J. Mol. Biol. (1994) **235**, 1560–1584

Refined Molecular Structure of Pig Pancreatic α-Amylase at 2·1 Å Resolution

Steven B. Larson, Aaron Greenwood, Duilio Cascio†, John Day and Alexander McPherson

Department of Biochemistry
University of California at Riverside
Riverside, CA 92521, U.S.A.

The structure of pig pancreatic α-amylase has been determined by X-ray diffraction analysis using multiple isomorphous replacement in a crystal of space group $P2_12_12_1$ ($a=70{\cdot}6$ Å, $b=114{\cdot}8$ Å, $c=118{\cdot}8$ Å) containing nearly 75% solvent. The structure was refined by simulated annealing and Powell minimization, as monitored by $2F_o-F_c$ difference Fourier syntheses, to a conventional R of 0·168 at 2·1 Å resolution. The final model consists of all 496 amino acid residues, a chloride and a calcium ion, 145 water molecules and an endogenous disaccharide molecule that contiguously links protein molecules related by the 2_1 crystallographic operator along x. The protein is composed of a large domain (amino acid residues 1 to 403) featuring a central α/β-barrel of eight parallel strands and connecting helices with a prominent excursion between strand β_3 and helix α_3 (amino acid residues 100 to 168). The final 93 amino acid residues at the carboxyl terminus form a second small domain consisting of a compact Greek key β-barrel. The domains are tightly associated through hydrophobic interfaces. The β_3/α_3 excursion and portions of the central α/β-barrel provide four protein ligands to the tightly bound Ca ion; three water molecules complete the coordination. The Cl^- ion is bound within one end of the α/β-barrel by two arginine residues in a manner suggesting a plausible mechanism for its allosteric activation of the enzyme.

A crystalline complex of the pancreatic α-amylase with α-cyclodextrin, a cyclic substrate analog of six glucose residues, reveals, in difference Fourier maps, three unique binding sites. One of the α-cyclodextrin sites is near the center of the long polysaccharide binding cleft that traverses one end of the α/β-barrel, another is at the extreme of this cleft. By symmetry this can also be considered as two half sites located at the extremes of the active site cleft. This latter α-cyclodextrin displaces the endogenous disaccharide when it binds and, along with the first sugar ring, delineates the extended starch binding site. The third α-cyclodextrin binds at an "accessory site" near the edge of the protein and is quite distant from the polysaccharide binding cleft. Its presence explains the multivalency of α-amylase binding to dextrins in solution.

The extended active site cleft is formed by large, sweeping, connecting loops at one end of the α/β-barrel. These include three sequence segments that are highly conserved among α-amylases. Residues suggested by kinetic and enzymological studies to participate in catalysis are found in the depths of the binding cleft close to the α-cyclodextrin bound near its center. Superposition of the lysozyme active site cleft upon that of the α-amylase produces virtual coincidence of the lysozyme catalytic residues, Glu35 and Asp52, upon Glu233 and Asp300 of α-amylase.

Keywords: α-amylase; X-ray structure; calcium ion; α/β-barrel; chloride ion

1. Introduction

Porcine pancreatic α-amylase (PPA‡, α-1,4-glucan-4-glucanohydrolase, EC 3.2.1.1) is a typical endo-amylase that catalyzes the hydrolysis of starch, glycogen and related polysaccharides by random, multiple attack hydrolysis of α-D-(1,4) glucan linkages (Robyt & French, 1970*a*,*b*; Thoma *et al.*, 1971). An extensive review of amylases and similar enzymes is found in Yamamoto *et al.* (1988).

The enzyme contains one chloride ion which serves as an allosteric effector causing an increase in V_{max} but having no effect on K_m (Muus *et al.*, 1956). α-Amylase contains at least one tightly bound

† Present address: Molecular Biology Institute, University of California, Los Angeles, CA 90024, U.S.A.

‡ Abbreviations used: PPA, porcine pancreatic α-amylase; MIR, multiple isomorphous replacement; CGM, conjugate gradient minimization; SA, simulated annealing.

0022–2836/94/051560–25 $08.00/0

Table 1
Crystal forms of procine pancreatic α-amylase

Crystal form	Crystal system	Space group	*a* (Å)	*b* (Å)	*c* (Å)	*V* ($\times 10^5$ Å^3)	V_m (Å^3/Da)	Solvent content	Ref.
I	Orthorhombic	$P2_12_12_1$	70·6	114·8	118·8	9·6	4·37	~70	This study
II	Orthorhombic	$P2_12_12_1$	56·3	87·8	103·4	5·1	2·32	~42	Buisson *et al.* (1987)
III	Orthorhombic	$P2_12_12$	70·6	110·0	64·0	5·0	2·27	~40	Shelley *et al.* (1980)

calcium ion that is essential for enzymatic activity and for maintenance of structural integrity (Fischer & Stein, 1960; Vallee *et al.*, 1959). The enzyme is a single N-acetylated polypeptide chain with five disulfide bonds having 496 amino acid residues of known sequence with a molecular weight of 55,000(±1000) Da (Pasero *et al.*, 1986).

The enzyme was first crystallized by Caldwell *et al.* (1931) and the first X-ray diffraction data was reported by McPherson & Rich (1972). They described the crystal form (form I) analyzed here as belonging to the orthorhombic space group $P2_12_12_1$ with $a=70$ Å, $b=110$ Å and $c=117$ Å. The crystals grow to sizes greater than 1 mm on an edge and yield diffraction intensities to about 2·0 Å Bragg spacing.

A crystal form (II) of PPA, also of space group $P2_12_12_1$ but otherwise quite different (Table 1) than that analyzed and described here, was reported by Payan *et al.* (1980). Buisson *et al.* (1987) later described the structure of PPA at 2·9 Å resolution based on a multiple isomorphous replacement (MIR) study of those crystals. We report here our analysis by MIR of the crystal form (I) reported by McPherson & Rich (1972), the refinement of the crystal structure at 2·1 Å resolution, and a description of the structure of the molecule in our unit cell. In addition, we describe the binding of α-cyclodextrin to the enzyme as determined by difference Fourier analyses of PPA-cyclodextrin crystals and its implications for substrate binding and catalytic function.

2. Materials and Methods

(a) *Enzyme preparation and crystallization*

Approximately 250 mg of crystalline α-amylase were prepared from 100 g of porcine pancreatin (Sigma grade VI) using a modification (Stankiewicz *et al.*, 1983) of the procedure of Loyter & Schramm (1962). The microcrystalline product of the preparative procedure was collected by centrifugation and suspended in 0·010 M cacodylate buffer (pH 6·75) containing 0·002 M $CaCl_2$ to yield a protein concentration of about 25 mg/ml. Concentrated NH_4OH was added to 20 μl samples of the suspension in each of 9 positions of Corning depression plates for crystallization by vapor equilibration as described by McPherson (1990). Crystallization was induced by the evaporation of NH_3 from the solution with concomitant decline in solubility of the protein as the pH decreased. Crystals of α-amylase appeared in 2 to 10 days; in many cases the crystals became as large as 1 mm in the smallest dimension and up to 2 or 3 mm in the largest.

(b) *Preparation of crystalline complexes*

For difference Fourier investigations of carbohydrate-enzyme complexes, the crystals were first passed through 4 to 6 successive washes of enzyme-free mother liquor. This was done to minimize enzymatic digestion of substrates or substrate analogs prior to their diffusion into the crystals. Compounds were then added in increasing concentrations directly to the mother liquor so that final levels substantially exceeded those required to saturate the enzyme. Among the compounds tested for complex formation were maltotetrose, heptomannose., maltoheptose, α, β and γ-cyclodextrin, and random limit dextrins from corn syrup.

(c) *Heavy-atom derivative search*

Isomorphous derivatives were identified by conventional trial and error procedures (Blundell & Johnson, 1976). Small samples (~2 μl) of a concentrated heavy-atom solution were added to the mother liquor and 24 to 48 h later X-ray diffraction photographs were recorded on Buerger precession cameras. Promising heavy-atom derivaties were identified by comparing native with complexed crystal photographs.

(d) *Data collection*

During the early stages of our analysis, data were collected using an Enraf-Nonius CAD-4 diffractometer and details of its correction, merging and processing are as described (Brayer & McPherson, 1983). From this data difference Patterson maps were calculated and major substitution sites of heavy atoms determined. Using diffractometer data and MIR phases (Dickerson *et al.*, 1961; Watenpaugh, 1985), a 5·0 Å structure determination was carried out and some initial studies of substrate analog binding were completed (Stankiewicz *et al.*, 1983).

X-ray diffraction data for the high resolution structure analysis were subsequently collected for both native crystals and 8 additional heavy-atom derivatives using a San Diego Multiwire Systems area detector system (Hamlin *et al.*, 1981; Xuong *et al.*, 1985) with frame sizes of 0·12° and 2 min exposure per frame. The X-ray source was a Rigaku RU-200 rotating anode X-ray generator operated at 175 mA and 45 kV with a Supper graphite crystal monochromator. Details of the data collection procedures are essentially the same as reported (Ko *et al.*, 1993). A highly redundant native data set in the resolution range 2·1 to 25 Å was collected using 9 crystals. Relevant statistics are presented in Table 2.

Each heavy-atom derivative X-ray data set was collected (see Table 2) from only a single crystal to a maximum resolution of 3·0 Å and these included anomalous pairs. Heavy-atom derivative intensities were scaled

Table 2

Native and heavy-atom derivative data statistics and heavy-atom refinement statistics

Data set	Resol. limit (Å)	No. of refls.	Completeness (%)	R_{sym}† (%)	R (der-nat) (%)	R_c (%)	Centric r.m.s. (F) / r.m.s. (ε)	Acentric r.m.s. (F) / r.m.s. (ε)	X	Y	Z	($Å^2$)	Relative occupancy	Nearest amino acid and distance (Å)		
Native no. 1	2·8–∞	22758	93·1	4·10	—	—	—	—	—	—	—	—	—	—		
Native no. 2	2·1–∞	43986	80·4	7·26	—	—	—	—	—	—	—	—	—	—		
Combined	2·1–∞	47149	82·6	5·02	—	—	—	—	—	—	—	—	—	—		
A. *Heavy-atom set no. 1*																
OCMP‡	3·2–14·0	15281	94·0	7·66	19·4	61	1·26	1·96	0·1359	0·0719	0·4710	11·3	1·44	Cys103	SG	1·95
									0·5121	0·3350	0·4556	4·8	0·92	Asp197	OD1	2·09
K_2PtCl_4	3·2–14·0	15062	92·6	9·62	10·1	68	0·54	0·69	0·8002	0·4202	0·4357	19·9	0·50	Cys70	SG	1·70
									0·2790	0·0995	0·4946	0·0	0·05	Thr113	O	1·66
K_2HgI_4	3·2–14·0	14932	91·8	6·80	23·2	67	0·95	1·36	0·1352	0·0721	0·4700	19·0	1·82	Cys103	SG	1·91
									0·2987	0·1907	0·4463	51·9	0·52	Lys261	NZ	3·15
									0·4095	0·3264	0·4791	32·0	0·82	His201	NE2	2·67
									0·4935	0·2553	0·4661	27·1	0·14	Asn301	N	1·91
									0·5078	0·3346	0·4550	5·1	0·55	Asp197	OD1	2·07
									0·4166	0·4212	0·4818	25·0	0·14	Ca^{++} site		
									0·7997	0·4168	0·4371	8·2	0·17	Cys70	SG	1·59
$PbNO_3$	3·2–14·0	9593	59·0	14·28	22·6	65	0·96	1·22	0·4168	0·4165	0·4800	6·6	0·99	Ca^{++} site		
									0·4938	0·3271	0·4544	5·4	0·81	Asp197	OD1	2·85
									0·0388	0·2506	0·4668	6·2	0·39	Asp356	OD1	1·84
									0·3071	0·2349	0·4533	6·2	0·11	Leu237	O	2·84
OCMP‡ & $PtBr_2$	3·2–14·0	15067	92·6	7·66	19·0	61	1·37	2·05	0·1361	0·0719	0·4710	20·4	1·79	Cys103	SG	1·96
									0·5129	0·3353	0·4561	20·1	1·03	Asp197	OD1	2·13
									0·5213	0·2902	0·4526	20·0	0·10	Asp300	OD2	2·37
									0·8043	0·4181	0·4311	20·0	0·09	Cys70	SG	2·01
$PbNO_3$§	3·5–14·0	9006	72·3	—	19·5	70	0·74	0·80	0·4192	0·4164	0·4805	12·6	1·02	Ca^{++} site		
									0·4956	0·3252	0·4543	15·4	0·70	Asp197	OD1	2·94
									0·0811	0·2530	0·4860	12·5	0·18	Asp356	OD1	1·84
									0·8007	0·4194	0·4291	0·0	0·17	Cys70	SG	2·01
B. *Heavy-atom set no. 2*																
$PtCl_4$§	5·0–14·0	2507	59·1	—	18·4	63	0·83	0·97	0·7962	0·4230	0·4384	19·9	1·25	Cys70	SG	1·61
									0·8330	0·0728	0·4171	20·0	0·14	thr143	O	1·23
K_2HgI_4§	5·0–14·0	2040	48·1	—	22·4	62	0·74	1·03	0·1365	0·0713	0·4684	19·0	1·06	Cys103	SG	2·11
									0·4091	0·3286	0·4742	21·9	1·09	His201	NE2	2·51
$HgAc_2$§	4·8–14·0	2450	50·9	—	22·5	69	0·47	0·65	0·5453	0·3117	0·3934	9·6	0·19	Gln41	NE2	1·20
									0·1239	0·0732	0·4660	9·4	0·90	Cys103	SG	1·57
U_2O_7§	4·8–14·0	3606	74·8	—	20·4	63	0·87	1·10	0·4913	0·3242	0·4469	8·4	0·59	Asp197	OD1	2·79
									0·5077	0·3044	0·4747	15·1	0·57	Asp300	OD2	2·03
									0·5881	0·1238	0·3271	0·0	0·14	Glu390	OE1	2·43
									0·0560	0·0409	0·5064	0·0	0·13	Met102	SD	3·46

† R_{sym} for heavy-atom derivatives are based on the full data set rather than the resolution ranges listed above for the heavy-atom refinement.

‡ o-chloromercuriphenol.

§ These data sets were collected on a diffractometer as described in the text.

to native by the Fourier-Bressel procedure as programmed by Weissman (Terwilliger *et al.*, 1982). Derivative-to-native residuals are found in Table 2. Difference Patterson and Fourier maps using 14 to 5·0 Å resolution diffractometer data were used to confirm previously determined substitution sites and locate new sites. Five of the putative isomorphous derivatives were found substituted and useful in MIR phasing. Inclusion of diffractometer data significantly contributed to the quality of low resolution phases as judged by statistical criteria.

(e) *Structure solution*

Optimization of heavy-atom coordinates, relative occupancies and thermal parameters were initial carried out by refinement against an origin-removed Patterson map (Rossmann, 1960) followed by conventional Blow & Crick (1959) refinement as programmed by Terwilliger (Terwilliger & Eisenberg, 1983). Two sets of heavy-atom derivatives were refined independently of each other. Set 1 contained the derivatives with data greater than 4·8 Å resolution and included 5 data sets collected on the area detector (3·2 Å) and 1 diffractometer data set (3·5 Å); set 2 contained 4 data sets collected on the diffractometer to a resolution of 4·8 Å. Hendrickson Lattman coefficients (Hendrickson & Lattman, 1970) were calculated for each set, and appropriately merged. The refinement statistics for all deriviatives are shown in Table 2. The overall figures-of-merit were 0·69 for set 1 and 0·50 for set 2.

The electron density map at 3·2 Å resolution provided a clearly demarcated envelope of the single α-amylase molecule in the asymmetric unit. Each molecule was virtually isolated from every other and appeared completely surrounded by water which comprises more than 70% of the unit cell. In the MIR map 7 helices and 8 strands of the major domain α/β-barrel were clear as were the 8 strands of the antiparallel β-barrel domain at the carboxyl terminus. Connecting loops and more extended structural elements protruding into the solvent were occasionally ambiguous though most could be traced. The most problematic electron density corresponded to a 69 amino acid residue insert between β-strand 3 and α-helix 3 of the α/β-barrel. This β_3/α_3 excursion corresponds to the B domain of the structure reported by Buisson *et al.* (1987).

Approximately 90% of the amino acid sequence could be confidently fitted to the MIR phased electron density including virtually all of the α/β-barrel domain and the small carboxyl terminal antiparallel β-barrel. The remaining 10% of the sequence was positioned according to our best estimation. Alternation of least-squares refinement and $2F_o-F_c$ difference Fourier maps, with questionable segments omitted from the phasing, ultimately yielded unambiguous placement of all 496 residues.

(f) *Structure refinement*

The initial model, M1, containing all 496 residues had an R factor value of 0·48 for 3·8 to 8 Å data. Refinement by conjugate gradient minimization (CGM) with the program X-PLOR (Brunger, 1988, 1991; Brunger *et al.*, 1987) produced model M6 with $R=0·39$ for 2·8 to 8 Å data. Refinement using simulated annealing (SA) was begun employing a temperature gradient of 3000°→300° K in 25° increments and a timestep of 0·0005 ps. Model M10, with an R value of 0·25, was compared to the original MIR map resulting in selection of 9 problematic regions (ranging from 18 to 47 residues) for which 9 $2F_o-F_c$ omit maps were calculated. The structural model was manually rebuilt to these maps using FRODO. A second set of native data to 2·1 Å resolution was collected and merged with data set 1 (see Table 2). This combined set was partitioned into a working set (~90%) and a test set (~10%) as described by Brunger (1992). The test set was used to calculate the statistical quantity R_{free} which was subsequently used to monitor the course of the refinement.

Continued SA and CGM refinement yielded model M17 with $R=0·218$ and $R_{free}=0·299$. Rebuilding to a $2F_o-F_c$ map required major changes in the region 116 to 160 with only minor changes elsewhere. Continued SA refinement produced model M19 with $R=0·192$ ($R_{free}=0·265$). A $2F_o-F_c$ map now yielded clearly interpretable density for residues 116 to 160 and this region was rebuilt accordingly along with minor adjustments elsewhere.

Refinement was continued by SA extending the resolution from 2·8 to 2·1 Å by 0·1 Å increments to yield model M30 with an R-factor of 0·215 ($R_{free}=0·264$). This and subsequent models were used to phase F_o-F_c and $2F_o-F_c$ maps from which water molecules, ions and a disaccharide fragment were added. The present model M48 consists of the 496 protein residues, 145 water molecules, a 1→4 disaccharide modeled as glucose units, a Ca ion and a Cl ion. The disaccharide occupancy was fixed at 50% to give refined B values in the B value range of the protein.

(g) *Computer programs*

Refinement with X-PLOR utilized the parameter set of Engh & Huber (1991) with variations in the force constants. Solvent accessible surface areas were calculated with a 1·4 Å probe using the procedure of Lee & Richards (1970) as implemented in X-PLOR. Three-dimensional Fourier maps were computed using the program FFT written by Ten Eyck *et al.* (1976). Visualization of density, model building, and subsequent computer graphics analysis of the structure were carried out on Evans and Sutherland PS390 and SGI graphics systems using the programs FRODO (Jones, 1985) and RIBBONS (Carson & Bugg, 1986), respectively. The programs PROCHECK (Laskowski *et al.*, 1992) and X-PLOR were used for model analysis. Computers used during this study included a VAX 8870, an SGI 420 DX, and the Cray Y-MP at the San Diego Supercomputer Center.

3. Results

(a) *Quality of the structure*

Model M48 resulted in an R-factor of 0·168 ($R_{free}=0·213$) with a correlation coefficient of 0·939 for 2·1 to 8 Å data and $F\geq 5\sigma_F$. The refinement statistics and measures of the quality of the geometry of the model are presented in Table 3. The peptide bonds are essentially planar ($\langle\Delta\omega\rangle=0·6°$). A Ramachandran plot (Ramachandran & Sasisekharan, 1968; Morris *et al.*, 1992) of the ϕ/ψ angles is shown in Figure 1. No residues are in disallowed regions and only Ser414 is in a generously allowed region (as determined by the program PROCHECK). The Luzatti (1952) plot in Figure 2 shows that the upper estimate of the error in the atomic positions is 0·25 Å. Figure 3 shows variation by residue number in the B values for (a) backbone

Table 3
Refinement statistics

	R			R_{free}		
Resolution (Å)	Number of refl.	Shell	Sphere	Number of refl.	Shell	Sphere
4·04–8·00	6055	0·1539	0·1539	686	0·2105	0·2105
3·28–4·04	5510	0·1385	0·1470	620	0·1873	0·1873
2·88–3·28	4578	0·1628	0·1505	468	0·1982	0·1996
2·63–2·88	4648	0·1891	0·1560	524	0·2264	0·2034
2·45–2·63	4558	0·1991	0·1606	533	0·2338	0·2067
2·31–2·45	4044	0·2124	0·1649	406	0·2577	0·2105
2·19–2·31	2799	0·2163	0·1676	357	0·2544	0·2133
2·10–2·19	552	0·2152	0·1681	69	0·2225	0·2134
Total	32,744			3663		
All data	42,400		0·2074	4749		0·2437

r.m.s.d. from ideal	Protein	Disaccharide
A. *Geometry*		
Bonds (Å)	0·010	0·018
Angles (°)	2·10	6·46
Dihedral (°)	25·4	20·7
Improper (°)	0·50	0·30
ω (0°, 180°)	0·61	—

Groups	min B	max B	$<B>$
B. *Temperature factors (in Å²)*			
Main-chain atoms (1–496)	2·0	72·5	17·5
Side-chain atoms (1–496)	2·0	91·7	19·6
Main-chain atoms (1–99 and 169–403)	2·0	72·5	16·6
Side-chain atoms (1–99 and 169–403)	2·0	91·7	18·6
Main-chain atoms (100–168)	12·2	44·3	24·0
Side-chain atoms (100–168)	7·3	78·6	26·0
Main-chain atoms (404–496)	3·5	53·0	17·6
Side-chain atoms (404–496)	3·6	77·5	20·3
Water molecules (all)	2·0	56·0	26·0
Disaccharide (Q=0·5)	24·5	37·4	29·2
Ca^{++}	-		2·0
Cl^-			14·2

X-PLOR weight W_A=440,000.
Total reflections with $F \geq 5\sigma_F$ is 36,407 reflections.
Working reflection set=32,744 reflections.
Test reflection set=3663 reflections.

atoms, (b) side-chain atoms and (c) in the distance between the C^α positions of the initial model and model M48. There are several areas of the structure that appear to be flexible or mobile, especially the region containing residues 100 to 168 and the loop from 347 to 369. Region 100 to 168 was the most difficult to build into the density and was only accomplished after considerable refinement. The loop region had poor or broken density for side-chains in the final $2F_o - F_c$ map. However, the electron density map was generally well defined as illustrated by the $2F_o - F_c$ omit map in Figure 4. Residues 53 to 60 have been superimposed on the density demonstrating the quality of the fit of model to density. The 145 water molecules have full occupancy and B values in the range 2 to 56 Å^2 ($\langle B \rangle$=26 Å^2). No attempt was made to refine the occupancies. The disaccharide, however, had individual atomic B values in excess of 100 Å^2 when occupancies were 1·0. Fixing the occupancies at 50% resulted in $\langle B \rangle$=29 Å^2.

(b) *The tertiary structure*

Diagrams of α-amylase emphasizing the structural elements are shown in Figure 5. The first 403 amino acid residues form a large domain (domain A) whose central components are a parallel β-barrel of eight strands interconnected serially by helices and extensive loops of extended chain. The topology is that of the commonly observed α/β or TIM barrel. There is a complex excursion (β_3/α_3 excursion) from the α/β-barrel of about 69 residues between strand 3 and helix 3 that contains both a small four-stranded antiparallel sheet and a helix. The tightly bound calcium ion is liganded in part by components of this excursion as well as residues from the domain core. Ca^{2+} clearly plays a structural role by tying otherwise extended chain into a single, cohesive unit. The poorly defined electron density and generally larger thermal factors (Table 3) associated with the β_3/α_3 excursion suggest it to be mobile.

Ninety-three amino acid residues at the carboxyl

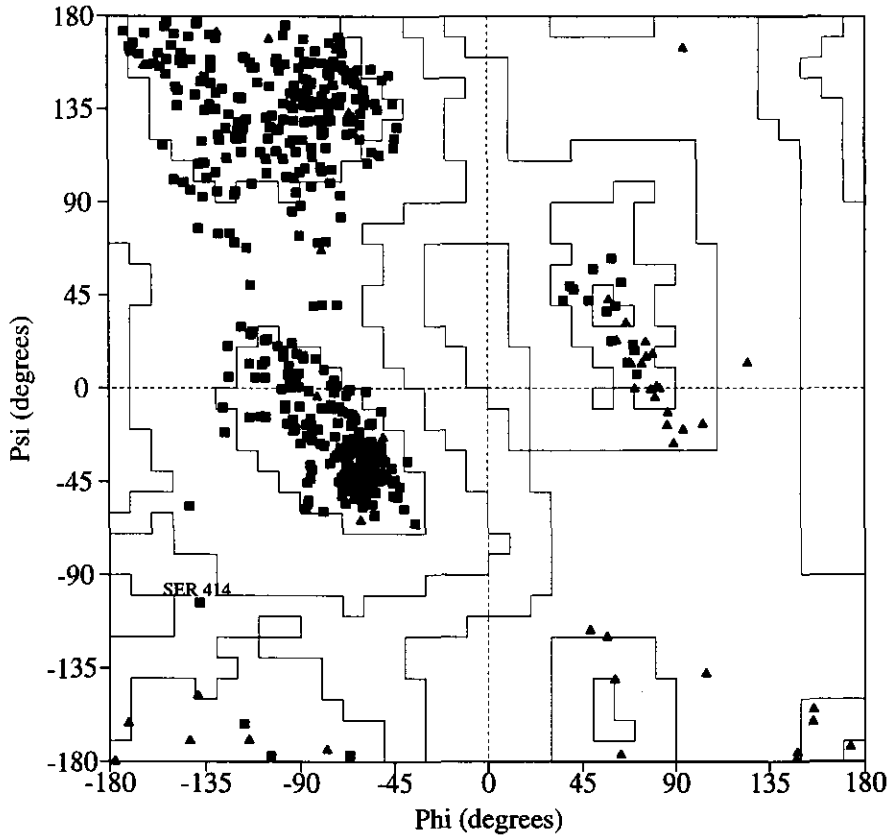

Figure 1. A Ramachandran plot (Ramachandran & Sasisekharan, 1968) of model M48 produced with PROCHECK (Laskowski *et al.*, 1992). Glycine residues are shown in triangles (▲). There are no residues in disallowed regions and only Ser414 is in a generously allowed region. More than 86% of the residues are in most favored regions.

terminus form the small, compact antiparallel β-barrel (domain B) of Greek Key topology (Richardson, 1981; Branden & Tooze, 1991). This domain lies very tightly against helices 7 and 9 of the α/β domain to give the entire α-amylase molecule an ellipsoidal appearance, with axes roughly 50 Å × 50 Å × 75 Å.

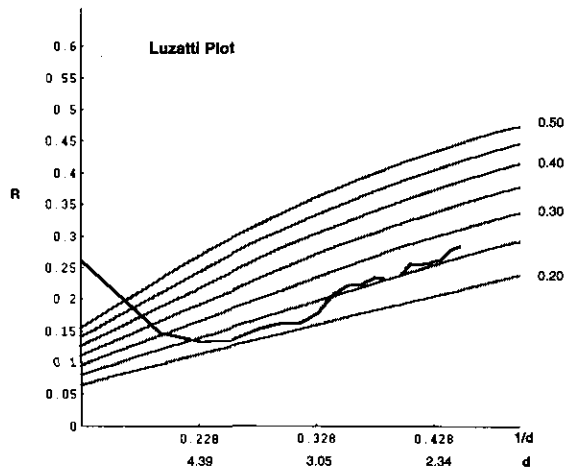

Figure 2. A Luzatti (1952) plot showing the variation of the crystallographic residual *R* with respect to resolution for the final model M48. The plot suggests that the upper estimate in the error in atomic positions is 0·25 Å.

As seen in Figure 5(a), the amino terminal end of the central β-barrel is rather blunt. The helix-to-strand connections are fairly brief with most loops having few residues. This is in striking contrast to the carboxyl end of the β-barrel where the extensive connecting loops are meandering, and flair out to form a long, deep cleft. As indicated by inhibitor complex studies, and other evidence presented below, this long trench is the polysaccharide binding region and contains the catalytic center.

The architecture of the molecule is such that the α/β-barrel forms a compact, stable and probably rigid superstructure. The long active site cleft, which is anticipated from substrate binding analyses (Robyt & French, 1970*a*,*b*; Prodanov *et al.*, 1984), is fashioned across one end by extended connecting loops, which, in several cases, appear to be mobile. The carboxyl terminus, B domain seems quite distal to the active site cleft, and if it participates at all in substrate binding, must do so at points well removed from the site of cleavage.

There are five disulfide bridges in PPA (Pasero *et al.*, 1986). The bridge between Cys28 and Cys86 connects helix 1 and helix 2 of the α/β-barrel. The β_3/α_3 excursion is linked to the A domain through a Cys70–Cys115 bridge. Cys141 is bonded to Cys160 to tie a meandering loop in the β_3/α_3 excusion to the Ca^{2+}-bound segment of the excursion at residue 158. Loop 378 to 384 is stabilized by the

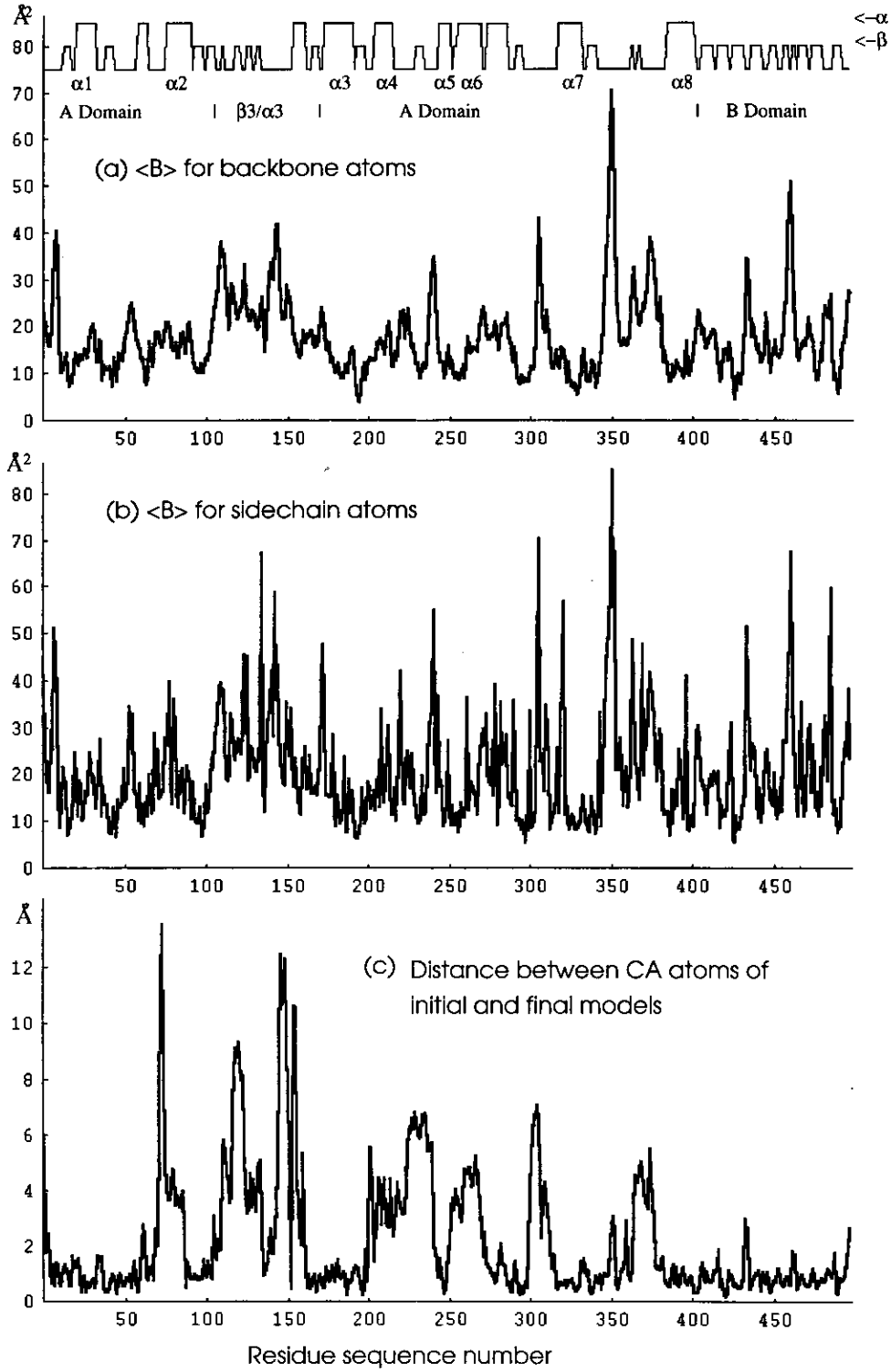

Figure 3. The variation in average atomic B values as a function of residue number is shown for (a) the backbone atoms and (b) the side-chains atoms. In (c) is shown the distance between C^{α} positions of the initial and final models. Secondary structure is indicated at the top of the Figure. Some correlation is seen between regions of large errors in the initial model and regions of large thermal parameters, particularly in the β_3/α_3 excursion.

Cys378–Cys384 bridge. In the small domain the Cys450–Cys462 bridge, the only *trans* disulfide bridge in PPA, holds a loop secure to the β-barrel. The only other cysteines in PPA are Cys103 and Cys119 in the β_3/α_3 excursion. These residues are situated such that the inter-sulfur distance is 4·7 Å. It is difficult to understand why these residues do not react to form a sixth disulfide bridge.

In the course of analyzing the water structure, we discovered a short oligosaccharide bound at the extreme edge of the active site cleft. Only a disaccharide is clearly defined in electron density maps

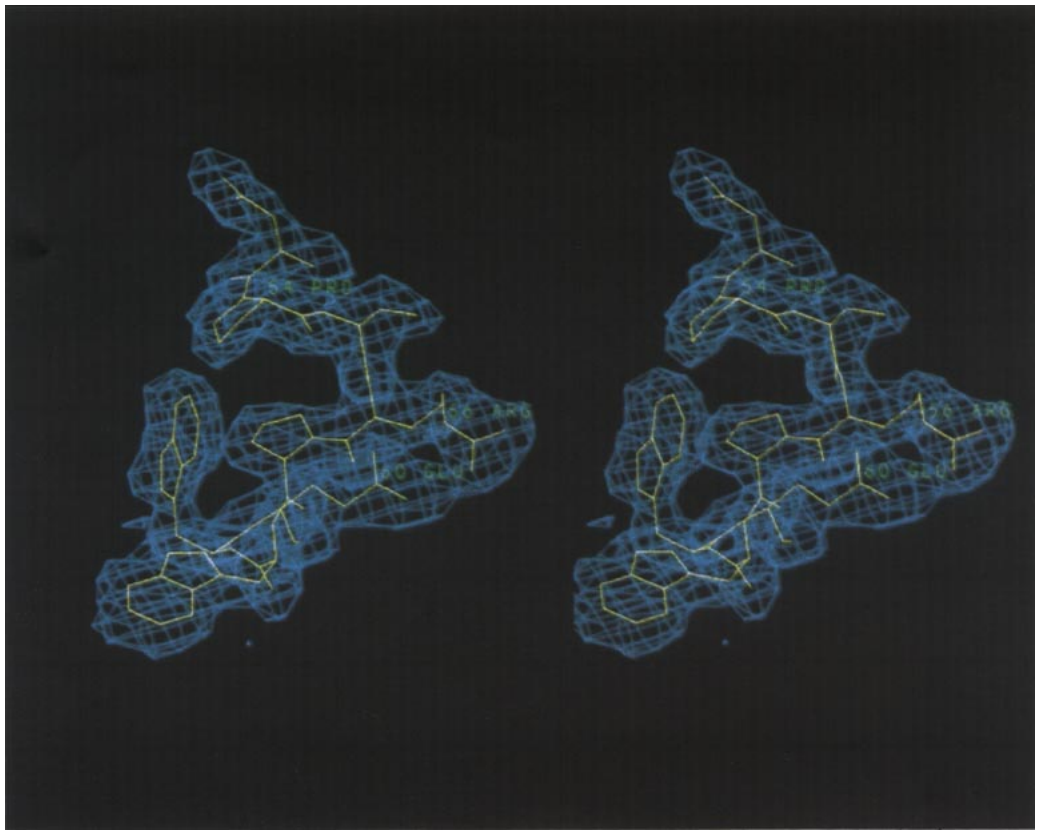

Figure 4. An example of a representative region of a $(2F_o - F_c)$ exp $i\alpha_c$ electron density map of PPA in which amino acid residues 50 to 60 have been omitted. Residues 53 to 60 of the refined model have been superimposed on the density. This region includes *cis*-Pro54 and the Arg56–Glu60 salt bridge. The side-chain densities are generally equally well-defined throughout the map, though of poorest quality in the β_3/α_3 excursion between amino acid residues 100 and 168.

but weak density is present that suggests a third, disordered sugar ring could be present and that the oligosaccharide may be a triose. As illustrated in Figure 6, each oligosaccharide is bound, in fact, not to a single enzyme molecule, but bridges the interval between two protein molecules related along x by a 2_1 symmetry operator. Two protein molecules contribute to the binding of both glucose residues comprising the carbohydrate. Thus, a composite substrate binding site is formed by screw-axis related PPA molecules in the crystalline lattice. Interestingly, the disaccharide lies at one end of the polysaccharide binding cleft and a 2_1-related disaccharide lies at the opposite end of that cleft. Thus, the bound oligosaccharide marks the extremes of the long starch binding region and fixes its overall course with respect to the protein.

(c) *Crystal packing*

The crystals of PPA (form I) used in this structure determination are somewhat unusual in that they contain nearly 75% solvent by volume (Table 1). The asymmetric unit could, in fact, easily accommodate an entire additional protein molecule and still fall near the center of the range for V_m (Matthews, 1968). Individual molecules in the lattice are quite separated from one another and intermolecular contacts are sparse. As seen in Table 4 there are only 11 hydrogen bonds likely present between any one α-amylase molecule and its neighbors in the crystal lattice. There are very large voids in the crystals of 40 Å or more across and vast networks of channels elsewhere. The carbohydrate binding sites are entirely accessible as evidenced by the difference Fourier experiments presented below.

A puzzling feature of the α-amylase crystals used for this study is that they are mechanically sturdy. They were among the first enzymes crystallized (Caldwell *et al.*, 1931), show good optical properties, and they diffract strongly to 2·0 Å resolution. This is unexpected given the tenuous links between protein molecules in the crystals. Ordered water may play a significant role in maintaining the physical integrity of the crystal but the non-covalent crosslinking of protein molecules along the x direction through mutually bound disaccharides undoubtedly contributes as well. Indeed, these maltose residues produce linear chains of α-amylase molecules extending along the x direction throughout the crystals. It is not surprising that these crystals of

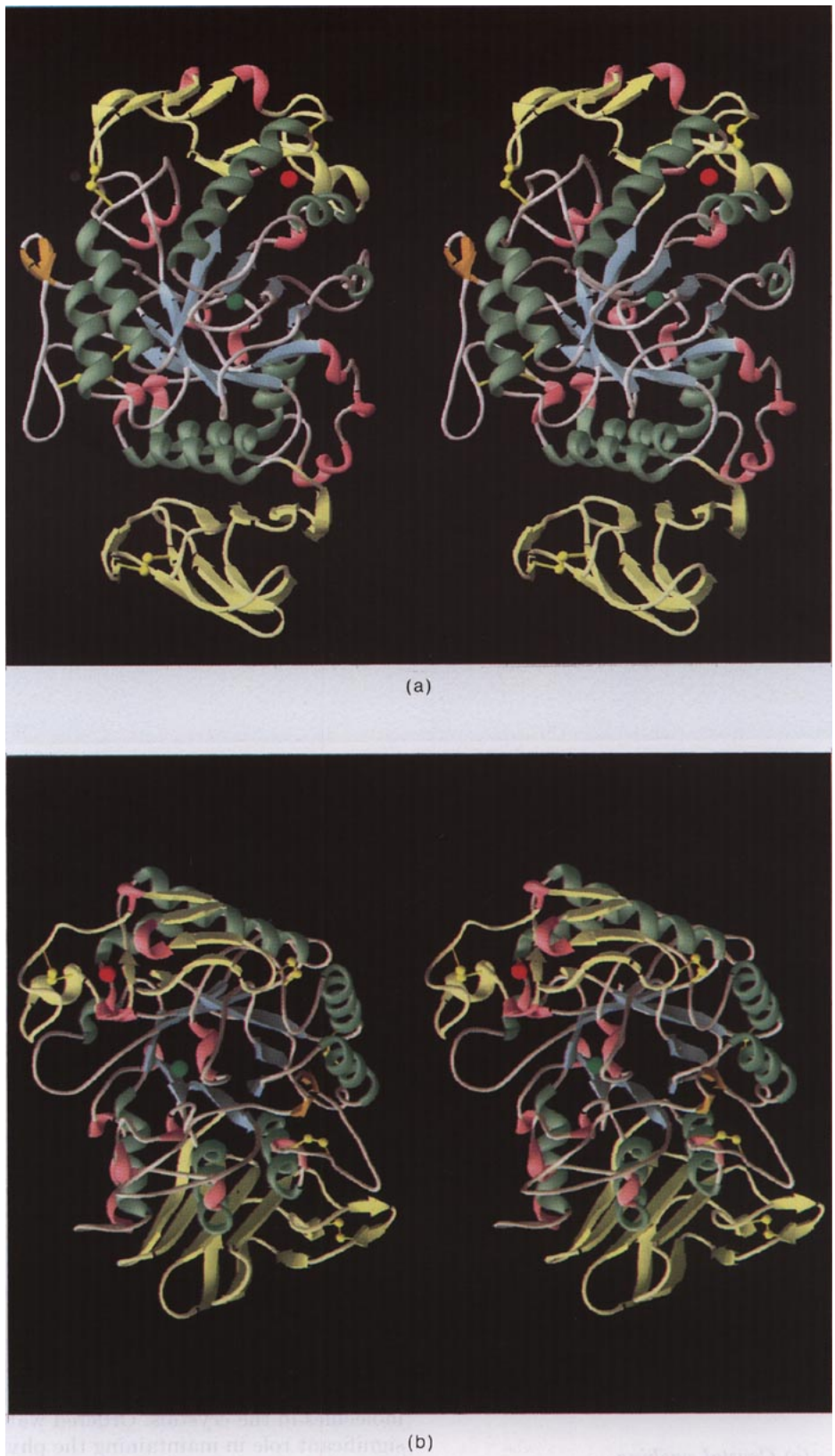

Figure 5. Schematic representations of pancreatic α-amylase using the program RIBBONS (Carson & Bugg, 1986). The helices and strands of the major domain α/β-barrel are shown in green and blue, respectively. The compact β-barrel domain at the carboxyl end is in yellow as is the calcium ion binding excursion between β-strand 3 and α-helix 3. The 5 disulfide bridges are yellow; the calcium is a red sphere; and the chloride ion is a green sphere. 3_{10}-Helix elements are pink. In (a) is featured the blunt, amino terminal end of the β barrel while in (b) the view is approximately along the axis of the α/β-barrel from the carboxyl terminal end.

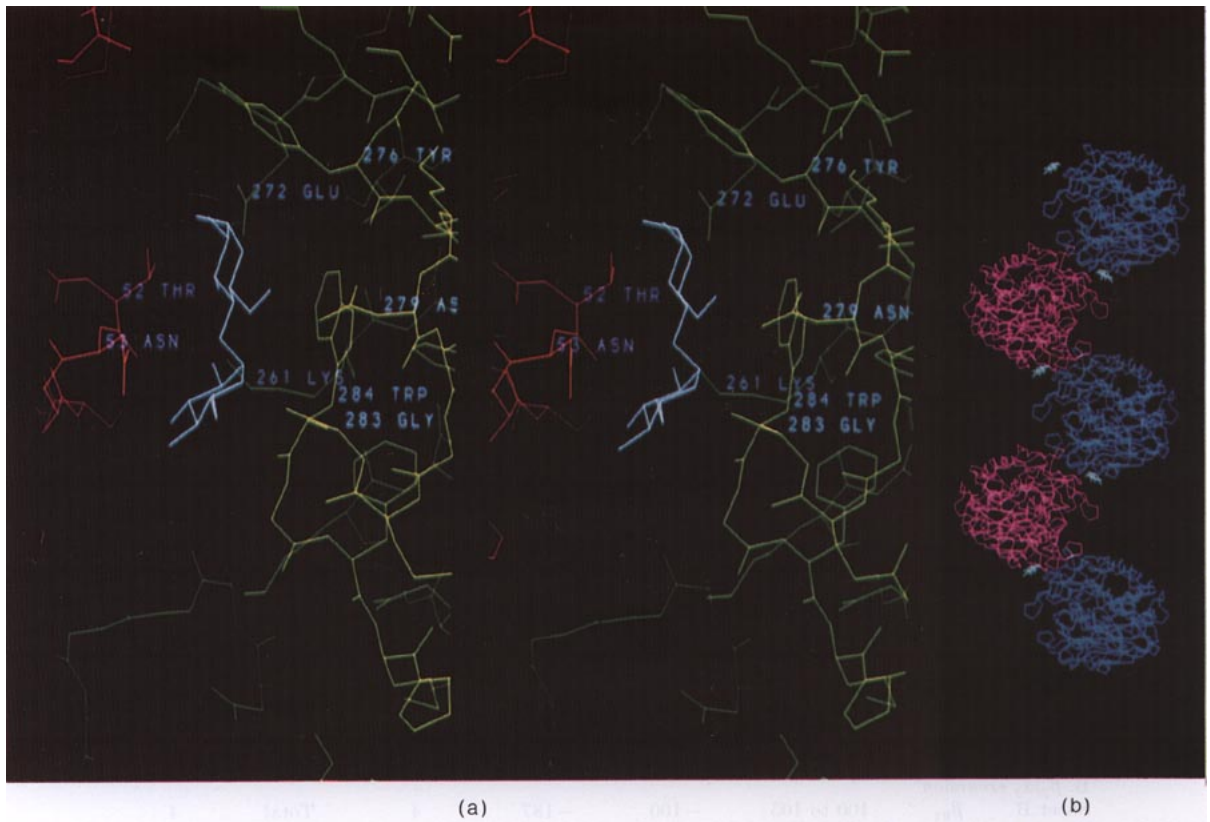

Figure 6. In (a) the disaccharide (in blue) bridges two 2_1-axis related enzyme molecules (in red and yellow). The amino acid residues in the 2 protein molecules that are responsible for saccharide binding are labeled. Some weaker density is found at one end of the disaccharide suggesting a 3rd sugar residue may be present, but if so it is less ordered. The disaccharide is bound by a composite arrangement in that both protein molecules contribute to the binding of both glucose residues. This point of contact in the crystal is responsible for the majority of the intermolecular interactions observed between all protein molecules in the lattice. By linkage through these disaccharide units, extended chains, like those seen in (b), are formed such that the polysaccharide binding clefts of the molecules are contiguous. Thus, continuous chains of enzyme molecules might be expected when α-amylase is associated with linear polysaccharides or large limit dextrins.

Table 4
Possible intermolecular hydrogen bonding interactions

Donor (D)		Acceptor (A)		d(D...A) (Å)	Symmetry operator applied to A
Ser471	N	Asp181	OD1	3·20	$1-x, -\frac{1}{2}+y, \frac{1}{2}-z$
Tyr468	OH	Ser177	O	3·33	$1-x, -\frac{1}{2}+y, \frac{1}{2}-z$
Tyr468	OH	Asp181	OD2	2·73	$1-x, -\frac{1}{2}+y, \frac{1}{2}-z$
Thr439	OG1	Asp188	OD1	2·83	$1-x, -\frac{1}{2}+y, \frac{1}{2}-z$
Lys474	NZ	His215	NE2	2·99	$1-x, -\frac{1}{2}+y, \frac{1}{2}-z$
Asn216	N	Gln476	OE1	2·86	$1-x, \frac{1}{2}+y, \frac{1}{2}-z$
Ala345	N	Asn152	OD1	3·38	$\frac{1}{2}+x, \frac{1}{2}-y, 1-z$
Asn364	ND2	Gly306	O	3·28	$\frac{1}{2}+x, \frac{1}{2}-y, 1-z$
Gly308	N	Asn364	OD1	2·86	$-\frac{1}{2}+x, \frac{1}{2}-y, 1-z$
Asn152	ND2	Trp344	O	3·07	$-\frac{1}{2}+x, \frac{1}{2}-y, 1-z$
Ser150	OG	Asp381	OD2	2·80	$-\frac{1}{2}+x, \frac{1}{2}-y, 1-z$

the native PPA contain residual bound carbohydrate in light of the glycogen digestion procedure used to prepare the enzyme.

(d) *Secondary structure*

Amino acid sequences making up the eight strands of the parallel β-barrel, and the attendant α-helices are presented in Table 5. Several of these strands maintain their approximately linear geometry upon completing participation in the β-barrel, and extend well into the connecting loops. This is particularly true of extended chains composing the active site cleft. The hydrogen bonding arrangement among these strands is shown in Figure 7(a). The slightly eliptical barrel has a mean diameter of about 13 to 14 Å and good strand geometry. The core of the β-barrel is filled with amino acid side-chains, principally hydrophobic in character, that would appear to exclude the instrusion of water.

The α-helices, similarly, display quite regular geometry and maintain good hydrogen bonding

Table 5
Secondary structural elements in PPA

Structure type		Region	ϕ	ψ	Number of bonds	d(N...O)(σ)	
A. *Domain A*							
Sheet I	β_1	12 to 16	−105	142	6		
Sheet I	β_2	38 to 42	−104	122	7		
Sheet I	β_3	92 to 97	−99	131	9		
Sheet I	β_4	192 to 199	−69	92	11		
Sheet I	β_5	229 to 233	−107	126	9		
Sheet I	β_6	252 to 255	−96	143	6		
Sheet I	β_7	291 to 294	−107	81	5		
Sheet I	β_8	335 to 340	−109	135	5		
		Average	−97	121	Total 29	2·91(13)	
α_R	α_1	20 to 31	−62	−43	7	2·96(9)	
α_R+3_{10}	α_2	75 to 90	−65	−39	12	2·95(9)	
α_R	α_3	172 to 190	−62	−41	15	2·94(11)	
α_R+3_{10}	α_4	203 to 214	−62	−36	10	3·01(14)	
3_{10}	α_5	243 to 250	−68	−22	3	3·06(14)	
$3_{10}+\alpha_R$	α_6	255 to 268	−65	−39	8	2·98(13)	
α_R	α_7	317 to 331	−60	−45	10	2·93(10)	
α_R	α_8	388 to 402	−65	−39	12	2·95(13)	
		Average			Total 77	2·96(12)	
3_{10}	α_A	57 to 63	−64	−21	3	3·07(8)	
3_{10}	α_B	273 to 280	−61	−23	3	2·88(4)	
3_{10}	α_C	281 to 285	−52	−30	2	2·83(13)	
α_R+3_{10}	α_D	298 to 304	−72	−24	3	3·19(9)	
Sheet A	β_{A1}	361 to 363	−136	144	2	Total	2
Sheet A	β_{A2}	366 to 368	−68	124	2	Ave 2·75(3)	
B. β_3/α_3 *excursion*							
Sheet B	β_{B1}	100 to 105	−100	−187	4	Total	4
Sheet B	β_{B2}	164 to 168	−73	119	4	Ave 2·90(10)	
Sheet C	β_{C1}	107 to 110	−120	147	2		
Sheet C	β_{C2}	117 to 120	−114	154	4	Total	6
Sheet C	β_{C3}	124 to 126	−93	72	4	Ave 2·99(13)	
Sheet C	β_{C4}	130 to 132	−138	154	2		
α_R	α_E	153 to 160	−70	−36	4	2·94(15)	
C. *Domain B*							
Sheet D	β_{D1}	405 to 412	−92	135	6		
Sheet D	β_{D2}	415 to 422	−140	144	12	Total	18
Sheet D	β_{D3}	424 to 431	−139	136	12	Ave 2·83(17)	
Sheet D	β_{D4}	486 to 492	−142	150	6		
Sheet E	β_{E1}	436 to 442	−115	127	7		
Sheet E	β_{E2}	447 to 451	−96	148	5	Total	14
Sheet E	β_{E3}	465 to 470	−100	131	7	Ave 2·84(9)	
Sheet E	β_{E4}	473 to 479	−121	115	9		
Sheet F	β_{F1}	456 to 459	−56	136	2	Total	2
Sheet F	β_{F2}	461 to 463	−104	160	2	Ave 2·95(16)	

over their entire lengths. Three of the eight begin or end with 3_{10}-helical geometry while a fourth, helix α_5 is a true 3_{10}-helix. To our knowledge this is a unique example of a 3_{10}-helix in an α/β-barrel. In addition, we find four other helical segments in domain A, three of which are true 3_{10}-helices (α_A, α_B, α_C in Table 5) and the fourth (helix α_D) ends with 3_{10}-helical geometry and hydrogen bonding pattern.

The β_3/α_3 excursion (amino acid residues 100 through 168) contains a small four-stranded antiparallel β-sheet (sheet C in Table 5) having a repetitive "up-down" topology. This sheet, though rather distant from the polysaccharide binding cleft, contributes substantially to the formation of an accessory carbohydrate binding site as described below. In addition, there is an α-helix in this excursion as well as a two-stranded antiparallel β-sheet (α_E and sheet B in Table 5).

The carboxyl terminal B domain of PPA is shown in Figure 5(b) and its hydrogen bonding arrangement is presented in Figure 7(b). It appears to be remote from the polysaccharide binding cleft. We have no evidence suggesting relevance to carbohydrate binding, and it is not apparent that it serves to maintain structural elements of the A domain. This domain is a flattened, antiparallel β-barrel comprised of two sheets (D and E in Table 5) of four strands rotated with respect to one another by about 40°. The cavity between the two sheets is filled with predominantly hydrophobic side-chains including Trp410, Phe426, Phe477 and Tyr449. The two-stranded β-sheet F resulting from

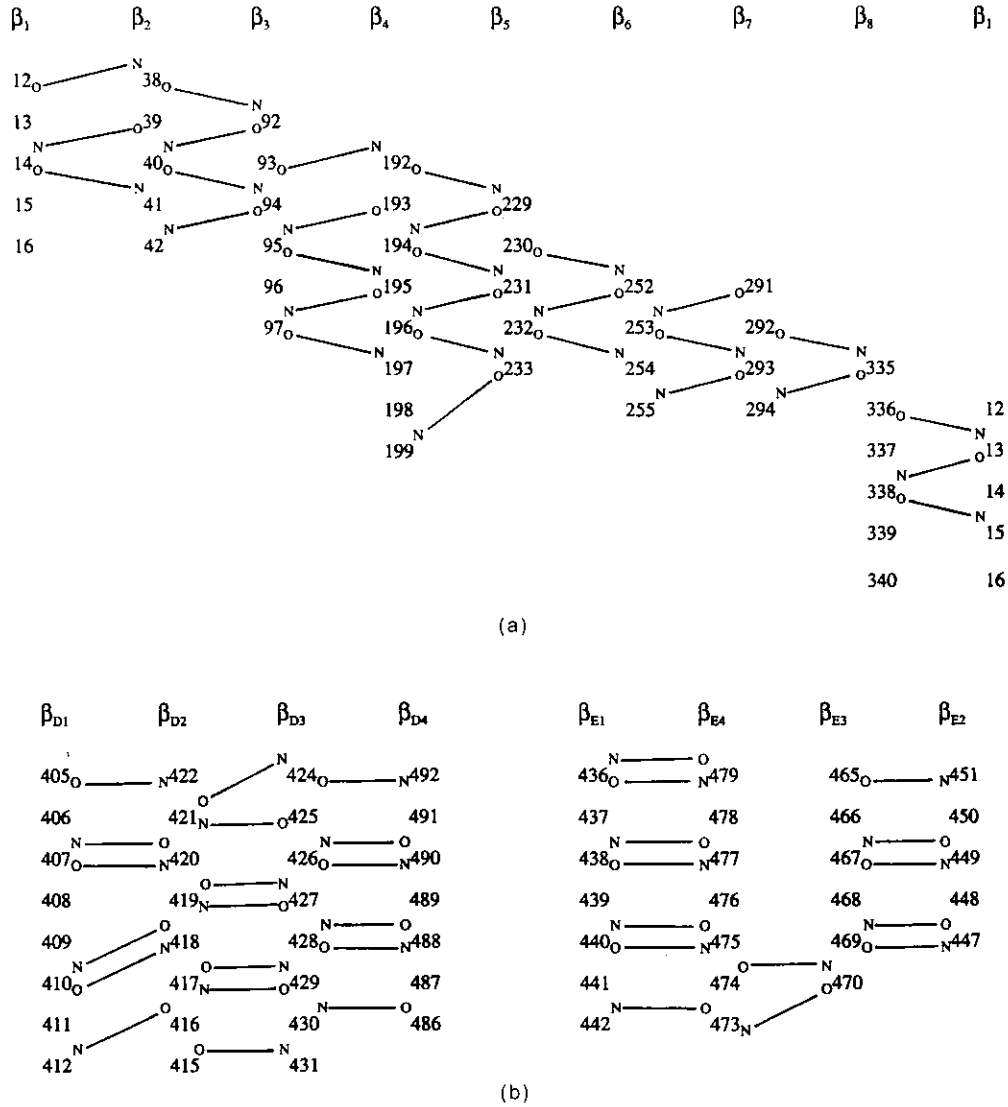

Figure 7. Schematic diagram of the hydrogen bonding pattern between strands of (a) the β-barrel of domain A and (b) the β-sheets D and E of domain B.

the tight turn in loop 452 to 464 is secured to the β-barrel by a disulfide bridge between Cys450 and Cys462.

The α-amylase structure contains 32 tight turns as defined by a hydrogen bonding pattern of $i \leftarrow i+3$ (except for the *cis*-Pro and γ-turns). These turns are listed along with the summary statistics in Table 6. The two *cis*-Pro turns exhibit very distinct ϕ/ψ angles as do the three γ-turns. The γ-turn is a tight three-residue turn with a hydrogen bonding pattern of $i+1 \leftarrow i+3$. The 364 to 366 and 459 to 461 turns make hairpin turns for two-stranded β-sheets such that the first hydrogen bond of the sheet is $i \rightarrow i+3$. The other γ-turn exhibits this same pattern of two hydrogen bonds but does not lead into a two-stranded sheet.

Two consecutive glycine residues occur four times in the sequence (38 to 39, 146 to 147, 238 to 239 and 308 to 309) and consecutive prolines twice (44 to 45 and 360 to 361). In all cases but one, these occur in regions of extended chain or large loops; the exception, 308 to 309, occurs in a turn. Proline-glycine, or the reverse pair, occurs four times in the structure, once at the beginning of a helix and three times in extended chain. The two *cis*-prolines (residues 54 and 130) in the molecule are in regions involved in carbohydrate binding.

(e) *The interaction between domains*

The small domain is joined to the major α/β domain by only a short hinge of residues Asp402-Gly-Glu-Pro-Phe-Ala407 that might be expected to confer independence to the small domain. There is, however, a broad, cohesive interface that would seem to mediate against movement. This interface is formed by juxtaposition of one sheet of the B domain directly against the surfaces of α-helices 7 and 8 of the A domain. Between polypeptide strands making up these secondary

Table 6
Summary of tight turns in PPA

Residues	Secondary elements	Type	ϕ_2 (°)	ψ_2 (°)	ϕ_3 (°)	ψ_3 (°)	$i \leftarrow i+3$ (Å)
7–10	$-\beta_1$	II	−45	138	94	−20	2·93
16–19	$\beta_1-\alpha_1$	II	−52	113	67	12	3·16
52–55	$\beta_2-\alpha_A$	*cis*-Pro(VI)	−136	93	−71	170	
71–74	$\alpha_A-\alpha_2$	I	−65	−13	−80	−27	3·06
104–107	$\beta_{B1}-\beta_{C1}$	I	−53	−33	−82	−4	2·94
113–116	$\beta_{C1}-\beta_{C2}$	I	−64	−25	−114	12	3·31
121–124	$\beta_{C2}-\beta_{C3}$	I	−50	−40	−92	4	2·78
127–130	$\beta_{C3}-\beta_{C4}$	*cis*-Pro(VI)	−75	−8	−87	−58	
132–135	$\beta_{C4}-\alpha_E$	I	−38	−35	−71	−9	2·70
137–140	$\beta_{C4}-\alpha_E$	I	−42	−48	−68	−34	2·83
143–146	$\beta_{C4}-\alpha_E$	I	−51	−35	−87	−17	2·82
150–153	$\beta_{C4}-\alpha_E$	I	−74	4	−71	−9	3·18
162–165	$\alpha_E-\beta_{B2}$	I′	39	49	68	12	2·82
169–172	$\beta_{B2}-\alpha_3$	I	−51	−11	−85	−15	2·84
198–201	$\beta_4-\alpha_4$	III	−57	−31	−61	−27	2·86
223–226	$\alpha_4-\beta_5$	II	−50	149	74	12	3·25
269–272	$\alpha_6-\alpha_B$	I′	36	43	75	22	2·89
288–291	$\alpha_C-\beta_7$	I	−49	−42	−73	1	2·79
306–309	$\alpha_D-\alpha_7$	II	−61	120	103	−17	2·88
308–311	$\alpha_D-\alpha_7$	II′	49	−117	−71	−29	3·13
314–317	$\alpha_D-\alpha_7$	III	−57	−18	−67	−25	2·90
349–352	$\beta_8-\beta_{A1}$	γ	−115	68	61	23	3·09
353–356	$\beta_8-\beta_{A1}$	I	−64	−14	−98	18	3·17
363–366	$\beta_{A1}-\beta_{A2}$	γ	64	51	86	−18	3·27
373–376	$\beta_{A2}-\alpha_8$	I	−51	−28	−89	16	2·60
379–382	$\beta_{A2}-\alpha_8$	I′	59	23	70	18	3·28
401–404	$\alpha_8-\beta_{D1}$	II	−43	126	86	−12	2·71
421–424	$\beta_{D2}-\beta_{D3}$	II′	57	−120	−124	20	3·07
458–461	$\beta_{F1}-\beta_{F2}$	γ	60	−141	−83	40	2·97
470–473	$\beta_{E3}-\beta_{E4}$	I	−57	−28	−94	22	2·93
480–483	$\beta_{E4}-\beta_{D4}$	I	−68	−3	−108	11	2·93
491–494	β_{D4}	I	−49	−35	−90	7	2·89

Turn type	No. of occurrences	$\langle\phi_2\rangle$	$\langle\psi_2\rangle$	$\langle\phi_3\rangle$	$\langle\psi_3\rangle$	
A. *Averages for tight turns by turn type*						
Type I	15	−55(10)	−26(15)	−88(13)	−2(17)	2·92(19)
Type I′	3	45(13)	38(14)	71(4)	17(5)	3·00(25)
Type II	5	−50(7)	129(14)	85(15)	−5(16)	2·99(22)
Type II′	2	53(6)	−118(2)	−98(37)	−4(35)	3·10(4)
Type III	2	−57(0)	−25(9)	−64(4)	−26(1)	2·88(3)

structural elements is a dense layer of hydrophobic residues provided by both domains. These include a large number of aromatic residues, five from the small domain (Phe406, Phe419, Phe429, Phe487, and Trp409) and three from the large (Phe327, Phe397 and Tyr319). This large hydrophobic interface is presumably responsible for the intimate association of the two domains and an overall molecular stability.

(f) *The chloride binding site*

PPA is allosterically activated by chloride ion, though the mechanism is uncertain (Yamamoto *et al.*, 1988; Thoma *et al.*, 1971). The bound chloride ion lies near the polysaccharide binding cleft end of the β-barrel and 2·8 Å off the barrel axis. Its environment, viewed approximately down the barrel axis, is shown in Figure 8. The carboxyl end of the barrel is segregated with polar side-chains (Arg195, Arg337 and Asn298) on one side, hydrophobic groups (Phe295, Phe256, Thr254 CH_3 and Glu233 CH_2) on the other and the Cl ion in between. The polar groups form a triangular array of ligands to the Cl^- with the Cl^- at the apex of a flattened tetrahedron. On the opposite side of the ligand triangle lies a single water molecule. The hydrogen bonding and hydrophobic interactions are given in Table 7.

The Cl^- is ~11 Å from the center of the starch binding cleft. The closest oxygen atom of the carboxylate groups of the suggested catalytic residues, Asp197, Glu233 and Asp300, are 7·1, 4·7 and 7·8 Å, respectively, from the Cl ion. Thoroughly embedded in the protein structure, the Cl^- has less than 1 $Å^2$ of solvent accessible surface area.

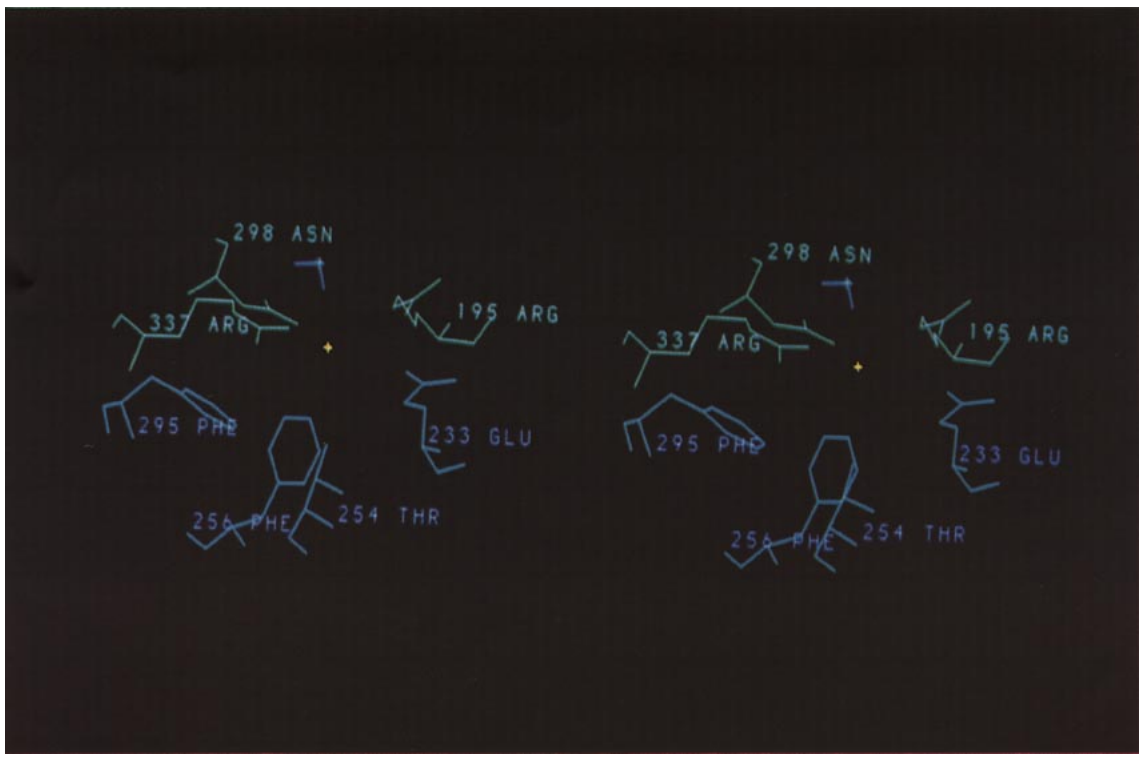

Figure 8. The chloride ion, which binds within the carboxyl end of the major domain β-barrel is neutralized by, and hydrogen bonds with, the guanidinium side-chains of arginine residues 195 and 337. In addition, Asn298 and a water molecule participate in hydrogen bonding. On the opposite side of the Cl^- from these groups is a closely packed array of hydrophobic side-chains. By virtue of the Cl^- interaction, the Glu233 side-chain is spared an otherwise probable interaction with the oppositely charged arginine groups. Because Glu233 is a likely catalytic residue, the arrangement suggests a plausible explanation for the activation of α-amylase by chloride ion.

Table 7

Interactions involving chloride and calcium ions and the disaccharide

Donor		Acceptor		Distance (Å)	Donor		Acceptor		Distance (Å)
A. *Chloride ion interactions*									
(i) *Hydrogen bonding*									
Arg195	NE	Cl^-		3·27	Arg195	NH2	Cl^-		3·37
Asn298	ND2	Cl^-		3·26	Arg337	NH1	Cl^-		3·39
Arg337	NH2	Cl		3·06	H_2O7	O	Cl		3·21
(ii) *Hydrophobic*									
Glu233	CG	Cl		3·90	Thr254	CG2	Cl		3·83
Phe256	CZ	Cl^-		4·24	Phe285	CZ	Cl^-		4·74
B. *Calcium ion interactions*									
Asn100	OD1	Ca^{++}		2·31	Asp167	OD1	Ca^{++}		2·67
Asp167	OD2	Ca^{++}		2·64	Arg158	O	Ca^{++}		2·32
His201	O	Ca^{++}		2·28	H_2O37	O	Ca^{++}		2·62
H_2O38	O	Ca^{++}		2·44	H_2O39	O	Ca^{++}		2·49
C. *Disaccharide-PPA interactions*									
glc1	O2	Thr52	O	3·22	glc1	O3	Thr52	O	3·34
glc1	O6	Gly283	O	2·95†	Thr52	OG1	glc2	O3	2·97
Thr52	OG1	glc2	O2	3·23	Asn53	ND2	glc1	O4	2·59
glc2	O2	Glu272	OE1	3·34†	glc2	O3	Glu272	OE1	3·20†
glc2	O2	Glu272	OE2	2·63†	Lys261	NZ	glc1	O2	2·78†
Asn279	ND2	glc2	O6	2·88					

† The symmetry operator for protein is $\frac{1}{2}+x, \frac{1}{2}-y, 1-z$. glc, glucose.

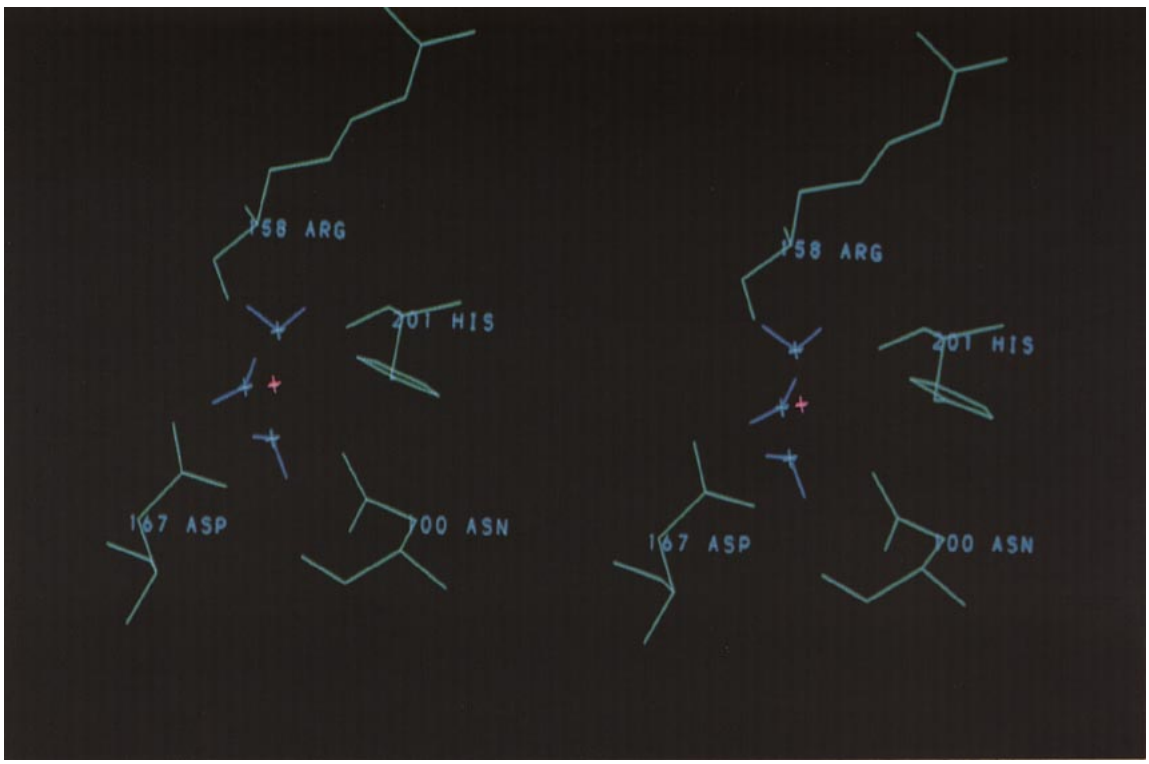

Figure 9. The coordination sphere of the tightly bound calcium ion in pancreatic α-amylase is shown here. An approximately square array is formed by 4 protein ligands, the main-chain carbonyl oxygen atoms of His201 and Arg158, and the side-chains of Asp167 and Asn100. Two water molecules reside on one side of the plane and one water on the other. The overall arrangement is octahedral. The side-chain of His201 from the major domain A has been suggested to participate in catalysis. The other 3 protein ligands are provided by the β_3/α_3 excursion.

(g) *The calcium binding site*

The environment of the tightly bound calcium ion is shown in Figure 9. It's ligand sphere involves groups from the core of the α/β domain, the β_3/α_3 excursion, and water molecules. Like the Cl^- in the α/β-barrel, the Ca^{2+} is tightly integrated into the protein structure and stitches together the α/β-barrel and the β_3/α_3 excursion to form part of the interior of the polysaccharide binding cleft. Four protein ligands are provided by two main-chain carbonyl oxygen atoms and by two side-chains. These form an approximately square planar array around the Ca ion. A water molecule occupies an apical position on one side of the square plane and two other water molecules share an equivalent position on the opposite side. These interactions are summarized in Table 7.

The carbonyl oxygen of His201 is 2·3 Å from the Ca^{2+}. This histidine residue is part of a 3_{10}-helical turn that occurs just prior to entry into α-helix 4 of the α/β-barrel. It is the only ligand provided by the core of the α/β-barrel. The carbonyl oxygen of Arg158 in the α-helix of the β_3/α_3 excursion is 2·3 Å from the Ca^{2+}. The side-chain ligands to the Ca^{2+} are the Asp167 carboxylate group, which has both oxygens atoms ~2·7 Å from the Ca ion, and the carbonyl oxygen atom of Asn100 which lies 2·3 Å from the Ca^{2+}. Asn100 and Asp167 are residues in a two-stranded β-ribbon which connects the β_3/α_3 excursion to the α/β-barrel. The single water on one side of the ligand plane is 2·6 Å from Ca^{2+}, while both of the water molecules on the opposite side are ~2·5 Å from the ion. The former water molecule lies between the Ca^{2+} and the polysaccharide binding cleft while the other two are between the Ca^{2+} and the protein exterior.

The Ca^{2+} is not directly a part of either the active site or the structure forming the polysaccharide-binding cleft, but it is clearly necessary to maintain that structure intact. While deeply enmeshed in the polypeptide structure, van der Waals surface representations of PPA show that it is partially visible both from the exterior of the enzyme and when viewed from the active site cleft on the opposite side. However, accessible surface area of the Ca^{2+} coordinated by the four protein ligands is only 3% of the unsolvated solvent accessible surface area; the area is reduced to zero when the water molecules are included in the coordination sphere. Though somewhat difficult to envision, it is perhaps not unreasonable that the Ca^{2+} could be substituted by other ions such as Pb^{3+} or various lanthanides (Buisson *et al.*, 1987). Indeed, one of our own heavy-atom derivatives (Table 2) was found to substitute Pb^{3+} for the Ca^{2+}, though this was attended by severe non-isomorphism.

Table 8
Summary of intramolecular hydrogen bonding in PPA

Donor	Acceptor							
	Peptide O	Asp/Glu	Asn/Gln	Ser/Thr	Tyr	His	H_2O	Average
Peptide N	2·92(15) [245]	2·96(11) [25]	2·84(14) [16]	3·01(14) [17]		3·01(4) [2]	2·99(15) [61]	2·94(15) [366]
Arg	2·91(21) [25]	2·94(19) [18]	3·02(26) [3]	2·55 [1]	2·82 [1]		2·95(11) [15]	2·93(19) [63]
Lys	2·79(20) [3]	3·16(19) [4]		*2·60* [*1*]			3·05(31) [5]	2·99(29) [13]
His	2·94(4) [3]	2·93(14) [3]					2·77(15) [5]	2·86(14) [11]
Tyr	2·77(17) [10]				2·94 [1]		2·79(13) [4]	2·79(16) [15]
Asn/Gln	2·91(17) [18]	2·98(11) [6]	3·25 [1]	2·98(21) [3]		2·64 [1]	2·90(18) [20]	2·92(18) [49]
Ser/Thr	2·79(11) [14]	2·62(18) [2]	2·59 [1]				2·99(11) [9]	2·84(17) [26]
Trp	2·90(23) [2]	2·82(4) [2]	2·84(7) [2]		3·20 [1]		2·92 [1]	2·90(16) [8]
H_2O	2·86(15) [109]	2·90(21) [24]	2·85(21) [12]	2·92(16) [14]	2·90(22) [6]	2·69(15) [3]	2·77(27) [43]	2·85(20) [211]
Average	2·90(16) [429]	2·94(18) [84]	2·86(19) [35]	2·95(18) [36]	2·93(20) [9]	2·79(17) [6]	2·91(21) [163]	2·91(18) [712]

Distances are in Å; standard deviation of the mean is in parentheses; and, number of occurances in brackets. Ambiguous assignments have been included in averages of both possible assignments. Hydrogen bonds were determined by distance limits of 3·40 Å for O-H . . . O(N) bonds and 3·45 Å for N-H . . . O(N) bonds.

(h) *Hydrogen bonding, salt bridges and water structure*

Intramolecular hydrogen bonding statistics by residue type are given in Table 8. Serine and/or threonine residues, in particular, are involved in several short hydrogen bonds. Included in these statistics are the salt bridge distances found in Table 9. Several interactions have distances greater than 3·3 Å. These include Arg195–Glu233, already discussed in relation to the Cl ion which obstructs strong salt bridge formation between them. The proximity of the Lys243–Glu246 side-chains is a consequence of two other hydrogen bonding interactions between these residues, namely the Lys243 N . . . Glu246 O (3·1 Å) α-helical hydrogen bond and the Lys243 N . . . Glu246 OE1 (3·0 Å) hydrogen bond. Arg389 and Glu390 are consecutive residues in helix α_8 and would require considerable contortion of their side-chains to achieve a strong salt bridge. As noted in Table 9, there are four salt bridges between helices α_3 and α_4. Helix pairs α_1/α_2, α_2/α_3 and α_6/α_7 are also connected by salt bridges. The only salt bridge involving a residue of the carboxyl domain B serves to anchor loop 452 to 464 to helix α_8 of the major domain A.

The overall distribution of water molecules is illustrated in Figure 10. Hydrogen bonding parameters for the water structure are summarized in Table 8. There are no unusual clusters of water molecules nor any prominent aggregation or ordered

Table 9
Salt bridges in PPA

Donor		Acceptor		Distance (Å)	Connections
Arg20	NH2	Asp23	OD1	2·76	Intrahelix α_1
Arg30	NE	Glu27	OE1	2·73	Intrahelix α_1
Arg56	*NH2*	*Glu60*	*OE2*	2·72	*Intrahelix* α_A
Arg61	NH2	Glu18	OE2	3·03	α_A to loop β_1-α_1
Arg80	NE	Glu76	OE1	2·82	Intrahelix α_2
Arg85	NH1	Glu29	OE2	2·92	α_1 to α_2
Arg85	NH2	Asp81	OD2	3·13	Intrahelix α_2
Arg124	*NH1*	*Asp138*	*OD1*	2·74	*β_{C3} to loop β_{C4}-α_E*
Arg158	NH1	Glu246	OE2	3·02	α_E to α_5
Arg176	NH1	Asp206	OD2	2·94	α_3 to α_4
Lys185	NZ	Glu76	OE2	3·25	α_2 to α_3
Arg195	NH1	Asp96	OD2	2·79	β_3 to β_4 at Cl^- site
Arg195	NH1	Asp197	OD1	3·07	Intrastrand β_4
Arg195	NH2	Glu233	OE1	3·43	β_4 to β_5 at Cl^- site
Lys200	NZ	Glu240	OE2	2·91	β_4 to loop β_5-α_5
Lys213	*NZ*	Asp173	OD2	3·13	α_3 to α_4
His215	ND1	Asp188	OD2	2·85	α_3 to α_4
Lys243	NZ	Glu246	OE1	3·34	Consequence of other interactions
Arg267	NH1	Asp317	OD2	3·35	α_6 to α_7
His305	ND1	Glu352	OE1	2·85	α_D to loop α_7-β_{A1}
Arg343	NE	Asp381	OD2	3·00	Intraloop β_8-α_8
His386	NE2	Glu27	OE2	2·77	α_1 to loop β_8-α_8
Arg387	NH2	Glu27	OE1	2·81	α_1 to loop β_8-α_8
Arg389	NH1	Glu390	OE2	3·33	Intrahelix α_8
Arg392	*NH1*	*Asp456*	*OD2*	3·06	*α_8 to β_{F1}*
Average				2·99(21)	

Salt bridges were considered present if the N . . . O distance was ≤3·45 Å.

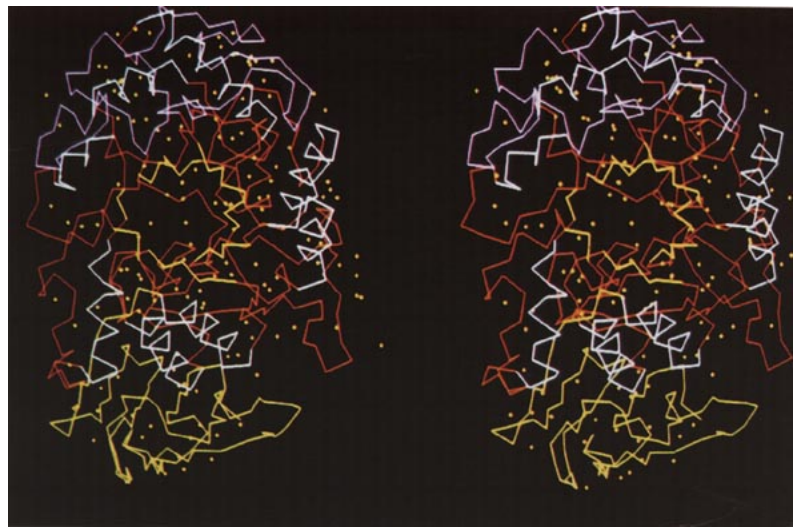

Figure 10. A stereo diagram of the α-carbon tracing of the pancreatic α-amylase molecule showing the distribution of bound water molecules surrounding the protein in the crystal. The α-helices and strands of the major domain α/β barrel are highlighted in white and yellow, respectively, the β_3/α_3 excursion is in purple, and the small domain β-barrel in yellow.

assembly of water molecules along the polysaccharide binding cleft. The water and the endogenous disaccharide molecule together reduce the accessible surface area of PPA by only 15%.

Arnold & Rossmann (1990) introduced a quantity, *Qualwat*, defined as:

$$Qualwat = 100\ Q\ \exp(-B/4d^2), \qquad (1)$$

where Q is the occupancy, B is the thermal parameter and d is the maximum resolution of the data in Å. This quantity is useful in assessing the quality of the water molecules in a model. In the final model of PPA, the *Qualwat* values of the 145 water molecules range from 89·1 to 4·2 (an ideal water molecule with $Q = 1{\cdot}0$ and $B = 0{\cdot}0$ Å^2 would have *Qualwat* of 100). The average B value for protein side-chain atoms is 19·6 Å^2 which equates to a *Qualwat* value of 32·9. Over 39% of the water molecules have values greater than this. No attempt was made to maximize the *Qualwat* values by refining the occupancies of the water molecules.

(i) *The active site cleft and polysaccharide binding*

A thorough analysis of all difference Fourier maps and refinement of complex structures is still underway. We present here only our study of the α-cyclodextrin complex crystals at 2·8 Å resolution. Our preliminary results from the studies of the non-cyclic polysaccharides as well as other cyclodextrins are consistent, however, with the observations presented here for the α-cyclodextrin results.

Both the $[F_o(\text{cmplx}) - F_o(\text{nat})]$ and $[F_o(\text{cmplx}) - F_c(\text{cmplx})]$ difference Fourier maps for PPA complexed with α-cyclodextrin were clear. These showed three independent α-cyclodextrin molecules (designated I, II and III) bound to each α-amylase molecule. An example of the difference electron density of α-cyclodextrin-I and surrounding protein structure is shown in Figure 11(a). Figure 12 shows the distribution and orientations of the three cyclodextrin molecules with respect to the protein. Two of these (α-cyclodextrin-I and -II) are closely adjacent in the long, deep binding cleft, shown in Figure 13, on the surface of the enzyme. α-Cyclodextrin-I is positioned between screw axis related protein molecules such that two of the six glucose residues superimpose exactly upon the endogenous disaccharide of the native crystalline protein. This is shown in Figure 11(b). α-Cyclodextrin-II is positioned closer to the center of the cleft and proximal to the putative catalytic residues (see below).

α-Cyclodextrin-III is bound at a site (henceforth referred to as the accessory carbohydrate site), ~15 Å from the active site cleft in a slight depression on the surface of the protein. This is formed by an edge of the four-stranded β-sheet C of the β_3/α_3 excursion (residues 129 to 136, including *cis*-Pro130) and the initial amino terminal turn of α-helix 3 of the major domain α/β-barrel (residues 172 to 190). α-Cyclodextrin-III participates in a number of interactions with amino acid side-chains. This accessory site is quite distinct from the first two, which are spatially associated, and clearly represents an independent, non-catalytically related, affinity site for the α-cyclodextrin.

The most prominent interactions between α-cyclodextrin-III and amino acid side-chains at the accessory site can be seen in Figure 14(a). The side-chain of Trp134 is directed toward the center of the cyclic polysaccharide cavity. Tyr174 is disposed with its side-chain flat against the exterior of the cyclodextrin ring. While hydrogen bonds between polysaccharide and protein likely occur through the side-chains of Ser132, Asp135 and Lys172, except for contacts with the Trp134 ring, only two of the

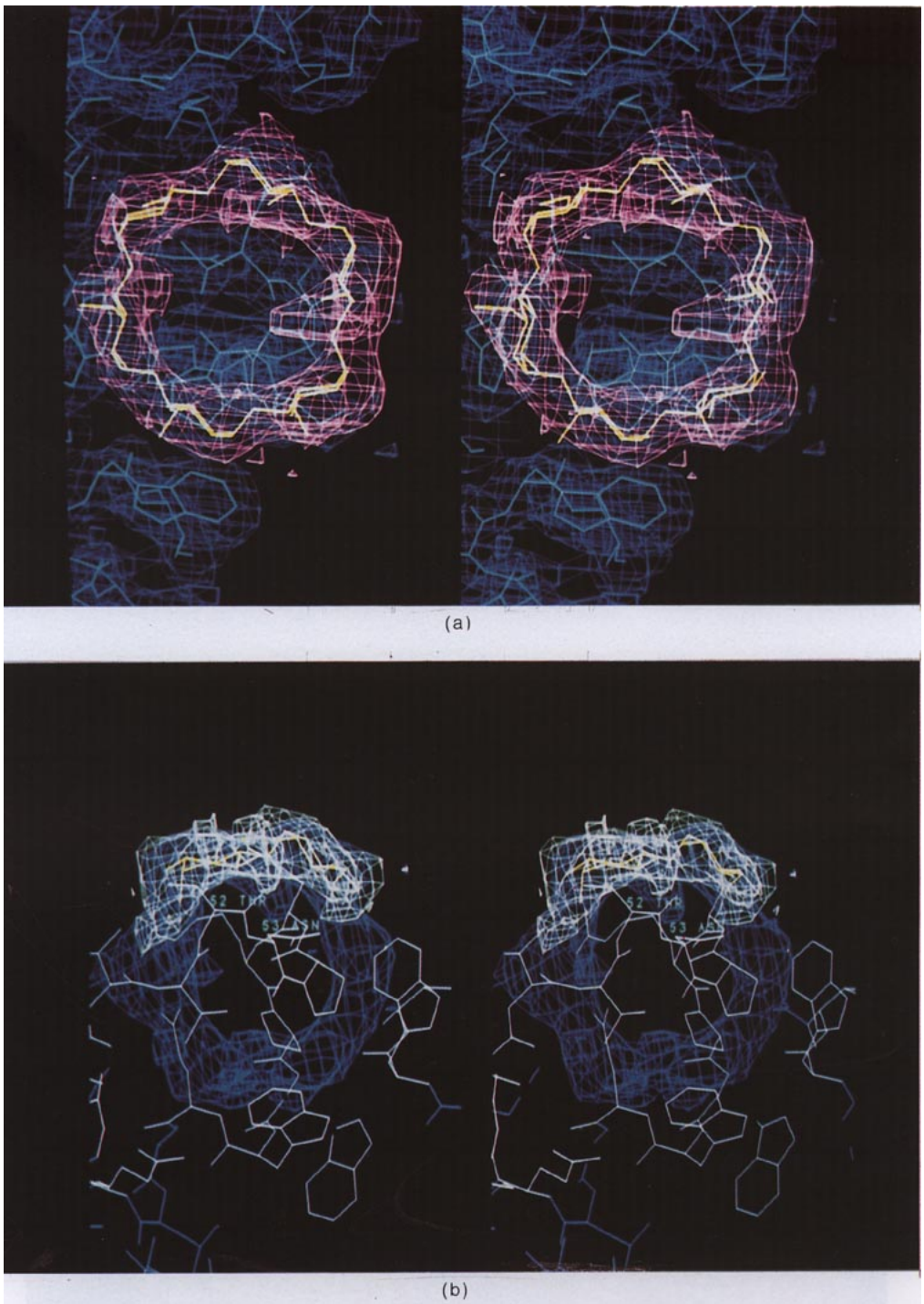

Figure 11. (a) Shows the $(F_o(\text{complex}) - F_o(\text{native}))\exp i\alpha_c(\text{native})$ electron density, in purple, corresponding to 1 of the 3 cyclodextrin molecules (I) bound to α-amylase in the complex crystals. Amino acid residues in its immediate environment are also shown superimposed on their $(2F_o - F_c)\exp i\alpha_c$ density. In (b) the $(F_o - F_c)\exp i\alpha_c$ electron density for the disaccharide and its skeletal structure are shown superimposed on the α-cyclodextrin difference density, shown here in a different orientation. The disaccharide density is virtually congruent with that corresponding to 2 sugar residues of the α-cyclodextrin ring. Some neighboring amino acid side-chains involved in their binding are shown as well. Particularly prominent is the side-chain of Asp53 which protudes directly into the center of the polysaccharide ring.

six sugar residues of α-cyclodextrin-III appear to be in direct contact with protein residues. The remaining four glucose residues in the ring protrude into the solvent.

The binding environment for α-cyclodextrin-II, near the center of the long polysaccharide binding cleft, is formed by residues from the short loop segment 147 to 148, by a turn of 3_{10}-helix comprised of amino acid residues 58 to 63, and another loop segment 304 to 305. As shown in

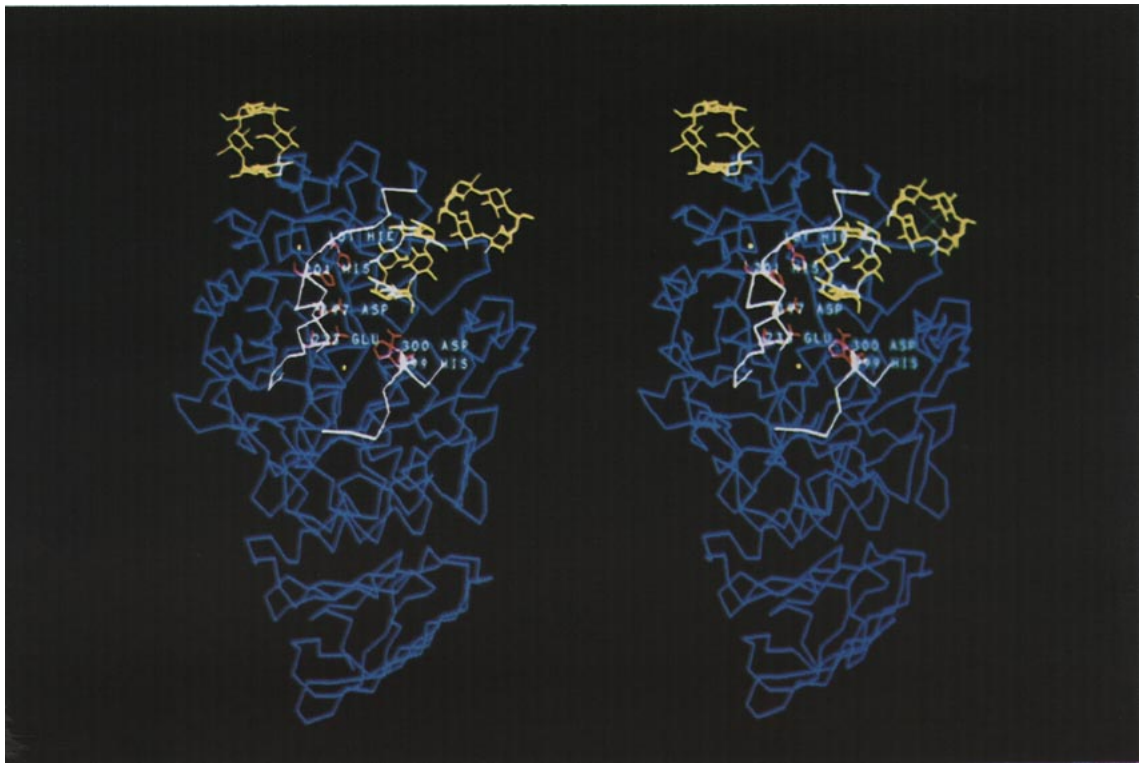

Figure 12. Stereo diagram of the α-amylase backbone in dark blue with the 3 bound α-cyclodextrin molecules in yellow. The view is into the polysaccharide binding cleft. Highlighted in white are the 3 peptide segments that are found to be highly conserved in several amylases. All 3 segments contribute to the formation of the active site structure. In red are 6 side-chains that have been proposed as catalytic residues. All fall in close proximity to the active site cleft.

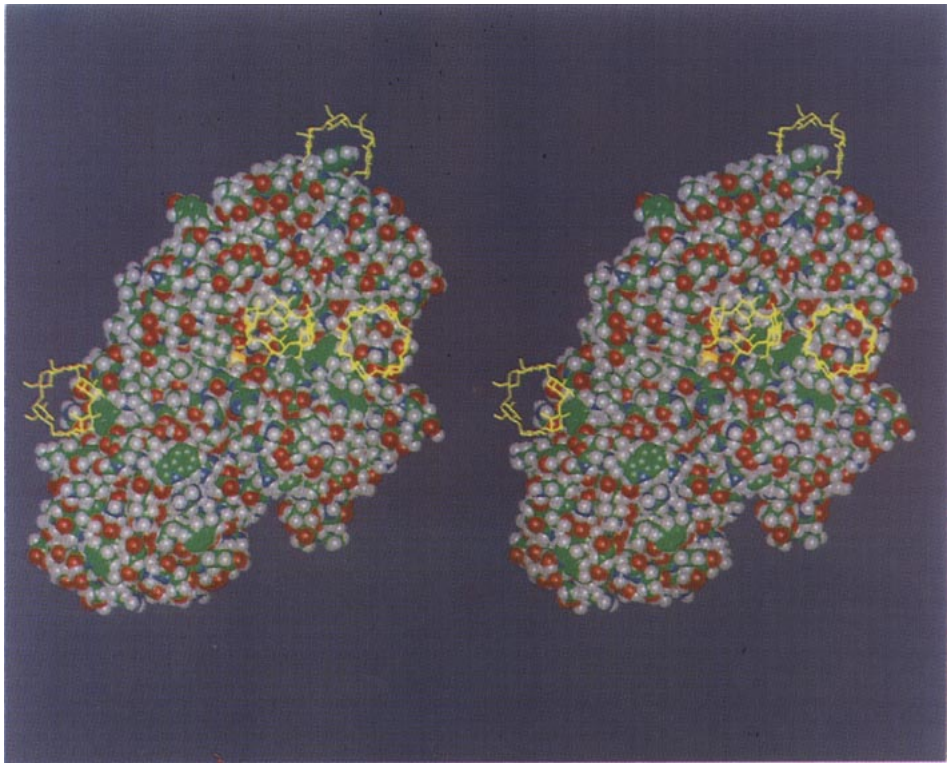

Figure 13. Van der Waals representation of the entire α-amylase molecule oriented to look into the polysaccharide binding cleft. Four α-cyclodextrin molecules, shown in yellow, are present. α-Cyclodextrin-III is partially seen at the top edge of the protein, α-cyclodextrin-II near the center of the starch binding cleft, and 2 symmetry related positions for α-cyclodextrin-I are seen at the extremes of the polysaccharide binding cleft. The deep polysaccharide binding cleft can be seen passing across the midsection of the enzyme from upper right to mid left. The allosteric activating Cl ion is the yellow sphere deep within the active site cleft.

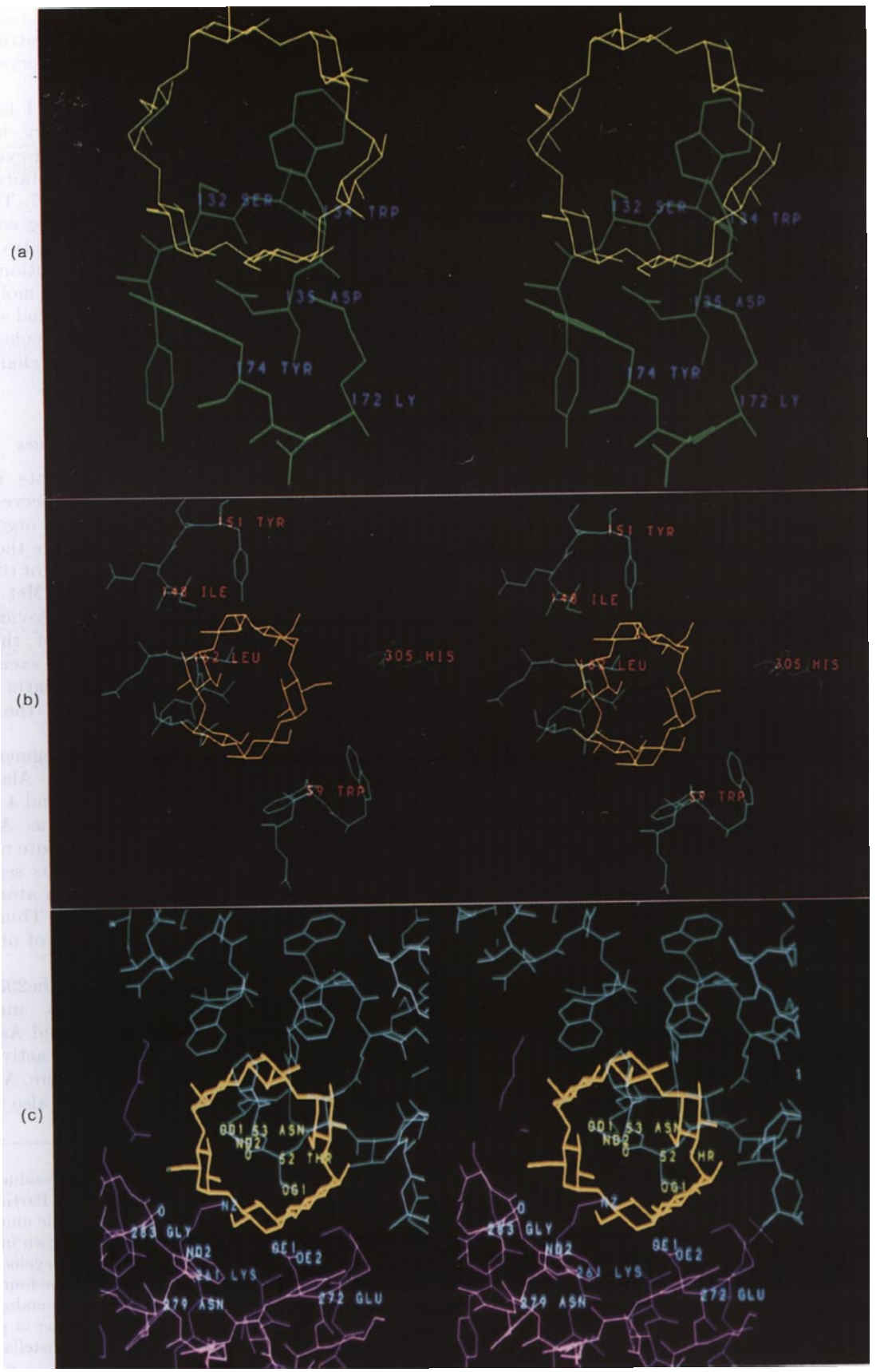
(a)
132 SER
134 TRP
135 ASP
174 TYR
172 LY
(b)
151 TYR
142 ILE
142 LEU
305 HIS
58 TRP
(c)
GD1 53 ASN
ND2
O
52 THR
OG1
N2
O
263 GLY
ND2
GE1
OE2
241 LYS
279 ASN
272 GLU

Figure 14(b), protruding directly into the top of the α-cyclodextrin ring is the propyl group of Val163. The Tyr151 side-chain is hydrogen bonded to sugar residues through its phenolic OH while Trp59 has its side-chain ring stacked against the exterior of the α-cyclodextrin. The α-cyclodextrin is also bound through hydrogen bonds to the main-chain carbonyl oxygen atoms of His305 and Val163. The side-chain of Leu162 is packed against one side of the polysaccharide ring. Three of the six sugar residues of α-cyclodextrin-II appear to interact directly with the protein while the remaining three project from the top of the binding cleft and are exposed to solvent. α-Cyclodextrin-I and -II are almost within contact distance, a contact mediated in part by the intercession of the side-chain of Trp59 at the point of closest approach.

α-Cyclodextrin-I, whose binding environment is shown in Figure 14(c), partially superimposes on the endogenous disaccharide and lies at an extreme of the polysaccharide binding cleft. By 2_1 symmetry along x, however, a symmetry equivalent also lies at the opposite end of this same cleft. In a manner similar to the endogenous, bound disaccharide, this α-cyclodextrin serves to tie together consecutive protein molecules, whose polysaccharide binding clefts are perfectly contiguous, along the x direction throughout the crystal. It might also be considered two half sites that mark the extremes of the polysaccharide binding cleft of an individual enzyme molecule in the crystal.

One of the most prominent features of the binding of this α-cyclodextrin-I at one end of the PPA cleft is the insertion of the side-chain of Asp53 into the cavity of the bound cyclodextrin giving the appearance of a ring on a peg. This is illustrated in Figure 14(c). The binding site is otherwise formed by residues from a large meandering loop that includes amino acid residues 51 to 55 (Pro54 is *cis*), and bounded by residues 106 to 108 of the loop entering the first strand of the four-stranded β-sheet C of the β_3/α_3 excursion (the strand on the opposite end of the sheet from that associated with the accessory α-cyclodextrin). Hydrogen bonds to the α-cyclodextrin are also probable from both the side-chain and the main-chain oxygen atoms of Thr52. At the opposite end of the binding cleft from the first α-cyclodextrin-I, a screw axis related α-cyclodextrin-I molecule is bound through hydrogen bonding interactions to at least four different amino acid residues. These include the main-chain carbonyl oxygen atom of Gly283, and the side-chains of Asn279, Glu272 and Lys261. Again, there is an aromatic side-chain (Trp284) abutting the exterior of the α-cyclodextrin ring as observed with sites II and III.

The composite nature of binding site I is particularly interesting because the primary binding interactions are the same for both the α-cyclodextrin and the endogenous disaccharide. Details of the disaccharide binding are given in Table 7. The site bridges two protein molecules, suggesting cooperative, contiguous binding sites. Both of the two sugar residues bound are fixed through contributions from both of the symmetry related protein molecules. These observations are consistent with, and suggest that, contiguously bound α-amylase molecules along a continuous, extended polysaccharide chain is a distinct possibility.

(j) *Conserved and active site residues*

There are three amino acid segments in pig pancreatic α-amylase that are highly conserved with respect to α-amylases from a variety of organisms (Yamamoto *et al.*, 1988). Figure 12 shows the locations of these segments in PPA. The first of these is: Val95 - Asp - Ala - Val - X - Asn100 - His - Met - Cys - Gly - Ser105 - Gly - Ala. This sequence provides the structural transition from the core of the α/β domain, leaving strand 3, into the β_3/α_3 excursion. It contains a Ca^{2+} ligand, Asn100, and parts of the segment contribute to the formation of the polysaccharide binding cleft.

The second conserved sequence segment is: Val191 - Ala - Gly - Phe - Arg195 - Ile - Asp - Ala - Ser - Lys200-His. This includes the entire strand 4 of the β-barrel, residues 193 to 196, as well as Asp197 which has been proposed to be an active site residue (Ishikawa *et al.*, 1990). In addition, this segment includes His201, whose carbonyl oxygen atom is a Ca^{2+} ligand, and Arg195, a Cl^- ligand. Thus, this sequence contributes to several features of obvious structural and catalytic significance.

The final conserved sequence is: Val-Phe295-Val-Asp-Asn-His-Asp300-Asn-Gln-Arg. This includes Val294, the last residue in β-strand 7, and Asp300, which has also been suggested to be an active site residue (Ishikawa *et al.*, 1990). Furthermore, Asn298 is involved in Cl^- binding and Phe295 is also in the

Figure 14. The non-active site, accessory α-cyclodextrin-III is shown in (a) with its associated binding residues. The sugar ring lies in a shallow depression on the edge of the protein roughly 15 Å from the active center. Particularly prominent here is Trp134 which protrudes into the top of the sugar ring and contacts several of the saccharide units. The skeletal structure of the α-cyclodextrin-II molecule near the center of the polysaccharide binding cleft is shown in (b) as it is bound by amino acid residues in its environment. Trp59, seen at lower right, lies exactly between this α-cyclodextrin and a second α-cyclodextrin-I at one end of the binding cleft. α-Cyclodextrin-I, in yellow, is seen in (c) to be found at a composite binding site formed by 2 symmetry related α-amylase molecules. This apparently displaces the endogenous disaccharide found in the native crystals upon complex formation. One protein molecule is in blue, the other in purple. The α-cyclodextrin makes numerous interactions with both protein molecules. Particularly striking is the constellation of 3 tryptophan residues on 1 side of the cyclodextrin ring provided by 1 enzyme molecule while on the opposite side, provided by the symmetry related protein, is another tryptophan and a tyrosine. Amino acid side-chains forming probable hydrogen bonds to the α-cyclodextrin are labeled.

vicinity of the Cl ion. This segment of polypeptide forms a large sweeping loop that comprises a lateral portion of the polysaccharide binding cleft.

4. Discussion

Three unique binding sites for α-cyclodextrin might have been predicted from earlier studies by Mora *et al.* (1974), Levitzki *et al.* (1964) and Loyter & Schramm (1962, 1966) who studied the binding of β-cyclodextrin and limit-dextrins to pancreatic α-amylase. Mora *et al.* (1974) used a variety of techniques to show that three β-cyclodextrin molecules bound to PPA and that complex formation could be characterized by a single dissociation constant of about 0·25 mM. Their evidence persuaded them to reject the possibility of three independent binding sites of equal affinity for β-cyclodextrin in favor of a highly cooperative mechanism involving a high affinity "trigger site".

The environments of our three α-cyclodextrin sites are quite different from each other, as are the binding modes. Thus, it seems doubtful they would have similar, let alone identical, binding constants. On the other hand, it is not clear how any one of the three could serve as a trigger site for cooperative binding.

That multiple, independent carbohydrate binding sites are not artifactual was demonstrated further by Levitzki *et al.* (1964) and by Loyter & Schramm (1962, 1966). They showed by equilibrium dialysis that α-amylase possessed two independent binding sites for maltotriose, and that α-amylase formed non-covalently crosslinked networks of limit dextrins. They too concluded that α-amylase contains at least two independent binding sites for polysaccharide and that it bound in a cooperative manner to dextrin fragments. In electron micrographs these complexes were both multiply branched and extended linear arrays like beads on a string.

In this regard composite disaccharide and cyclodextrin binding between screw axis related molecules is of particular interest, as it does not appear to be fortuitous or an artifact of crystal packing. It serves to promote fairly extensive interactions between the linked protein molecules and could explain the contiguous strings of protein molecules seen when PPA binds to dextrins in solution (Loyter & Schramm, 1966).

In this particular crystal lattice there are few direct contacts, except those mediated by water, between protein molecules along two crystallographic directions, *y* and *z*. The only extensive protein–protein interactions are between the 2_1 related molecules in the *x* direction and these exclusively in the immediate neighborhood of the composite carbohydrate binding site. Therefore, through shared binding to carbohydrate, contiguous protein–protein interactions, continuous with respect to polysaccharide, are established and promoted. This is exactly the basis for cooperative binding of protein to any extended, generally polymeric substrate (von Hippel *et al.*, 1977) and is perhaps not an unreasonable expectation, as one of the functions of PPA is not only to cleave starch, but to disrupt its intricate helical structure thereby promoting more efficient enzyme attack (Leloup *et al.*, 1992).

The structural environment of the bound chloride ion suggests a possible mechanism for the observed activation of pancreatic α-amylase (Muus *et al.*, 1956). In Figure 8 it can be seen that the guanidinium groups of Arg195 and Arg337 are involved in bridges to the chloride ion, while the side-chain of Glu233 is turned away from the chloride and projects into the active site cleft. One can well imagine that were the Cl ion not present to neutralize the charge on the guanidinium groups, the mobile side-chain of Glu233 would adjust its orientation so that its negatively charged carboxyl group would interact with the positively charged arginine side-chains. Thus, the effect of integrating a chloride ion would be to disrupt Glu-Arg salt bridges and release the glutamic acid side-chain into the active site. This is significant because Glu233 has been proposed as a key catalytic residue along with an aspartic acid and a histidine (Ishikawa *et al.*, 1990). Chloride ion, therefore, might provide a release mechanism for an otherwise restrained catalytic glutamic acid carboxyl group.

Comparison of PPA to the amylases of *Aspergillus niger* (Brady *et al.*, 1991; Boel *et al.*, 1990) and *A. oryzae* (Matsuura *et al.*, 1984; Boel *et al.*, 1990; Swift *et al.*, 1991) show that the overall topologies are very similar. The two globular domains A and B are composed of a $(\alpha/\beta)_8$-barrel and an eight-stranded β-barrel respectively. The β_3/α_3 excursion of PPA has its counterpart in the bacterial amylases as well. This subdomain in all three amylases is held to the $(\alpha/\beta)_8$-barrel *via* coordination to a Ca ion. In each case the ligands are a bidentate Asp, a unidentate Asn O^{δ}, two carbonyl oxygen atoms and three water molecules identically arranged. In the bacterial amylases a secondary Ca^{2+} site is observed within the active site. No counterpart is observed in PPA. PPA, on the other hand, is the only molecule of the three that contains a Cl ion.

Ishikawa *et al.* (1990), based on the structures of taka-amylase and of PPA described by Buisson *et al.* (1987), proposed that the catalytic residues for PPA involved a histidine (likely to be 101, 201 or 299) and two aspartic acid and/or glutamic acid residues believed to be Asp197, Asp300 and/or Glu233, Indeed, all of these residues, as seen in Figure 12, are proximal to the active site cleft and could conceivably be involved in catalysis.

Several crystal forms of hen egg white lysozyme have been studied by X-ray diffraction (triclinic: Ramanadhan *et al.*, 1987; monoclinic: Rao *et al.*, 1983; tetragonal: Kodandapani *et al.*, 1990; binding in tetragonal form: Kelly *et al.*, 1979) and the catalytic residues identified. These residues consist of glutamic acid 35 and aspartic acid 52 which have

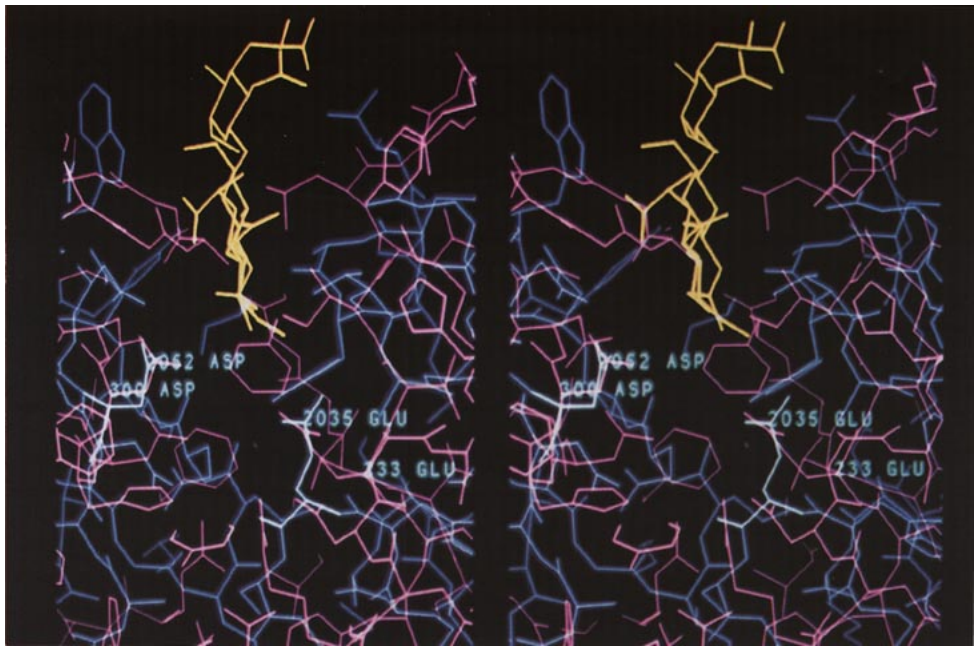

Figure 15. The known catalytic residues, Asp52 and Glu35, of lysozyme, highlighted here in white and labeled 2052 and 2035, can be superimposed with excellent precision upon putative α-amylase catalytic residues Asp300 and Glu233, also labeled, in such a way that the polysaccharide binding cleft of lysozyme (in blue) superimposes with good correspondence upon the extended starch binding cleft of pancreatic α-amylase, shown in pink. Shown in yellow, for reference, is the trisaccharide inhibitor of lysozyme, NAG-NAG-NAM, as it binds at the lysozyme active site.

the carboxylate carbon atoms separated by 7·1 Å in the crystalline state. Superposition of these residues of lysozyme on Glu233 and Asp300, respectively, (carboxylate carbon atoms are separated by 7·0 Å) of PPA and alignment of the binding clefts, illustrated in Figure 15, demonstrates the remarkable resemblance of the active sites of the two polysaccharide hydrolyzing enzymes. This similarity further supports the proposition that Glu233 is a catalytic residue and may, as in lysozyme, function as the proton donor during catalysis.

This research was supported by grants from the NIH, from NSF and from NASA. The authors wish to acknowledge P. M. D. Fitzgerald, and P. J. Stankiewicz who contributed significantly to the early stages of the work. They also wish to acknowledge Ms M. Greene and Mr S. Singer of the UCR computer graphics facility for their assistance. We thank the San Diego Supercomputer Center for a grant of time on the Cray Y-MP.

References

Arnold, E. & Rossmann, M. G. (1990). Analysis of the structure of a common cold virus, human rhinovirus 14, refined at a resolution of 3·0 Å. *J. Mol. Biol.* **211**, 763–801.

Blow, D. M. & Crick, F. H. C. (1959). The treatment of errors in the isomorphous replacement method. *Acta Crystallogr.* **12**, 794–804.

Blundell, T. L. & Johnson, L. N. (1976). *Protein Crystallography*, pp. 334–336, Academic Press, New York.

Boel, E., Brady, L., Brzozowski, A. M., Derewenda, Z., Dodson, G. G., Jensen, V. J., Petersen, S. B., Swift, H., Thim, L. & Woldike, H. F. (1990). Calcium binding in α-amylases: An X-ray diffraction study at 2·1-Å resolution of two enzymes from *Aspergillus*. *Biochemistry*, **29**, 6244–6249.

Braden, C. & Tooze, J. (1991). *Introduction to Protein Structure*, Garland Publishing Inc., New York.

Brady, R. L., Brzozowski, A. M., Derewenda, Z. S., Dodson, E. J. & Dodson, G. G. (1991). Solution of the structure of *Aspergillus niger* acid α-amylase by combined molecular replacement and multiple isomorphous replacement methods. *Acta Crystallogr. sect. B*, **47**, 527–535.

Brayer, G. D. & McPherson, A. (1983). Refined structure of the gene 5 DNA binding protein from bacteriophage fd. *J. Mol. Biol.* **169**, 565–596.

Brunger, A. T. (1988). Crystallographic refinement by simulated annealing; application to a 2·8 Å resolution structure of aspartate aminotransferase. *J. Mol. Biol.* **203**, 803–816.

Brunger, A. T. (1991). Simulated annealing in crystallography. *Annu. Rev. Phys. Chem.* **42**, 197–223.

Brunger, A. T. (1992). The free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature (London)*, **355**, 472–474.

Brunger, A. T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science*, **235**, 458–460.

Buisson, G., Duée, E., Haser, R. & Payan, F. (1987). Three dimensional structure of procine pancreatic α-amylase at 2·9 Å resolution. Role of calcium in structure and activity. *EMBO J.* **6**, 3909–3916.

Caldwell, M. L., Booker, L. E. & Sherman, H. C. (1931). Crystalline amylase. *Science*, **74**, 37–42.

Carson, M. & Bugg, C. E. (1986). An algorithm for ribbon models of proteins. *J. Mol. Graph.* **4**, 121–122.

Dickerson, R. E., Kendrew, J. C. & Strandberg, B. E. (1961). The crystal structure of myoglobin: phase determination to a resolution of 2 Å by the method of isomorphous replacement. *Acta Crystallogr.* **14**, 1188–1195.

Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystollagr. sect. A*, **47**. 392–400.

Fischer, E. H. & Stein, E. A. (1960). Alpha amylases. In *The Enzymes* (Boyer, P. D., Lardy, H. & Myrbäck, K., eds), 2nd edit., vol. 4, chapt. 6, pp. 333–343, Academic Press, New York.

Hamlin, R., Cork, C., Howard, A., Nielson, C., Vernon, W., Matthews, D. & Xuong, N.-H. (1981). Characteristics of a flat multiwire area detector for protein crystallography. *J. Appl. Crystallogr.* **14**, 85–89.

Hendrickson, W. A. & Lattman, E. E. (1970). Representation of phase probability distribution for simplified combination of independent phase information. *Acta Crystallogr. sect. B*, **26**, 136–143.

Ishikawa, K., Matsui, I. & Honda, K. (1990). Substrate dependent shift of optimum pH in procine pancreatic α-amylase-catalyzed reactions. *Biochemistry*, **9**, 7119–7123.

Jones, T. A. (1985). Interactive computer graphics: FRODO. In *Methods in Enzymology* (Wyckoff, H. W., Hirs, C. H. W. & Timasheff, S. N., eds), vol. 115, chapt. 12, pp. 157–171, Academic Press, New York.

Kelly, J. A., Sielecki, A. R., Sykes, B. D., James, M. N. G. & Phillips, D. C. (1979). X-ray crystallography of the binding of the bacterial cell wall trisaccharide/NAM-/NAG-/NAM to lysozyme. *Nature (London)*, **282**, 875–878.

Ko, T. P., Ng, J. & McPherson, A. (1993). The three-dimensional structures of canavalin from Jack bean *(Canavalia ensiformis)*. *Plant Physiol.* **101**, 729–744.

Kodandapani, R., Suresh, C. G. & Vijayan, M. (1990). Crystal structure of low humidity tetragonal lysozyme at 2·1 Å resolution. Variability in hydration shell and its structural consequences. *J. Biol. Chem.* **265**, 16126–16131.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1992). *PROCHECK v.2.: Programs to Check the Stereochemical Quality of Protein Structures*, Oxford Molecular Ltd., Oxford, England.

Lee, B. & Richards, F. M. (1971). The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55**, 379–400.

Leloup, V. M., Colonna, P. & Marchis-Mouren, G. (1992). Mechanisms of the adsorption of pancreatic α-amylase on starch crystallites. *Carbohydrate Res.* **232**, 367–374.

Levitzki, A., Heller, J. & Schramm, M. (1964). Specific precipitation of enzyme by its substrate: the α-amylase-macrodextrin complex. *Biochem. Biophys. Acta*, **81**, 101–107.

Loyter, A. & Schramm, M. (1962). The glycogen-amylase complex as a means of obtaining highly purified α-amylases. *Biochim. Biophys. Acta*, **65**, 200–206.

Loyter, A. & Schramm, M. (1966). Multimolecular complexes of α-amylase with glycogen limit dextrin. *J. Biol. Chem.* **241**, 2611–2617.

Luzatti, P. V. (1952). Traitement statistique des erreurs dans la determination des structures cristallines. *Acta Crystallogr.* **5**, 802–810.

Matsuura, Y., Kusunoki, M., Harada, W. & Kakudo, M. (1984). Structure and possible catalytic residues of Taka-amylase A. *J. Biochem. (Tokyo)*, **95**, 697–702.

Matthews, B. W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–502.

McPherson, A. (1990). Current approaches to macromolecular crystallization. *Eur. J. Biochem.* **189**, 1–23.

McPherson, A. & Rich, A. (1972). X-ray crystallographic analysis of swine pancrease α-amylase. *Biochim. Biophys. Acta*, **285**, 493–497.

Mora, S., Simon, I. & Elödi, P. (1974). Studies on the active center of pancreatic amylase: binding of β-cyclodextrin. *Mol. Cell. Biochem.* **4**, 205–209.

Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. *Proteins*, **12**, 345–364.

Muus, J., Brockett, F. P. & Connelley, C. C. (1956). The effect of various ions on the stability of crystalline salivary amylase in solution. *Arch. Biochim. Biophys.* **65**, 268–277.

Pasero, L., Mazzei-Pierron, Y., Abadie, B., Chicheportiche, Y. & Marchis-Mouren, G. (1986). Complete amino acid sequence and location of the 5 disulfide bridges in porcine pancreatic α-amylase. *Biochem. Biophys. Acta*, **869**, 147–157.

Payan, F., Haser, R., Pierrot, M., Frey, M., Astier, J. P., Abadie, B., Duée, E. & Buisson, G. (1980). The three dimensional structure of α-amylase from porcine pancreas at 5 Å resolution. The active site location. *Acta Crystallogr. sect. B*, **36**, 416–421.

Prodanov, E., Seigner, C. & Marchis-Mouren, G. (1984). Subsite profile of the active center of porcine pancreatic α-amylase. Kinetic studies using maltooligosaccharides as substrates. *Biochem. Biophys. Res. Commun.* **122**, 75–81.

Ramachandran, G. N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Advan. Protein Chem.* **23**, 283–437.

Ramanadham, M., Sieker, L. C. & Jensen, L. H. (1987). Stereochemically restrained least-squares refinement of triclinic lysozyme. *Acta Crystallogr. sect. A*, **43**, 13.

Rao, S. T., Hogle, J. & Sundaralingam, M. (1983). Studies of monoclinic hen egg white lysozyme. The refinement at 2·5 Å resolution: conformational variability between the two independent molecules. *Acta Crystallogr. sect. C*, **39**, 237–240.

Richardson, J. S. (1981). The anatomy and taxonomy of protein structure. *Advan. Protein Chem.* **34**, 167–339.

Robyt, J. F. & French, D. (1970*a*). The action pattern of porcine pancreatic α-amylase in relationship to the substrate binding site of the enzyme. *J. Biol. Chem.* **245**, 3917–3927.

Robyt, J. F. & French, D. (1970*b*). Multiple attack and polarity of action of porcine pancreatic α-amylase. *Arch. Biochem. Biophys.* **138**, 662–670.

Rossmann, M. G. (1960). The accurate determination of the position and shape of heavy-atom replacement groups in proteins. *Acta Crystallogr.* **13**, 221–226.

Shelley, K., Hillman, B. & McPherson, A. (1980). Spatial filtering of electron micrographs of negatively stained α-amylase crystals. *Ultramicroscopy*, **5**, 281–295.

Stankiewicz, P. J., Cascio, D. & McPherson, A. (1983). β-thimaltosides as active site probes for α-amylase. *J. Appl. Biochem.* **5**, 388–398.

Swift, H. J., Brady, L., Derewenda, Z. S., Dodson, E. J., Dodson, G. G., Turkenburg, J. P. & Wilkinson, A. J. (1991). Structure and molecular model refinement of *Aspergillus oryzae* (TAKA) α-amylase: an application of the simulated-annealing method. *Acta Crystallogr. sect. B*, **47**, 535–544.

Ten Eyck, L., Weaver, L. A. & Matthews, B. W. (1976). A method of obtaining a stereochemically acceptable

protein model which fits a set of atomic coordinates. *Acta Crystallogr. sect. A*, **32**, 349–355.

Terwilliger, T. & Eisenberg, D. (1983). Unbiased three dimensional refinement of heavy atom parameters by correlation of origin removed Patterson functions. *Acta Crystallogr. sect. A*, **39**, 813–817.

Terwilliger, T. C., Weissman, L. & Eisenberg, D. (1982). The structure of melittin in the form I crystals and its implication for melittin's lytic and surface activities. *Biophys. J.* **37**, 353–361.

Thoma, J. A., Spradlin, J. E. & Dygert, S. (1971). Plant and animal amylases. In *The Enzymes* (Boyer, P. D. ed.), 3rd edit., vol. 5, chapt. 6, pp. 115–199, Academic Press, New York.

Vallee, B. L., Stein, E. A., Sumerwell, W. N. & Fischer, E. H. (1959). Metal content of α-amylases of various origins. *J. Biol. Chem.* **234**, 2901–2905.

von Hippel, P. H., Jensen, D. E., Kelly, R. C. & McGhee, J. D. (1977). Molecular approaches to the interaction of nucleic acids with "melting" proteins. In *Nucleic Acid–Protein Recognition* (Vogel, H. J., ed.), pp. 66–89, Academic Press, New York.

Watenpaugh, K. D. (1985). Overview of phasing by isomorphous replacement. In *Methods in Enzymology* (Wyckoff, H. W., Hirs, C. H. W. & Timasheff, S. N., eds), vol. 115, chapt. 1, pp. 3–15, Academic Press, New York.

Yamamoto, T., Kitahata, S., Hiromi, K., Ohnishi, M., Minamiura, N., Shinke, R., Okada, S. & Komaki, T. (1988). Editors of *Handbook of Amylases and Related Enzymes; Their Sources, Isolation Methods, Properties and Applications*. Pergamon Press, New York.

Xuong, N.-H., Nielson, C., Hamlin, R. & Anderson, D. (1985). Strategy for data collection from protein crystals using a multiwire counter detector diffractometer. *J. Appl. Crystalogr.* **18**, 342–360.

Edited by B. W. Matthews

(Received 23 July 1993; accepted 7 October 1993)