the γ atom towards its own carbonyl group.

For α-helical residues, the $\chi_1 = 60°$ position is 'forbidden', since in this configuration the atom of the *N*th helix residue would approach too close to the carbonyl oxygen atom of residue N-4. Residues C4 (Thr) and F7 (Ser) form exceptions to this rule (see FIGURES 7.7 and 7.8), since their side-chain hydroxyl groups form bonds with main-chain oxygen atoms which distort the regular C and F helix structures. The same close contact restricts those α-helical residues which have two γ atoms (the forked β residues, valine, threonine and isoleucine) to one basic configuration with $\chi_1 = 300°$. One residue, isoleucine G8 (see FIGURE 7.10), is allowed to take up a 'forbidden' position ($\chi_1 \approx 180°$) because serine G9 distorts the helix by forming a hydrogen bond with the carbonyl oxygen of G4.

Allowing for the restricted configuration of the β forked residues, the staggered direction of the Cβ–Cγ bond which is *trans* to the Cα–C bond ($\chi_1 = 300°$) seems preferred to the position which is *trans* to the N–C bond ($\chi_1 = 180°$) for the helical residues alone (39 from a possible 66). A similar preference exists amongst the non-helical residues (the first in each helix is also included, since its main-chain dihedral angles differ significantly from those of the remaining helical residues), where the number of residues with χ_1 values close to the three staggered positions are 7, 5 and 18, respectively.

Similar restrictions apply to the χ_2 angles of side-chains isoleucine and leucine. For example, all the helical isoleucines, with the exception of G2 (which occurs at the amino-end of a helix), favour the $\chi_2 \simeq 180°$ configuration. For the helical leucine residues, all the χ_1 angles are similar to the χ_2 angles (i.e. $\chi_1 \simeq \chi_2 \simeq 180°$ or $300°$), but no marked preference is shown for either of these two configurations. With the longer side chains of methionine, glutamic acid, glutamine, lysine and arginine, no dihedral-angle restrictions apply beyond the tendency for connected atoms to be correctly staggered and for the side chains to favour, where possible, the fully extended configuration.

Care must be exercised in comparing the individual χ_2 angles of phenylalanine, tyrosine and aspartic acid, and also the χ_3 angles of glutamic acid, since an ambiguity in the atomic assignment results in equally valid χ angles separated by 180°. The same ambiguity also exists for the χ_2 angles of histidine and asparagine and for the χ_3 angles of glutamine where the respective dihedral angles, as defined by Edsall *et al.* (1966), depend on the correct assignment of all the atoms of a side chain. Given this limitation, it is surprising to find that the helical histidine residues (with the exception of C1) all favour the configuration with χ_2 closest to 60°.

5 FORCES STABILIZING THE STRUCTURE

A. *Specific interactions*

With the exception of the one covalent link between the haem group's iron atom and the ε-nitrogen of histidine F8, all the specific intramolecular and intermolecular interactions are formed by hydrogen bonds and by salt bridges.

The acidic and basic side chains are fairly evenly distributed over the surface of the molecule (see FIGURE 7.12), but only a small number are involved in side-chain/main-chain or side-chain/side-chain interactions (see long-range interactions in TABLE 5). Unless the molecular packing arrangement within the crystal excludes such a possibility, these side chains prefer to interact with ions in the space between molecules rather than with their oppositely-charged neighbours.

The haem-linked imidazole group of histidine F8 is an obvious exception to this generalization, as also are the side chains of aspartic and glutamic acid which, in the absence of neighbouring serine or threonine residues, bond to exposed nitrogens at the amino ends of α-helices (for examples, see residues B1, C3, and H2 in TABLE 5).

Of the remaining polar side chains, only two lie buried within the molecule, and these appear to have some special significance, since they are amongst the few residues which occur at equivalent positions in all normal globin chains (Perutz *et al.*, 1965). In sperm whale myoglobin both these residues, threonine C4 and tyrosine H22, form internal hydrogen bonds with main-chain carbonyl groups (see FIGURES 7.7 and 7.10). Apart from the serine and threonine hydroxyl groups which invariably interact with either main-chain NH or CO groups, the uncharged polar side chains appear to have little influence over the overall or local configuration of the molecule.

The local configuration of the main chain is 'stabilized' to some extent by the considerable number of imino–carbonyl-type hydrogen bonds. In TABLE 5 these are represented in graphic form by joining the respective NH and CO positions. The regular α-helical bond pattern is thus illustrated by a series of parallel sloping lines. The omission of one carbonyl group at the end of an α-helical region, followed by a change in slope of the bond line, indicates a change of residue configuration to that of the $3{\cdot}0_{10}$ helix (see Section 4B). At the carbonyl end of the terminal helix (H), the crossing of the bond lines merely indicates that the helix is terminated by one residue in the $3{\cdot}0_{10}$ configuration followed immediately by one in a configuration which, if repeated, would propagate a regular left-handed α-helix (see FIGURE 7.10).

Where NH and CO groups are not involved in hydrogen-bond formation, they appear to interact with ions in the solution unless the groups are sterically inaccessible.

B. *Hydrophobic interactions*

The most important and perhaps the most striking feature of the myoglobin structure is the way in which the non-polar side chains interlock to form a hydrophobic core within which calculations have shown that there is little, if any, free space (see FIGURE 7.13). With the special exceptions detailed in the last Section, all the hydrophilic groups are in contact with the ambient solution. The converse is not true, however, since hydrophobic residues occur both inside and on the surface of the molecule, although in general the bulkier side chains such as methionine and phenylalanine tend to occur at internal positions.

Simple distance calculations show that this form of molecular packing results in a very large number of interatomic contacts. Although the van der Waals' attractive forces are small for individual hydrophobic interactions, their total effect on the final configuration of the molecule

could well be appreciable. A quantitative analysis of these interactions has been undertaken, but so far no satisfactory method has been devised which permits the information contained in these calculations to be presented in a simple comprehensible form.

C. *Non-polar bonding*

The need for the large hydrophobic groups to avoid contact with water is clearly indicated by the haem group and by the tryptophan and tyrosine side chains which manage to bury their hydrophobic sections whilst cleverly arranging for their hydrophilic groups to form part of the molecular surface. These features of the myoglobin structure, together with the strikingly hydrophobic core, provide clear confirmation of Kauzmann's (1959) predictions concerning the 'non-polar bonding' of protein molecules. The newly synthesized myoglobin chain obviously folds so as to form the structure with the least free energy but maximum entropy of both protein and the surrounding medium.

I would like to thank Dr. J. C. Kendrew, F.R.S. for permission to present details of the myoglobin structure which remain unpublished, and for his help and encouragement since I joined the myoglobin structure project in 1959. I would also like to thank Dr. R. Diamond for refining the original co-ordinate data using his mathematical model-building system, Dr. C. K. Johnson for preparing the stereo-drawings using a modified version of his OR.TEP computer program, and my colleagues both past and present who have helped me in so many ways.

APPENDIX. THE MYOGLOBIN CO-ORDINATE DATA

Although attempts have been made to improve the accuracy of the myoglobin structure determination, the poor parameter to observation ratio (1:4) and the magnitude of the computational problem have so far thwarted efforts to apply conventional least-squares refinement techniques.

The co-ordinates presented in Table A1 have been derived by an elaborate mathematical tidying procedure (Diamond, 1966) in which a set of standard groups are fitted to the experimentally determined structure. As used for the myoglobin structure, this procedure allows free rotation about single bonds but constrains all interatomic bond lengths and angles to standard values. Details of the standard groups and of possible variations between refined and stereochemically accepted bond angles and bond lengths are beyond the scope of this Chapter and will be published elsewhere.

TABLE A1. MYOGLOBIN CO-ORDINATE AND DIHEDRAL ANGLE SET

Residues of the same kind are grouped together. Atoms are labelled and angles defined according to the convention proposed by Edsall *et al.* (1966). The haem group atoms are labelled: iron, Fe; methine carbons, CH1–4; pyrrol ring, N, C1, C2, C3, C4; methyl carbons, CM; vinyl carbons, Cα, Cβ; propionic acid residues Cα, Cβ, Cγ, O1, O2.

The co-ordinates (Å) are relative to the crystal axes *a*, *b* and *c**. Atoms marked * are uncertain, whilst those marked ** have been obtained by assuming some standard configuration.

Main-chain ψ and ϕ and side-chain dihedral angles, χ_1, χ_2, etc. are listed in that order below the *z* co-ordinates of the residues.

GLY		O	C	N	CA
	X	4.4	3.2	1.1	2.2
5	Y	10.3	10.2	10.8	9.8
	Z	22.5	22.2	23.4	23.3
		125	126		
	X	25.3	25.2	26.0	26.4
23	Y	18.0	18.2	19.9	18.6
	Z	11.2	12.4	13.8	13.3
		118	119		
	X	22.5	23.0	22.4	22.1
25	Y	20.2	20.4	19.8	20.7
	Z	8.0	9.1	11.4	10.3
		126	125		
	X	20.2	21.3	22.6	21.9
65	Y	24.4	24.7	25.4	24.2
	Z	11.3	10.9	9.0	9.6
		137	123		
	X	10.5	11.5	12.8	12.6
73	Y	23.1	23.6	23.2	22.8
	Z	18.4	17.8	15.8	17.2
		131	123		
	X	-0.2	0.5	2.2	1.1
80	Y	23.5	23.1	21.3	21.7
	Z	15.9	16.8	17.8	16.8
		164	273		
	X	16.7	17.8	19.8	19.0
121	Y	3.1	3.6	3.2	2.7
	Z	19.8	19.6	18.1	19.2
		133	110		
	X	11.6	11.8	13.8	13.1
124	Y	4.5	3.3	3.8	2.7
	Z	15.4	15.4	16.6	16.0
		12	106		
	X	5.5	5.9	7.7	6.4
129	Y	10.0	9.1	7.3	7.7
	Z	12.4	13.1	13.0	12.5
		132	109		
	X	0.7	0.1	-1.2	-1.1
150	Y	31.4	31.2	32.3	32.0
	Z	-6.0	-4.8	-3.0	-4.4
		212	270		
		**	**	**	**
	X	-0.9	0.3	1.5	1.1
153	Y	22.8	22.7	24.7	24.0
	Z	-7.1	-7.5	-6.4	-7.6
		78	116	136	

ALA		O	C	N	CA	CB
	X	18.8	17.7	15.4	16.3	16.1
15	Y	16.8	16.2	16.6	16.8	16.0
	Z	23.0	22.6	21.8	22.9	24.2
		144	139			
	X	25.2	24.1	21.7	23.0	23.1
19	Y	18.7	18.2	18.0	17.7	16.2
	Z	21.3	20.8	21.2	21.7	22.0
		145	109			
	X	25.8	25.7	24.5	25.3	26.1
22	Y	20.9	20.9	21.9	22.2	23.2
	Z	11.8	13.1	15.1	13.9	14.6
		145	119			
	X	33.8	33.9	33.0	34.1	35.5
53	Y	19.7	19.3	17.3	17.8	17.7
	Z	2.8	1.6	0.6	1.3	0.7
		120	125			
	X	34.1	33.7	33.1	34.2	35.2
57	Y	24.3	23.7	21.3	22.3	22.7
	Z	6.6	5.5	5.0	5.2	4.1
		126	119			
	X	11.3	12.4	14.5	13.2	13.2
71	Y	26.1	26.1	27.1	27.4	28.1
	Z	14.6	14.0	14.3	13.9	12.5
		133	120			
	X	8.4	9.2	11.6	10.7	11.4
74	Y	25.4	25.4	24.9	25.9	27.1
	Z	18.7	17.8	17.6	18.1	17.5
		130	138			
	X	5.1	4.7	2.6	3.5	3.7
84	Y	33.0	31.9	30.8	31.9	32.1
	Z	14.3	14.9	15.5	15.8	17.3
		127	138			
	X	6.1	5.7	6.4	5.6	4.1
90	Y	31.3	31.5	30.6	30.3	30.2
	Z	2.8	4.0	6.2	5.0	5.4
		128	128			
	X	7.4	7.1	7.6	6.7	5.2
94	Y	32.7	32.8	31.3	31.5	31.3
	Z	-2.6	-1.4	0.6	-0.6	-0.4
		124	106			
	X	17.8	17.0	15.7	17.1	17.5
110	Y	13.3	13.9	13.8	13.9	15.3
	Z	7.1	6.4	4.3	4.8	4.4
		139	137			

ALA		C	C	N	CA	CB
	X	8.1	8.6	10.9	9.6	9.5
125	Y	4.7	3.7	2.4	2.8	1.3
	Z	14.5	15.1	15.0	14.5	13.9
		132	132			
	X	9.3	9.7	9.3	9.8	11.3
127	Y	9.1	8.0	6.1	7.4	7.2
	Z	15.3	15.6	17.2	17.0	17.4
		132	109			
	X	5.7	6.4	6.1	5.7	6.1
130	Y	12.3	11.4	9.1	10.3	10.0
	Z	14.0	14.4	14.4	15.3	16.7
		120	132			
	X	4.8	5.3	4.6	5.0	6.1
134	Y	17.8	16.7	14.5	15.9	15.7
	Z	11.5	11.7	12.6	12.9	14.1
		125	142			
	X	-0.5	-0.3	1.7	0.9	0.7
143	Y	26.0	25.4	24.7	24.5	23.0
	Z	1.3	2.3	3.7	2.4	2.1
		129	127			
	X	-2.7	-2.1	-1.1	-2.3	-3.0
144	Y	28.2	27.6	25.4	26.2	26.0
	Z	2.2	3.1	3.3	3.4	4.8
		125	126			

VAL		O	C	N	CA	CB	CG1	CG2
		**	**	**	**	**	**	**
	X	-3.7	-3.0	-2.9	-3.6	-3.5	-4.6	-2.1
1	Y	14.7	15.3	17.6	16.4	16.0	14.9	15.7
	Z	17.0	16.2	15.5	15.3	13.8	13.4	13.3
		123	133	196				
	X	10.9	9.8	8.3	8.8	8.1	7.6	8.9
10	Y	14.5	14.2	12.2	13.3	14.2	13.4	15.4
	Z	19.0	19.5	19.6	18.7	17.7	16.5	17.3
		131	128	294				
	X	16.3	15.1	13.6	14.5	14.4	13.5	15.8
13	Y	14.2	14.0	12.7	12.6	12.1	10.9	11.9
	Z	19.4	19.6	21.0	19.8	18.4	18.3	17.7
		136	114	338				
	X	20.9	19.9	18.8	19.3	18.2	17.3	18.7
17	Y	18.1	17.6	15.5	16.3	17.0	16.0	18.1
	Z	18.1	18.7	19.4	18.2	17.5	16.8	16.5
		148	119	330				
	X	22.8	23.3	23.4	22.7	22.2	21.3	21.4
21	Y	20.9	21.2	20.3	21.0	22.3	22.0	23.3
	Z	14.0	15.0	17.4	16.4	17.0	18.2	16.1
		124	132	293				
	X	19.6	20.5	22.2	21.8	22.7	24.0	22.1
66	Y	26.7	26.9	25.5	26.1	26.9	26.1	27.4
	Z	13.6	12.8	11.6	12.9	13.8	14.1	15.1
		125	124	314				

VAL		O	C	N	CA	CB	CG1	CG2
		**	**	**	**	**	**	**
	X	15.4	16.6	18.2	17.1	17.7	18.1	16.9
68	Y	25.2	25.2	27.0	26.1	25.2	26.0	23.9
	Z	11.9	11.6	10.6	10.4	9.3	8.1	9.0
		124	127	291				
	X	20.5	19.5	18.7	19.7	20.3	21.1	21.1
114	Y	10.9	11.4	11.3	11.9	13.3	13.7	13.5
	Z	11.4	10.7	8.4	9.3	9.4	8.2	10.8
		126	138	297				

LEU		O	C	N	CA	CB	CG	CD1	CD2
				*			*	*	*
	X	-0.9	-1.0	-1.7	-0.9	0.6	1.1	2.6	0.4
2	Y	14.9	13.9	15.1	14.1	14.2	14.3	14.4	15.5
	Z	19.0	18.3	16.0	16.7	16.5	15.1	15.0	14.4
		314	133	313	177				
							*	*	*
	X	10.4	9.1	7.4	8.4	7.9	7.9	6.6	8.2
9	Y	11.5	11.4	11.0	10.4	9.6	8.1	7.6	7.7
	Z	20.2	20.3	21.9	21.1	19.9	19.9	19.5	21.4
		133	116	288	130				
	X	12.5	11.6	9.4	10.2	9.5	8.7	9.6	7.7
11	Y	15.8	15.0	14.5	15.4	16.3	17.4	18.5	18.0
	Z	22.3	22.2	20.7	21.6	22.5	21.9	21.4	22.9
		139	117	284	284				
	X	21.9	22.5	21.5	21.9	22.3	21.5	20.2	22.3
29	Y	19.2	19.2	19.2	20.0	21.4	22.2	22.7	23.3
	Z	2.4	3.5	5.7	4.6	5.1	6.1	5.5	6.7
		116	131	301	286				
	X	19.6	20.1	20.3	19.4	18.6	17.5	16.9	16.4
32	Y	17.5	17.8	16.7	17.4	18.6	19.2	20.4	18.2
	Z	-1.4	-0.3	1.9	1.0	1.6	0.8	1.5	0.6
		125	124	186	187				
	X	21.3	21.0	19.6	20.9	22.0	23.4	24.4	23.8
40	Y	26.4	25.2	23.3	24.1	23.2	23.8	22.8	24.3
	Z	-4.0	-4.4	-3.6	-3.3	-3.9	-4.0	-4.5	-2.6
		140	99	196	175				
	X	28.6	29.7	29.4	29.7	28.9	29.4	28.4	29.6
49	Y	22.9	23.4	25.9	24.6	24.2	24.6	24.2	26.1
	Z	-4.5	-4.1	-3.8	-3.2	-2.0	-0.6	0.5	-0.6
		290	122	274	176				
	X	25.1	26.3	27.6	26.7	27.5	26.8	27.0	27.5
61	Y	25.7	25.8	27.3	26.1	24.9	24.1	22.6	24.3
	Z	7.7	7.3	6.0	5.9	5.4	4.3	4.7	3.0
		126	122	230	229				
	X	15.4	16.5	17.6	17.3	18.7	19.1	18.0	19.6
69	Y	23.8	24.3	24.5	23.5	22.9	21.7	21.3	20.6
	Z	14.7	14.3	12.1	13.2	13.5	12.7	11.8	13.6
		128	133	264	351				
	X	10.7	11.9	12.9	12.4	13.0	12.2	13.0	10.9
72	Y	22.8	23.2	25.0	23.7	22.6	21.3	20.3	21.6
	Z	15.1	14.9	13.5	13.5	12.7	12.5	11.7	11.7
		123	130	185	185				

LEU		O	C	N	CA	CB	CG	CD1	CD2
				*			*	*	*
	X	6.3	7.5	8.2	8.0	9.4	9.8	8.7	11.0
76	Y	20.8	21.1	22.2	20.9	20.2	19.9	19.3	19.0
	Z	19.1	18.8	16.7	17.4	17.1	15.7	14.9	15.7
		111	130	280	313				
	X	5.3	4.8	5.5	5.2	4.2	4.6	3.8	6.1
86	Y	30.0	30.1	29.4	29.1	28.0	26.8	25.6	26.5
	Z	8.7	9.8	12.3	10.9	10.5	9.8	10.1	10.1
		134	126	246	151				
								*	*
	X	8.3	7.6	7.5	8.3	8.4	9.4	8.8	10.7
89	Y	31.0	30.9	32.1	31.2	29.9	28.8	27.4	28.9
	Z	5.1	6.1	8.3	7.5	8.3	7.8	8.0	8.6
		128	113	186	221				
								*	*
	X	10.1	9.7	10.1	9.7	8.5	8.6	10.1	7.8
104	Y	18.7	19.2	20.9	20.7	21.5	23.0	23.4	23.5
	Z	3.9	2.8	1.2	2.6	2.8	3.0	3.1	4.2
		140	124	269	352				
	X	18.6	18.2	18.3	17.9	16.4	16.1	17.3	15.1
115	Y	9.1	9.6	11.5	11.1	11.4	12.9	13.7	13.0
	Z	13.8	12.8	11.2	12.5	12.8	13.0	13.3	14.2
		139	121	288	341				
	X	5.5	5.6	6.2	6.6	7.9	9.0	9.2	10.3
135	Y	18.3	17.2	16.1	16.8	15.9	16.0	17.5	15.5
	Z	8.1	8.6	10.9	9.6	9.4	10.4	10.8	9.7
		127	125	273	319				
	X	1.7	2.5	2.7	2.0	1.4	1.0	1.5	-0.6
137	Y	20.9	20.0	17.6	18.6	18.5	19.8	19.8	19.9
	Z	9.1	9.5	9.1	9.8	11.2	11.9	13.3	12.0
		128	119	157	227				
						*	**	**	**
	X	-0.1	-0.8	-2.0	-1.0	0.2	0.1	0.1	1.3
149	Y	34.3	33.4	32.6	33.4	33.1	33.4	34.9	32.8
	Z	-3.0	-2.4	-0.1	-0.9	0.0	1.5	1.7	2.2
		166	115	279	292				

ILE		O	C	N	CA	CB	CG1	CG2	CD1
	X	20.4	20.8	21.7	20.5	19.4	19.0	18.3	18.5
28	Y	17.6	18.0	17.1	17.3	17.9	17.0	18.3	17.7
	Z	4.5	5.6	7.7	6.9	7.8	9.0	6.8	10.2
		130	138	284	152				
	X	23.5	23.5	23.6	24.4	25.7	26.4	26.5	27.7
30	Y	16.6	16.7	18.6	17.8	17.6	19.0	16.6	18.8
	Z	1.0	2.2	3.8	2.9	3.7	3.9	2.8	4.8
		130	130	275	178				
	X	6.2	7.3	9.1	7.8	8.0	8.4	6.7	8.9
75	Y	23.1	23.1	25.0	24.4	24.3	25.7	23.7	25.6
	Z	17.2	16.7	16.5	16.1	14.5	13.9	14.1	12.4
		127	117	287	169				
	X	8.6	9.6	10.7	10.8	11.4	12.8	11.2	13.4
99	Y	26.2	26.1	28.3	26.9	26.9	27.5	25.5	27.8
	Z	-1.2	-2.1	-2.4	-2.0	-0.6	-0.5	-0.2	-1.9
		290	90	281	358				

ILE		O	C	N	CA	CB	CG1	CG2	CD1
	X	7.5	7.2	6.9	6.4	4.9	4.0	4.5	4.0
101	Y	20.2	20.8	23.1	22.1	21.8	23.0	21.2	24.1
	Z	0.0	-1.0	-1.9	-0.8	-1.0	-1.4	0.4	-0.4
		146	138	324	295				
	X	14.6	13.8	13.5	14.2	14.5	15.8	13.1	15.8
107	Y	16.7	17.3	18.0	18.3	19.8	20.0	20.2	21.4
	Z	5.2	4.5	2.2	3.4	3.8	4.7	4.5	5.4
		128	141	201	196				
	X	16.5	15.7	16.0	15.7	14.6	14.8	14.5	14.3
111	Y	13.3	13.5	14.7	14.9	15.8	17.3	15.4	18.2
	Z	9.9	9.0	6.8	8.3	8.8	8.5	10.3	9.6
		119	133	293	215				
									*
	X	16.1	15.7	14.8	14.6	13.1	12.0	13.1	10.6
112	Y	9.6	10.4	12.7	11.4	11.0	11.8	9.5	11.7
	Z	9.8	8.9	8.6	9.2	8.7	9.4	8.8	8.8
		141	114	267	169				
	X	2.0	2.2	2.1	2.9	4.3	5.2	4.8	6.7
142	Y	26.9	25.9	25.0	25.9	25.2	25.3	26.0	25.3
	Z	3.4	4.0	6.2	5.4	5.1	6.4	3.9	6.1
		134	115	292	206				

SER		C	C	N	CA	CB	OG
	X	1.1	-0.1	-1.1	-1.1	-1.1	0.2
3	Y	12.8	12.6	12.6	12.2	10.8	10.1
	Z	20.9	21.2	18.6	20.0	20.5	20.3
		329	130	82			
	X	18.5	19.0	21.1	19.9	19.1	19.9
35	Y	14.3	14.9	14.7	14.1	13.7	12.8
	Z	-4.9	-4.0	-2.5	-3.0	-1.8	-0.9
		141	110	292			
	X	30.4	31.4	32.8	32.2	31.0	30.3
58	Y	25.1	25.8	24.2	25.5	25.6	26.8
	Z	6.5	6.2	4.7	4.9	4.0	4.2
		303	120	198			
	X	10.2	9.7	8.0	9.4	10.5	10.6
92	Y	33.5	33.0	33.9	34.0	33.9	32.6
	Z	1.5	2.6	4.1	3.8	4.9	5.4
		120	131	69			
	X	12.5	12.4	12.5	11.8	10.3	9.6
108	Y	14.2	14.9	17.2	16.3	16.0	17.1
	Z	6.7	5.6	4.7	5.6	5.5	5.9
		146	131	298			
							*
	X	23.3	22.4	20.2	21.7	22.4	21.6
117	Y	7.0	7.8	7.6	7.6	8.0	8.0
	Z	12.3	11.9	10.7	10.6	9.3	8.2
		128	148	331			

THR		O	C	N	CA	CB	CG2	CG1
	X	18.2	18.4	17.4	17.3	17.2	16.1	18.4
39	Y	24.7	23.7	22.0	22.7	21.9	20.8	21.3
	Z	-2.7	-3.3	-5.0	-3.7	-2.4	-2.6	-2.3
		146	128	82				
	X	30.9	31.2	31.7	31.8	33.3	33.9	33.9
51	Y	19.5	18.5	20.0	18.6	18.3	18.3	19.4
	Z	-1.0	-1.7	-3.7	-3.1	-3.1	-4.4	-2.3
		350	62	59				
	X	17.3	18.2	20.5	19.4	19.8	18.7	20.8
67	Y	28.1	27.9	27.9	28.8	29.9	30.8	30.7
	Z	12.5	11.6	11.9	11.7	10.7	10.2	11.4
		117	134	279				
	X	14.1	14.9	17.0	16.4	17.3	17.1	18.6
70	Y	26.4	26.6	25.5	26.3	27.5	28.2	27.0
	Z	16.4	15.4	14.6	15.6	15.8	17.2	15.8
		125	132	311				
	X	9.0	8.9	7.2	7.6	6.9	5.5	7.7
95	Y	36.9	35.8	33.9	35.2	36.3	36.0	36.5
	Z	-2.2	-1.6	-0.7	-1.2	-0.5	-0.1	0.7
		173	101	77				

ASP		O	C	N	CA	CB	CG	OD1	OD2
	X	24.1	24.0	23.8	24.8	26.1	26.7	27.0	26.8
20	Y	18.5	19.2	18.1	18.6	17.9	18.2	17.3	19.3
	Z	16.1	17.2	19.5	18.5	18.3	16.9	16.1	16.6
		249	31	149	239				
	X	23.1	22.9	24.6	24.0	24.1	25.5	26.3	25.8
27	Y	16.5	16.7	17.9	16.6	15.8	15.1	15.1	14.7
	Z	5.9	7.1	8.6	8.1	9.4	9.5	8.7	10.6
		102	140	254	351				
	X	25.3	24.3	21.9	22.9	22.7	22.0	22.3	20.9
44	Y	31.5	31.7	31.0	31.8	33.3	33.7	33.3	34.3
	Z	-3.6	-3.0	-3.0	-3.7	-3.8	-5.1	-6.2	-5.0
		138	119	277	311				
	X	26.2	27.3	29.6	28.4	28.6	27.5	27.1	27.1
60	Y	28.6	28.4	28.7	29.5	30.3	31.3	31.7	31.8
	Z	7.2	6.6	6.9	6.7	5.4	5.2	4.2	6.3
		139	116	187	215				
		*	*			*	*	*	*
	X	14.8	15.9	18.1	17.1	18.0	19.4	19.6	20.3
122	Y	6.4	6.0	4.9	5.9	7.1	6.8	6.6	6.7
	Z	19.7	19.2	19.7	20.1	20.3	20.7	22.1	20.0
		149	121	317	274				
							*	*	*
	X	7.1	8.0	8.5	7.6	7.2	8.3	8.4	9.2
126	Y	6.7	5.8	3.5	4.4	3.5	3.1	3.1	2.6
	Z	16.9	17.1	16.4	17.3	18.5	19.3	20.5	18.7
		134	128	41	132				
	X	5.4	5.6	7.6	7.0	7.0	6.2	6.9	5.1
132	Y	13.5	12.6	11.6	11.8	10.4	10.4	10.9	10.0
	Z	9.4	10.2	11.4	10.1	9.5	8.2	7.0	8.1
		135	121	157	82				

ASP		O	C	N	CA	CB	CG	OD1	OD2
	X	0.3	0.8	0.3	0.1	0.4	-0.7	-1.7	-0.5
141	Y	26.5	25.4	22.9	24.4	24.5	23.8	23.3	23.7
	Z	6.2	6.5	7.1	7.4	8.9	9.7	9.3	10.9
		142	118	283	352				

GLU		O	C	N	CA	CB	CG	CD	OE1	OE2
								*	*	*
	X	2.4	1.3	-0.7	0.0	-0.3	-0.0	0.3	-0.6	1.2
4	Y	12.6	12.1	12.6	12.9	12.8	14.0	15.4	15.5	16.0
	Z	23.6	23.5	22.4	23.6	25.1	26.0	25.2	24.3	25.4
		125	132	251	8	52				
	X	5.5	4.4	2.7	3.5	2.6	2.0	C.9	0.7	0.4
6	Y	11.9	11.9	10.4	10.7	10.7	9.4	9.5	10.6	8.5
	Z	19.4	19.9	21.0	19.8	18.6	18.1	17.C	16.7	16.6
		140	127	281	169	1				
							*	**	**	**
	X	8.7	7.6	5.4	6.3	6.3	7.6	7.7	8.8	6.6
8	Y	12.3	11.9	12.9	12.3	12.2	12.2	12.2	11.9	12.5
	Z	23.2	22.9	22.7	23.6	25.1	25.7	27.2	27.8	27.8
		123	140	178	176	165				
	X	21.8	21.2	19.4	19.8	18.8	18.5	17.0	16.6	16.5
18	Y	20.2	19.2	18.1	19.3	19.9	21.4	21.9	22.5	21.6
	Z	21.5	21.1	19.8	20.5	21.4	21.2	21.5	20.5	22.5
		171	111	225	147	232				
	X	24.5	24.8	24.3	25.3	26.7	27.8	29.2	30.1	29.3
26	Y	18.9	18.9	20.4	20.1	20.4	20.4	20.7	21.0	20.6
	Z	6.5	7.8	9.4	8.4	9.0	8.1	8.6	7.8	9.9
		121	132	180	182	192				
	X	17.8	17.6	18.6	17.7	16.4	15.6	15.7	14.6	16.8
38	Y	23.8	22.5	20.4	21.6	20.8	20.1	18.5	18.0	18.C
	Z	-6.3	-6.2	-7.1	-7.3	-7.9	-6.9	-7.C	-7.C	-7.0
		131	123	52	253	228				
	X	20.4	20.0	20.7	20.7	20.1	20.0	21.4	22.3	21.4
41	Y	28.2	27.1	24.9	25.8	25.3	26.3	26.4	26.0	26.9
	Z	-6.6	-6.2	-5.6	-6.7	-8.1	-9.2	-10.0	-9.3	-11.1
		146	134	179	80	34C				
	X	31.4	31.7	31.2	30.7	31.0	30.2	29.1	28.5	28.9
52	Y	18.0	17.4	17.2	16.9	15.4	14.7	13.7	13.3	13.5
	Z	2.1	1.0	-1.3	0.1	0.4	1.5	1.0	2.0	-0.1
		139	113	224	1C4	183				
	X	32.7	32.4	33.8	33.5	33.4	34.6	34.4	35.5	33.3
54	Y	22.8	21.9	20.0	21.5	22.1	21.9	22.5	22.7	22.8
	Z	2.3	1.5	0.5	0.5	-0.9	-1.9	-3.3	-3.8	-3.7
		128	133	292	184	158				
	X	29.0	29.8	31.8	31.2	32.1	33.1	32.4	32.0	32.3
59	Y	28.0	28.0	26.9	27.3	28.3	27.7	26.9	27.7	25.7
	Z	9.0	8.1	6.9	8.1	8.8	9.9	11.0	11.9	10.9
		152	106	263	306	274				
						*	*	*	*	*
	X	2.7	2.3	1.6	1.3	-0.2	-0.8	-0.9	-0.7	-1.1
83	Y	30.9	30.4	28.1	29.2	28.9	29.6	28.8	27.6	29.3
	Z	13.2	14.2	15.0	14.2	14.5	15.7	17.0	16.7	18.1
		131	138	59	258	345				

GLU		O	C	N	CA	CB	CG	CD	OE1	OE2
								**	**	**
	X	7.1	6.5	5.4	6.6	7.6	7.2	7.7	7.8	7.8
85	Y	30.9	30.3	30.8	30.6	29.6	29.4	30.6	30.1	31.7
	Z	11.7	12.6	14.9	14.1	14.9	16.4	17.3	18.5	16.9
		133	101	331	283	205				
							*	*	*	*
	X	7.2	7.1	5.5	5.6	5.3	3.9	3.5	2.3	4.5
91	Y	34.2	34.0	32.7	33.9	35.0	34.9	35.8	36.0	36.4
	Z	2.0	3.2	4.5	3.8	4.8	5.3	6.4	6.6	7.1
		126	123	293	174	163				
	X	11.0	10.5	9.2	9.1	8.2	8.8	7.8	7.0	7.9
105	Y	15.8	16.6	18.5	17.1	16.3	15.1	14.1	13.7	13.9
	Z	2.7	1.7	1.8	1.8	0.7	0.1	-0.7	0.2	-1.8
		120	113	144	163	296				
								*	*	*
	X	14.9	14.8	12.8	13.4	13.3	12.1	11.2	11.1	10.8
109	Y	12.3	13.1	14.5	13.2	12.7	11.8	11.3	10.1	12.1
	Z	5.9	4.9	4.4	4.2	2.7	2.4	3.6	3.6	4.4
		112	113	251	358	230				
								*	*	*
	X	8.9	8.8	10.1	10.1	10.5	10.8	11.9	11.8	12.9
128	Y	9.0	8.0	7.1	7.4	6.2	6.4	5.6	4.3	6.2
	Z	11.8	12.6	14.6	13.2	12.4	11.0	10.4	10.4	9.9
		132	131	188	146	61				
	X	2.2	2.8	4.7	3.6	2.7	3.2	2.3	1.3	2.6
136	Y	18.3	17.5	16.2	16.4	15.2	13.8	12.8	12.6	12.3
	Z	7.0	7.8	8.2	7.2	7.6	7.5	6.7	7.4	5.7
		147	122	307	230	293				
	X	-3.9	-3.3	-3.5	-4.0	-3.7	-3.8	-3.4	-3.2	-3.4
148	Y	33.4	32.7	30.4	31.7	32.1	33.6	33.9	35.1	33.0
	Z	-1.1	-0.3	0.3	0.7	2.2	2.5	4.0	4.2	4.9
		121	148	161	182	167				
		*	*	*	*	*	**	**	**	**
	X	-0.6	0.7	1.8	1.3	0.3	0.7	-0.4	-0.0	-1.6
152	Y	25.1	25.1	27.1	25.7	25.0	24.7	24.1	23.5	24.2
	Z	-5.6	-5.5	-4.3	-4.2	-3.4	-2.1	-1.2	-0.1	-1.5
		263	262	307	176	165				

MET		O	C	N	CA	CB	CG	SD	CE
	X	30.4	30.5	31.3	30.1	29.1	28.7	27.2	27.4
55	Y	22.5	21.4	21.3	21.5	20.5	20.8	20.1	20.0
	Z	4.4	3.7	1.3	2.2	1.6	0.2	-0.3	-2.0
		115	128	279	203	206			
	X	7.7	7.9	7.7	8.5	10.0	10.9	10.9	11.8
131	Y	13.9	12.6	11.3	12.2	11.8	12.8	14.4	13.8
	Z	11.9	12.2	14.3	13.6	13.6	12.9	13.8	15.2
		126	137	172	69	67			

LYS		O	C	N	CA	CB	CG	CD	CE	NZ
							*	*	**	**
	X	20.9	19.6	17.7	18.9	18.3	18.4	17.8	17.6	17.0
16	Y	15.3	15.1	15.1	14.4	13.0	12.1	10.7	9.9	8.6
	Z	20.4	20.4	21.9	21.5	21.3	22.5	22.3	23.5	23.3
		134	107	251	180	189	179			

LYS		O	C	N	CA	CB	CG	CD	CE	NZ
								*	*	**
	X	22.3	22.2	22.9	23.3	24.1	25.4	26.6	27.0	28.3
34	Y	14.6	14.9	16.8	15.6	14.9	15.6	14.9	15.3	14.6
	Z	-4.5	-3.2	-1.8	-2.5	-1.4	-1.0	-1.4	-2.8	-3.2
		133	111	273	250	85	179			
									*	*
	X	18.4	18.9	19.0	18.3	17.2	15.9	15.0	14.3	12.9
42	Y	30.1	28.9	26.9	28.0	27.5	27.1	28.3	28.2	28.6
	Z	-3.5	-3.7	-5.4	-4.8	-3.8	-4.4	-4.7	-6.0	-5.8
		157	110	289	280	141	135			
							*	**	**	**
	X	27.5	27.2	25.6	25.7	25.3	24.6	24.1	23.6	23.2
47	Y	25.7	26.9	27.7	27.2	28.0	29.3	30.0	31.4	32.0
	Z	-5.8	-5.4	-3.7	-5.1	-6.4	-6.1	-7.4	-7.1	-8.4
		127	142	6	184	189	179			
								*	*	**
	X	30.4	31.0	30.8	31.0	31.9	31.8	30.4	29.5	28.2
50	Y	19.3	20.2	22.8	21.7	21.8	23.2	23.4	24.2	24.4
	Z	-5.5	-4.8	-4.4	-5.3	-6.5	-7.2	-7.8	-6.9	-7.5
		136	92	309	293	97	179			
								**	**	**
	X	32.3	32.3	30.9	31.2	32.1	31.5	32.3	31.9	32.7
56	Y	21.4	21.0	20.2	20.0	18.7	17.4	16.2	15.0	13.8
	Z	7.1	5.9	4.1	5.5	5.6	5.1	5.4	4.6	4.9
		143	128	325	184	189	179			
									*	*
	X	25.0	26.1	27.4	27.3	28.3	28.0	29.1	30.5	31.6
62	Y	25.7	26.1	25.6	25.2	24.9	23.6	23.3	23.5	23.0
	Z	10.2	10.0	8.1	9.5	10.6	11.3	12.4	11.8	12.7
		101	138	236	176	307	187			
							*	*	**	**
	X	23.2	24.2	26.5	25.6	26.4	27.5	27.7	28.9	29.1
63	Y	28.2	28.2	27.4	28.4	29.7	30.2	31.7	32.2	33.6
	Z	11.0	10.2	10.2	10.7	10.2	11.1	11.0	11.7	11.5
		129	150	259	208	189	179			
	X	6.2	6.7	8.4	8.1	9.2	10.6	11.7	13.1	14.1
77	Y	22.5	22.5	21.6	21.9	22.7	22.2	23.0	22.6	23.6
	Z	22.4	21.3	19.6	21.0	21.7	21.4	22.1	21.8	22.2
		160	139	314	177	181	167			
								**	**	**
	X	2.5	3.7	6.2	4.8	5.0	5.6	5.5	6.3	6.3
78	Y	23.1	22.8	23.1	23.7	25.1	26.2	27.5	28.6	29.9
	Z	20.2	20.1	20.2	20.3	19.5	20.4	19.6	20.3	19.5
		175	97	262	184	189	179			
	X	0.8	1.9	4.1	3.1	2.9	4.2	3.9	3.0	2.6
79	Y	20.4	20.7	21.5	20.4	19.8	19.5	19.3	18.0	17.9
	Z	19.2	18.9	19.9	19.7	21.1	21.9	23.4	23.6	25.0
		225	223	300	161	66	186			
								*	**	**
	X	4.7	4.5	3.9	3.4	2.2	1.6	0.1	-0.6	-2.0
87	Y	32.5	32.7	31.0	32.0	32.8	33.8	33.8	34.7	34.6
	Z	7.4	8.6	10.2	9.4	10.0	9.0	9.1	8.1	8.2
		107	133	172	138	189	179			
								*	*	*
	X	12.7	12.2	10.0	11.4	12.2	11.9	12.8	12.3	13.2
96	Y	34.1	34.2	35.1	35.5	36.1	37.6	38.2	39.6	40.2
	Z	-3.0	-1.9	-1.2	-1.5	-0.3	-0.2	0.9	1.3	2.3
		115	42	286	184	189	179			

LYS		O	C	N	CA	CB	CG	CD	CE	NZ
							**	**	**	**
98	X	11.1	10.9	11.3	10.8	11.1	10.7	10.9	10.7	11.0
	Y	28.1	28.8	31.1	30.3	30.5	31.9	32.0	33.4	33.5
	Z	-4.6	-3.6	-2.6	-3.7	-5.2	-5.7	-7.3	-7.8	-9.2
		218	242	303	184	189	179			
102	X	10.2	9.6	7.5	8.3	8.4	8.6	7.9	8.1	7.1
	Y	18.3	19.4	20.6	19.4	19.3	17.9	16.9	15.5	14.5
	Z	-1.4	-1.8	-2.2	-2.6	-4.1	-4.7	-3.9	-4.4	-4.0
		146	125	191	398	182	203			
								*	*	**
133	X	3.3	3.8	4.8	3.5	2.7	2.0	1.4	1.0	0.4
	Y	15.1	14.2	12.2	12.8	12.1	10.8	10.1	8.6	8.0
	Z	10.8	11.6	11.1	11.4	12.5	12.1	13.3	13.0	14.1
		118	125	279	172	195	179			
									*	*
140	X	-0.6	-0.1	1.6	0.2	-0.6	-2.1	-2.8	-3.9	-4.7
	Y	23.0	22.4	20.5	20.9	19.9	19.9	18.6	18.7	19.9
	Z	5.0	6.0	6.0	5.9	6.7	6.4	6.6	7.6	7.3
		119	134	179	152	126	36			
							*	*	**	**
145	X	-0.6	-0.3	-1.1	-0.7	0.4	0.6	-0.7	-0.5	-1.7
	Y	31.1	30.0	28.2	29.6	29.8	31.2	32.0	33.5	34.3
	Z	1.9	2.4	3.9	3.8	4.8	5.3	5.2	5.4	5.4
		142	126	194	390	189	179			
								*	*	*
147	X	-3.1	-3.1	-1.3	-2.6	-3.5	-4.7	-5.4	-6.5	-7.2
	Y	30.9	30.1	28.6	28.7	27.5	27.3	26.0	25.6	24.3
	Z	-1.9	-0.9	-0.4	-1.1	-0.6	-1.4	-1.0	-1.9	-1.5
		108	141	184	171	189	179			

ARG		O	C	N	CA	CB	CG	CD	NE	CZ	NH2	NH1
								*	*	*	*	*
31	X	20.8	20.9	22.8	21.8	21.1	20.2	21.0	20.3	20.8	22.0	20.1
	Y	15.0	15.6	16.0	15.0	14.1	13.0	12.4	11.2	9.9	9.7	8.9
	Z	0.4	1.5	3.1	2.6	3.6	3.0	1.8	1.3	1.5	2.1	1.1
		125	129	179	42	168	252	359				
45	X	27.1	26.0	24.2	25.3	25.4	24.4	24.1	24.5	23.7	22.4	24.1
	Y	30.2	30.3	31.8	31.7	32.3	31.7	32.8	32.3	32.2	32.5	31.8
	Z	-0.1	-0.7	-1.6	-0.7	0.7	1.7	2.7	4.1	5.1	5.0	6.3
		171	113	58	207	242	237	5				
118	X	22.8	22.1	22.0	22.6	22.6	23.9	23.9	24.3	25.3	26.0	25.7
	Y	9.1	8.8	8.9	9.3	10.8	11.3	12.9	13.4	14.3	14.7	14.7
	Z	16.3	15.3	12.6	13.9	13.7	13.1	13.0	14.3	14.5	13.4	15.7
		171	95	254	184	80	124	2				
139	X	2.4	2.6	4.3	4.0	4.3	5.7	5.9	6.7	6.4	5.3	7.2
	Y	22.1	21.1	20.6	20.6	19.2	19.0	17.8	16.8	15.4	15.0	14.6
	Z	4.6	5.4	7.1	5.7	4.9	4.6	3.6	4.3	4.1	3.4	4.6
		125	124	274	191	242	143	358				

PHE		O	C	N	CA	CB	CG	CD1	CE1	CZ	CE2	CD2
33	X	22.4	22.5	21.2	22.0	23.2	22.8	21.5	21.1	22.0	23.2	23.7
	Y	18.1	17.9	18.5	19.0	19.8	21.2	21.6	22.9	23.6	23.2	22.0
	Z	-3.5	-2.3	-0.2	-1.3	-0.9	-0.3	-0.5	0.0	0.7	0.9	0.4
		135	127	286	350							

PHE		O	C	N	CA	CB	CG	CD1	CE1	CZ	CE2	CD2
43	X	22.9	21.9	19.9	20.7	21.2	20.4	20.7	20.0	19.0	18.6	19.3
	Y	29.0	29.7	28.4	29.2	27.9	27.7	28.4	28.2	27.3	26.7	26.8
	Z	-3.2	-2.8	-3.1	-2.1	-1.3	-0.1	1.0	2.3	2.3	1.2	-0.0
		288	82	246	281							
46	X	27.1	26.2	25.4	25.9	25.3	25.0	26.0	25.7	24.5	23.5	23.7
	Y	26.3	27.3	29.3	28.0	26.9	27.5	28.1	28.6	28.5	28.0	27.4
	Z	-2.8	-2.7	-1.3	-1.4	-0.4	0.9	1.6	2.9	3.5	2.7	1.5
		154	60	315	306							
106	X	13.7	13.3	11.2	12.6	13.1	12.8	13.2	12.9	12.2	11.9	12.1
	Y	15.7	16.7	17.0	16.6	17.5	17.0	17.7	17.3	16.1	15.4	15.8
	Z	2.4	1.8	0.7	0.4	-0.7	-2.1	-3.1	-4.5	-4.7	-3.6	-2.3
		106	138	274	179							
123	X	14.4	14.4	16.1	15.0	15.8	14.9	14.7	13.9	13.2	13.4	14.3
	Y	4.8	4.7	5.7	5.8	6.7	7.8	7.7	8.7	9.7	9.7	8.8
	Z	14.7	16.0	17.9	16.9	16.0	15.3	14.0	13.3	14.1	15.4	16.1
		241	63	203	248							
138	X	3.9	4.2	3.8	4.5	5.9	7.0	7.5	8.6	9.0	8.4	7.4
	Y	22.8	21.7	20.2	21.4	21.1	21.9	21.6	22.3	23.5	23.9	23.1
	Z	7.4	7.8	9.6	9.3	9.9	9.3	8.1	7.5	8.1	9.2	9.9
		137	121	143	277							

TYR		O	C	N	CA	CB	CG	CD1	CE1	CZ	CE2	CD2	OH
103	X	12.0	11.2	10.1	11.4	11.7	12.2	13.4	13.8	13.2	12.0	11.6	13.6
	Y	19.7	20.4	20.6	20.8	22.1	22.0	22.7	22.6	21.9	21.1	21.2	21.8
	Z	1.3	0.7	-1.5	-0.8	-1.4	-2.9	-3.3	-4.6	-5.5	-5.1	-3.8	-6.8
		133	111	264	224								
146	X	-0.5	-0.4	0.4	0.9	1.6	3.1	3.9	5.3	5.9	5.1	3.7	7.2
	Y	30.6	29.5	29.1	29.2	27.9	28.0	27.2	27.3	27.9	28.7	28.6	27.9
	Z	-1.0	-0.4	1.8	0.4	-0.1	0.1	-0.8	-0.7	0.3	1.2	1.1	0.4
		119	123	263	211								
151	X	-0.2*	1.0	0.6	1.7	3.0	3.4	4.5	4.8	4.0	2.9	2.6	4.3
	Y	28.1	28.2	30.4	29.5	29.9	31.4	31.8	33.2	34.1	33.7	32.3	35.4
	Z	-3.9	-4.1	-3.9	-4.2	-3.4	-3.5	-4.3	-4.3	-3.6	-2.8	-2.8	-3.6
		359	62	310	248								

HIS		O	C	N	CA	CB	CG	CD2	NE2	CE1	ND1
12	X	15.1	13.9	11.6	12.8	12.3	13.4	14.0	15.0	14.9	14.0
	Y	13.2	13.0	13.7	13.1	11.7	10.8	9.7	9.2	9.9	10.9
	Z	22.6	22.2	22.6	23.2	23.5	24.0	23.4	24.3	25.4	25.3
		145	114	169	73						
24	X	22.1	22.4	24.1	22.8	21.9	21.7	20.8	21.2	22.3	22.6
	Y	17.9	18.4	18.0	17.6	17.8	16.5	15.5	14.7	15.1	16.2
	Z	10.1	11.2	13.1	12.4	13.6	14.5	14.5	15.5	16.0	15.5
		128	131	257	84						
36	X	18.8	18.9	18.9	18.1	16.8	16.1	16.0	15.2	14.9	15.5
	Y	19.4	18.3	16.2	17.1	17.2	16.0	15.3	14.2	14.1	15.2
	Z	-4.3	-4.9	-3.7	-4.6	-3.7	-3.4	-2.2	-2.4	-3.7	-4.3
		264	49	293	290						
48	X	31.0	30.0	28.0	29.5	30.2	30.4	31.4	31.1	30.0	29.5
	Y	25.8	26.4	27.9	27.7	28.9	29.0	28.4	28.7	29.4	29.6
	Z	-5.5	-4.9	-5.1	-5.3	-4.8	-3.3	-2.5	-1.2	-1.2	-2.4
		145	139	74	269						
64	X	20.8	22.0	24.1	22.9	22.6	21.2	20.3	19.2	19.4	20.6
	Y	26.8	26.6	27.9	27.7	27.8	28.1	27.3	28.1	29.2	29.3
	Z	9.0	8.8	8.9	8.2	6.7	6.3	5.6	5.5	6.2	6.6
		129	129	199	64						
81	X	1.8	1.3	0.6	0.1	-0.5	-1.5	-2.1	-3.0	-3.1	-2.2
	Y	26.7	25.9	23.7	25.1	25.5	24.7	24.6	23.6	23.1	23.7
	Z	18.7	17.9	18.0	18.2	19.6	20.2	21.4	21.4	20.2	19.4
		235	84	321	343						

HIS		O	C	N	CA	CB	CG	CD2	NE2	CE1	ND1
	X	3.9	2.8	1.8	2.9	3.7	2.8	1.6	1.3	2.2	3.1
82	Y	28.0	27.6	25.7	26.4	25.5	24.6	24.7	23.5	22.6	23.2
	Z	14.6	15.2	16.7	16.1	15.0	14.3	13.6	13.1	13.4	14.1
		188	83	49	231						
	X	9.4	8.8	9.4	9.6	9.7	10.9	12.1	12.8	12.0	10.9
93	Y	30.5	30.9	31.8	30.7	29.4	29.3	28.8	28.9	29.4	29.7
	Z	−0.6	0.5	2.8	1.8	2.5	3.4	3.2	4.4	5.3	4.8
		141	120	279	90						
	X	13.5	12.6	12.3	13.0	13.4	13.6	13.2	13.7	14.4	14.4
97	Y	30.4	31.1	33.3	32.0	31.3	32.3	32.2	33.3	34.0	33.4
	Z	−2.9	−2.3	−1.0	−1.1	0.2	1.3	2.6	3.2	2.4	1.2
		145	126	318	314						
	X	18.9	18.4	16.3	17.4	18.2	18.9	20.3	20.4	19.2	18.3
113	Y	9.2	10.0	10.5	9.6	10.1	9.0	8.8	7.6	7.1	7.9
	Z	9.1	8.4	7.7	7.3	6.1	5.4	5.2	4.5	4.3	4.8
		120	121	234	64						
							*	*	*	*	*
	X	20.1	19.6	17.8	18.0	17.2	17.4	18.4	18.1	17.0	16.5
116	Y	6.8	7.2	8.8	7.4	6.5	5.0	4.2	2.9	2.9	4.2
	Z	12.8	11.8	11.7	11.7	10.8	11.0	10.6	11.0	11.7	11.7
		111	116	178	103						
		*									
	X	18.6	19.8	21.0	20.4	19.4	20.0	20.5	20.9	20.7	20.1
119	Y	6.0	6.2	8.1	7.5	8.5	9.8	10.2	11.5	11.9	10.9
	Z	16.0	16.2	15.3	16.5	17.1	17.6	18.7	18.6	17.4	16.7
		259	36	297	88						

TRY		O	C	N	CA	CB	CG	CD1	NE	CE1	CZ1	CH	CZ2	CE2	CD2
	X	6.9	5.7	3.8	4.5	3.7	2.8	1.5	1.1	2.2	2.3	3.6	4.7	4.6	3.3
7	Y	14.0	13.7	13.0	14.2	15.4	16.1	16.2	16.9	17.3	18.1	18.4	17.9	17.1	16.8
	Z	21.4	21.7	20.6	20.8	21.3	20.2	20.1	18.9	18.3	17.2	16.8	17.5	18.6	19.1
		131	118	283	117										
	X	16.8	15.7	14.2	14.6	13.4	13.9	14.0	14.4	14.5	14.9	15.0	14.6	14.2	14.2
14	Y	17.0	16.7	15.0	16.4	17.4	18.8	19.9	21.0	20.5	21.3	20.5	19.2	18.4	19.2
	Z	20.1	20.5	19.7	19.6	19.2	18.7	19.4	18.5	17.4	16.2	15.1	15.1	16.2	17.4
		116	129	179	271										

PRO		O	C	N	CA	CB	CG	CD
	X	20.4	19.8	19.7	20.5	21.5	21.2	19.9
37	Y	21.6	20.5	13.1	19.2	18.5	17.0	16.9
	Z	−6.6	−6.8	−6.0	−6.5	−7.4	−7.5	−6.6
		148	130					
	X	7.9	7.4	5.3	6.5	7.0	6.2	5.2
88	Y	33.9	33.4	33.5	34.2	34.9	34.5	33.5
	Z	6.9	7.9	9.4	8.9	10.1	11.4	10.8
		142	138					
	X	9.1	8.1	9.5	8.4	8.6	9.9	10.5
100	Y	22.6	23.2	25.2	24.2	23.7	24.3	25.1
	Z	−1.6	−2.1	−3.0	−3.2	−4.6	−5.2	−4.1
		317	117					
						*	*	*
	X	18.3	19.4	20.7	20.4	21.7	22.8	22.1
120	Y	2.8	3.2	5.2	3.8	3.1	4.0	5.3
	Z	16.5	16.8	16.1	15.8	15.8	16.1	16.5
		119	122					

HAEM		FE	H O		CH1	CH2	CH3	CH4
	X	14.0	16.7		16.3	14.0	13.8	16.1
	Y	28.1	27.1		31.3	28.2	24.8	27.9
	Z	4.8	4.8		5.1	8.2	4.7	1.6

HAEM		N	C1	C2	C3	C4	CM	CA	CB	CG	O1	O2
	X	15.1	15.7	15.5	14.9	14.6	14.5	16.0	15.0	14.1	14.4	12.9
1	Y	29.4	30.7	31.5	30.5	29.3	30.7	32.8	34.0	33.7	33.9	33.2
	Z	6.3	6.3	7.5	8.4	7.7	9.8	7.7	8.1	7.0	5.8	7.3
	X	14.1	13.6	13.0	13.1	13.7	12.6	12.4	12.8			
2	Y	26.8	25.4	24.8	25.8	27.0	25.8	23.5	22.5			
	Z	6.1	5.9	7.1	8.1	7.5	9.5	7.1	7.6			
	X	14.9	15.4	15.2	14.5	14.4	14.0	15.6	14.6			
3	Y	26.6	26.7	25.5	24.6	25.3	23.2	25.4	24.4			
	Z	3.4	2.1	1.3	2.3	3.5	2.0	0.0	-0.7			
	X	16.0	16.4	17.1	17.0	16.3	17.5	17.6	18.8	19.7	20.4	19.6
4	Y	29.3	30.6	31.2	30.2	29.1	30.3	32.6	31.6	32.8	33.3	33.3
	Z	3.6	3.8	2.7	1.7	2.3	0.3	2.7	2.6	3.0	2.1	4.1

REFERENCES

Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. T. C., Phillips, D. C. and Sarma, V. R. (1965) *Nature, Lond.* **206,** 757

Blow, D. M. and Crick, F. H. C. (1959) *Acta crystallogr.* **12,** 794

Bodo, G., Dintzis, H. M., Kendrew, J. C. and Wyckoff, H. W. (1959) *Proc. R. Soc.* **A253,** 70

Diamond, R. (1966) *Acta crystallogr.* **21,** 253

Dickerson, R. E., Kendrew, J. C. and Strandberg, B. E. (1961) *Acta crystallogr.* **14,** 1188

Dunnill, P. (1965) *Sci. Prog., Lond.* **53,** 609

Edsall, J. T., Flory, P. J., Kendrew, J. C., Liquori, A. M., Némethy, G., Ramachandran, G. N. and Scheraga, H. A. (1966) *J. molec. Biol.* **15,** 399

Edmundson, A. B. (1965) *Nature, Lond.* **205,** 883

Green, D. W., Ingram, V. M. and Perutz, M. F. (1954) *Proc. R. Soc.* **A225,** 287

Huggins, M. L. (1943) *Chem. Rev.* **32,** 195

Kauzmann, W. (1959) *Adv. Protein Chem.* **14,** 1

Kendrew, J. C. (1962) *Brookhaven Symp. Biol.* **15,** 216

— and Parrish, R. G. (1956) *Proc. R. Soc.* **A238,** 305

— Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C. and Shore, V. C. (1960) *Nature, Lond.* **185,** 422

— Watson, H. C., Strandberg, B. E., Dickerson, R. E., Phillips, D. C. and Shore, V. C. (1962) *Nature, Lond.* **190,** 666

Lipson, H. and Cochran, W. (1966) *The Determination of Crystal Structures*, London, Bell

Nobbs, C. L., Kendrew, J. C. and Watson, H. C. (1966) *Nature, Lond.* **209,** 339

Pauling, L. and Corey, R. B. (1951) *Proc. natn. Acad. Sci. U.S.A.* **37,** 729

Perutz, M. F., Kendrew, J. C. and Watson, H. C. (1965) *J. molec. Biol.* **13,** 669

Ramachandran, G. N., Ramakrishnan, C. and Sasisekharan, V. (1963) *J. molec. Biol.* **7,** 955; (1965) *Biophys. J.* **5,** 909

Schellman, J. A. and C. (1964) in *The Proteins*, 2nd Edn, Vol. 2, p. 1, New York, Academic Press

Watson, H. C., Kendrew, J. C. *et al.* (1964) *Acta crystallogr.* **16A,** 81

— and Nobbs, C. L. (1968) in *Biochemie des Sauerstoffs*, p. 37, Springer, Berlin