Acta Cryst. (1954). 7, 353

The Structure of Horse Hemoglobin in the Light of the Intensity Map of the Horse Methemoglobin Crystal*

By Dorothy Wrinch

Department of Physics, Smith College, Northampton, Mass., U.S.A.

(*Received* 22 *May* 1953 *and in revised form* 4 *November* 1953)

The results of a preliminary examination of Perutz's intensity map for normal wet monoclinic horse methemoglobin are described. A new interpretation is proposed in which the 8–12 Å and 4–5 Å ranges of pronounced scattering are taken to be symptomatic of the first and second regions of relatively high density in the intensity maps of individual structures within the crystal. It is seen that such structures would presumably be more or less equi-dimensional and have diameters of the order of 11–15 Å and that they may tentatively be identified with molecular skeletons within the horse hemoglobin entity, made up of polymerised amino acid backbones. Reasons are given for anticipating that these structures, of which there would be several, possibly 6, pairs per 66,700 horse hemoglobin entity, would be cage-like, with interiors of relatively low density.

1. Introduction

The possibility that some general information regarding the nature of individual structures in protein crystals can be obtained by the direct study of the X-ray data, without the need of presuppositions, has long been under study (Wrinch, 1946, 1952*a*, 1953*a*). On this occasion we test this viewpoint directly on the very extensive intensity data obtained by Perutz (1949) for the normal wet monoclinic horse methemoglobin crystal†. So far our examination of the data, studied in the form of the published vector maps (Wrinch, 1953*b*, *c*), has been principally concerned with disproving, for this particular globular protein, the traditional type of structure comprising bundles of polypeptide chains, which has apparently been adopted for globular proteins in general. In this paper we look for positive evidence regarding the nature of the horse hemoglobin structure in the three-dimensional intensity map which comprises entries at 62,700 reciprocal lattice points out to spacings of 2·8 Å.

As a first step in the preliminary analysis of this unprecedented wealth of data for a protein crystal, we may study the mean intensity curve as a function of distance from the origin in reciprocal space, which Perutz has derived from his intensity map. In this curve which is given in Fig. 24 of Perutz's 1949 paper, we see the two ranges of pronounced scattering, with spacings say 8–12 Å and 4–5 Å, which are often regarded as typical of globular proteins as a class; the maximum in the first range is at about 1/10 Å and in the second at about 1/4½ Å. The question then arises as to the significance of these two ranges in the case of this particular protein crystal.

To answer this question we begin by taking the (*h*0*l*), the (0*kl*) and the (*hk*0) intensities lying on central planes normal to the axes of the crystal, and draw the

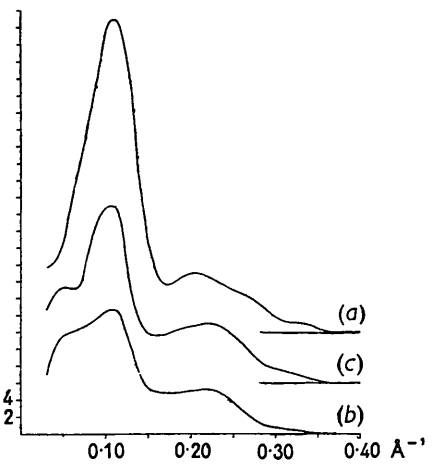

Fig. 1. The (0*kl*), (*h*0*l*), (*hk*0) intensities for the 'normal wet' monoclinic horse methemoglobin crystal, each circularly smoothed about the origin in reciprocal space. Each curve is drawn through points at 0·03 Å^{-1}, 0·05 Å^{-1}, ···, 0·39 Å^{-1} obtained by averaging the given $|F|^2$ values over a range of 0·02 Å^{-1} on either side of the point. (The intensities are taken from Perutz's unpublished three-dimensional intensity map and are used with his permission.)

curves obtained by circularly smoothing each of these three sets of intensities about the origin (Wrinch, 1952*c*). The result† is seen in Fig. 1. In each of the three curves, we remark the presence of a maximum

* This work is supported by the Office of Naval Research, U.S. Navy.

† The author is very much indebted to Dr Perutz for permitting her to see these unpublished data.

† Since this preliminary examination concerns only the intensities for the 12–8Å and 5–4Å range of spacings, we do not now discuss the third maxima clearly to be seen in the *a*, *b* and *c* curves. That this feature may have some relationship to hydrogen-bridged arrays of water molecules in the crystal seems very likely.

in the 8–12 Å range of spacings and in the 4–5 Å range, actually at $r^* = 0{\cdot}11$ Å$^{-1} = 1/9{\cdot}1$ Å and at $r^* = 0{\cdot}21$ Å$^{-1} = 1/4{\cdot}8$ Å. We have also pursued the question of the positions in space of the intensities responsible for the two maxima one step further. In Fig. 2 we record the result of circularly smoothing about

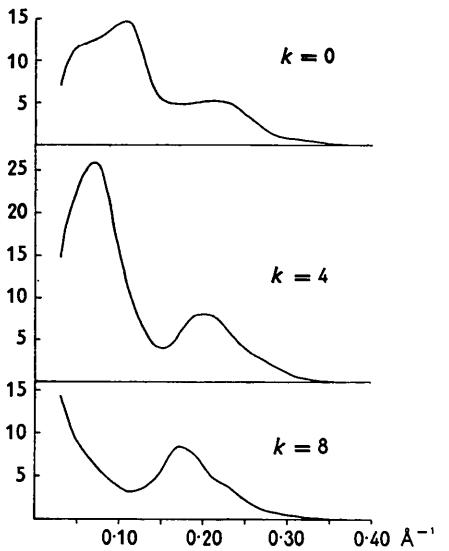

Fig. 2. The ($h0l$), ($h4l$) and ($h8l$) observed intensities for the crystal, each circularly smoothed about the dyad axis. The curves are obtained as in the previous case. (The intensities are taken from Perutz's unpublished three-dimensional intensity map and are used with his permission.)

the k-axis the intensities in the planes $k = 4$ and $k = 8$, repeating the curve for the intensities in the plane $k = 0$ for comparison. For the intensities in the plane $k = 8$, distances from the origin are $8/63{\cdot}2$ Å $= 1/7{\cdot}9$ Å and up, and we look only for the presence, or absence, of the minor maximum. We remark the maximum at $\varrho^* = 0{\cdot}17$ Å^{-1}, for which, since $8/b = 0{\cdot}13$ Å^{-1}, $r^* = 0{\cdot}21$ Å$^{-1} = 1/4{\cdot}7$ Å. For the intensities in the plane $k = 4$, we remark maxima at $\varrho^* = 0{\cdot}07$ Å^{-1} and at $\varrho^* = 0{\cdot}20$ Å^{-1}. For the former $r^* = 0{\cdot}094$ Å^{-1} $= 1/10{\cdot}6$ Å; for the latter $r^* = 0{\cdot}21$ Å$^{-1} = 1/4{\cdot}8$ Å.

Already we have a vivid picture of the spread of specially large intensities in the 1/12 Å to 1/8 Å and the 1/5 Å to 1/4 Å spherical shells about mean values of (say) $r^* = 0.11$ Å^{-1} and $0{\cdot}21$ Å^{-1}, and it is worth while to press it to its logical conclusion. An examination of the incidence of specially large intensities within the two spherical shells, for all the relevant sections of the intensity map normal to the dyad axis, shows that such intensities occur for each set on every one of the relevant sections. The situation may be summed up by saying that, for each set of intensities, the specially large values are widely spread through the spherical shell. Certainly there is no one direction from the origin nor is there one central plane in which specially large intensities in either shell lie. In other words, neither distribution has a uniaxial or a biaxial character: both are essentially three-dimensional. From this analysis of the main regions of scattering in the mean intensity curve for hemoglobin, important consequences follow: the structures in the hemoglobin crystal also are not basically uniaxial and are not basically biaxial, but, on the contrary, must be essentially three-dimensional in nature.

2. Indications concerning the structures in the crystal

In order to interpret these findings, we suggest that both ranges of specially large intensities are essentially intra-structural in origin, being symptomatic of the

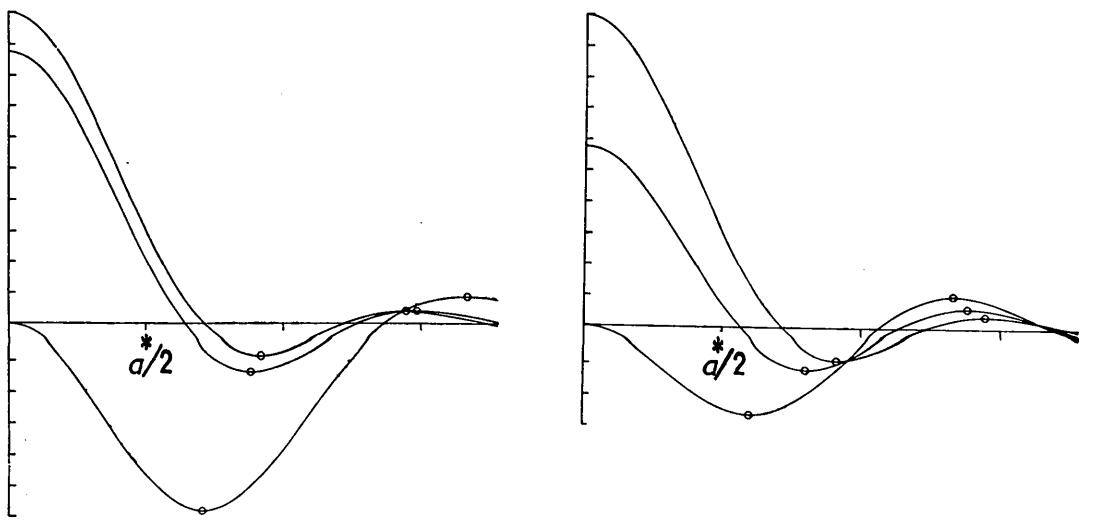

Fig. 3. The transforms of a spherical shell of radius a and unit density from which a concentric shell of radius b and various densities k have been deleted, in the case of the '$\frac{1}{2}$ shell' (on the left), $b = \frac{1}{2}a$; in the case of the '$\frac{3}{4}$ shell' (on the right), $b = \frac{3}{4}a$. In each case the transform for $k = 0$ (the solid sphere of unit density) is shown. The next curve in each case is for the empty shell ($k = 1$). The last curve shows the transform in the case when $k = (a/b)^3$ and the total weight is zero. The positions of the first and second turning points of each transform are marked.

first and second high-density regions in the individual intensity maps of more or less equi-dimensional structures in the crystal. To implement this idea we have looked for a simple type of spherically symmetric structure for which the transform $T(r^*)$ has first and second turning points, $r_1^* \sim 0{\cdot}11$ Å^{-1}, $r_2^* \sim 0{\cdot}21$ Å^{-1}, giving the ratio $r_2^*/r_1^* \sim 1{\cdot}91$, with a reasonably large value of the ratio $|T_2/T_1|$ (Wrinch, 1946).

From a forthcoming report, we take the results for a 'b/a shell' in which, from a spherical volume of radius a and unit density, a central sphere of radius b and density s is subtracted. So long as the density in the interior remains positive, it proves impossible to attain the ratio 1·91 for any shell, a result of interest in view of the significance of negative densities in hydrated protein crystals (Wrinch, 1939, 1950). The situation is readily understood if we scan the results for a few shells as s increases from zero to the value for which the weight of the shell is reduced to zero. Accordingly for three cases, the $\frac{1}{2}$ and $\frac{3}{4}$ shells and the spherical surface, we give the range of values of the two ratios r_2^*/r_1^* and $|T_2/T_1|$ and record, in addition, the range of values of the outside diameter of the shell, as $2a_1$ when r_1^* is actually 0·11 Å^{-1} and as $2a_2$ when r_2^* is actually 0·21 Å^{-1} (Table 1).

Table 1

	$\frac{1}{2}$ shell	$\frac{3}{4}$ shell	Spherical surface
ar_1^*	0·883–0·706	0·806–0·606	0·715–0·531
ar_2^*	1·483–1·667	1·383–1·333	1·229–1·161
r_2^*/r_1^*	1·67–2·36	1·715–2·20	1·72–2·19
$\|T_2/T_1\|$	0·28–0·15	0·51–0·39	0·59–0·46
$2a_1$ (Å)	16·1–12·8	14·7–11·0	13·0–9·7
$2a_2$ (Å)	14·1–15·8	13·2–12·7	11·7–11·0

Looking at these results and at Fig. 3, we see that larger values of the ratio r_2^*/r_1^* can be obtained by taking thicker shells, with sufficiently negative interiors. But, on the other hand, we see that to maintain relatively large values of $|T_2/T_1|$, the shell has to be relatively thin.

3. A hypothesis regarding the structure of horse hemoglobin

A striking result of this study of very simple spherically symmetric structures and their transforms is the relatively small size of the structures which yield intensity maxima at about the required distances from the origin. If we consider what structures in the horse hemoglobin entity are likely to be responsible for the main contributions of the protein component to the intensity map of the crystal, the first hypothesis would be that they are the rigid skeletons of the individual molecules, i.e. the N–C_α–C amino acid backbones interlocked into characteristic spatial patterns. Visualizing shells made up of such monomers, we should require a certain minimum thickness (corresponding presumably to a polymer one amino acid backbone thick). Hence we may perhaps take the dimensions found for the $\frac{3}{4}$ shell as suggestive and picture a thin shell of diameter say 11–15 Å.

But, corresponding to one horse hemoglobin entity (with molecular weight about 66,700) and its share of water, the volume in the crystal under study amounts to 174,000 Å^3; of this volume 52·4% is water: the volume left for the protein entity is then 83,000 Å^3 (Boyes-Watson, Davidson & Perutz, 1947). This figure corresponds to the volume of a sphere of diameter 54·1 Å. How then are we to reconcile with these facts molecular skeletons of diameter 15 Å or less or, allowing up to (say) 10 Å for the emerging R-groups, a diameter no larger than (say) 25 Å for the actual molecules? Evidently by postulating that to each of these 66,700 protein entities there corresponds a considerable number of actual molecules.

It is already known from physico-chemical investigations that this 66,700 protein entity is not a single chemical entity, i.e. an entity in which every atom is covalently bonded to every other either directly or by means of covalently bonded atoms as intermediates. How otherwise can we interpret the fact that by means of various mild treatments, including even dilution, the 66,700 entities in solution can be dissociated into what the physical chemist evaluates as 'halves'? Interestingly enough, the space group of the crystal gives the 66,700 entity symmetry C_2 and there can then be little doubt that the dissociation in question is actually into pairs of identical sub-units. Thus we may proceed at once on the hypothesis that the 66,700 protein entity corresponds to two anorthic 'half-entities' A, which may be written A^+ and A^- to indicate that they are mutually oriented to form a diad about the diad axis of the crystal. To each A, there is then allocated a volume of 41,500 Å^3, corresponding (in volume) to one sphere of diameter 43·0 Å, to two spheres of diameter 34·1 Å to three of diameter 29.8 Å and to 4, 5, 6, 7, ... of diameters 27·1 Å, 25·1 Å, 23·6 Å, 22·5 Å, ... respectively. The possibility of as many as six separate molecular skeletons in each A is thus suggested, with an average of 290/6, say about 48, amino acid backbones per skeleton, and one heme to three skeletons, if the number is actually 6.

According to this view, the horse hemoglobin entity would be a dimer of A particles; each A would comprise an array of n (possibly 6) skeletons, each carrying its own characteristically arranged R-groups rooted in its C_α atoms. We may call the n points at which the skeletons are placed the 'molecular pattern' of the A particle: the molecular pattern of the '66,700' horse hemoglobin entity would then be the diad of this pattern of n points, corresponding to the (A^+, A^-) diad, a network of $2n$ points. We picture arrays of such networks of structures at cell vertices and C face-centers, spreading through the monoclinic unit cells. According to this analysis, contributions from the protein entity to the structure factor of the crystal would come from a three-dimensional array of the,

globulite molecules. Thus the structure factor—in the case of the simplified picture of identical structures in parallel displacements—would be the product of two terms of basically different types. On the one hand would be the structure factor of a single individual small structure; on the other we would have the structure factor of the molecular pattern of the entity, i.e. of the set of $2n$ points. We started out with the idea that the existence of the 8–12 Å and 4–5 Å ranges of pronounced scattering is, in some sense, a direct manifestation of the structural type of the molecules in the crystal. We may continue to assume this, having shown that there are simple (spherically symmetric) structures which can account qualitatively for these features. But now we see a ready interpretation of the irregularity and of the intermittent character of the distributions of specially large intensities in the two spherical shells. This we may now attribute to the other component in the structure factor, the structure factor of the set of $2n$ points. With the present picture, an extremely interesting situation has developed. The fact that the pronounced scattering is in the two particular ranges is due, we presume, to the individual structures having high density regions in their intensity distributions within these ranges. The general irregularity within these ranges we now attribute to the intensity distribution for some three-dimensional network of points. Then the pregnant observation may be made that, despite the different nature of the second component of the intensity distribution, the main features of the intensity distribution characteristic of the individual skeletons are still detectable; in other words, the first component of the intensity distribution still makes its presence manifest, even though the second component may obscure its details.

It is to be emphasized that, throughout this discussion, we have been attempting only a preliminary interpretation of the facts and have attempted this only in terms of a very simple spherically symmetric structure. It is also to be emphasized that there are no indications, in the facts already discussed, that the small structures are in fact spherically symmetric, though, with a sufficiently 'irregular' network of points in the molecular pattern, such spherical symmetry, if present, might be difficult to detect. On the other hand, the fact that it is now possible to entertain the possibility that there are, in the protein particle, small structures whose individual intensity maps have first maxima in the first range and second maxima in the second leads to two general suggestions. Thus, this fact may mean that these structures are essentially globulite or roughly equi-dimensional in their geometry, with their gross structure not differing greatly in dimensions in various directions. It need hardly be emphasized that there is no question at this time of saying anything about the undoubtedly highly specific, intricate and precise atomic patterns of the small structures, and that we are using the term 'gross structure' in the sense in which we could say, for example, that the gross structure of $(PW_{12}O_{40})^{3-}$, the 12-tungstophosphate anion (Keggin, 1934; Bradley & Illingworth, 1936), is globulite in character while that of $(P_2W_{18}O_{62})^{6-}$, the 18-tungstophosphate anion (Dawson, 1953), with its markedly oblate shape, is not. Further, it may be suggested that these globulite structures are distributions through shells and are not, like both the tungstophosphate anions just mentioned, volume distributions.

Actually, a study of the transforms of various globulite types of structures has already been given, in anticipation of their present usefulness (Wrinch, 1946, pp. 72–77, 1952*a*, 1953*a*). In these examples of transforms of volume, shell and surface distributions of cube, octahedral and tetrahedral types, and of one 216 point-atom distribution on a thin polyhedral shell (Wrinch, 1952*b*), it is interesting to study the morphologies of the first and second high-density regions. All the indications are, as in the case of the spherically symmetric distributions, that the distributions must form relatively thin shells and furthermore that they must have interiors of relatively low density. It is premature at this time to attempt any general statement as to how low the densities have to be or, indeed, as to the thickness the shells must have, for several reasons. Thus no attempt has as yet been made to consider contributions to the intensity map of the water clusters which in all possibility exist in the crystal. Further it is readily seen that even the relation between the intensity map of any such structure and the Fourier transform of its spherically smoothed vector map, which would be useful in using experimental data, is exceedingly complex. In addition, the nature of the globulite shells cannot be fully investigated without investigating the molecular pattern of the complete dyadic particle, i.e. the networks on which the individual molecules are built into the half entities, two of which, set on a diad axis, make up the complete protein entity. Some indications of the nature of this particular problem can be gleaned from studying the various transforms of finite networks of points already recorded for this purpose (Wrinch, 1946, pp. 21–34, 38–40, 79–88). It is, furthermore, to be emphasized that the present formulation, complex as it is, still pictures the particle as made up of identical molecules in parallel translations. Only in such circumstances is the structure factor composite, i.e. analysable into two factors, one for the molecule and the other network (Wrinch, 1946, pp. 13–14). This simplifying assumption will certainly need modification in due course. Meanwhile it enables us to appreciate the shape of the problem and to proceed with the planning of a more comprehensive attempt to interpret the X-ray data.

References

Boyes-Watson, J., Davidson, E. & Perutz, M. F. (1947). *Proc. Roy. Soc.* A, **191**, 83.

Bradley, A. J. & Illingworth, J. W. (1936). *Proc. Roy. Soc.* A, **157**, 113.

DAWSON, B. (1953). *Acta Cryst.* **6**,113.
KEGGIN, J. F. (1934). *Proc. Roy. Soc.* A, **144**, 75.
PERUTZ, M. F. (1949). *Proc. Roy. Soc.* A, **195**, 474.
WRINCH, D. (1939). *Phil. Mag.* **27**, 490.
WRINCH, D. (1946). *Fourier Transforms and Structure Factors*. ASXRED Monograph No. 2.
WRINCH, D. (1950). *Phys. Rev.* **79**, 203.
WRINCH, D. (1952*a*). *Phys. Rev.* **86**, 611.
WRINCH, D. (1952*b*). *Phil. Mag.* **43**, 801.
WRINCH, D. (1952*c*). *Acta Cryst.* **5**, 694.
WRINCH, D. (1953*a*). *Phys. Rev.* **91**, 471.
WRINCH, D. (1953*b*). *Acta Cryst.* **6**, 562.
WRINCH, D. (1953*c*). *Acta Cryst.* **6**, 638.

Acta Cryst. (1954). **7**, 357

The Structure of Analcite and Ion-Exchanged Forms of Analcite

BY I. R. BEATTIE

Department of Chemistry, The University, Sheffield 10, *England*

(*Received* 21 *September* 1953 *and in revised form* 22 *January* 1954)

From a consideration of the accepted structure of analcite it is shown that this mineral need not necessarily be considered as having an averaged lattice. The structure of leucite, and other ion-exchanged forms of analcite, is discussed in relation to the known data on pollucite, leucite and analcite. The importance of the degree of hydration of the various ion-exchanged analcites is mentioned.

Analcite and pollucite give almost identical spacings of the diffraction arcs obtained by means of X-ray powder photography, since the alumino-silicate network is essentially the same for each of these minerals. Their symmetry is cubic O_h^{10}, and the length of the unit cell is 13·7 Å. The structure of analcite originally proposed by Taylor (1930) placed the sodium ions in the positions (0·125, 0, 0·25), of which there are 24 in the unit cell. Since only 16 sodium ions are available to fill these 24 positions Taylor regarded the resultant lattice as an averaged structure with the 16 sodium ions arranged statistically among 24 sites. Náray-Szabó (1938) on the other hand suggested that the sodium ions in analcite occupied the same lattice points as the caesium ions in pollucite. The positions of the caesium ions in pollucite, namely (0·125, 0·125, 0·125), were not disputed.

Taylor (1930) studied the effect of dehydration of analcite on the X-ray diffraction patterns obtained from single crystals. In a discussion of the observations of Náray-Szabó (1938), Taylor (1938) pointed out that his experiments showed that the changes observed in line intensities on dehydration of analcite were of the same order as those calculated using the Taylor model as a basis. The alterations predicted for the Náray-Szabó structure were however in several cases the opposite of those observed in practice. Further, the water molecules in Taylor's structure are tetrahedrally bonded to two sodiums and two oxygens, in agreement with the known properties of the water molecule (Taylor, 1934). The Náray-Szabó structure results in the water molecules being surrounded by six nearest neighbours, four oxygens and two sodiums completing the octahedral group.

The formula of anhydrous analcite is $Na_2O, Al_2O_3, 4SiO_2$. From an examination of the structure proposed by Taylor, it is easily seen that each sodium ion is surrounded by four oxygens, the whole array being approximately situated in one plane. The four oxygens are made up of two pairs, both oxygens in each pair being attached to the same adjacent silicon or aluminium. This is shown in Fig. 1, which may then be said to represent the formula of analcite if the sodiums are taken to be present in only two-thirds of the possible positions. This may be written $\frac{2}{3}$Na, 2(Si or Al), 4O, or more accurately as $\frac{2}{3}$Na, $\frac{2}{3}$Al, $\frac{4}{3}$Si, 4O. In

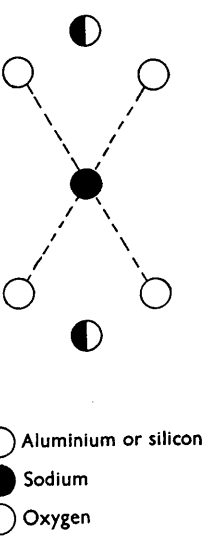

Fig. 1. Representation of the lattice position of a sodium ion in analcite.