Structure of the Concanavalin A–Methyl α-D-Mannopyranoside Complex at 6-Å Resolution†

Karl D. Hardman* and Clinton F. Ainsworth

ABSTRACT: The carbohydrate binding site of concanavalin A has been identified in crystals of the concanavalin A–methyl α-D-mannopyranoside complex and is 35 Å from the iodophenol binding site (K. D. Hardman and C. F. Ainsworth (1973), *Biochemistry* 12, 4442), which has been postulated to be adjacent to the carbohydrate-specific binding site (Edelman et al. (1972), *Proc. Natl. Acad. Sci. U.S.A. 69,* 2580). The crystals are orthorhombic in space group $C222_1$ and crystal density measurements indicate a protein mass of four monomers (molecular weight of 104 000) per asymmetric unit. However, the electron density map contains eight monomers/asymmetric unit, revealing lattice disorder. The electron density map with a nominal resolution of 6 Å has been solved using three heavy-atom derivatives and the position and orientation of each monomer established. Atomic coordinates of the native protein which has previously been determined (K. D. Hardman (1973), *Adv. Exp. Med. Biol. 40,* 103) were transposed into this new space group and the gross conformations of the monomers, dimers, and tetramers were found to be very similar to the previous structure. However, some minor differences were apparent even at this resolution. After crystal growth, the methyl α-D-mannopyranoside was replaced by *o*-iodophenyl β-D-glucopyranoside or methyl 2-iodoacetimido-2-deoxy-α-D-glucopyranoside in separate experiments, and difference electron density maps were calculated. The highest peaks for both iodinated sugar derivatives associated with each monomer agreed within a few angstroms of each other and were found near side chains Tyr-12 and -100 and Asp-16 and -208. This region is 10–14 Å from the manganese, in good agreement with nuclear magnetic resonance (NMR) studies in solution (C. F. Brewer et al. (1973), *Biochemistry 12,* 4448) and with the site predicted from cross-linked I222 crystal studies (K. D. Hardman (1973), *Adv. Exp. Med. Biol. 40,* 103).

The lectin Con A[1] exhibits numerous biological activities related to its ability to bind specific carbohydrates.[2] The three-dimensional structure of the native protein has been determined by x-ray crystallographic methods (Hardman and Ainsworth, 1972, 1973; Hardman, 1973; Becker et al., 1975; Reeke et al., 1975; Edelman et al., 1972); atomic coordinates are available from the Protein Data Bank at Brookhaven National Laboratory. Nevertheless, the carbohydrate-binding site had not, until now, been unambiguously identified. βIphGlcp was found to bind to crystals of native Con A(*I*222) (Becker et al., 1971) and a difference electron density peak was assumed to indicate the carbohydrate-binding region (Edelman et al., 1972). However, it was subsequently shown that many iodophenyl compounds bind to this same site, including βIphGalp, even though Con A does not bind galactose (Goldstein et al., 1965). Therefore, the carbohydrate moiety was not responsible for the binding to this site (Hardman and Ainsworth, 1973). Additionally, native crystals of Con A are cracked and dissolved by concentrations of αMeManp of 1 mM and above while lower concentrations produce no significant changes in difference electron density maps (Hardman and Ainsworth, 1973), suggesting the specific carbohydrate-binding site is elsewhere, presumably near intermolecular contacts that stabilize the crystal lattice (Hardman, 1973).

The carbohydrate-binding site was predicted on the basis of experiments on *I*222 crystals of native Con A (no sugar present during crystallization) which were cross-linked with glutaraldehyde so that subsequent binding of sugars would not destroy the crystal lattice (Hardman, 1973). Binding of sugars produced only marginal difference peaks before the mosaic quality of crystals was markedly decreased in spite of the covalent cross-links. The highest peak found, however, did correspond to a Mn^{2+} to carbohydrate distance of 10–12 Å (Hardman, 1973).

Corroboratively, NMR studies of carbohydrate bound to Con A indicate the carbohydrate–Mn^{2+} distance to be about 10–12 Å (Brewer et al., 1973a,b; Villafranca and Viola, 1974; Alter and Magnuson, 1974), which conflicts notably with the greater than 23 Å distances from Mn^{2+} to the iodophenyl site.

Therefore, to confirm the carbohydrate binding site we have crystallized the Con A–αMeManp complex, solved the structure at 6-Å resolution by phasing with three heavy-atom derivatives, and located the carbohydrate-specific binding site by difference electron density techniques. Growth of the new crystal form was preceded by reacting the protein with 10–50 mM αMeManp, 200–1000 times the concentrations necessary to give ½ occupancy in solution. After crystal growth, the αMeManp was replaced by either βIphGlcp or IAcGlcn with no adverse effects on the crystalline lattice. The space group is $C222_1$ with unit-cell dimensions of 118.6, 102.6, and 253.1 Å, for *a*, *b*, and *c*, respectively, and a unit-cell volume of 3.077×10^6 Å^3 (or 0.3846×10^6 Å^3 per asymmetric unit). The crystal density indi-

†From the General Sciences Department, IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598 (K.D.H.), and the Division of Biological and Medical Research, Argonne National Laboratory, Argonne, Illinois 60439 (C.F.A.). *Received July 28, 1975.* This work was supported by the Research Division of IBM Corporation and in part, prior to July 1974, by the Divison of Biomedical and Environmental Research of the U.S. Atomic Energy Commission. A summary of this work was presented at the 10th International Congress of Crystallography, August 1975, Amsterdam, The Netherlands.

[1]Abbreviations used are: Con A, concanavalin A; αMeManp, methyl α-D-mannopyranoside; βIphGlcp, *o*-iodophenyl β-D-glucopyranoside; βIphGalp, *o*-iodophenyl β-D-galactopyranoside; IAcGlcn, methyl 2-iodoacetamido-2-deoxy-α-D-glucopyranoside.

[2]For a recent review of the biological effects of Con A and other lectins, see Lis and Sharon (1973).

cates a protein mass of 104 000 daltons (four monomers) per asymmetric unit; however, the electron density map at 6-Å resolution clearly outlines eight unique monomers, which implies a disordered lattice. Regardless of the nature of this disorder, the location of the nonpolar binding cavity (Hardman and Ainsworth, 1973) and the carbohydrate-binding site for each monomer are confirmed and are about 35 Å apart.

Methods

Procedures for Con A purification and crystallization were as previously described (Hardman and Ainsworth, 1972, 1973; Hardman et al., 1971a,b) except as noted. After adsorption onto Sephadex from the aqueous bean meal extract, the Con A was eluted with 10 mM αMeManp in 0.01 M Tris–0.35 M NaCl and 10 mM Mn^{2+} and Ca^{2+} (pH 7.4). The eluted protein solutions were used directly for growing crystals by one of the following two methods: (1) protein solutions were vacuum dialyzed against 2.0 M sodium potassium phosphate (pH 7.4) containing 10–50 mM αMeManp, 10 mM Mn^{2+}, and 10 mM Ca^{2+} with inverted microdiffusion tubes; (2) after concentration by pressure or vacuum dialysis solutions of Con A (20–60 mg/ml) in the Tris–NaCl buffer containing 50 mM αMeManp were diluted with 4 M sodium potassium phosphate (pH 7.4) to give final phosphate concentrations ranging from 1.6 to 2.2 M. Crystals were harvested when growth to appropriate size had been reached, 3 days to 2 weeks, by replacing the mother liquor with 2.0–2.2 M phosphate containing 10–100 mM αMeManp.

Heavy-atom derivatives were prepared by soaking the crystals in 2.0 M phosphate (pH 7.4) with αMeManp plus the appropriate heavy-atom compound. The concentrations of K_2PtCl_4, $HgCl_2$, and *o*-iodophenol were 0.5, 0.25, and 13 mM, respectively, and reacting times were 2, 6, and 14 days. Crystals for the K_2PtCl_4 and $HgCl_2$ derivatives were washed free of the heavy-atom compounds at the end of the specified time. αMeManp was removed from the crystals for the iodo sugar derivatives by washing and soaking with three or more changes of phosphate buffer containing βIphGlcp (saturated, about 4 mM) or 10 mM IAcGlcn for 8 and 20 days, respectively. The crystals for the *o*-iodophenol derivative also contained 100 mM αMeManp. Additionally, for control purposes, a 6-Å data set was collected on a crystal in the presence of 4 mM βIphGlcp and 25 mM αMeManp.

Density Measurements. Densities of crystals were determined with two different procedures: (1) by flotation in a series of sodium potassium phosphate solutions (pH 7.4) of known density similarly to the manner described for the *I*222 native crystals (Hardman et al., 1971a); and (2) by flotation in solutions containing chloroform and mineral oil. Both procedures were basically the same except for the flotation solvents used. The crystals were first changed to 0.05 M phosphate (density of 1.015) in several steps and allowed to stand several hours. The crystals remained stable under these conditions; however, we additionally used several crystals which had first been cross-linked with glutaraldehyde (0.1–0.25% for 18 h at room temperature) as controls. Large crystals (all dimensions more than 0.5 mm, generally with at least one dimension over 1.0 mm) were removed from the 0.05 M phosphate with tungsten hooks, blotted quickly, and plunged into the flotation solutions. This procedure was used rather than the microsyringe method described by Matthews (1974) so that the crystals could be handled in the same manner for both the phosphate and the chloroform–mineral oil solvents. Buoyancy of the crystals was observed immediately upon addition of the crystal to the flotation solvent. Results with cross-linked crystals were not significantly different from the uncross-linked ones, which is consistent with previous observations of lightly cross-linked crystals (for further discussion, see Matthews, 1974). The average densities for flotation in phosphate and chloroform–mineral oil solvents were 1.137 and 1.130, respectively. The partial specific volume, 0.73 cm^3/g (Sumner et al., 1938), was used for the calculation of protein fractional volume and is very close to the partial specific volume calculated from the amino acid composition, 0.727 cm^3/g.

Data Collection. Preliminary space group determination was made from precession photographs and confirmed by thoroughly sampling reciprocal space to *d*-spacing limits of 3.0 Å. Data sets were collected on a computer-controlled diffractometer with techniques similar to those previously described (Hardman et al., 1971a). The x-ray tube was a Picker standard spot-focus copper target operating at 43 kV and 16 mA using a nickel filter and a take-off angle of 3°. The target to crystal distance was 210 mm with the incident beam collimation of 1.0 mm diameter placed 12 mm from the crystal. The crystal to detector distance was 545 mm with the diffracted beam collimation of 1.0 mm placed 12 mm from the crystal and receiving aperture slits of 1.4 × 4.4 mm, horizontal and vertical, respectively. The diffracted beam collimator was purged with He to eliminate air scattering of the x-rays.

Half-intensity widths under data-collect conditions between 0.10 and 0.11° were characteristic of those found throughout reciprocal space. Additionally, the half-intensity width of this crystal, as measured with open detector slits and with a 1° take-off angle, was 0.08 ± 0.01, and was representative of the mosaic character of this crystal form.

Reproducibility in data collection is shown by agreement residuals between identical data sets from different native crystals. These were calculated as R_F, as defined in Table II, except the numerators were the sums of the absolute differences of F_P between two different crystals. For 6-Å data sets, about 4200 reflections, the agreement residuals were typically 0.044 as compared to 0.034 for two data sets from the same crystal. For comparison, residuals calculated from different native *I*222 crystals fell between 0.031 and 0.033 (Hardman et al., 1971a). Average intensities of a set of standard reflections of native $C222_1$ crystals typically decreased 5% during collection of a 6-Å data set (45–50-h x-ray exposure) while intensities for derivative crystals decreased 8–10%.

In order to resolve reflections along the *z* axis (c = 253.1 Å), the crystals were oriented on the diffractometer with the *z* axis nearly equatorial ($\chi \simeq 0$) and the ω step-scan procedure of Wyckoff et al. (1967) was utilized with 0.02° ω intervals. A sample set of adjacent reflections is shown in Figure 1. These data were taken from a K_2PtCl_6 derivative crystal after 80 h of x-ray exposure (the average intensity of the set of standard reflections had decreased by 15%). The mosaic quality and the resolution of reflections had not appreciably changed from the start of x-ray exposure. The $9,\bar{7},l$ reflection was set to $\chi = 90$, placing c^* near $\chi = 0$, necessary for best resolution along the *l* rows. Routinely, for data collection a reflection such as this was set polar so that the χ limits for the reflections collected were about ±45°. In Figure 1, 64 stationary counts (40 s each) were

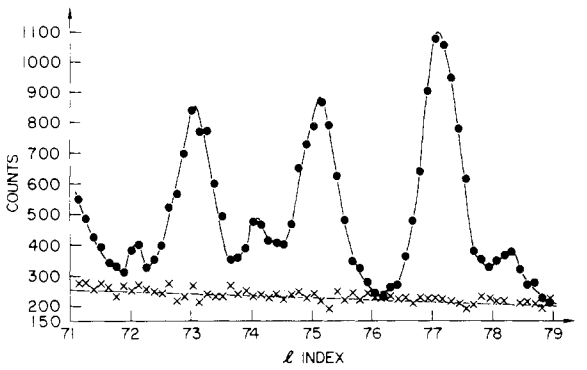

FIGURE 1: Diffractometer data scan of a portion of the 2,0,*l* row. This demonstrates the ability to resolve individual reflections in the most difficult region of reciprocal space (see text). (●) Positions for stationary counts along the row; (X) background as collected by offsetting ω by 0.8°.

collected along the 2,0,*l* row, from 2,0,71 to 2,0,79. All angles were incremented proportionally for each step in order to pass through each reflection. The background counts were established by an identical scan except that ω had been offset by 0.8°. It is apparent that a data collection scheme scanning 2θ over the entire reflection and integrating the intensity would *not* be possible; however, Wyckoff-type scans summing the highest three ω steps with 0.02° increments are suitable. In Figure 1, since the intensity drops to background at the 2,0,76 reflection, it is demonstrated that even the moderately strong adjacent reflections, the 2,0,75 and 2,0,77, do not elevate the count at the 2,0,76 center.

For calculation of the electron density map at 6-Å resolution, three sets of native data (about 4200 unique reflections per set), one set of Friedel pairs for $PtCl_4^{2-}$, and one set each for the *o*-iodophenol and the $HgCl_2$ derivatives were collected. Each data set was collected on a different crystal. Additional methods of data collection, correction, and reduction have been previously described (Hardman et al., 1971a,b; Hardman and Ainsworth, 1972).

Determination and Refinement of Heavy-Atom Positions, Protein Phase Refinement and Electron Density Map Calculation. The heavy-atom positions of the $PtCl_4^{2-}$ derivative were located by examination of three-dimensional difference Patterson maps, $|F_{PH} - F_P|^2$, followed by initial least-squares refinement and trial heavy-atom Fourier and difference Fourier maps, recycling the process several times in a conventional manner. Protein phases for this derivative were then used to calculate trial sites for both the $HgCl_2$ and *o*-iodophenol derivatives. Sites for all three derivatives were refined and protein phases were calculated in alternate cycles using Dickerson's program (Dickerson et al., 1968) which we had modified to accommodate the $C222_1$ space group and up to 30 heavy-atom sites per derivative.

Electron density maps were calculated with a conventional Fourier program written by S. T. Rao and a fast-Fourier program (CHAFF) written by D. Bantz and M. Zwick of the University of Chicago. Contour maps were computer drawn by linear interpolation from a 30,30,60 array of *X*, *Y*, and *Z*, respectively, in fractions of unit-cell dimensions from 0.0 to 0.5 and were copied onto infrared projector transparencies.

The molecular symmetry, although predicted by the symmetry of the $PtCl_4^{2-}$ derivative alone, was confirmed by inspection of the 6-Å electron density maps. The densities for each monomer were outlined and the atomic coordinates of the native *I*222 structure were translated and rotated so that prominent densities of the $C222_1$ map corresponded to known features, such as the pronounced β-structure regions, and the maximum symmetry of the Con A molecule was observed.

Table I: Heavy-Atom Coordinates.

Derivative	Site	Unit-Cell Fractions *X*	*Y*	*Z*	A^a	B^b
$PtCl_4^{2-}$	1	0.2528	0.0241	0.2326	56.4	5.8
	2	0.1703	0.0184	0.2634	102.1	17.3
	3	0.1591	0.0032	0.2405	43.0	1.8
	4	0.2536	0.0105	0.2599	124.0	20.8
	5	0.0675	−0.0042	0.4871	105.4	22.4
	6	0.1530	0.0067	0.4883	63.1	5.9
	7	0.4818	0.0029	0.4878	78.7	9.8
	8	0.5677	0.0178	0.4883	114.0	23.9
$HgCl_2$	1	0.2481	0.0158	0.2313	62.6	15.5
	2	0.1779	0.0175	0.2635	191.3	43.4
	3	0.1733	0.0167	0.2348	54.9	22.6
	4	0.2407	0.0125	0.2599	116.0	28.6
	5	0.0823	0.0030	0.4861	98.7	24.5
	6	0.1431	0.0008	0.4866	81.5	25.4
	7	0.4879	0.0101	0.4875	115.8	32.2
	8	0.5552	0.0144	0.4894	100.8	29.6
	9	0.1062	0.0134	0.2554	88.5	35.5
	10	0.2267	−0.0091	0.5028	42.0	36.3
	11	0.4165	−0.0051	0.4933	71.5	36.8
	12	0.4838	0.1888	0.2686	75.5	17.9
	13	0.1012	0.1825	0.2342	95.6	29.9
	14	0.2917	0.3346	0.0085	75.2	19.7
	15	0.3030	0.3352	0.4928	58.4	16.7
	16	0.3222	0.3168	0.4035	79.6	41.4
o-Iodophenol	1	0.1884	0.1397	0.2038	43.4	8.1
	2	0.3135	0.3889	0.3000	64.9	15.5
	3	0.2863	0.3927	0.2038	37.8	13.1
	4	0.2192	0.1316	0.2991	33.1	2.9
	5	0.1140	0.1384	0.4630	30.5	14.6
	6	0.4766	0.1438	0.0392	69.5	11.8
	7	0.0254	0.4005	0.4436	84.0	14.1
	8	0.3965	0.1369	0.4451	22.4	32.8
	9	0.2968	0.1326	0.2610	20.5	17.8

[a] *A*, occupancy, on an arbitrary scale. [b] *B*, temperature factor ($Å^2$).

Results and Discussion

Location and Refinement of Heavy-Atom Positions. In the native *I*222 crystals, the $PtCl_4^{2-}$ derivative produced one major site per monomer (Hardman and Ainsworth, 1972; Becker et al., 1975), which lies *between* the side chains of His-127[3] and Met-129. The $PtCl_4^{2-}$ is bound irreversibly at nearly 100% occupancy, presumably because of chelation to a histidyl N and the methionyl S. This site (x,y,z = 3.0, 0.5, 5.2 Å, respectively) is 6.0 Å from the origin of the Con A tetramer (0,0,0), a point of crystallographic 222 symmetry which produces a rectangular cluster of 4 Pt atoms, 6.0 Å by 10.4 Å (distorted slightly by $y = \pm 0.5$ Å). Inspection of the three-dimensional Patterson map of this derivative in the new crystal form in the presence of αMeManp, $C222_1$, readily showed a cluster of 4 noncrystallographically equivalent sites closely resembling that in the *I*222 crystals, Table I, sites 1–4. The four sites are found within 6.5 Å of the origin of tetramer A (Figure 2). However, inspection of F_H and ΔF_H Fourier maps and further

[3] Amino acid residue numbers correspond to the revised sequence of Cunningham et al. (1975).

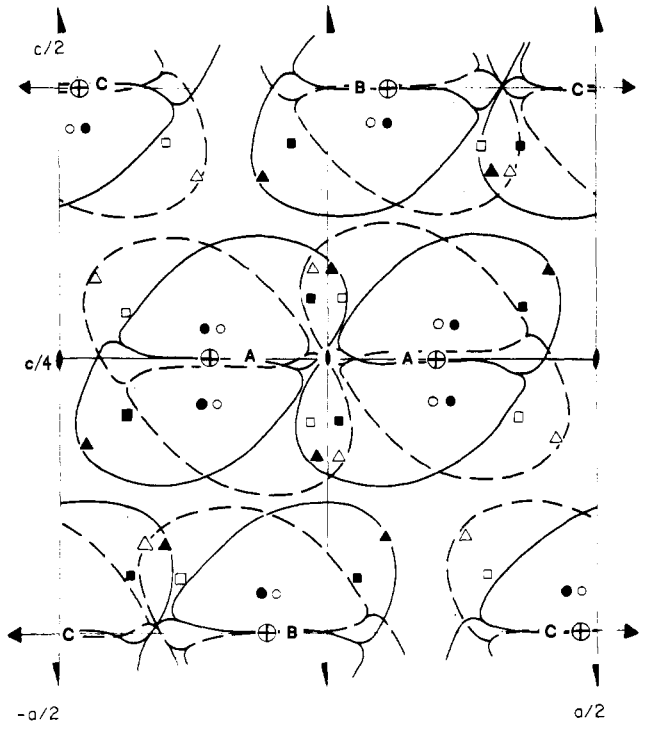

FIGURE 2: General outlines of Con A molecules in the $C222_1$ crystals. (triangle) Carbohydrate-binding sites; (squares) calculated Mn^{2+} positions; (circles) nonpolar binding sites; and (circles with +) points of *noncrystallographic* 222 symmetry. Crystallographic symmetry elements are marked with standard symbols (Lonsdale, 1952). Open symbols correspond to sites on monomers of Figure 3 and solid symbols to those in Figure 4. $+y$ points below the plane of the paper and $-y$ above. Carbohydrate binding sites which lie close together in this projection are actually about 50 Å apart, and the nonpolar sites are separated by 28 Å. Unique tetramers are arbitrarily labeled A, B, and C. Three crystallographic twofold rotation axes appear: (1) parallel to the y axis in the center of the figure $(0,y,\frac{1}{4})$, which relates the two A tetramers, (2) horizontal at the bottom $(x,0,0)$, and (3) horizontal at the top $(x,0,\frac{1}{2})$. Therefore, tetramers B and C each consist of two crystallographically identical dimers while the four monomers of tetramer A are related only by noncrystallographic symmetry. The 2-monomer unit of the A tetramer in the right-center of this figure outlined with the dashed line (also Figure 3) corresponds to the stereogram shown in Figure 5 and has the same orientation.

inspection of the Patterson map revealed four additional sites, Table I, sites 5–8. These four sites *cannot* possibly represent additional *second* sites on the original tetramer since they are found near $z = 0$, or $\frac{1}{2}$, over 65 Å from the first four sites, near $z = \frac{1}{4}$ (see Figure 2), and the longest dimension of a monomer in the native crystal is only 45 Å. Least-squares refinement of these eight sites produced an R value for the $PtCl_4^{2+}$ derivative of 0.29 in contrast to 0.52 for the first four sites alone.

The crystallographic twofold rotation axis at $(x,0,\frac{1}{2})$ generates two more sites related to $PtCl_4^{2-}$ sites 5 and 6, completing the four-atom cluster for the B tetramer, and the same occurs for the C tetramer, sites 7 and 8. [The identical situation is repeated at the twofold rotation axis at $(x,0,0)$ as a result of the crystallographic twofold screw axes parallel to c.] When these three four-atom clusters, which are all similar to that found in the native $I222$ crystals, were first observed, they not only suggested the molecular origin for each tetramer, but they also suggested orientation of the axes. This orientation was consistent with the fact that unit-cell dimension c for the $C222_1$ crystals is 4 × the a for the $I222$ native crystals. Thus, the location of a total of eight major sites for the $PtCl_4^{2-}$ derivative at that

Table II: Refinement Parameters at 6-Å Resolution.

	$PtCl_4^{2-}$	$HgCl_2$	*o*-Iodophenol
R_K (all data)[a]	0.085	0.155	0.111
R_K (centric)	0.108	0.204	0.141
R (all data)[b]	0.479	0.704	0.665
R (centric)	0.426	0.627	0.643
R_F[c]	0.178	0.219	0.167
K scale[d]	1.007	1.021	1.020
ΔB scale[e]	0.601	0.400	0.100
rmse[f]	161	336	204
Figure of Merit all data = 0.632, centric = 0.629.			

[a] $R_K = \Sigma|F_{PH(obsd)} - F_{PH(calcd)}|/\Sigma F_{PH(obsd)}$, the Kraut R factor where F = structure factor amplitudes. The subscripts P, PH, H (obsd), and (calcd) refer to terms for protein, protein plus heavy atom, heavy atom, observed, and calculated, respectively. [b] $R = \Sigma|F_{PH(obsd)} - F_{PH(calcd)}|/\Sigma|F_{PH(obsd)} - F_{P(obsd)}|$. [c] $R_F = \Sigma|F_{PH} - F_P|/\Sigma F_P$. [d] K scale = scale-factors between native and derivative data sets. [e] ΔB = difference temperature factor between native and derivative data sets. [f] rmse = root mean square error of closure, on same relative scale as A in Table I.

point suggested the presence of eight monomers per asymmetric unit. This was again suggested by the other derivatives, including βIphGlcp and IAcGlcn, and finally confirmed by the protein electron density map. These results, which directly conflict with the unambiguous density measurements, indicate the crystalline lattice is disordered (see electron density map section).

Protein phases calculated from the $PtCl_4^{2-}$ derivative, including anomalous scattering data, provided trial positions for both the $HgCl_2$ and *o*-iodophenol derivatives. The original F_H electron density map for the $HgCl_2$ derivative suggested approximately 12 sites which remained on initial refinement. The noncrystallographic symmetry indicated by the $PtCl_4^{2-}$ derivative was then used to generate additional trial $HgCl_2$ sites, and of about 30 tried, 16 remained convincingly throughout the entire refinement procedure. Trial *o*-iodophenol sites for least-squares refinement were *not* predicted from the molecular symmetry and orientation suggested by the $PtCl_4^{2-}$ derivative, but were obtained directly from *o*-iodophenol F_H maps produced with phases assigned from the $PtCl_4^{2-}$ derivative alone. Results suggested 12 possible sites of which 9 remained after refinement (Table I).

During phase calculation and refinement, convergence was slow, presumably because of large numbers of sites per derivative and the number of reflections resulting from the large unit cell. Approximately 10 runs were made with four cycles per run; F_H or ΔF_H Fourier maps were calculated between runs to look for additional heavy-atom sites. Sites which were known to be incorrect were added for test purposes and in all cases were eliminated by the refinement procedure. The protein phase refinement data for the three derivatives used for calculation of the 6-Å map are shown in Table II.

The noncrystallographic symmetry of the $HgCl_2$, *o*-iodophenol, βIphGlcp, and IAcGlcn derivatives all support the molecular symmetry indicated by the $PtCl_4^{2-}$ derivative and confirmed by the protein electron density maps (Figure 2 and Table III). Transforming the eight sites of the $PtCl_4^{2-}$ derivative to the corresponding site in the $I222$ crystal, the mean difference and root mean square difference are 0.71 and 0.97 Å, respectively. Similarly for the first eight $HgCl_2$ sites, the values are 0.67 and 0.86 Å respectively. The values for *o*-iodophenol, βIphGlcp, and

Table III: Tetramer Origins.

Tetramer	$C222_1$ Coordinates (Unit-Cell Fractions) X	Y	Z
A	0.21	0.01	0.25
B	0.11	0.00	0.50
C	0.52	0.00	0.50

Transformation of Axes		
Con A–αMeMan Complex Crystals ($C222_1$)		Native Con A Crystals ($I222$)
X	=	Z
Y	=	Y
Z	=	$-X$

IAcGlcn are found in Table IV. Small shifts of molecular origins and noncrystallographic axes have not been attempted on this low resolution data and might improve the agreement of the different heavy-atom sites. Such improved placement of the molecular symmetry elements may also more accurately predict symmetry related sites for the $HgCl_2$ derivative which did not remain on refinement.

The 6-Å Electron Density Map. Analysis of the protein electron density map confirmed the symmetry proposed by the $PtCl_4^{2-}$ derivative. Figures 3 and 4 show some sample sections 10 Å thick through the main portions of the monomers which were superimposed to produce the outlines in Figure 2. Volumes equivalent to two asymmetric units are projected along the y axis for Figures 3 and 4. Atomic positions were calculated for all monomers assuming isomorphous structures with the $I222$ form. Monomer outlines were marked and were consistent with boundaries of each monomer observed by inspection of this map. Boundaries of monomers from neighboring tetramers came into contact but, as well as could be discerned at this resolution, did not overlap. In these projections, the monomers are oriented so the six strands of the β sheet region I (Hardman and Ainsworth, 1972) run almost parallel to the noncrystallographic twofold at $c/4$ and the crystallographic twofold at $c/2$ (see Figures 5 and 6). The strands are stacked vertically (normal to the plane of the paper) to form the continuous 12-strand β sheet (Hardman and Ainsworth, 1972; Hardman, 1973) which links two different monomers to form the dimers (one monomer below in Figure 3 and the other above in Figure 4). Figures 3 and 4 together thus represent four tetramers in two asymmetric units or eight monomers per asymmetric unit (Figure 2). Although there are eight nonequivalent monomers per asymmetric unit, Figures 3 and 4 clearly show there are *three* unique tetramers and their origins and transformations from the $I222$ axes to the $C222_1$ space group are listed in Table III. The monomers of the A tetramer are related only by noncrystallographic twofold rotation axes; however, tetramers B and C are each split by the crystallographic twofold at $(x,0,0)$ and $(x,0,\frac{1}{2})$.

In Figures 3–5 the electron density contours for the β-sheet regions and the Mn^{2+} and Ca^{2+} binding regions (Hardman and Ainsworth, 1972; Hardman, 1973) are the most pronounced (compare the lower right monomer of Figure 5 with Figure 6). Two additional monomers, obtained by rotating the 2 in Figure 5 around the noncrystallographic twofold at $(x,0,\frac{1}{4})$, complete the tetramer. Electron density contours for the B and C tetramers appear very similar to these, particularly in the more dense regions, but distinct differences are present.

The paradox between the eight monomers found per asymmetric unit in the protein electron density map and the crystal density, which indicates four monomers, suggests a disordered crystal lattice. Initially, on measuring unit-cell dimensions of this crystal form, protein volume fractions of 0.327 and 0.654 were calculated, assuming four and eight monomers per asymmetric unit, respectively. These values correspond to crystal volume per unit of protein molecular weight, Vm (Matthews, 1968) of 3.67 and 1.84 $Å^3$, respectively. These values alone, since they lie at either extreme for those found for proteins (Matthews, 1968), cannot resolve the number of monomers per asymmetric unit. Therefore, the crystal density was determined. After exchange of the $C222_1$ crystals to 0.05 M phosphate, average crystal densities of 1.130 and 1.137 were found in chloroform–mineral mixtures and phosphate solutions, respectively. (Theoretical densities for four and eight monomers per asymmetric unit are 1.131 and 1.254, respectively.) Protein volume fractions calculated from these densities are 0.339 and 0.359, which compare very closely to theoretical value of 0.327. Furthermore, wet and dry crystal weights were measured on several large crystals and the average weight loss was 52%. Even disregarding incomplete drying and presence of some salt, this value indicates that eight monomers per asymmetric unit *cannot* be present. It was therefore unambiguously concluded that the protein mass per asymmetric unit is four monomers and the doubling of this number in the protein electron density map apparently results from a disordered lattice. Although similar disorder is found occasionally in crystals of relatively small molecules (for example, Karle and Karle, 1972), to our knowledge this type has not previously been reported for protein crystals. In this case the different molecules do not appear to overlap in the electron density maps, as generally occurs. However, slight overlap cannot be excluded at this low resolution, and if found at high resolution, would confirm the disordered lattice. The possibility should be considered that this lattice disorder may be partially a function of the large unit-cell dimensions and self-assembling properties of Con A and that perhaps similar disorder may begin to appear as more oligomeric proteins are studied.

The Carbohydrate Binding Site. Carbohydrates with the minimum structure for specific binding by Con A contain hexose residues with the D-*arabino*-pyranoside configuration at C-3, C-4, and C-5 (Goldstein et al., 1965; So and Goldstein, 1969). The monosaccharide with the highest affinity to Con A is αMeManp ($K_a = 2 \times 10^4$), with αMeGlcp about 25-fold less (Poretz and Goldstein, 1970). The affinities of the β anomers are about 40-fold less than the α anomers (Poretz and Goldstein, 1970). Con A has generally been found to bind polysaccharides containing mannosyl or glucosyl residues at the nonreducing termini (Goldstein et al., 1965; Poretz and Goldstein, 1970); however, in some instances nonterminal mannosyl residues appear to be involved (Goldstein et al., 1973). Galactose, the C-4 epimer of glucose, however, does not bind (Goldstein et al., 1965).

It has been reported previously that βIphGlcp binds to a cavity in Con A and since this compound inhibits dextran precipitation it was argued that this cavity included the site of specific carbohydrate binding (Becker et al., 1971, 1975; Edelman et al., 1972). However, this cavity has also been shown to bind a variety of relatively nonpolar molecules, such as *o*-iodoaniline, *o*-iodobenzoic acid, phenyl phosphate, and methyl *p*-hydroxybenzoate, completely indepen-

Table IV: Coordinates of Carbohydrate and Nonpolar Binding Sites.[e]

Monomer	Carbohydrate Binding Sites											Nonpolar Sites						
	Calcd			βIphGlcp Found			Δ[a]	IAcGlcn Found			Δ[a]	Calcd			*o*-Iodophenol Found			Δ
	X	*Y*	*Z*	*X*	*Y*	*Z*	(A)				(A)							(A)
A1	0.424	0.231	0.157	0.404	0.254	0.122	5	0.367	0.270	0.125	5	0.196	0.141	0.202	0.188	0.140	0.204	1.1
A2	0.494	0.297	0.157	0.450	0.290	0.150	1	0.467	0.253	0.161	1	0.278	0.387	0.298	0.313	0.389	0.300	4.2
A3	0.076	0.297	0.159	0.042	0.285	0.150	0	0.040	0.263	0.154	1	0.304	0.387	0.204	0.286	0.393	0.203	2.2
A4	0.006	0.231	0.341	0.015	0.260	0.377	5	0.041	0.267	0.370	1	0.222	0.141	0.296	0.219	0.132	0.299	1.2
B1	0.325	0.217	0.408	0.330	0.250	0.400	1	0.322	0.296	0.403	5[b]	0.097	0.127	0.453	0.114	0.138	0.463	3.4
B2	0.385	0.283	0.408	0.383	0.317	0.375	4	0.322	0.296	0.403	3	0.377	0.373	0.047				[c]
C1	0.261	0.217	0.092	0.247	0.277	0.092	1	0.247	0.290	0.083	3[b]	0.489	0.127	0.047	0.477	0.144	0.039	3.0
C2	0.191	0.283	0.092	0.209	0.310	0.126	4	0.247	0.290	0.083	2	0.035	0.373	0.453	0.025	0.401	0.444	3.8
Mean difference[d]							4.97				5.80							2.11
RMS difference							5.59				7.11							2.41

[a] Differences for βIphGlcp and IAcGlcn are adjusted by 5 A, about the smallest possible distance from center of the iodine atom to the center of the glucopyranose ring as determined by space filling models. [b] Only one site appears near the predicted carbohydrate sites for B1 and B2 and could be associated with either. The same is true for the C monomers. [c] No *o*-iodophenol difference peak was found corresponding to the nonpolar binding site for this monomer. [d] Mean and root mean square differences indicate agreement of sites related by molecular symmetry after transformation to coordinates of the original native *I*222 monomer. [e] Coordinates are given in unit-cell fractions.

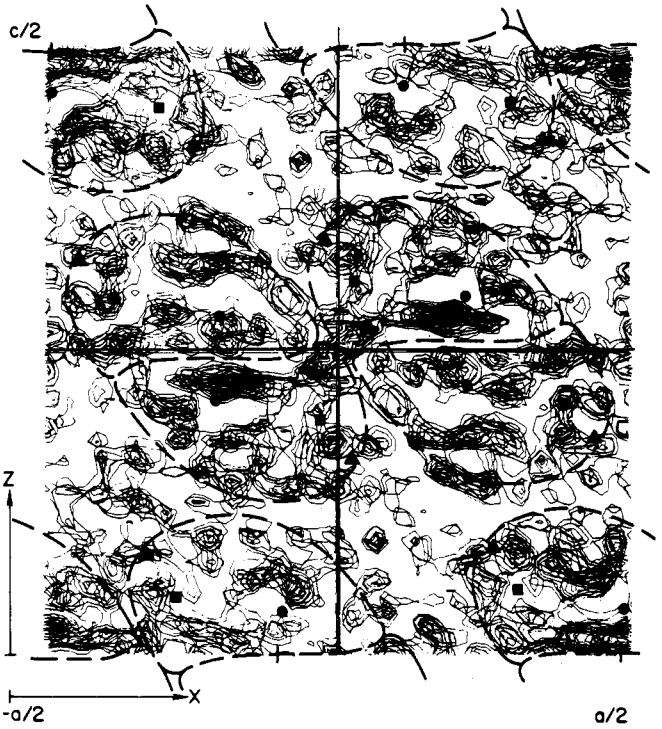

FIGURE 3: Sections of the 6-Å electron density map from $y/12$ to $y/6$. Six contour levels, 10 Å thick, are projected together with outlines of each monomer. The two-monomer units in this figure are associated with those in Figure 4 to produce tetramers (Figure 2). Symbols are equivalent to open symbols in Figure 2.

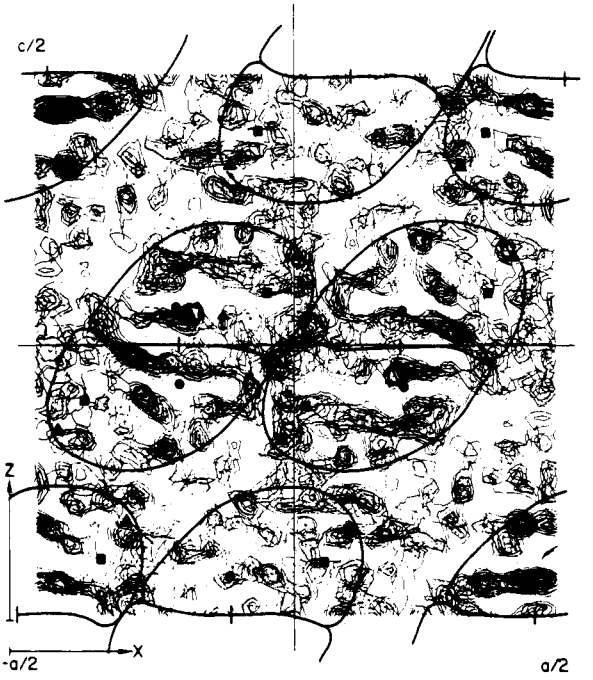

FIGURE 4: Sections of the 6-Å electron density map from $-y/12$ to $-y/6$.

dent of any carbohydrate binding (Hardman and Ainsworth, 1973). In addition, it was also shown to strongly bind βIphGalp although neither βIphGalp nor galactose will inhibit dextran precipitation (Hardman and Ainsworth, 1973). Therefore, this cavity was shown to be a general site for binding nonpolar molecules and the binding of βIphGlcp and βIphGalp is due to the iodophenyl moiety and *not* the carbohydrate. This is confirmed by the identification of the carbohydrate specific site in the $C222_1$ crystals as indicated by the binding of both IAcGlcn and βIphGlcp.

Regardless of the nature of the lattice disorder and the microscopic packing of molecules in the crystal, the outlines and orientations of each monomer in the electron density map are clear and the location of the carbohydrate- and the nonpolar-binding sites are identified. The difference electron density maps for the two iodo sugars, βIphGlcp and IAcGlcn, produced peaks shown in Table IV. The predicted carbohydrate positions for all monomers, based on the molecular-symmetry elements in Table III, are also listed. The coordinates in Table IV are not identical with sites in Figures 2–4 because crystallographically equivalent positions were chosen so the sites would fall in the same asymmetric unit, rather than on the same tetramer. For the βIphGlcp derivative, the largest difference between the calculated and observed carbohydrate binding positions was 9.4 Å (monomer A1). However, since the predicted positions were calculated from the center of the hexopyranose moiety, 5 Å were subtracted from the observed differences to correct for the distance from the iodine to the center of the glucopyranose rings for both IAcGlcn and βIphGlcp. The predicted posi-

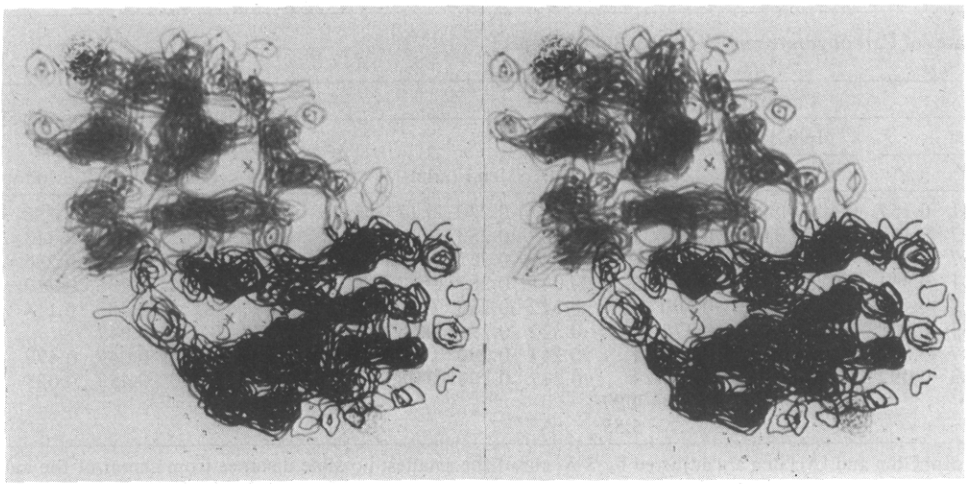

FIGURE 5: Stereogram of monomers A1 (lower right) and A4. This section (25-Å thick and 80-Å diagonally) contains most of each monomer. The dashed contours (extreme lower right and upper left) represent the positive contours of the βIphGlcp difference map associated with these monomers. The x's mark the position of *o*-iodophenol binding which is the same site the binds βIphGlcp and βIphGalp the native (*I*222) crystals. The carbohydrate-binding site and *o*-iodophenol binding site for the same monomer are 35 Å apart.

FIGURE 6: Stereogram of the Con A monomer backbone. Represented are all atoms in the backbone of the *I*222 crystal form (Hardman and Ainsworth, 1972; Hardman, 1973). The orientation of this monomer is identical with the A1 monomer in the $C222_1$ form, shown in Figure 5.

tions of the nonpolar binding sites are also included in Table IV and compared with those found for the *o*-iodophenol derivative. Mean differences and root mean square differences from the average molecular positions of βIphGlcp, IAcGlcn, and *o*-iodophenol are listed in Table IV.

The positions found for these two iodo sugars and *o*-iodophenol corroborate the molecular symmetry suggested by the $PtCl_4^{2-}$ heavy-atom positions and confirmed by the protein electron density map. Most importantly, however, they indicate an average distance from the carbohydrate binding site to the predicted Mn^{2+} position of 12 Å. The shortest distance in the crystal structure from *any* found carbohydrate position to *any* *o*-iodophenol position is 22 Å and the shortest distance for those sites in the *same monomer* is about 35 Å. Positions of the difference peaks for βIphGlcp and IAcGlcn (Table IV) correspond to the region near Asp-16 and -208 and Tyr-12 and -100 for each monomer. In the *I*222 crystals this region involves a large number of intermolecular contacts (Hardman and Ainsworth, 1972; Hardman, 1973) and there is not sufficient space for a hexose to bind without disrupting these contacts and/or altering the conformation of the protein in this region. This readily accounts for the destruction of the *I*222 crystals. Figure 7 shows a stereo view of the carbohydrate-binding region as determined in the *I*222 crystals (Hardman and Ainsworth, 1972) in the absence of sugar. The distances from this region to the Ca^{2+} and Mn^{2+} ions are about 7 and 12 Å, respectively. This is in good agreement with a number of reported NMR studies (Brewer et al., 1973a,b; Villafranca and Viola, 1974; Alter and Magnuson, 1974) which have calculated carbohydrate to Mn^{2+} distances. Average C–Mn^{2+} distances for the six non-methyl carbons of αMeGlcp from ^{13}C enriched NMR studies of Brewer et al. (1973b) were 10.3 and 11.0 Å for the α and β anomers, respectively. Similarly, Villafranca and Viola (1974) reported an average of 10.0 Å for the non-methyl carbons of the α anomer and a distance of 13.8 Å for the methyl carbon from natural abundance ^{13}C NMR results.

Alter and Magnuson (1974) report average ^{19}F–Mn^{2+} distances for *N*-trifluorcetyl-D-glucosamine binding to both dimeric (pH 5.1) and tetrameric (pH 7.0) Con A of 12.2 and 14.0 Å for the α and β anomers, respectively. Their results additionally suggest that no major difference occurs in sugar binding between the dimeric and the tetrameric forms. The ^{13}C NMR studies (Brewer et al., 1973b; Villafranca and Viola, 1974) also indicate that the plane of the pyranose ring is nearly normal to a line drawn between the sugar and the Mn^{2+} and that C-6 is closer to the Mn^{2+} than C-1. The 6-Å electron density map is not of sufficient resolution to allow us to identify or orient the pyranose ring in the binding site; however, the positions of the maximum difference density, which should be the iodine atoms, suggest likewise that C-1 extends outward from the protein, away from the Mn^{2+} and Ca^{2+}.

The position of the carboxyl side chain of Asp-208 (Figure 7) is indirectly linked to Ca^{2+} by a H_2O bridge and is

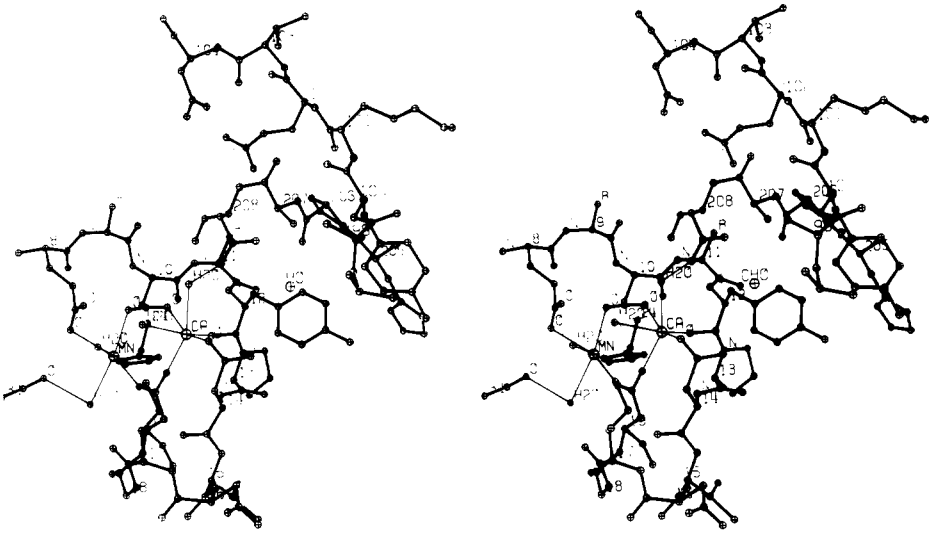

FIGURE 7: Stereogram of the carbohydrate binding site. Atomic coordinates are from the *I*222 crystal form and the position labeled CHO represents the approximate center for the pyranose ring predicted from the $C222_1$ crystal studies. Asp-16 and -208, Asp-14, Tyr-12 and -100, Leu-99, and Arg-228 are the closest residues. The distance from Mn^{2+} to this position is 11 Å. Atomic positions of the bound carbohydrate and conformational changes of the Con A must await higher resolution studies of the $C222_1$ complex. Figure 7 was drawn by the ORTEP program written by Johnson (1965).

suitably placed for hydrogen binding to the sugar. Asp-16 is also nearby and could interact directly with the sugar. Chemical modification studies of carboxyl groups with carbodiimide and hydrogen ion titration studies of Hassing et al. (1971) in the presence and absence of αMeManp have suggested involvement of two carboxyl groups per monomer. This is consistent with the crystal structure. Glu-8, Asp-10, and 19 remain involved with Mn^{2+} and Ca^{2+} binding and should be unavailable during chemical modification and titration. Additionally, we predict from the *I*222 model that after a saccharide has bound, the positions of the atoms in the amino acid residues in this region will be stabilized, particularly residues 12–19, Figure 7, and therefore the dissociation of Mn^{2+} and Ca^{2+} should be significantly decreased.

Figure 7 shows the presence of two tyrosyl residues in the carbohydrate binding region. The binding of *p*-nitrophenyl α-D-mannopyranoside is approximately twice as strong as αMeManp (Bessler et al., 1974) and other aryl pyranosides have been shown to bind more strongly than their methyl analogues (Poretz and Goldstein, 1970). Tyr-12 and -100 could easily account for these increased affinities by π–π interactions with the aryl groups. Additionally, studies of Doyle and Roholt (1968) and Hassing and Goldstein (1972) have implicated tyrosine involvement in carbohydrate binding.

It should also be noted that for a reasonable fit of the polypeptide chain in this region, the peptide bond between Ala-207 and Asp-208 must be in the cis configuration (Hardman, 1973). This is the only cis peptide bond in the structure. Also adjacent is Glu-102, which is the only charged residue in the *interior* of the monomer (other than those directly involved in the Mn^{2+} and Ca^{2+} binding). These two features, which appear to destabilize the structure in this region, are likely to be involved in conformational changes during carbohydrate binding.

The *o*-iodophenol has been shown to bind to the identical position as βIphGlcp in the native *I*222 crystals (Hardman and Ainsworth, unpublished observations; Becker and Reeke, unpublished observations, as reported in Becker et al., 1975). However, in the $C222_1$ crystals the *o*-iodophenol binding serves as a control, clearly indicating that this site and the carbohydrate-specific site cannot be contiguous. Another control was performed by calculating a difference map of a $C222_1$ crystal in 4 mM βIphGlcp and 25 mM αMeManp. No significant peaks were observed at any of the carbohydrate binding sites or at the nonpolar sites. Becker et al. (1975) have reported that in solution only one βIphGlcp molecule binds per monomer (which competes with αMeGlcp binding) although binding sites with K_a of 100 or less would not have been detected. Therefore, it is not apparent why βIphGlcp binds to the nonpolar site of the *I*222 crystals at concentration of 1–5 mM and fails in solution and in the $C222_1$ crystals. Con A binding properties in the $C222_1$ crystal form therefore appear to be indistinguishable from those in solution even though these both indeed differ from behavior in the *I*222 crystals. These differences, however, can be readily explained by constraints imposed by molecular packing in the *I*222 lattice and by slight differences in the nonpolar cavity between the two space groups and do not otherwise indicate any profound differences in three-dimensional structure between the crystalline and soluble states.

Acknowledgments

The authors thank M. Mandel for scientific support, S. Krasney for preparation of the interactive contouring package, C. F. Brewer (Albert Einstein Medical School) and I. J. Goldstein (University of Michigan Medical School) for kindly providing samples βIphGlcp and IAcGlcn, respectively, and D. Sayre, S. Koenig, R. Brown, C. F. Brewer, and I. J. Goldstein for numerous valuable discussions.

References

Alter, G. M., and Magnuson, J. A. (1974), *Biochemistry 13,* 4038–4045.

Becker, J. W., Reeke, G. N., Jr., and Edelman, G. M. (1971), *J. Biol. Chem. 246,* 6123–6125.

Becker, J. W., Reeke, G. N., Jr., Wang, J. L., Cunningham, B. A., and Edelman, G. M. (1975), *J. Biol. Chem. 250,* 1513–1524.

Bessler, W., Shafer, J. A., and Goldstein, I. J. (1974), *J. Biol Chem. 249,* 2819–2822.
Brewer, C. F., Sternlicht, H., Marcus, D. M., and Grollman, A. P. (1973a), *Proc. Natl. Acad. Sci. U.S.A. 70* 1007–1011.
Brewer, C. F., Sternlicht, H., Marcus, D. M., and Grollman, A. P. (1973b), *Biochemistry 12,* 4448–4457.
Cunningham, B. A., Wang, J. L., Waxdal, M. J., and Edelman, G. M. (1975), *J. Biol. Chem. 250,* 1503–1512.
Dickerson, R. E., Weinzierl, J. E., and Palmer, R. A. (1968), *Acta Crystallogr., Sect B 24,* 997–1003.
Doyle, R. J., and Roholt, O. A. (1968), *Life Sci. 7,* 841–846.
Edelman, G. M., Cunningham, B. A., Reeke, G. N., Jr., Becker, J. W., Waxdal, M. J., and Wang, J. L. (1972), *Proc. Natl. Acad. Sci. U.S.A. 69,* 2580–2584.
Goldstein, I. J., Hollerman, C. E., and Smith, E. E. (1965), *Biochemistry 4,* 876–883.
Goldstein, I. J., Reichert, C. M., Misaki, A., and Gorin, P. A. J. (1973), *Biochim. Biophys. Acta 317* 500–504.
Hardman, K. D. (1973), *Adv. Exp. Med. Biol. 40,* 103–123.
Hardman, K. D., and Ainsworth, C. F. (1972), *Biochemistry 11,* 4910–4919.
Hardman, K. D., and Ainsworth, C. F. (1973), *Biochemistry 12,* 4442–4448.
Hardman, K. D., Wood, M. K., Schiffer, M., Edmundson, A. B., and Ainsworth, C. F. (1971a), *Proc. Natl. Acad. Sci. U.S.A. 68* 1393–1397.
Hardman, K. D., Wood, M. K., Schiffer, M., Edmundson, A. B., and Ainsworth, C. F. (1971b), *Cold Spring Harbor Symp. Quant. Biol. 36,* 271–276.
Hassing, G. S., and Goldstein, I. J. (1972), *Biochim. Biophys. Acta 271,* 388–399.
Hassing, G. S., Goldstein, I. J., and Marini, M. (1971), *Biochim. Biophys. Acta 243,* 90–97.
Johnson, C. K. (1965), *Oak Ridge Tech. Manual,* 3794.
Karle, J. M., and Karle, I. L. (1972), *J. Am. Chem. Soc. 94,* 9182–9189.
Lis, H., and Sharon, N. (1973), *Annu, Rev. Biochem. 42* 541–574.
Lonsdale, K., Ed. (1952), International Tables for X-ray Crystallography, Vol. I, Birmingham, England, Kynoch Press.
Matthews, B. W. (1968), *J. Mol. Biol. 33,* 491–497.
Matthews, B. W. (1974), *J. Mol. Biol. 84,* 512–526.
Poretz, R. D., and Goldstein, I. J. (1970), *Biochemistry 9,* 2890–2896.
Reeke, G. N., Jr., Becker, J. W., and Edelman, G. M. (1975), *J. Biol. Chem. 250,* 1525–1547.
So, L. L., and Goldstein, I. J. (1969), *Carbohyd. Res. 10,* 231–244.
Sumner, J. B., Gralén, N. and Eriksson-Quensel, I. B. (1938), *J. Biol. Chem. 125,* 45–48.
Villafranca, J. J., and Viola, R. E. (1974), *Arch. Biochem. Biophys. 160,* 465–468.
Wyckoff, H. W., Doscher, M., Tsernoglou, D., Inagami, T., Johnson, L. N., Hardman, K. D., Allewell, N. M., Kelley, D. M., and Richards, F. M. (1967), *J. Mol. Biol. 27,* 563–578.

The Primary Structure of Myohemerythrin†

Gerald L. Klippenstein,* Joseph L. Cote, and Suzanne E. Ludlam

ABSTRACT: The complete amino acid sequence of muscle hemerythrin (myohemerythrin) from the sipunculid *Themiste* (syn. *Dendrostomum*) *pyroides* has been determined by analysis of tryptic, chymotryptic, and cyanogen bromide peptides. The primary structure of myohemerythrin differs substantially from that of coelomic hemerythrins of *Phascolopsis* (syn. *Golfingia*) *gouldii* and *Themiste pyroides,* the amino acid sequence of the muscle protein being only 46 and 45% homologous with the respective coelomic hemerythrins. The most extensive regions of homology between muscle and coelomic proteins occur near the terminii. These and other shorter regions of homology are interpreted in terms of the essential iron ligand residues of the active center.

A number of chemical, spectroscopic, and magnetic techniques have been used in the characterization of the oxygen-binding site of the non-heme iron protein hemerythrin. Chemical modification experiments directed toward identification of the amino acid side chains which provide ligands to the two iron atoms have indicated that lysyl (Fan and York, 1969, 1972), tryptophanyl (York and Fan, 1971), and carboxyl (Klippenstein, 1972a) groups are not iron linked while perhaps four histidyl (Fan and York, 1969) and two or three tyrosyl (Rill and Klotz, 1970, 1971; York and Fan, 1971; Fan and York, 1972) residues may be iron ligands. To confirm and extend these results, analysis and comparison of the amino acid sequences of hemerythrins from a variety of sources are being done. A limitation in this approach has been the small number of amino acid interchanges between hemerythrins studied, e.g., five among variants of *Phascolopsis* (syn. *Golfingia*) *gouldii* coelomic hemerythrin (Klippenstein, 1972b) and four between *P. gouldii* and *Themiste* (syn. *Dendrostomum*) *pyroides* coelomic hemerythrins (Ferrell and Kitto, 1971).

The isolation and characterization of myohemerythrin

† From the Department of Biochemistry, Spaulding Life Science Building, University of New Hampshire, Durham, New Hampshire 03824. *Received July 24, 1975.* Supported in part by a grant from the National Science Foundation (GB 35610) and a grant from the Central University Research Fund at the University of New Hampshire. Published with the approval of the Director of the New Hampshire Agricultural Experiment Station as Scientific Contribution No. 788.