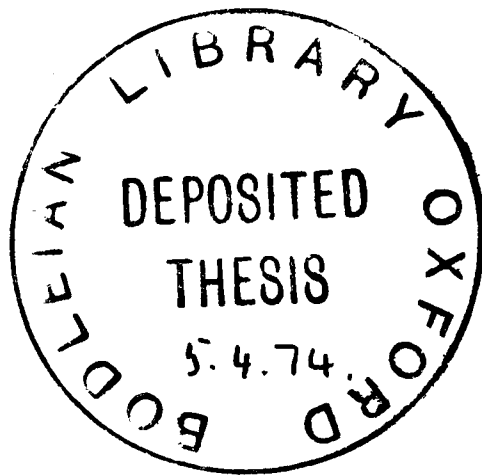


Structural Analysis of Vitamin A
transport system.

A thesis submitted for the degree
of Doctor of Philosophy.



M.J.GEISOW 1973

LINCOLN COLLEGE
OXFORD.

THE LABORATORY OF
MOLECULAR BIOPHYSICS.

ABSTRACT.

The method of isomorphous replacement has been used to determine the crystal structure of human plasma prealbumin (Mwt. 56 000). Three derivatives were measured to a resolution of $2.5 \text{ \AA}$ and a mercury derivative was re-measured with a long x-ray count time to improve the quality of the estimates of anomalous scatter.

A protein Fourier synthesis has been calculated using phases determined from the anomalous and isomorphous differences. The conformation of the two polypeptide chains in the asymmetric unit of $P2_12_12$ (the crystal space group) has been determined.

The chains in each monomer have an apparently identical conformation. They are nearly fully extended and form an extensive network of β -structure. Eight strands of β -sheet in each monomer form a flattened cup shape the interior of which appears to be entirely hydrophobic. The ends and bottom of the cup are closed by loops in a random coil conformation. One loop contains a right-handed α -helix of about two turns. The top of the cup is closed by the second monomer, so that the front of this unit is a continuous 8-stranded β -pleated sheet. The back of this is also formed by an 8-stranded β -sheet, but this is less regular. The dimers formed in this way interact so that their front sheets are opposed at a distance of $10 \text{ \AA}$.

The thyroxine-binding site has been identified within the channel formed by the β -sheets of opposed dimers. There are two possible binding sites, only one of which may be occupied in one molecule (on chemical evidence).

ACKNOWLEDGEMENTS

I am indebted to my research supervisor, Dr.C.C.F.Blake for his guidance and encouragement while I have been working in the Laboratory of Molecular Biophysics. I also express my sincere thanks to my colleagues for helpful discussions and good fellowship. I would particularly like to acknowledge Dr.A.C.Bloomer, Mr.P.R.Evans and Mr.K.Wilson for software facilities, and Mr.S.J.Oatley for his continued assistance with the long task of transferring contours to acetate and typing atomic co-ordinates.

My thanks are due to Dr.H.Geisow for typing most of this thesis, and to the photographic technicians of the Zoology Dept. for the photographs.

I am grateful to the Science Research Council for their financial support of this work.

List of symbols used.

*	refers to reciprocal space quantity
a, b, c, a^*, b^*, c^*	unit cell translations
V, V^*	unit cell volume
h, k, l	Miller indices of reflections
x, y, z	real space co-ordinates
d, d^*	inter-planar spacing
X, Y, Z	vector space co-ordinates
λ, λ^*	wavelength in $\text{\AA}$
θ	Bragg angle
$\{1\ 1\ 0\}$	the set of crystal planes 1 1 0
$(1\ 1\ 0)$	the reflection 1 1 0
$\Phi, \chi, \omega, 2\theta, \Sigma$	circles on five-circle diffractometer
μ	precession angle
ϕ, ψ	spherical polar co-ordinates
ζ, ξ	cylindrical polar co-ordinates, axial and radial
τ	detector angle in cylindrical polar co-ordinates
ϕ	azimuthal angle in " " "
α_p	phase angle of subscript vector
β, γ, ϕ	angles in the phase triangle
I_p	intensity of subscript source
$F_p, F_p $	amplitude of subscript source
$\underline{F}_p$	structure factor of subscript source
$\Delta F, \Delta F $	difference amplitude
$\underline{F}_{\text{calc}}$	calculated structure factor
$\underline{F}_{\text{obs}}$	observed structure factor
Z	occupancy
B	temperature factor
k	scale factor
k'	ratio of real to anomalous scattering
F^+, F^-	Friedel related derivative amplitudes
ΔF_{PH}	anomalous scattering observed difference
OD	optical density
HRS	Hamilton, Rollett and Sparks (scaling)
P, PH	protein, protein and heavy atom(s)
H	heavy atom
$\sum_h$	sum over all hkl
$F(h)$	shortened form of $F(hkl)$
$\langle A, B \rangle$	the mean value of A and B
RMS()	root mean square quantity

CONTENTS

	Page.
Chapter 1. Introduction and history.	1
Chapter 2. Crystallographic method and crystallography of native prealbumin.	10
Chapter 3. Heavy atom derivatives of prealbumin.	22
Chapter 4. Measurement and processing of 3D data from prealbumin crystals.	52
Chapter 5. Solution of 3D Patterson, refinement and phasing.	80
Chapter 6. The tertiary structure of prealbumin.	103
Chapter 7. The interaction of T3 and T4 with prealbumin.	113
Appendix 1. Statistical quantities used in the data processing.	120
Appendix 2. Centric phase determination.	121
Table of α co-ordinates.	122
Table of structure-factors and phases.	125
Bibliography.	139

AMINO ACID.	ESTIMATED NO. RESIDUES PER PROTEIN MOLECULE.		
	PREALBUMIN (Gonzalez)	PREALBUMIN (Raz)	RBP (Kanai)
LYS	35	30	11
HIS	26	15	2
ARG	16	14	15
ASP	33	33	29
THR	45	42	10
SER	39	40	11
GLU	54	46	21
PRO	34	30	6
GLY	40	37	12
ALA	49	45	15
CYS	4	4	6
VAL	51	41	14
MET	4	4	4
ILE	20	18	4
LEU	28	25	14
TYR	19	18	8
PHE	19	19	11
TRP	17	8	-

Table 1. Published amino acid compositions for prealbumin and RBP.

CHAPTER 1.

Introduction and history.

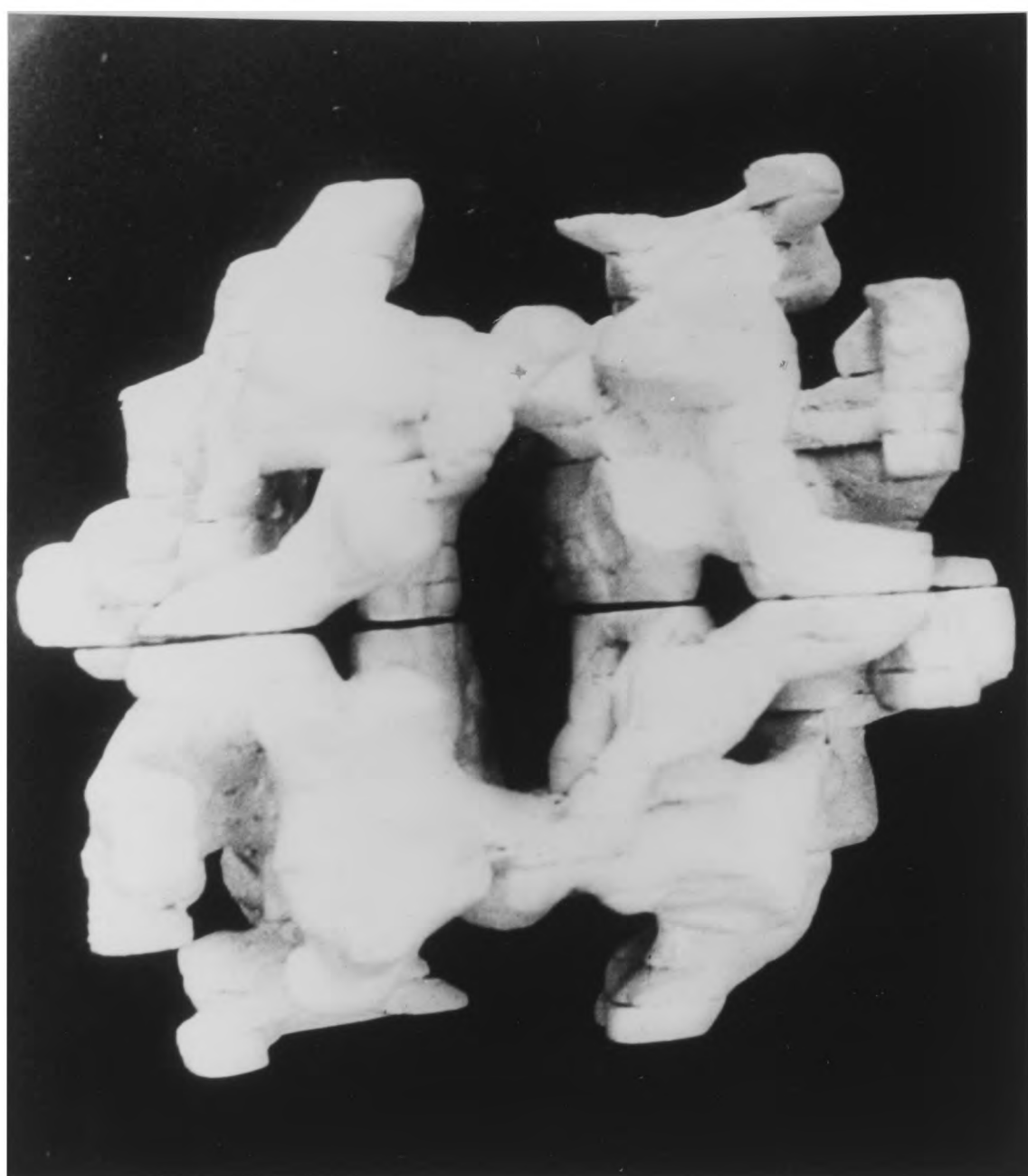
This work on the structure of human serum prealbumin has been undertaken because it has been established that this protein is responsible for the plasma transport of vitamin A in man and probably in most animals. Apart from this role, conclusive evidence has been presented that the protein is capable of binding the hormone L-thyroxine. These two phenomena make the knowledge of the 3-dimensional structure of prealbumin a matter of considerable interest and a preliminary to the eventual complete understanding of the interactions involved. The survey presented below is intentionally detailed as the work of numerous workers has not been adequately reviewed, and this study is only a starting point for further analysis of a very complex system.

The first isolation of prealbumin was reported in 1956 (Schultze, Schonenburger and Schwick). Its name derives from its ability to move faster than serum albumin on electrophoresis. The molecular weight of prealbumin reported in the literature varies greatly, but the best estimate, and that which agrees with published amino acid compositions, appears to be 56 000 Daltons. (This is also the mean of the published values).

The amino acid composition analyses for prealbumin of Raz and Goodman (1969) and Gonzalez and Offord (1971), which are in general agreement, are shown in table 1. The moderately high content of tryptophan (there is a greater-than-average content of aromatic amino acids in prealbumin) gives rise to the old name of tryptophan-rich prealbumin in the literature.

Quaternary structure of prealbumin.

In 1971 several papers appeared, all presenting experimental evidence that prealbumin was a tetrameric protein consisting of four



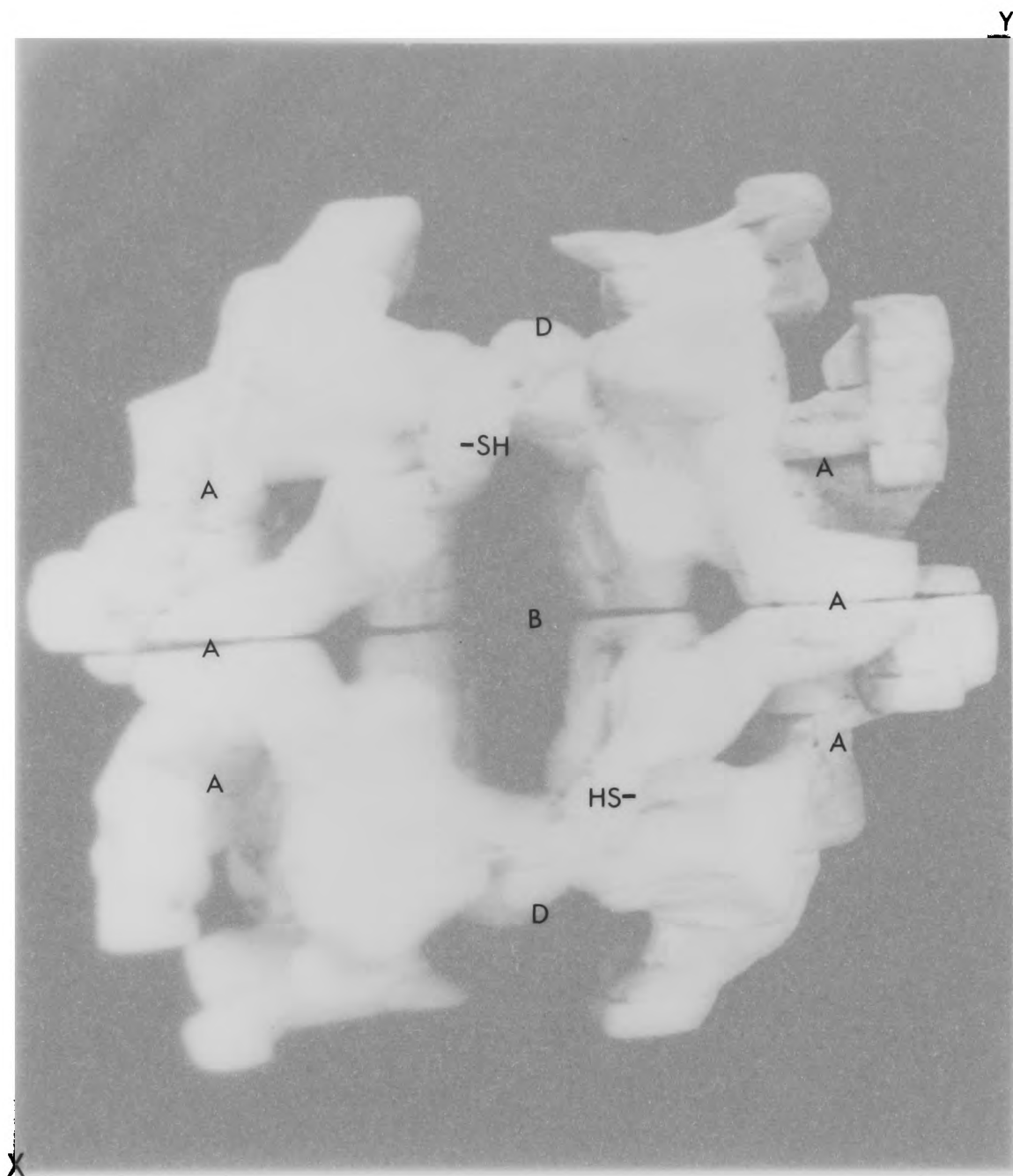


Fig. 1 The balsa wood model of prealbumin at 6 Å resolution.

The overlay shows the heavy atom sites referred to by table 4. -SH marks the assumed position of the single subunit cysteine residue (from mercury binding evidence)

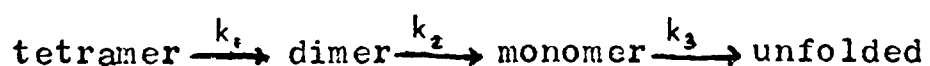
identical subunits. All of these put the subunit molecular weight at about 14 000 Daltons. Gonzalez and Offord (1971) compared the sum of the amino acids of the tryptic peptides with the amino acid composition of the entire protein and found four phenylthiohydantoin amino acids (glycine) from the Edman degradation of the N-termini of the protein. Similar results were obtained by Rask and others (1971).

Results of a different kind came from an examination of the diffraction pattern of crystalline prealbumin (Blake and Swan, 1971). The diffraction pattern indicated the primitive space group $P2_12_12_1$. However, there was extreme pseudo-symmetry in the pattern. This is usually an indication of molecular symmetry, commonly displayed by subunit-containing proteins. As is explained in greater detail in chapter 2, the crystallographic results could only be interpreted in terms of a tetrameric protein with closely similar subunits. One dimer is in the crystallographic asymmetric unit and the two constituent monomers are related by local symmetry.

A 6 Å resolution Fourier synthesis of prealbumin, based on five isomorphous derivatives (Blake, 1972) revealed a roughly spherical molecule (50 x 55 x 45 Å) with a very pronounced slot along the two-fold (z) axis of the molecule (fig.1). At this resolution, it is often possible to distinguish secondary structures such as α-helix or β-sheet. In this map of prealbumin no α-helical regions of any length can be distinguished. Either side of the slot, however, there are dense wall-like structures which might arise from β-sheet although it is not obvious how the strands are arranged within the walls. Secondly, it is not possible to distinguish the individual subunits, which can usually be achieved at this resolution. All that may be surmised is that the subunits run the full length (vertically in fig.1) of the molecular view presented, because the x-repeat corresponds to a distance of only 43 Å.

Many workers have commented on the unusual stability of prealbumin. The molecule undergoes only slow dissociation at neutral pH in 8 M urea and is undissociated in strong acid or alkali or 0.1% sodium dodecyl sulphate (Branch and Edelhoch 1971). These workers have studied prealbumin in water and guanidine solutions. In water the far ultraviolet circular dichroism spectra indicated that about half the peptide groups were incorporated into β -structure and that the remainder had essentially no secondary structure. They report a constancy of the β -structure over a pH range of 3.5 to 12.0. Below pH 3.5 there was a structural transition which involved exposure of all the buried aromatic chromophores. Gotto et al (1972) also agree that prealbumin contains little or no α -structure. Rask et al (1972) suggest that 60-80% of prealbumin is a rigid, non-periodic structure and the remainder random coil.

In guanidinium hydrochloride solutions below 6 M and in urea at this concentration, prealbumin was indefinitely stable (Branch and Edelhoch, 1972). At higher concentrations a slow dissociation occurred, increasingly rapid at pH values below 3.5. This stability is unique to prealbumin compared with many common proteins. A comparison of amino acid content of several globular proteins led to the implication of β -structure to explain this dissociation stability because only in this structural aspect was prealbumin atypical of the proteins examined. The kinetics of denaturation showed at least three different rate constants. This suggests a possible dissociation sequence:



where k_2 is slow enough to allow an appreciable concentration of dimer to exist. Further evidence for the dissociation sequence as shown comes from SDS-gel electrophoresis (Branch and Edelhoch, 1971) where a band of 32 000 molecular weight appeared in addition to bands

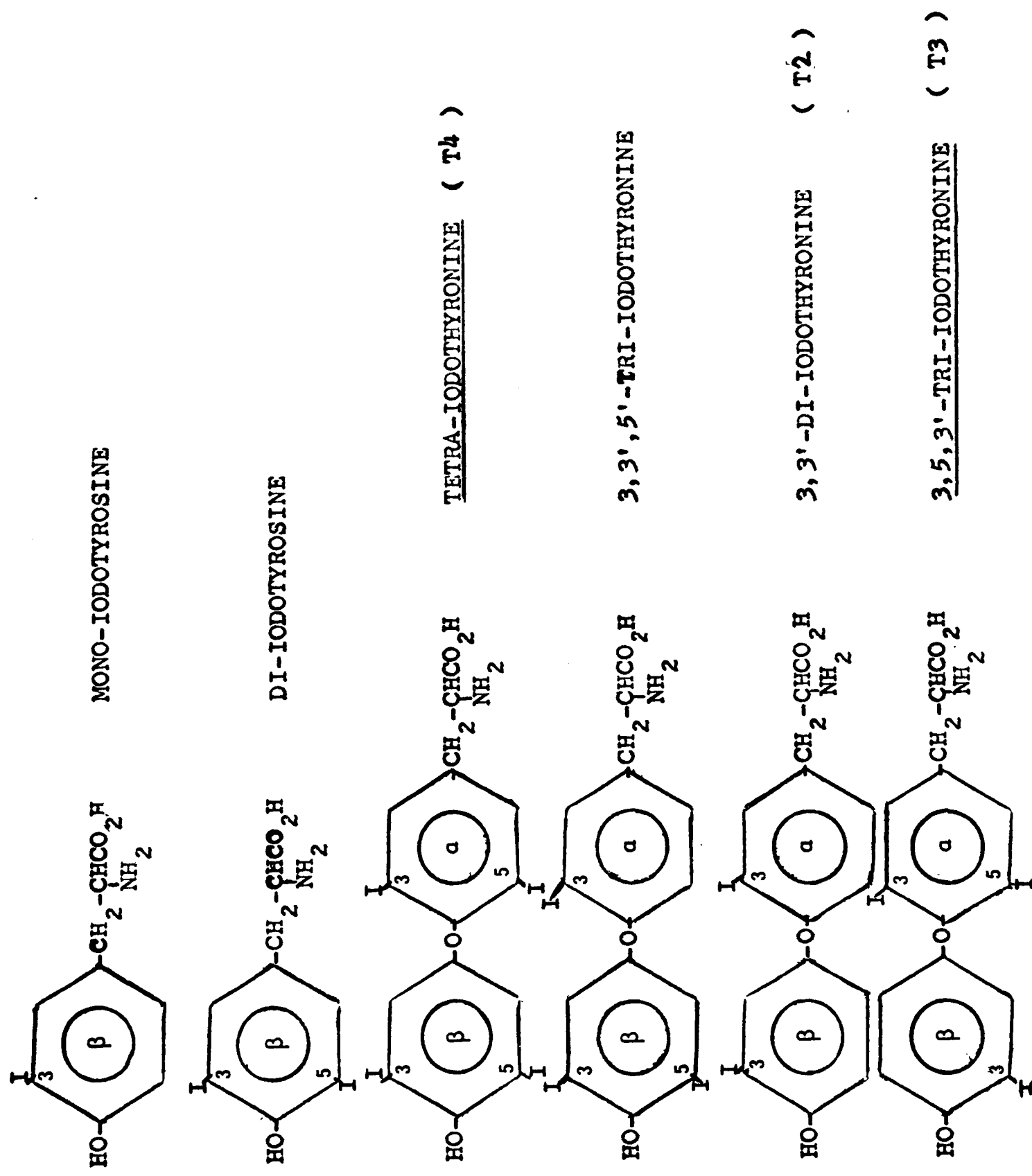


Fig. 2 The iodo-aminoacids.

of 56 000 (tetramer) and 14 000 (monomer).

Most workers report that the cysteine residues of prealbumin are unreactive except under highly denaturing conditions e.g. (Rask et al., 1971), suggesting they are in a highly hydrophobic environment. The positions of the sulphhydryl groups in the prealbumin monomer were determined crystallographically (Blake, 1971) by their interaction with mercuric acetate and these are marked on the model in figure 1. It now appears that only the bare mercuric ion will react, organic mercurials especially aromatic ones, being excluded in some way.

T3 and T₄ binding.

It has long been known that prealbumin is one of three proteins in human serum which are capable of binding the thyroid hormone L-thyroxine (T₄). Figure 2 shows the chemical structure of the thyroid hormones and table 2 the affinities of the plasma proteins for them. It is clear from table 2 that prealbumin cannot be considered a major carrier of T₄ in the presence of thyroxine-binding globulin (TBG) in serum. However it has been pointed out, (Davis et al, 1972) that in TBG-deficient patients, prealbumin may well be important in this role.

T₄ is only sparingly soluble in human serum at physiological pH, so the binding studies of T₄ to prealbumin and other proteins have made use of ¹³¹I-labelled T₄. Much of the early work of T₄ binding was done with unpurified, whole serum. It has now been established (Davis, 1972) that this gives rise to complex reactions not observed with purified protein preparations. Also the T₃ and T₄ binding experiments were carried out in a barbitone buffer system. Ingbar (1963) observed that sodium barbiturate competed with T₄ for common binding sites on prealbumin.

Ross and Tapley (1956) tested various inhibitors of T₄-binding to prealbumin, TBG and serum albumin. They showed that for maximal

Thyroxine-binding protein.	Normal serum levels. mg/100 ml	Binding constants. (M ⁻¹)		
		T4	T3	T2
Thyroxine - binding prealbumin.	20 - 30	2.35 x 10 ¹⁰	-	-
Prealbumin.	20 - 30	1.3 x 10 ⁸	1.2 x 10 ⁷	6.2 x 10 ⁴
Serum albumin.	-	1.4 x 10 ⁶	-	-
Retinol - binding protein.	3 - 4	1.0 x 10 ⁷ (for prealbumin)		
Retinol	0.5	-		

Table 2. Some recent estimates of the normal serum levels of prealbumin, RBP and other thyroxine-binding proteins, with their associated binding constants.

binding to prealbumin in whole serum, a di-iodo substituted phenolic ring with a 4'-hydroxyl group was essential, (see figure 2). The binding of T3, T2 and T1 was weaker than that of T4. Raz and Goodman* (1969) measured the association constant of prealbumin for T4 as $1.6 \times 10^7 \cdot \text{M}^{-1}$ at pH 7.4, using a pure preparation of prealbumin. They could identify only one binding site per molecule. The original reaction solutions were isotonic with blood, but removal of the salt did not affect the binding of T4.

Nilsson and Peterson (1971) have presented evidence for one major binding site for T4 and three others of low affinity. This observation would be the one expected for a tetrameric protein displaying a negative homotropic effect. Nilsson and Peterson also give the affinity constants of prealbumin for T4, T3 and T2 as 1.1×10^7 , 1.3×10^6 , and $6.2 \times 10^4 \text{ M}^{-1}$ respectively. Variations of the pH exerted a marked effect on T4 binding. At pH 7.4, where the phenolic hydroxyl group of T4 is ionized, they achieved optimum binding. At pH values lower than this the binding decreased markedly. They suggest that the ionized phenol is involved in the binding of T4 to prealbumin.

Gotto et al (1972) followed the binding of T4 to prealbumin by monitoring the circular dichroism spectrum of the protein. They found no appreciable perturbation of the spectrum at T4 concentrations known to saturate the protein. If the model of homotropic negative effects is to be believed, this result is surprising, as it suggests minimal or no conformational changes on T4 binding.

Green (1972) was able to bind 1,8-anilidonaphthalene sulphonic acid (ANS) to prealbumin. ANS and related dyes adsorb to hydrophobic sites on proteins. Addition of T4 caused quenching of the ANS fluorescence through its release into the solvent, which was complete upon the addition of one mole of T4 per mole of protein.

* These workers recently (1973) re-measured the association constants and their new published values are shown in Table 2.

In view of the reported lack of conformational change on T₄ binding, ANS probably binds at or close to the T₄ binding site, showing this site to be a relatively hydrophobic one. Green's evidence also points to a 1:1 complex of T₄ with prealbumin.

No worker has yet connected particular amino acid residues with T₄ binding, but Nilsson and Petersen (1971) reported that alkylated (carboxymethyl cysteine) prealbumin was unable to bind T₄. Raz et al. (1970) were unable to detect loss of T₄ binding capacity using different alkylating conditions. (In view of the inaccessibility of the cysteine there must be some doubt as to whether this group is involved in T₄ binding, but alkylation of it would probably cause extensive conformational changes in the protein).

Investigations of T₃ binding to prealbumin are relatively recent and have been assisted by the recognition that the thyroid hormones are displaced from the serum binding proteins by various small organic molecules, (Christensen, 1959; Ross and Tapley, 1966). The most frequently reported inhibitors are salicylate, 2,4-dinitrophenol and diphenylhydantoin (Osorio, 1962). Larsen (1972) reported that during aspirin therapy, the concentrations of salicylate in the blood serum of patients caused a 100-200% rise in free T₄.

Davis et al, (1972) using ¹²⁵I-T₃, found that prealbumin in whole serum bound 7-9% of T₃, but that prealbumin in serum deficient in TBG bound more than 30% of T₃. They went on to show that although additions of T₄ to ¹²⁵I-T₃-TBG displaced label from TBG to prealbumin and serum albumin, loading of T₄ did not lead to significant loss of labelled T₃ from prealbumin even at T₄ levels known to saturate prealbumin. Loading of serum with unlabelled T₃ caused loss of 50% of ¹²⁵I-T₃ from TBG, prealbumin and serum albumin. Inhibitors of T₃ binding to TBG such as diphenylhydantoin and salicylate showed displacement of labelled T₃ to prealbumin and serum albumin.

1. NH_2 -GLY-PRO-THR-GLY-THR-GLY-GLU-SER-LYS-CYS-PRO-LEU-MET-VAL-LYS-VAL-LEU-ASP-ALA-VAL-ARG-GLY-SER-PRO-ALA-ILE-
-ASN-VAL-ALA-VAL-HIS-(ASP, THR, SER, GLY)-LYS
 2. LYS-THR-SER-GLU-SER-GLY-GLU-LEU-THR-GLU-GLN-LEU-HIS-GLY-PHE-VAL-GLU-GLY-ILE-TYR-LYS-VAL-GLX-ILE-ASX-THR-LYS-
-SER-TYR-TRP-LYS
 3. LYS-ALA-LEU-GLY-ILE-SER-PRO-PHE-HIS-GLN-HIS-ALA-GLU-VAL-VAL-PHE-THR-ALA-ASX-SER-GLY-PRO-ARG-ARG-TYR
 4. ALA-LEU-LEU-SER-TYR-SER-TYR-SER-THR-THR-ALA-VAL-VAL-GLX-ASX-PRO-THR-LYS
 - a. Tryptic peptides of prealbumin.
 - b. Peptide of RBP.
- NH_2 -GLU-ARG-ASP-CYS-ARG-VAL-SER-SER-PHE-ARG-VAL-LYS-GLU-ASN-PHE-

Fig. 3. Amino acid sequences of prealbumin and RBP (Gonzalez and Offord, 1972 ; Morgan et al., 1971)

Although these inhibitors also displace T₄ from prealbumin, they had only minimal effects on T₃ bound to prealbumin. These results seem to be in conflict with the generally accepted association constants for the binding of the two hormones to prealbumin.

Retinol binding protein and prealbumin.

In 1966 Alvsaker and others isolated a protein fraction from serum which contained vitamin A (retinol). They identified the protein as prealbumin and proposed that this was the hitherto unknown plasma transport protein for vitamin A. Subsequently it was shown by Kanai et al. (1968) that retinol is bound to a peptide of molecular weight 21 000 Daltons. This is in turn bound to prealbumin to give the protein-protein-vitamin complex of molecular weight of 71 000 Daltons, which was first observed by Alvsaker. The retinol-binding protein (RBP) is composed of a single polypeptide chain. Partial sequence information is available (fig.3) and shows no resemblance to that of prealbumin. The studies of Raz and Goodman (1969) showed that 1 mole of RBP associated with 1 mole of prealbumin and that 1 mole of retinol associates with 1 mole of RBP. The retinol binding was non-covalent, and retinol could be extracted into heptane from the RBP-retinol complex. The ultraviolet spectrum of bound retinol was little different from that obtained in normal organic solvents. Several experiments were designed to determine the effect of the presence of RBP-retinol on the binding of T₄ to prealbumin and vice versa.

Raz and Goodman (1969) found that these two functions of prealbumin were independent. In circular dichroism studies Gotto et al. (1972) found that the spectra of RBP and prealbumin are additive. That is, no detectable perturbation of structure occurs upon the formation of a complex. In view of this, and the evidence that prealbumin is a tetramer of four identical subunits, it would

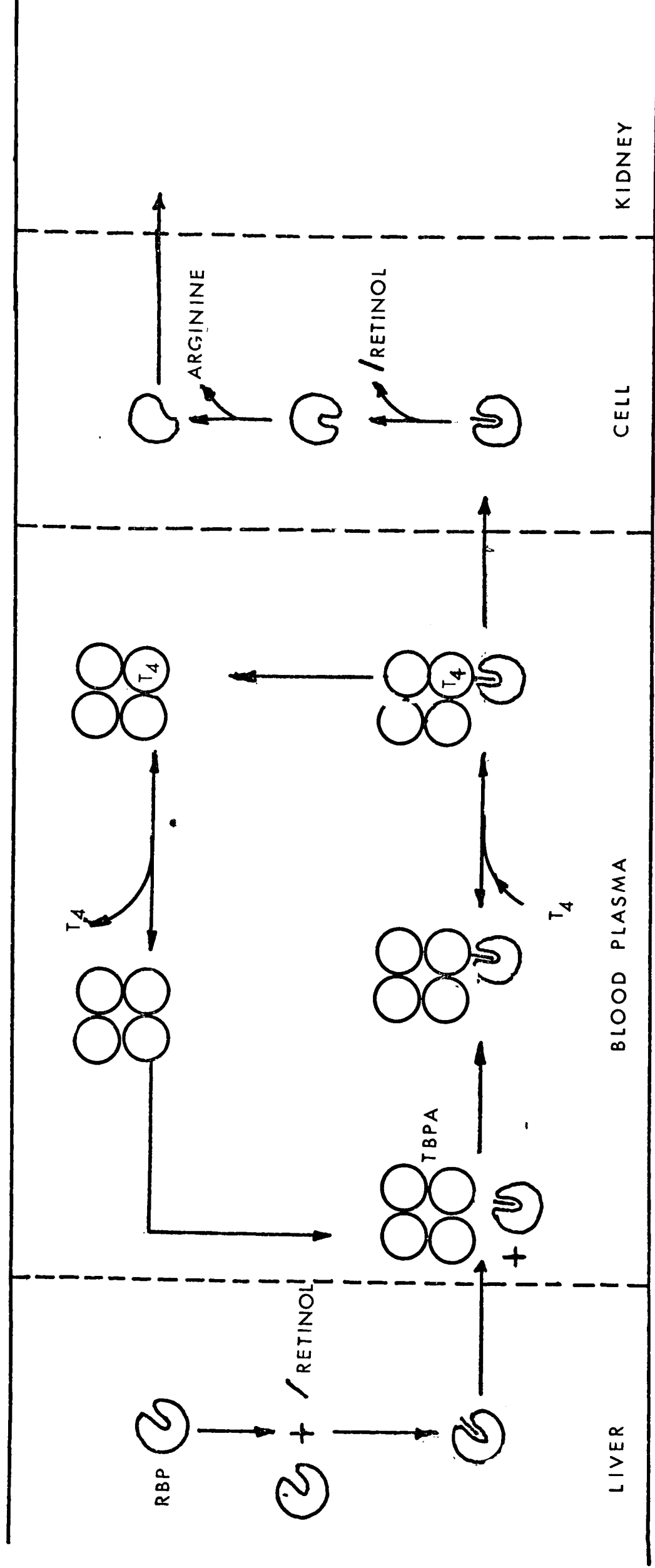


Fig. 4 Interactions of prealbumin(TBPA), RBP, T₄ and Retinol.

be surprising if there were only one binding site for RBP. Gotto and co-workers also showed that no alteration occurred in the circular dichroism spectrum when T_4 was bound to the prealbumin-RBP complex. The recent work of Jaarsveld et al., (1973a,b) has demonstrated four independent binding sites for RBP on both human and Rhesus monkey serum prealbumin.

The complex formation stabilizes the retinol-RBP association. Goodman and Leslie (1971) showed that when RBP-retinol complexed with prealbumin, the retinol was much less readily extractable by organic solvents, and fluorescence measurements showed that the retinol had removed from an aqueous environment. Furthermore the thermal stability of the RBP-retinol complex is enhanced in the presence of prealbumin. Peterson (1971) has shown that the binding of RBP to prealbumin is dependent upon the ionic strength of the environment and has isolated a catabolite of RBP which has a missing N-terminal arginine. This modified RBP does not display binding to prealbumin.

Summary.

A summary of the interactions of prealbumin and RBP is shown in fig. 4. The fate of RBP in the cell compartment is speculative, based on the observations by Rask et al., (1971), and Kirby et al., (1971).

Retinol combines with RBP in the liver and this complex is picked up in the plasma by prealbumin. The prealbumin stabilizes the RBP-retinol complex and prevents loss of the vitamin through the kidneys. At the receptor site the RBP dissociates from prealbumin and releases the retinol. A peptidase then removes the N-terminal arginine residue from the RBP, which is unable to recombine with prealbumin in this form. This is important since it leaves the prealbumin free to recombine with more RBP-retinol complex.

In this model, the RBP is seen to act at a cellular level, making a one-way trip with its vitamin and prealbumin acts as a shuttle in

the plasma which is continuously available for transport.

The observations on prealbumin and RBP pose the following questions:

1. What structural features give prealbumin its unusual stability towards denaturants ?
2. Where and of what nature are the binding sites for T₄, T₃ and the inhibitors such as salicylate and 2,4-dinitrophenol ?
3. Where and how does RBP interact with prealbumin ?
4. Why does this tetrameric protein bind only one molecule of T₃ or T₄, and what is the connection between the binding of the vitamin through RBP and the binding of the hormone ?

CHAPTER 2.

Introduction and definition of general crystallographic concepts used in this thesis.

The purpose of this section is to outline the concepts employed in crystal structure analyses. In particular an attempt is made here to introduce some of the terminology used throughout this thesis. No attempt will be made to prove any of the major, stated theorems, but the reader will be referred to a standard work on crystallography such as Woolfson (1972) or Stout and Jensen (1968). Each relevant aspect of the crystallographic method is treated with greater detail in subsequent chapters.

Each crystalline substance possesses, by way of a permanent property, a characteristic diffraction pattern with respect to irradiation by X-radiation. This pattern arises from the scattering of X-rays by the constituent electrons of the molecules within the crystal. The molecules within the crystal, by definition, are highly ordered and may be considered to be associated with a lattice which acts both as a reference frame, and can express the repetitiveness of order in the crystal. In their most simple form 14 such space lattices can exist. The simplest translations in each of the three 3-dimensional lattice directions which traverse a unique part of the crystal space, define a unit cell. A repetition of this unit cell would generate the entire crystal. The molecules themselves may possess symmetry, such that a smaller part of the intact molecule, by the operation of a rotation or reflection, might generate the whole molecule. A description of such a molecular symmetry is termed the point group of the molecule. The combination of these point groups with the 14 space lattices leads to 230 possible space groups. For proteins certain point groups, comprising the symmetry elements of mirror plane and inversion centre, are excluded

as the protein is an enantiomorphic structure built of left-handed amino acids.

The existence of such molecular symmetry allows the smallest unit from which the crystal may be built, to be smaller than the unit cell. It is termed the asymmetric unit of the crystal. It is to this unit of structure that all the calculations of crystallographers refer, since this only is mathematically unique.

It is well known from optical diffraction theory that a grating of slits gives rise to a series of fringes, when bathed in visible radiation. The separation of these fringes is inversely proportional to that of the slits. Similarly the 3-dimensional array of scattering matter in a crystal gives rise to a 3-dimensional set of fringes on irradiation with X-rays. It can be shown that the resulting pattern is only observable at points where the fringes intersect, in other words, on the points of a space lattice. This lattice shows the same phenomenon of reciprocity as the optical case and is termed the reciprocal lattice. The separation of the slits of the grating correspond to the three unique unit cell translations of the crystal.

As for optical diffraction, the distribution of intensity in the diffraction pattern is a function solely of the nature of the scattering object. In this case the object is the electronic configuration within one asymmetric unit of the crystal. Only part of this pattern (called the molecular transform) can be observed; that at the points of the reciprocal lattice.

The extent to which the diffracted radiation diverges from the incident radiation depends upon the closeness of diffracting planes within the crystal. This is such that the highest angles of diffraction correspond to the closest planes. Thus these diffracted rays contain information about fine structure within the asymmetric

unit, whilst low angles of diffraction yield information of a coarser nature. The crystallographer measures the intensity of the molecular transform at the reciprocal lattice points. He states the resolution of his intensity data as the finest spacings in Angstrom units which have given rise to that data. The conditions for obtaining diffraction from a crystal have been stated by Bragg (1913) and Laue. Ewald has contributed a geometrical construction to relate these conditions to the reciprocal lattice. The instruments used by the crystallographer; diffractometers and cameras have geometries which exploit the construction of Ewald to enable the automatic recording of the intensities at points of the reciprocal lattice. From these intensities the crystallographer derives a structure amplitude F_{hkl} where h, k, l , are integers used to denote a set of similar diffracting planes within the crystal. Let the unit cell translations be a, b , and c Angstroms respectively and distance between a set of diffracting planes be d . Then the set of planes giving rise to a diffracted amplitude F_{HKL} are those, spaced d Angstroms apart, which make intercepts of a/h , b/k , c/l with the unit cell axes.

It can be seen that as h , k , or l increases, the separation of the planes decreases and the information is of an increasingly high resolution. Clearly the magnitude of F_{hkl} is dependent on the number of contributing planes $\{hkl\}$ and if a , b , and c are large, as is the case for protein crystals, the crystal must be correspondingly large for adequate diffraction to take place.

Individual variations of the size of F_{hkl} from a single crystal depend upon the population of the plane $\{hkl\}$ by electrons. If the plane contains a heavy atom (uranium or lead) the contributions from this atom's electrons may lead to a larger F than a plane containing only the electrons from light atoms (carbon or oxygen).

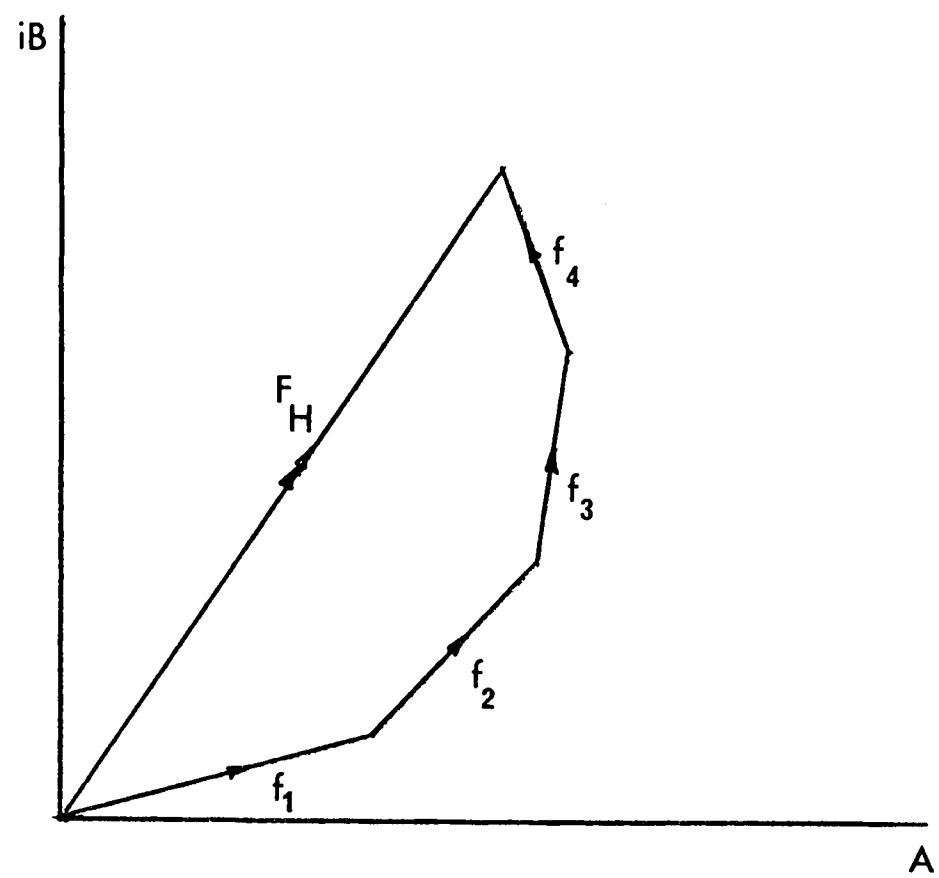


Fig. 5a. Structure factor summation.

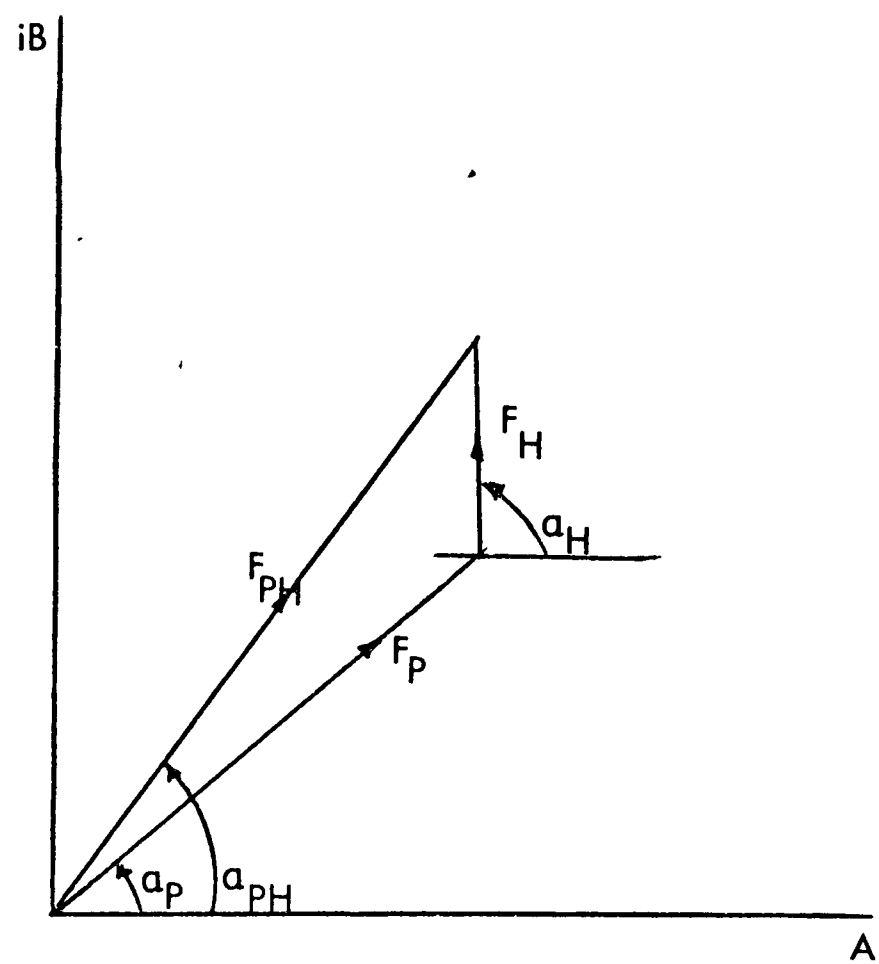


Fig. 5b. The protein-protein plus heavy atom vector triangle.

The scattering of X-rays from these individual atoms is also dependent upon the angle of scattering, the diffraction being less effective at higher angles. The scattering of these atoms expressed as a function of angle is known as a form factor f_j , and is given the value corresponding to the number of electrons in the atom (j) at zero angles of diffraction.

An origin is chosen at one corner of the unit cell which contains N atoms (j), where the fractional co-ordinates of the atomic positions are x_j , y_j and z_j along the unit cell translations. It can be shown that

$$F(hkl) = \sum_j f_j \exp [2\pi i (hx_j + ky_j + lz_j)] \quad (1)$$

The Argand diagram (fig.5) is a convenient graphical representation of the result of equation (1). For the reflection from a plane $\{hkl\}$ the scattering from each individual atom j in the unit cell is represented by a vector of magnitude $|f_j|$ and phase $2\pi (hx_j + ky_j + lz_j)$. This construction shows how the structure factor F_{hkl} is formed. In protein crystallography the positions of the atoms (j) remains unknown until a high resolution map is produced and interpreted. The only atomic positions known in the early stages of a protein crystal analysis are those of introduced heavy atoms. At all stages prior to actual structure determination, the protein crystallographer is concerned only with the simplified Argand construction in fig. 5b. Here $|F_P|$ and $|F_{PH}|$ are the measured amplitudes of a single reflection from the Protein and Protein plus Heavy atom crystals respectively. The vector F_H is the structure factor from the heavy atom solution alone which may involve one or more different atoms. The phase angles α_P and α_{PH} are unknown.



Fig. 6 A PAQ crystal after soaking in di-thiothreitol.

Table 3. Some of the cell parameters of crystals of human serum prealbumin.

Parameter	Quantity		
	Real space	Reciprocal space	Reciprocal space (r.l.u.)
a	$43.3 \pm 0.2 \text{ \AA}$	$0.02309 \text{ \AA}^{-1}$	0.0356
b	$85.6 \pm 0.5 \text{ \AA}$	$0.01167 \text{ \AA}^{-1}$	0.0179
c	$65.8 \pm 0.3 \text{ \AA}$	$0.01518 \text{ \AA}^{-1}$	0.0235
	90°	90°	90°
Cell volume	$244434 \text{ \AA}^3$	$0.409 \times 10^{-5} \text{ \AA}^{-3}$	$1.495 \times 10^{-5} \text{ \AA}^{-3}$
Crystal system	Orthorhombic		
Space group	$P2_12_12$		
Water content	$30.0 \pm 2.0\%$ by weight		
Crystal density	$1.261 \pm 0.006 \text{ g/cm}^3$		
Mol.wt. in cell	108 000 Daltons		
No. of mols. in cell	2		
No. of mols. asymm/u	$\frac{1}{2}$ (27 000 Daltons)		

To obtain an image of the electronic structure of the molecules in the crystal, the Fourier inversion is used :

$$\rho(xyz) = \frac{1}{V} \cdot \sum_h \sum_k \sum_l \cdot F_{hkl} \exp[-2\pi i (hx + ky + lz)] \quad (2)$$

where $\rho(xyz)$ is the electron density (intensity) at the point xyz in the unit cell, V is the unit cell volume in cubic Ångströms.

The difficulty lies in the fact that the quantity F_{hkl} is complex and the crystallographer only knows the magnitude. All the techniques of data collection measure the intensity which is a scalar quantity. It is with overcoming this difficulty ("the phase problem") that this thesis is largely concerned.

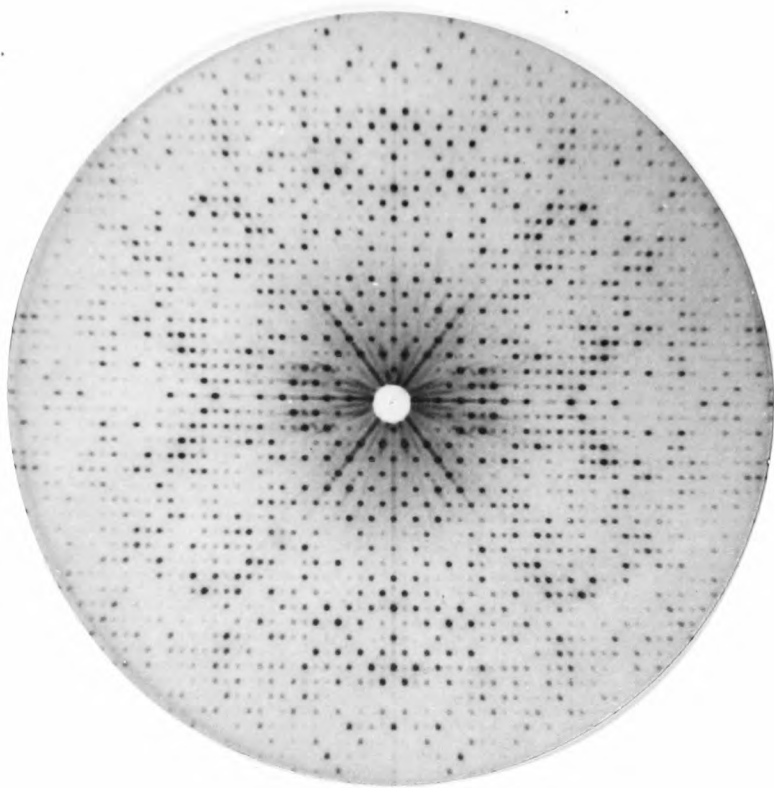
Crystallography of native human serum prealbumin.

Human serum prealbumin can be crystallised from phosphate buffers over a pH range of 3.7-9.0, with a wide variety of precipitants (Haupt and Heide, 1966).

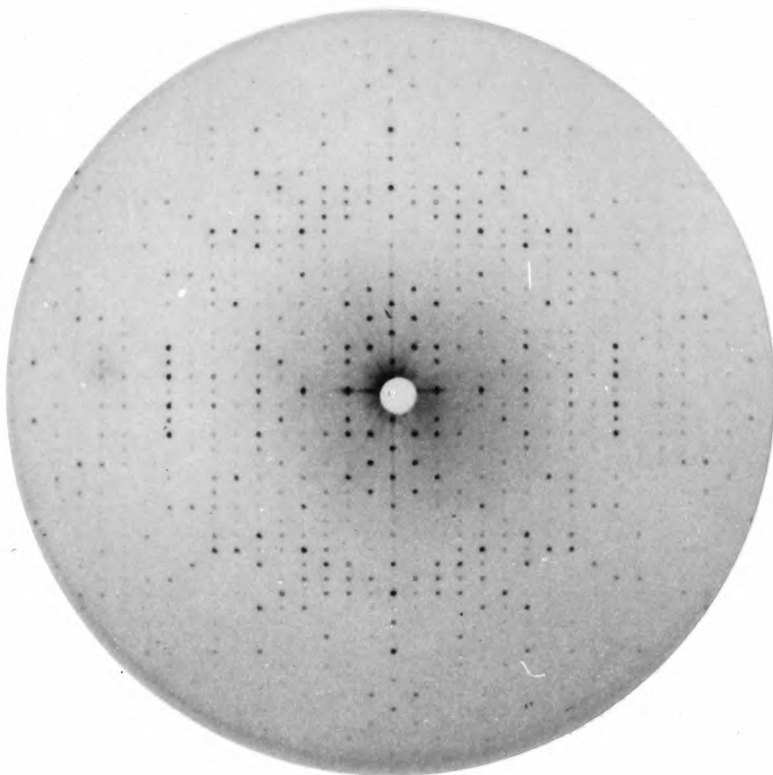
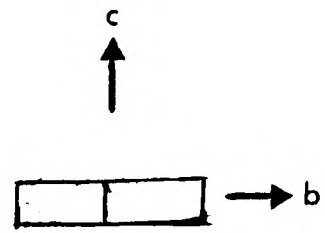
Native protein crystals suitable for X-ray data collection would be 0.5 mm on edge for photography and approximately 1 mm on edge for diffractometry. Native prealbumin crystals of this size have proved exceedingly difficult to grow. However, crystals of prealbumin co-crystallised with mercuric acetate in the form of diamond-shaped plates of up to 2 mm in length have been grown (fig.6). These were obtained as a gift from Behringwerke A.G., Marburg am Lahn, West Germany. Only this crystal form has been investigated crystallographically.

Figure 7 shows the three principal zones of the crystal diffraction pattern. The pictures were taken with two hours exposure, on a Buerger precession camera with film to crystal distance of 7.5 cm and a precession angle of 18° (see chapter 3 for details of photographic data collection). Table 3 lists some of the cell parameters of human serum prealbumin.

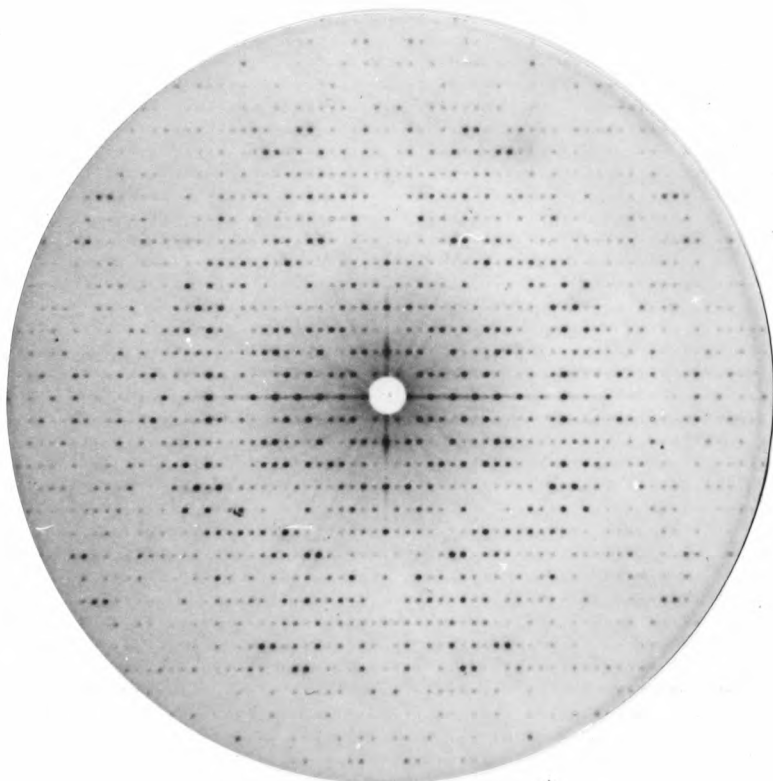
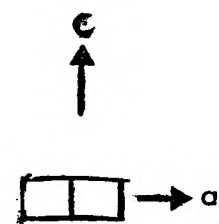
The determination of the crystal space group is due to



OKL



HOL



HKO

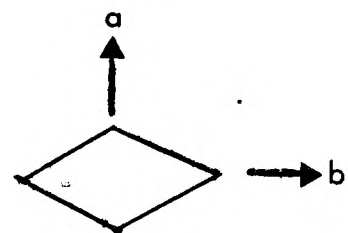


Fig. 7 The 3 principal zones of prealbumin native protein crystals. $a = 43.3 \text{ \AA}$ $b = 85.6 \text{ \AA}$ $c = 65.6 \text{ \AA}$.

Drs. C.C.F. Blake and I.D.A. Swan (Blake and Swan, 1971). Estimations of the water and salt content, together with the measured density of the crystals gave an estimate of 108 000 Daltons molecular weight in the unit cell (Swan 1970). The published values of the protein molecular weight then available were 50 000, 61 000 and 73 000 Daltons. The first two are in satisfactory agreement with the crystallographic data and give a value of 2 molecules in the cell.

Preliminary examination of the diffraction patterns showed that the odd-ordered axial reflections are absent on all three principal axes. The initial assignment of the space group was $P2_12_12_1$ (Rerat and Schwick, 1967), based on this observation. However, this space group has no positions on which two finite molecules can be located as the symmetry of the screws would lead to an infinite polymer.

Any model for the location of the protein molecules in the unit cell had also to account for the strong body centring apparent in the diffraction pattern. Examination of fig.7 shows that the reflections for which $h + k + l = 2n + 1$, near the centre of the diffraction pattern, are absent or weak. The effect also extends to the resolution limit of the pattern ($2.8 \text{ \AA}$) along the c^* axis. This pattern is characteristic of the body-centred space groups $I222$ and $I2_12_12_1$. This observation meant that the true space group had to be a subgroup of one of these pseudo-space groups.

The space groups $I222$ and $I2_12_12_1$ are indistinguishable on the basis of their diffraction patterns. However, the preparation of a single isomorphous derivative resolves the ambiguity, by the symmetry of the resulting Patterson function. Such a derivative was to hand in the PAQ crystals. Fig.8 shows the two sections, $U=0$ (a) and $U=\frac{1}{2}$ (b) of the $6 \text{ \AA}$ 3-dimensional difference Patterson. These two sections contain the only peaks. This shows that the heavy atom vectors are arranged in a pseudo body-centred array related by an

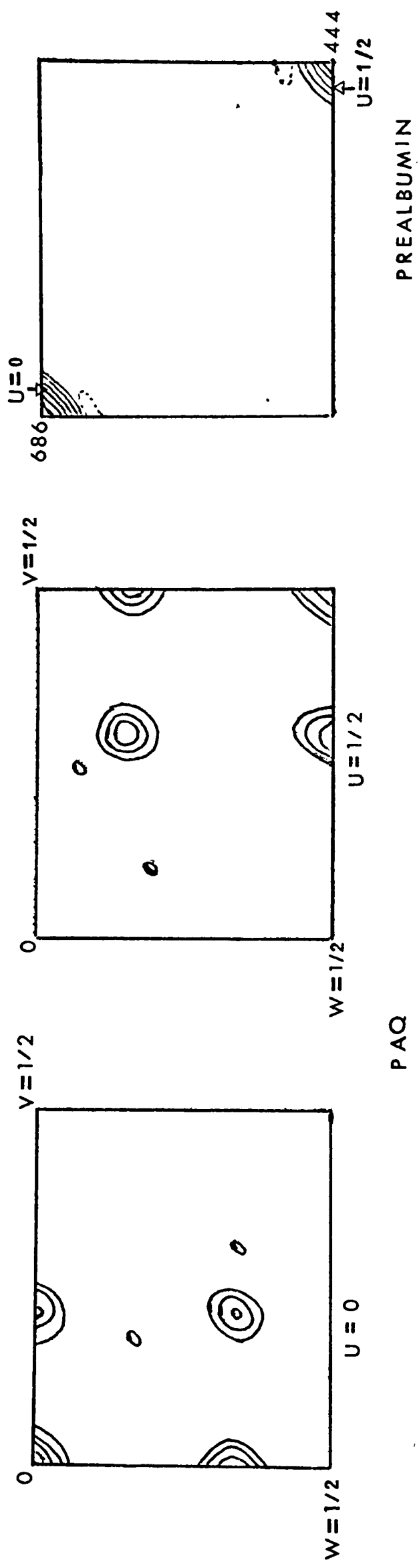


Fig. 8 Sections of the 3D difference Patterson of PAQ at 6 Å and the super-imposed sections U=0 and U=1/2 of the native protein self-Patterson at 6 Å.

inversion centre at $1/4, 1/4, 1/4$. The Harker peaks in the sections $W = 0$ and $V = 0$ have the same U co-ordinate, while the Harker peaks in the $W = 1/2$ and $V = 1/2$ have a co-ordinate $1/2-U$. This symmetry is characteristic of $I222$. The size of the peak at $1/2, 1/2, 1/2$ is an indication of the degree to which the heavy atoms are related by self-centring. From this it follows that the space group is either $P222$ or $P2_12_12_1$.

Two molecules can be accommodated in $I222$ if the molecules are located on 222 points. This suggests that the molecule is a tetramer, but obviously the full tetrahedral symmetry is not expressed in the cell, as this is primitive. Two conditions would lead to the necessary reduction in symmetry :

- 1). The molecule itself only possesses pseudo- 222 symmetry.
- 2). The two-fold axes of the tetramer are not co-incident with the cell axes.

From a consideration of model structures, Blake and Swan (1971) came to the conclusion that explanation 2. best fitted the experimental data. Figure 9 shows the projected arrangement of the prealbumin tetramer in $P2_12_12_1$. Here the molecule is located on a crystallographic two-fold axis so that a local two-fold axis is co-incident with it. In the diagram the shaded molecule is located at $z = 3/4$ and the unshaded one, at $z = 1/4$. The other local axes are offset by a small angle from the cell axes, leading to the necessary reduction in symmetry which allows the molecules to crystallize in a primitive space group. The effect of this offset of two of the local axes from the crystal axes explains the progressive loss of body-centring in the diffraction pattern and the appearance of odd-ordered terms. At low resolution the information contained in the diffraction pattern comes from gross structural elements such as large sections of density or protein/liquid boundaries. Such coarse structure may be expected

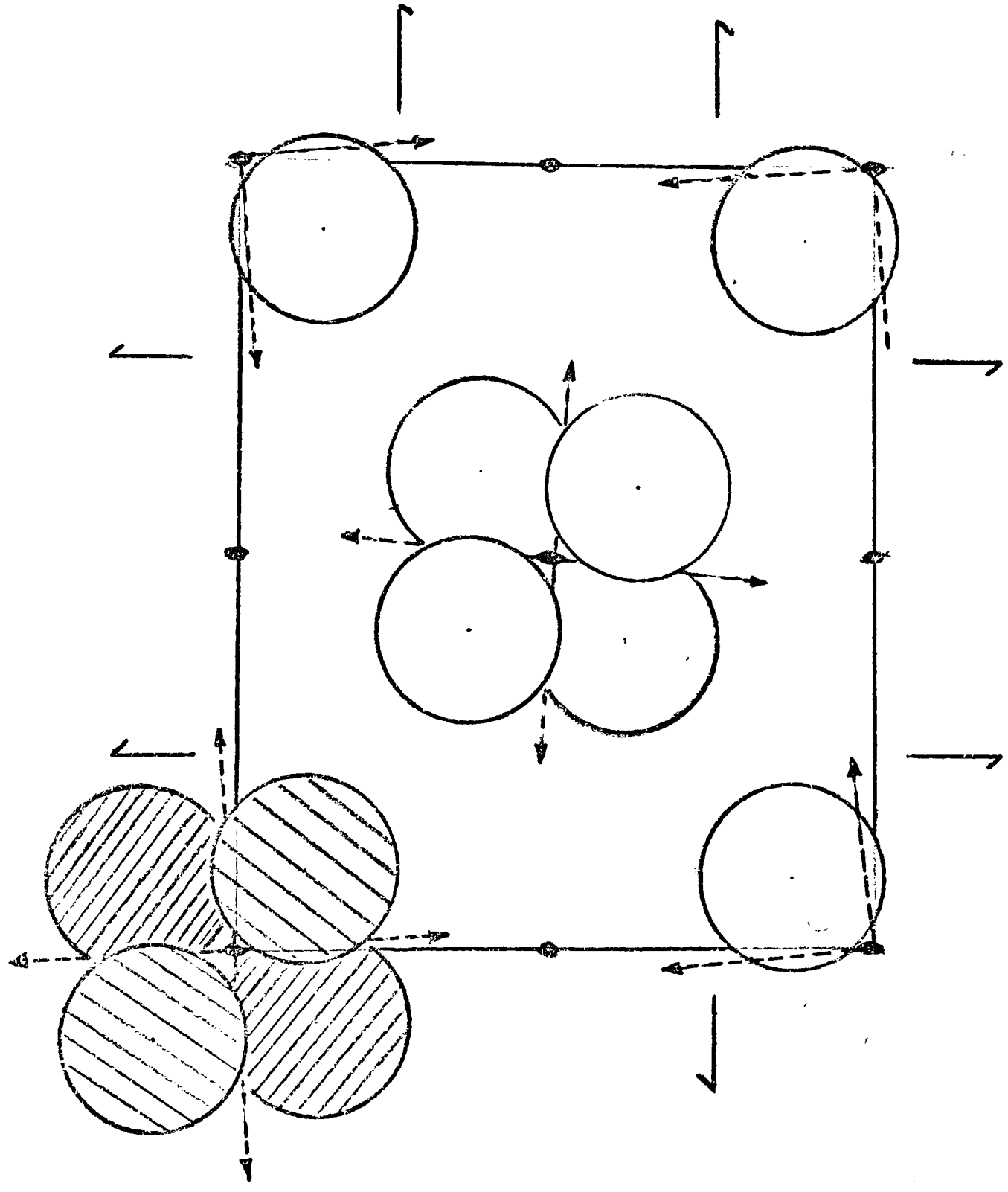


Fig. 9 The crystallographic model of Blake and Swan.

to be approximately related by a pseudo-twofold axis along the cell axis, leading to the weakening or extinction of the low angle terms. At a higher resolution, information in the diffraction pattern comes from structural details such as amino acid side chains and elements of tertiary structure, which are not related by the pseudo-twofold axis co-incident with that of the real cell. In this case there are no restrictions on the diffraction pattern and that of the true primitive cell appears.

The observed lack of centring to high resolution in the vicinity of the c^* axis, suggests that this axis is co-incident with one molecular two-fold (out of the plane of the paper in fig.9). An observation of the 0 0 17 reflection in a long exposure photograph of the OKL zone confirms this.

This model means that the two subunits in the asymmetric unit are related by a local symmetry which is not required to be exact and that the resulting dimer is exactly related to another by the two-fold symmetry of the crystal. The solution of the PAQ difference Patterson gives two sites in the asymmetric unit, presumably mercury bound to the single cysteine residues of each subunit. These sites are to a good approximation related by a local two-fold axis. Blake and Swan (1971) pointed out that if the molecular centres were located on the crystallographic two-fold axes at 0, 1/2, z and 1/2, 0, z, the angle that these local axes make with the a^* and b^* axes is $1^\circ 50'$. Otherwise if the molecular centres were at 0, 0, z and 1/2, 1/2, z, the offset would be $6^\circ 45'$. In view of the loss of centring at spacings less than $6 \overset{\circ}{\text{Å}}$, they suggest that the latter angle is the more likely. A Fourier synthesis of the protein at $6 \overset{\circ}{\text{Å}}$ resolution (Swan, 1971) showed this to be the case.

The presence of this degree of pseudo-body centring has had profound effects on the solution of the crystal structure of prealbumin.

It led to the prediction by Swan (1971) that serious difficulties would arise, particularly with the refinement of heavy atom positional parameters. This was indeed the case, with $6 \text{ \AA}$ data from prealbumin crystals and projection data to $3 \text{ \AA}$. The problem arises in these cases because of the weakness of the terms for which $h + k + l = 2n + 1$. This reduces the accuracy with which these terms can be measured, when clearly these terms are of extreme importance, because they contain information about the asymmetry of the structure with respect to the crystal axes. Neglect or poor phase determination of these reflections might lead to a final Fourier image containing false detail due to the presence of extraneous symmetry. The initial problems encountered by Swan and to a lesser extent by the author, were connected with the solution of heavy atom positions and refinement of these in just such a centred situation. These problems are discussed in greater detail in chapter 3.

Now that a full three-dimensional native data set has been collected to a resolution of $2.5 \text{ \AA}$, it can be seen that this centring ceases to be a problem. Figure 10 shows a plot of the mean native protein intensities for the odd and even ordered reflections against $\sin^2 \theta / \lambda^2$. An enormous discrepancy in these two intensity distributions is obvious to a resolution of approximately $5 \text{ \AA}$, but at spacings of $5 \text{ \AA}$ or less the two distributions are essentially the same.

The protein self Patterson has only two observable peaks, which conform to that expected of a highly pseudo-centred cell. These are an origin peak and one at $1/2, 1/2, 1/2$. When terms from planes with a spacing of not less than $6 \text{ \AA}$ are included in this synthesis, the peak at $1/2, 1/2, 1/2$ is 65% of the height of the origin peak (fig.8). When terms to $2.5 \text{ \AA}$ spacings are included, this peak

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17H
0	0	1196	676	498	121	256	242	127	109	56								
1	255	120	101	133	191	142	121	55	126	137	165	101	280	100	176	76	85	
2	448	0	359	186	404	159	221	106	65	254	223	339	389	75	101	199	248	58
3	388	297	167	196	340	222	206	537	295	212	84	108	93	70	68	198	55	
4	579	258	77	311	340	138	154	53	209	414	262	292	85	192	59	191	89	96
5	274	115	447	109	157	41	212	286	301	171	89	178	208	201	49	99	94	
6	953	85	499	186	181	0	59	713	95	212	341	467	88	153	113	38	66	119
7	577	67	450	42	330	128	605	300	518	136	112	539	120	99	82	164	0	
8	917	87	221	0	220	0	282	97	189	176	267	329	508	217	97	112	109	0
9	532	221	539	75	120	651	331	121	370	309	137	85	99	73	70	79	0	
10	942	191	672	365	103	0	165	373	212	81	118	495	71	148	182	76	53	0
11	63	394	394	146	171	399	244	85	235	89	615	221	0	66	63	51		
12	532	0	264	164	179	70	261	96	246	275	106	140	337	107	123	69	66	
13	300	277	211	40	190	383	480	161	48	174	226	143	150	138	108	229	102	
14	361	116	59	76	228	242	202	199	184	94	43	43	168	68	103	89	0	
15	361	124	223	709	81	413	203	454	287	188	202	202	132	104	96	0	0	
16	981	833	267	738	326	559	378	69	70	146	52	66	100	76	100	69		
17	223	350	75	831	64	110	204	153	0	144	120	222	222	66	175	29		
18	269	144	334	606	190	570	141	171	105	276	115	122	157	91	77	0		
19	254	275	369	328	43	103	147	47	93	75	83	149	87	159				
20	489	121	214	249	0	173	178	177	107	38	84	63	217	128	102			
21	571	53	56	86	197	124	216	111	100	137	84	142	58	0				
22	107	343	247	231	37	172	138	326	261	104	173	74	51	90				
23	115	70	316	367	54	0	114	197	79	59	146	162	0					
24	113	369	373	92	120	238	104	33	100	104	247	140	80					
25	149	85	170	191	206	219	127	159	383	124	10	0						
26	176	114	101	74	256	211	155	59	64	364	87	0						
27	256	109	93	91	106	158	401	222	165	75	0							
28	337	460	215	186	15	64	168	252	72	52	0							
29	103	281	86	186	23	109	31	97	97	0								
30	284	97	151	149	164	135	66	83	64									
31	285	87	135	119	97	123	169											
32	142	111	202	164	69	112	91											
33	0	0	0	0	0	0												
34	0	0	0															

Fig. 11a The HKO zone of native prealbumin structure amplitudes to 2.5 Å resolution.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	H
0	0	34	42	139	117	197	238	176	427	250	711	141	277	253	81	163	98	83	
1	0	55	521	82	463	64	157	164	253	111	370	254	116	161	88	153	129	86	
2	101	274	168	309	224	207	156	161	239	82	364	112	329	149	57	87	85	113	
3	199	131	128	0	232	80	53	189	223	110	378	284	454	96	77	202	94	111	
4	186	643	120	104	58	162	59	176	302	259	124	141	160	109	91	139	134	113	
5	188	0	564	484	146	216	132	388	44	429	147	62	173	147	73	92	81	62	
6	125	69	210	550	125	204	225	130	379	420	201	117	462	92	94	170	106	60	
7	752	186	893	141	317	109	360	155	90	169	109	260	285	180	102	147	101	0	
8	252	1080	212	523	199	168	106	281	104	251	247	235	201	146	91	67	121	0	
9	927	117	215	213	87	191	300	286	289	216	107	207	240	147	108	115	133	0	
10	63	452	367	300	27	101	60	119	217	106	149	199	96	97	114	130	61	0	
11	229	139	236	162	240	285	130	316	56	274	119	173	212	86	179	100	60		
12	149	322	93	217	262	221	84	146	108	72	214	332	134	290	196	138	81		
13	88	159	271	305	357	71	277	152	189	91	194	119	118	182	144	81	97		
14	199	263	112	311	77	54	233	168	273	73	76	117	158	100	90	142	0		
15	276	249	171	213	144	611	120	133	149	177	374	73	0	159	121	55	0		
16	222	324	128	78	356	182	434	265	200	182	0	146	208	11	62	0			
17	708	449	371	131	113	535	152	149	140	173	68	175	195	167	121	92			
18	119	414	203	202	183	153	152	311	325	92	40	141	138	106	80	0			
19	171	268	270	404	286	264	124	194	231	254	207	77	142	145	70				
20	82	408	446	190	156	238	135	204	102	70	98	222	179	74	98				
21	166	134	118	116	172	70	414	279	125	188	126	141	196	94	0				
22	46	114	154	216	254	105	144	340	145	58	130	0	103	0					
23	120	153	219	249	305	139	234	195	209	102	95	170	62	0					
24	110	110	200	171	111	243	283	195	127	186	81	102	152						
25	201	114	176	172	167	103	0	122	69	114	230	41	0						
26	135	269	169	124	245	92	88	0	109	72	102	0							
27	53	155	265	105	107	99	119	98	33	122	142	0							
28	191	219	225	106	99	231	144	230	175	55	0								
29	200	175	142	188	177	74	127	86	0	75									
30	120	269	254	110	60	95	165	152	120										
31	218	132	190	75	41	134	127	76											
32	154	166	60	103	85	92	110												
33	0	0	0	67	0	0													
34	0	0	0	0															

Fig 11b The level L = 1 of native prealbumin structure amplitudes to 2.5 Å resolution.

diminishes to 1%.

It should be noted, however, that the curves in figure 10 hide the fact that centring persists to high resolution in a cylinder of reciprocal space containing the c^* axis. As a consequence, terms in this region are likely to remain relatively poorly phased.

The unique set of data for orthorhombic crystals is one octant of reciprocal space. After structure amplitudes were measured from the native protein crystals a programme was written to arrange this data as it appears at points of a 3-dimensional lattice. Figure 11 shows examples of this format for lattice planes to spacings of $2.5 \text{ \AA}$ where $l = 0$ and $l = 1$. Both of these representations show the effect of centring on the low order terms (weak terms for $h + k$ odd in 11a and for $h + k$ even in 11b). The other feature shown by these diagrams is the smoother distribution of intensity in the upper level in (11b), than in the centric HKO zone in 11a. This form of representing the data was found to be extremely useful in locating reference reflections for the checking and alignment of subsequent crystals for data collection.

Absolute scale and temperature factors.

In small molecule crystallography it is necessary to place the measurements of structure amplitude on an absolute scale as soon as possible. This is because any refinement of trial structures involves a direct comparison of F_{obs} with an F_{calc} term calculated from the known positions and scattering factors f_j of the constituent atoms j .

$$\text{Hence} \quad k \cdot F_{\text{obs}} = F_{\text{calc}} \quad (3)$$

where k is the absolute scale factor.

In protein crystallography, the F_{obs} term is obtained by the vector difference of the parent and derivative structure amplitudes F_P and F_{PH} . It is not necessary for these to be on an absolute scale, as the F_{calc} term from a trial heavy atom structure can be multiplied

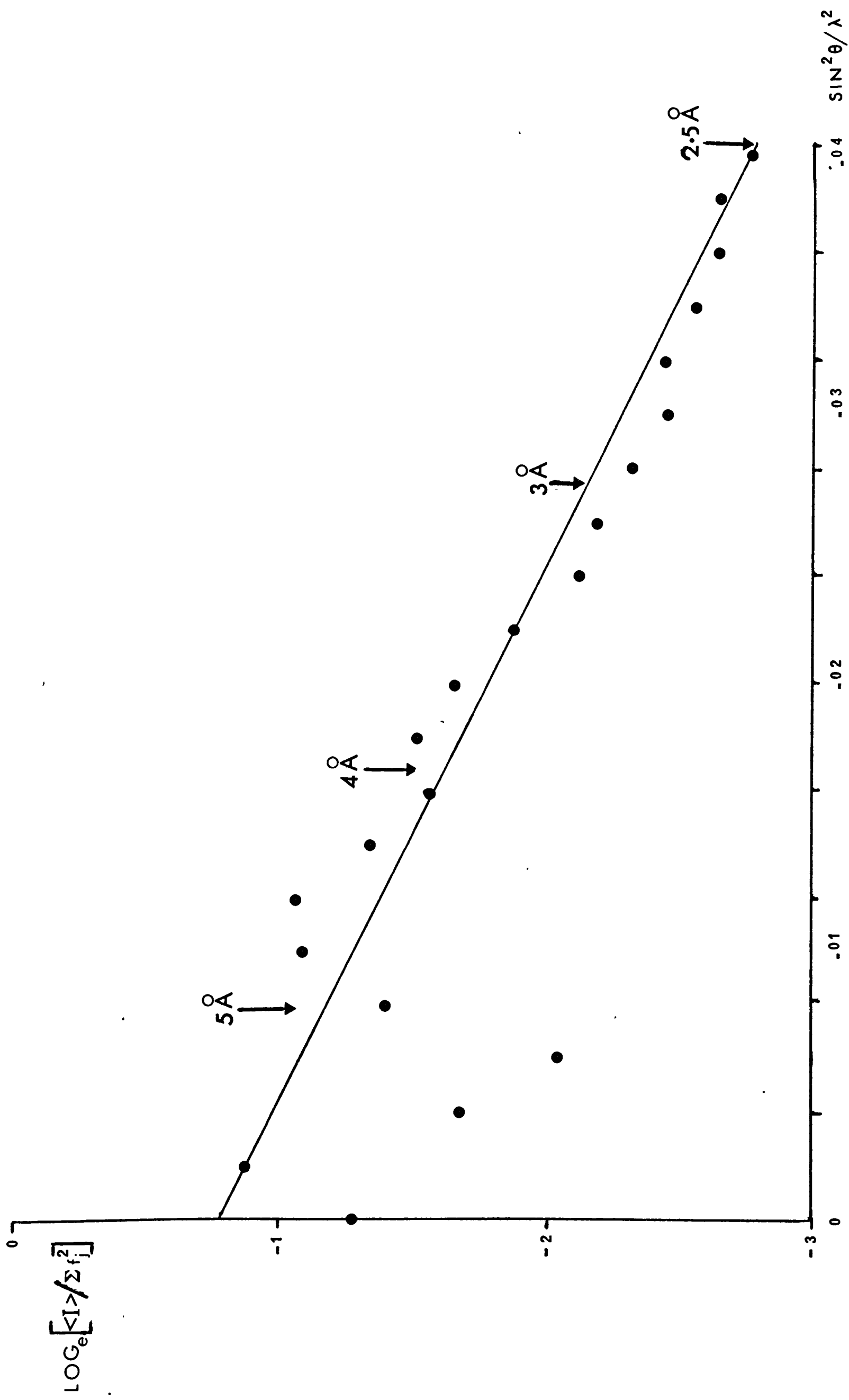


Fig. 12 Wilson Plot for native prealbumin data to 2.5 Å. resolution.

by an occupancy factor Z , for comparison purposes.

$$\underline{F}_{\text{obs}} = \underline{F}_P - \underline{F}_{\text{PH}} = Z \cdot \underline{F}_{\text{calc}} \quad (4)$$

If it is desired to know the true occupancy of the heavy atom sites in electrons and to be able to assign values to difference Fourier maps it is useful to have an idea of the value of k .

Wilson (1942) derived the result :

$$\sum f_j^2(\theta) = \langle I_{\text{abs}}(\theta) \rangle = k^2 \langle I_{\text{obs}} \rangle \exp(-2B \sin^2 \theta / \lambda^2) \quad (5)$$

where I_{abs} and I_{obs} are the absolute and observed intensities respectively and B is the temperature factor in $\text{\AA}^2$.

A plot of $\log_e \left[\langle I_{\text{obs}} \rangle / \sum f_j^2 \right]_{\theta}$ against $\sin^2 \theta / \lambda^2$

would then give $\log_e k$ as the intercept, and $2B$ as the slope. This derivation assumes (a) that terms to high (atomic) resolution are included in the plot and (b) that the atomic configuration in the cell is random. The distribution intensity given by equation (5) has been shown to be a poor approximation to that observed from protein crystals (Luzzatti, 1955 and Harker, 1953). This is worst at a resolution of $6 \text{ \AA}$ or lower where a high degree of order is obtained from secondary structure in the protein.

The Wilson plot for prealbumin is shown in fig.12. The terms $\sum f_j^2$ were calculated for prealbumin using a molecular weight of 108 000 in the cell, which gives an equivalent of 8000 nitrogen atoms, neglecting the effects of bound water and sulphate ions. The scattering curve for nitrogen atoms was calculated from the coefficients given by Vand et al. (1957). The least squares line of the total data ($\infty - 2.5 \text{ \AA}$ spacings) gave a value of k of 2.2 ± 0.3 and a temperature factor of $24 \pm 3 \text{ \AA}^2$. However, it is apparent that the plot is highly non-linear below $6 \text{ \AA}$ resolution. A least squares line fit to the data between $5 \text{ \AA}$ and $2.5 \text{ \AA}$ spacings gave an absolute scale factor of 1.8 ± 0.4 and a temperature factor of $26 \pm 5 \text{ \AA}^2$.

This temperature factor corresponds to an r.m.s. displacement of $0.55 \text{ \AA}^0$ of the protein molecule. A similar plot for the intensities as a function of resolution in each of the three axial directions suggested that the native protein temperature factor is anisotropic. The temperature factor for data in the c^* direction was significantly lower than that in the b^* direction. An approximate overall scale factor of 2 was assumed for subsequent work.

CHAPTER 3.

Heavy atom derivatives of prealbumin.Isomorphous derivative quality.

The preparation of isomorphous derivatives is a process of part intention, part luck. Useful ligands have been reviewed by Blake (1968 - 1972), Phillips (1966) and North and Phillips (1969). The most widely successful compounds, most likely because of their high specificity and small size, have been Pt^{IV} as a reagent for methionine, Hg^{II} as a cysteine specific and Sm^{III} and UO_2^{II} as carboxylic acid reagents. Recent work (Petsko, 1973) has clarified a confused literature on Pt^{IV} binding. Various criteria determine the usefulness of a heavy atom derivative, and most of these have been applied in the survey for isomorphous derivatives of prealbumin. These are set out below.

1) Mean fractional change in intensity expressed as

$$\text{MFCI} = \frac{\langle \Delta I \rangle}{\langle I_p \rangle}$$

where $\Delta I = |I_{\text{PH}} - I_p|$ taken over all the data set.

Crick and Magdoff (1956) have given an expression which allows an estimate of this quantity, given a known degree of replacement by a heavy metal. It had been implicitly assumed that since the standard error in measuring I is likely to be at least 10%, an MFCI of much greater than this would be required for useful data. However many successful structure determinations have resulted from quite light atom replacements, barium ions replacing calcium ions in one case, giving only 36 electrons (Arnone et al., 1971). This is probably because of the great redundancy of data over the small number of parameters to be determined. In the case of prealbumin there are nearly 9000 reflections to $2.5 \text{ \AA}$. Unfortunately, survey work was usually carried out either in projection (300 - 400 reflections) or at

6 Å (700 reflections). Because of this, a MFCI of 15%, giving easily visible changes on photographs, was looked for.

2) Crystallinity.

Some replacement leads to disordering of the crystal, perhaps by movements of the molecules as a whole within the unit cell. The effect upon the diffraction pattern is seen as a relative increase in intensity of those reflections from spacings lower than 6-8 Å over that of the reflections from high angles. (Figure 2D shows this effect present in a trial $\text{KAu}(\text{CN})_2$ derivative). Disordering may be obscured in a 9° precession picture (6 Å) so wherever possible 3 Å or 2.8 Å pictures were taken to test potential derivatives. The derivative need not necessarily be discarded, as an artificial temperature factor B in the form

$$I^{\text{corr}} = I_{\text{obs}} \exp \{ +2B \sin^2 \theta / \lambda^2 \}$$

may be applied to the data. Here I^{corr} is the corrected form of the initial intensity I_{obs} . Unless this is done, individual intensity changes ΔI between native and derivative compounds will tend to be biased to negative values, when clearly the mean value of ΔI should be zero. An artificial temperature factor has been applied to all the data in the routine scaling of the derivative to the native data. Only temperature factors leading to changes in I of greater than 20% at the highest values of $\sin^2 \theta / \lambda^2$ have caused the derivative to be treated with suspicion, or rejected.

3) Non-isomorphism.

The basic assumption of the method of isomorphous replacement is that additional electrons are added to each molecule without disturbance at positions remote from the site of replacement. In practice this assumption is always violated to some degree. Some of the most highly isomorphous structures reported, concern derivatives formed by replacement of an atom, or grouping on a natural substrate

or inhibitor of the protein molecule. In these cases the parent structure is the inhibited molecule, rather than the native protein (Sigler et al., 1968) and (Fridborg et al., 1972). In most other cases, however, some local movements of the protein to accommodate the artificial ligand are inevitable. These may be steric effects, for example, where a cysteine residue in a close-packed environment binds a mercurial. Alternatively the ligand might force a new conformation upon the molecule, causing changes remote from the site of replacement. This latter may have functional significance in the case of some proteins, including prealbumin, where several related ligands have been observed to cause local and remote changes in the protein structure.

The effects of non-isomorphism may not be readily apparent, even upon an examination of a high resolution photograph. There may be no apparent increased temperature factor due to loss of crystallinity. The intensity changes that are observed, may arise equally well from the substitution of a heavy atom at a specific site, as from additional shifts in protein structure, although the MFCI may be higher than expected in the latter instance. Some indication of lack of isomorphism may be available from a change in the cell dimensions and consequently these have been compared with those of the native structure, for all trial derivatives of prealbumin.

Several workers have proposed tests of non-isomorphism that can be applied at an early stage in the survey work. The best known is that of Crick and Magdoff (1956). They suggested that a plot of $\langle |\Delta I| \rangle$ should decrease with $\sin \theta$, in the case of an isomorphous replacement, as the scattering of any additional atom falls monotonically with scattering angle. If the quantity $\langle |\Delta I| \rangle$ does not fall off, or even increases, then this is an indication that other structural differences exist between the parent and the derivative. This test was applied to several of the potential derivatives for prealbumin, and

detailed results of these analyses are discussed in the next section.

If phase information is available, the most simple test of derivative quality is to use the phases together with the measured differences to give a difference Fourier. A good derivative should give peaks several times higher than the background. Negative peaks in such difference syntheses, provided that the derivative compound has been correctly scaled to the native, may indicate that a portion of the protein has moved in the derivative structure.

It can happen that negative features may represent the loss of an ion from the native protein during the formation of a derivative. This occurred in the case of 4-Zn insulin, where zinc ions diffused out in the preparation of a derivative (Mercola, 1972). It is theoretically possible to model this situation, providing useful phase information, but in practice difficulties arise with such negative occupancies and no case is known to the author where this information has been used successfully.

Derivatives of prealbumin.

Table 4 lists the heavy atom ligands which have been tried as potential heavy atom derivatives for prealbumin. Three species of parent compound have been used. PAQ was obtained by co-crystallising with mercury (Q from the German word 'Quecksilber' for mercury). PAN is a native crystal crystallised by Dr.J.Berthou in this laboratory. PAQDTT is a native crystal obtained by removing mercury from the large PAQ crystals with di-thiothreitol. Fig.20a shows a 2.8 Å⁰ precession picture of PAQ superimposed upon the same zone of the native protein. Individual intensity changes can clearly be distinguished at low and high angles. The native is the lower of the superimposed patterns.

The first nine derivatives were produced and investigated by Drs.Swan, Rérat and Berthou. All the remaining compounds have been produced and investigated by the author. The abbreviations given will

be used when referring to a particular derivative, individual crystals of a derivative being distinguished by a number following the abbreviation. Unless otherwise stated in the conditions shown, all the derivatives were prepared by soaking the crystals directly in 0.2 M $\text{Na}_2\text{HPO}_3\text{-NaH}_2\text{PO}_3$ buffer containing the heavy metal compound at the pH indicated, at room temperature. The first nine experiments were unbuffered. If binding occurred, the site of binding letter refers to the diagram of the molecule in fig.1. Other information shown is the type of data collected, the number of sites of replacement per asymmetric unit and the centric crystallographic reliability index R, discussed in the next section.

Several entries in the table need further clarification. The derivative IODOAC was unintentionally prepared. It had been noticed that prolonged soaking of prealbumin crystals in high concentrations of mercuric acetate produced a derivative with additional sites to those already seen in PAQ. An attempt was made to alkylate the sulphhydryl groups of the protein by soaking the crystals in iodoacetic acid at pH 7.5 prior to addition of the mercuric acetate. A derivative resulted which did not possess the characteristic diffraction pattern shown by substitution at the sites in PAQ. No information on this derivative could be obtained however, due to intensity changes which occurred during irradiation of the crystal. When the effect of iodoacetic acid alone was investigated, very strong binding occurred giving peaks near the molecular centre (sites B in fig.1). This was interesting because a wide variety of ligands such as mercury m-hydroxy-toluene sulphonic acid, gold cyanide and potassium mercuri-iodide gave nearly identical substitutions as deduced from both the appearance of their diffraction patterns and their projection Fourier syntheses (fig.51). Unfortunately the presence of negative peaks of equal height to the positive peaks,

as well as a preliminary refinement on gold cyanide (AUCY) suggested that these derivatives were not isomorphous with the native.

The strong binding of these ligands near the molecular centre suggested an experiment based on the ideas of working with an inhibited enzyme as the parent structure (Sigler, 1968). If the light atom congener of mercury m-hydroxy-toluene sulphonic acid (o-xylene sulphonic acid) bound, possibly these two structures would be highly isomorphous. The experiment was actually tried with p-toluene sulphonic acid, as p-chloromercuri-benzene sulphonic acid has been bound to prealbumin. No visible changes occurred in the diffraction pattern and the binding of o-xylene sulphonic acid has not been attempted.

Experimental.

All the photographic work on prealbumin crystals has been carried out on Supper precession cameras attached to a rotating anode X-ray source. Centric zones such as those in fig.7 have been used exclusively. The crystals used were smaller than those employed for diffractometry. The largest size used had a b axis of 0.5 mm. The 18° pictures in fig.7 were produced from 2 hour exposures at 1.6 KW using a 0.25 mm collimator. An incident beam collimator of 0.45 mm internal diameter was used for larger crystals. A film to crystal distance of 75 mm was found satisfactory for small crystals. Generally 3 exposures of 1/4, 3/4 and 2 hours were taken for the OKL zone and 2 exposures of 1/4 and 2 hours for the HKO zone. This allows a reasonable estimation of intensities to the $2.8 \overset{\circ}{\text{\AA}}$ resolution limit to be made over the linear range of the emulsions used. Ilford Industrial G X-ray film was used for intensity measurement and Kodak KODIREX was used for setting photographs. Films were developed for 5 minutes in Kodak 'Phenisol' under nitrogen agitation.

Intensity measurement.

Two methods of intensity measurement of photographic data were used. In most of the derivative survey work, use was made of a Joyce-Löbel["] scanning microdensitometer. This instrument scanned part or all of a reciprocal lattice row and gave a paper trace of peak heights. The machine was manually reset between reciprocal lattice rows. A rectangular aperture was used, such that all of the spot was focussed onto the photomultiplier in a scan across the spot. This instrument has one principal disadvantage. It can only measure peak height, so that unless the spot shape is constant across the film, there may be error in relative measurement of intensities. Some derivatives and ligands caused splitting of the crystals, such that individual portions of the spots became separated to an increasing extent with the Bragg angle θ . This effect prevented easy measurement on the Joyce-Löbel["] machine. Some of the derivative photographs and several prealbumin-light atom ligand pictures were measured using an Optronics International P 1000 drum scanner. This instrument measures optical densities of points on the film separated by spacings of 50, 100 or 200 μ . These are written sequentially on a magnetic tape as 8-bit numbers, 0 to 255, so that the full scale reading is approximately OD 2. A set of programmes, written by Mr. P. R. Evans of this laboratory, process the scanner output to intensities. This machine probably produces significantly better intensities than the Joyce-Löbel["]. There may be 25 readings of OD within one spot, so that an integrated intensity is produced. If the spot is split, the integration area can be extended to cover both parts of the spot.

Location of the heavy atom in the cell.

For the solution of any crystal structure a number of the constituent atoms must be unambiguously located with respect to the

chosen unit cell origin. These are generally those atoms of highest mass number Z and which consequently dominate the X-ray scattering by the molecule. The standard method of doing this is due to Patterson (1935). Patterson replaced the coefficients F_{hkl} (a complex number) in equation (2.1) by the squares of the structure factors (phaseless intensities). This equation then simplifies to

$$P(X,Y,Z) = \frac{1}{V} \cdot \sum_h \sum_k \sum_l I(hkl) \cos 2\pi(hX + kY + lZ) \quad (1)$$

where $P(X,Y,Z)$ is an intensity in $(\text{electrons})^2 \text{ \AA}^{-3}$ in Patterson space at the position (X,Y,Z) . $I(hkl)$ is a measured intensity for the reflection (h,k,l) . It should be noted that the geometric form of this expression leads to a centre of symmetry at the unit cell origin. The peaks in the Patterson synthesis represent vectors between pairs of atoms in the unit cell, of a height proportional to the mass numbers of the atoms. Because of the additional symmetry of an inversion centre, in protein crystallography the Patterson function always leads to a smaller asymmetric unit than that of the true space group. This smaller unique space then contains N^2 peaks from N atoms. Unless the resolution of the data is high, even molecules with a small number of atoms will have Patterson functions complicated by overlapping peaks. If a heavy atom is present in the structure, however, vectors between the heavy atoms will dominate this function, with lesser contributions from vectors between heavy and light atoms. The solution of such a Patterson is the start of a small-molecule structure determination. With proteins, the low resolution of the data, combined with the enormous number of the atoms present, causes the Patterson functions to be featureless, whether or not a heavy atom is present. In the $6 \text{ \AA}$ prealbumin self-Patterson, the single peak at $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ is due to the presence of molecular symmetry.

The most universally successful method of locating a heavy atom

in the protein unit cell is the 'difference Patterson synthesis'

using coefficients $F^2 = (|F_{PH}| - |F_P|)^2$ in equation (1)

(Blow, 1958). The structure factors F_P and F_{PH} are related by expression (2), where k is a scaling constant :-

$$F_{PH} = k \cdot F_P + F_H \quad (2)$$

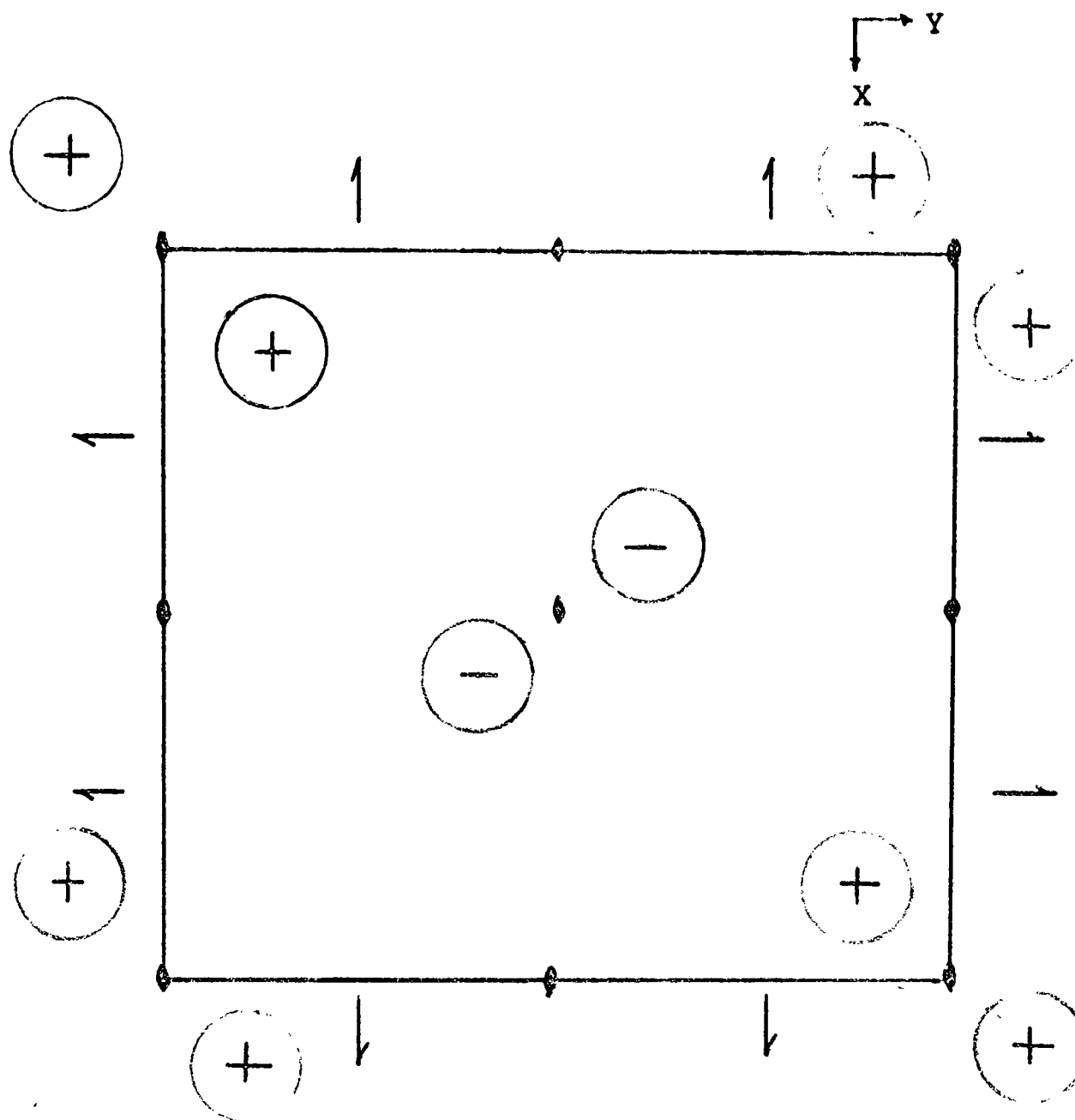
This expression is shown vectorially in fig.5b.

The space group $P2_12_12_1$ of prealbumin has 3 centric projections (OKL,HOL,HKO), where the phases of the structure factors F_H differ by 0 or π from those of the protein. In cases like this equation (2) may be written as ;

$$|F_H| = |F_{PH}| \pm |F_P| \quad (3)$$

A centric zone difference Patterson synthesis should represent information from the heavy atoms alone, provided that the sign of the phase of F_P and F_{PH} are the same. It occasionally happens that F_{PH} has opposite sign to F_P if $|F_P|$ is small. This accounts for the presence of the plus sign in equation (3). Such a term is called a crossover. It has been possible in prealbumin projection Pattersons to avoid crossovers by including phase information in the Patterson to the extent of choosing the correct sign in equation (3).

In the case of acentric reflections the equality (3) no longer holds and the estimate of $|F_H|$ that is obtained from it is generally too low. The features to be expected in an acentric Patterson synthesis have been fully discussed by Phillips (1966) who showed that the vectors between heavy atoms will appear with only half their predicted weight. However, the general experience is that the 3D Patterson synthesis is superior to that in projection, because of the great increase in the number of terms included. More recently Dodson and Vijayan (1971) have shown both from an expansion of equation (2) and by using model derivatives, that the difference



Origin at 112 in the plane of $2_1 2_1$

$$A = 4 \cos 2\pi \left(hx + \frac{h+k}{4} \right) \cdot \cos 2\pi \left(ky - \frac{h+k}{4} \right) \cdot \cos 2\pi lz = 0 \text{ if } l=0$$

$$B = -4 \sin 2\pi \left(hx + \frac{h+k}{4} \right) \cdot \sin 2\pi \left(ky - \frac{h+k}{4} \right) \cdot \sin 2\pi lz$$

$$\underline{h + k = 2n}$$

$$A = 4 \cos 2\pi hx \cdot \cos 2\pi ky \cdot \cos 2\pi lz$$

$$B = -4 \sin 2\pi hx \cdot \sin 2\pi ky \cdot \sin 2\pi lz = 0 \text{ if } h=0, \text{ or } k=0$$

$$\underline{h + k = 2n+1}$$

$$A = -4 \sin 2\pi hx \cdot \sin 2\pi ky \cdot \cos 2\pi lz = 0 \text{ if } h=0, \text{ or } k=0$$

$$B = 4 \cos 2\pi hx \cdot \cos 2\pi ky \cdot \sin 2\pi lz$$

Fig. 13 Space group equations for $P2_1 2_1 2_1$.

Patterson synthesis is truly representative even with a large number of sites. This result has been used to justify interpretations of the complex Pattersons given by uranyl derivatives of prealbumin.

The asymmetric unit of $P2_12_12$ is $1/4$ of the unit cell, whilst that of the Patterson space group is $1/8$ of the cell. Any atom in the asymmetric unit is related to three others by the operations of the screw diads and the rotation axis. These operations are summarised by the equivalent positions shown below, for an atom located at a general position (x,y,z) within the cell.

1. x, y, z
2. $\bar{x}, \bar{y}, z$
3. $\frac{1}{2} - x, \frac{1}{2} + y, \bar{z}$
4. $\frac{1}{2} + x, \frac{1}{2} - y, \bar{z}$

This symmetry allows the simplification of equation (1.1) so that structure factors relating to the scattering from one unit cell may be calculated from the atoms in one asymmetric unit. International tables for X-ray crystallography vol.1 lists the forms of equation (2.1) for each space group, and those for $P2_12_12$ are shown in fig.13. It should be noted that certain positions lead to degeneracy of the four vectors given above. For example, the positions $(0,0,z)$ and $(\frac{1}{2}, \frac{1}{2}, z)$. Atomic positions may be close to, or at a rational fraction of one or more of the unit cell translations $(1/2, 1/3, 1/4)$. Examination of the equations in fig.13 will show that in these cases whole classes of reflections exist for which the geometric form of the structure factor is zero. These points correspond to the nodes in the structure factor graphs for these orders of reflection. This observation is significant in the case of protein crystallography where the introduced heavy metal ion often chooses such a position of high symmetry. Frequently, these positions lead to difficulty with initial location and refinement because of their

reduced contribution to the diffraction pattern. These positions are often termed 'special'.

The presence of a single atom at a general position x, y, z in the asymmetric unit of $P2_12_12$ gives rise to the following vectors in the Patterson function of the data.

$$1. \quad 0, \quad 0, \quad 0 \quad (4)$$

$$2. \quad 2x, \quad 2y, \quad 0 \quad (2)$$

$$3. \quad \frac{1}{2}, \quad \frac{1}{2} + 2y, \quad 2z \quad (2)$$

$$4. \quad \frac{1}{2} + 2x, \quad \frac{1}{2}, \quad 3z \quad (2)$$

The numbers in parentheses indicate the total number of degenerate vectors of each type. Such vectors are known as self-vectors since they arise solely from vectors between crystallographically identical atoms. The symmetry of the space group is expressed in these vectors and was used in the determination of that of prealbumin, (chapter 2). Clearly vector 2. has arisen from two atoms with identical z co-ordinates and which were hence related by a two-fold rotation. (This eliminates the possibility of the space group being $P2_12_12_1$). These vectors always appear in sections of Patterson space with one of X, Y , or Z equal to 0 or $\frac{1}{2}$. The so-called Harker sections of $P2_12_12$ ($X, Y, 0$), ($\frac{1}{2}, Y, Z$) and ($X, \frac{1}{2}, Z$) are shown for three prealbumin derivatives at $2.5 \overset{0}{\text{\AA}}$ resolution in fig.35.

If there is more than one atom in the asymmetric unit, a set of self-vectors will appear for each and, in addition, a set of cross-vectors, relating atoms which are not crystallographically identical.

There are no constraints upon the positions of these vectors, unless they arise from atoms in special positions. Some can occur co-incidentally in the Harker sections, and may be confused with self-vectors. As the number of positions rises, it rapidly becomes difficult or impossible to solve the Patterson function,

particularly as vectors tend to overlap each other at the resolutions employed.

Patterson ambiguities.

Proteins are enantiomorphic molecules, composed of L-amino acids and asymmetric secondary structures. No protein cell has a centre of symmetry, so that the position x, y, z is not equivalent to its inverse $\bar{x}, \bar{y}, \bar{z}$. This information is lost due to the imposed symmetry of the Patterson function and must be determined at a later stage. In the solution of the Patterson function an arbitrary choice is made, corresponding to an initially unknown hand. A problem is then presented in relating the position of heavy atoms in different derivatives to a common origin. Phillips (1966) reviewed methods for overcoming this difficulty. In prealbumin the relation of sites in a mercury derivative (PAQ) and a uranyl derivative (PAUN) were solved by the preparation of a double derivative (PAQUN) where cross-vectors between the two different sets of sites gave their relative co-ordinates. If a phase set exists, then the application of these phases to difference amplitudes from an unknown derivative is sufficient to fix the relative positions of the sites in that derivative, with respect to a common origin. This cross-phase procedure was also used on prealbumin to check the validity of common origins.

Solution of prealbumin Patterson functions.

The first prealbumin Patterson functions to be examined at high resolution were difference Pattersons of the OKL and HKO zones containing 400 and 230 terms to $3 \overset{0}{\text{\AA}}$ respectively. These had been solved in general terms (Swan 1970), but several problems remained as an obstacle to the complete description of the derivatives.

Two projection difference Pattersons of the mercury derivative PAQ are shown in fig.14 c,d. The mercury atoms in prealbumin form a nearly body-centred array in the cell. This weakens the odd-ordered

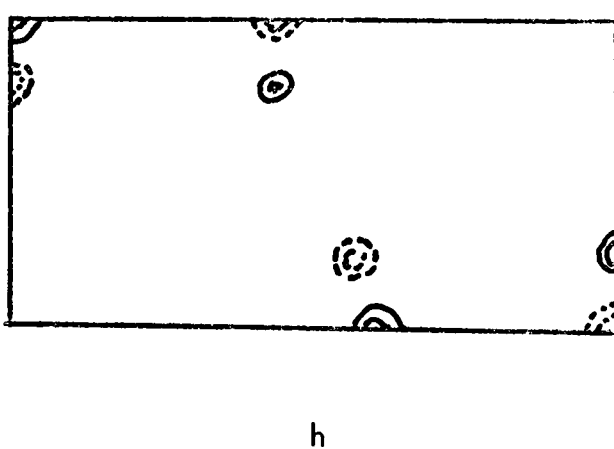
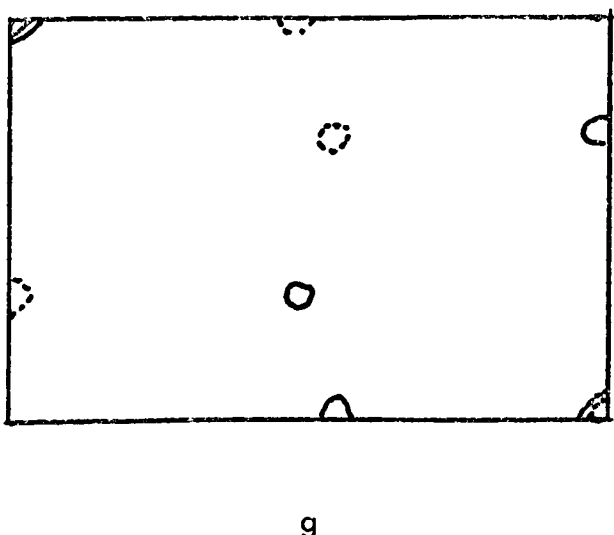
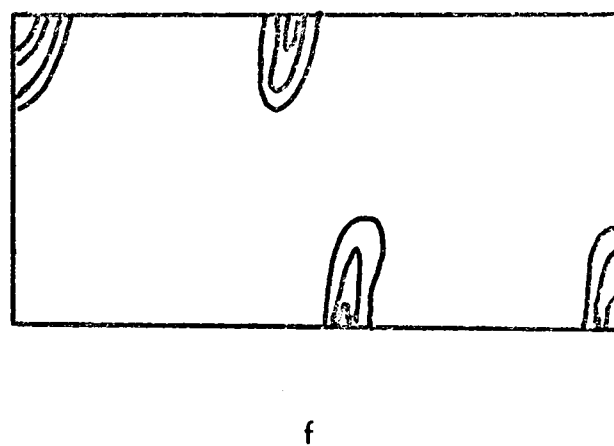
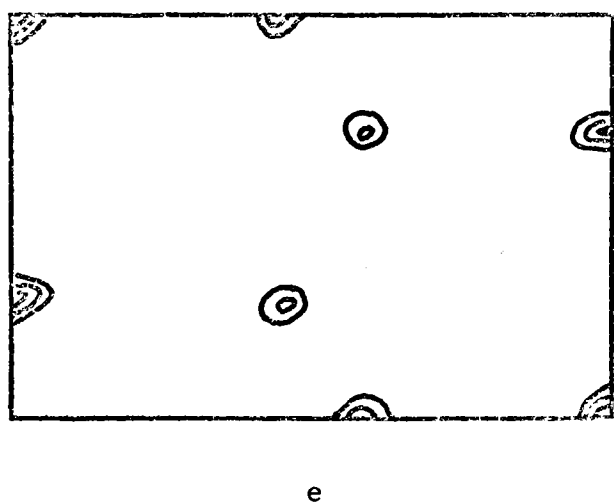
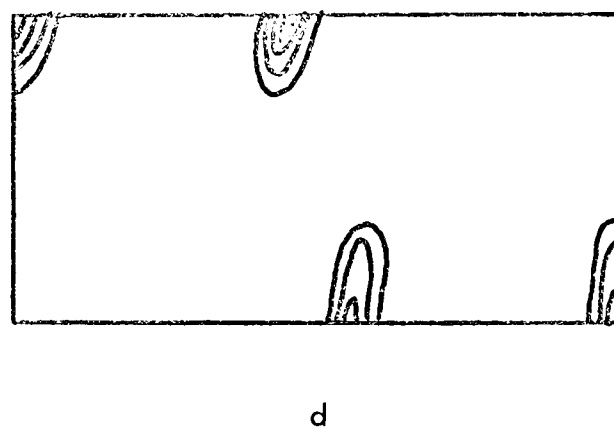
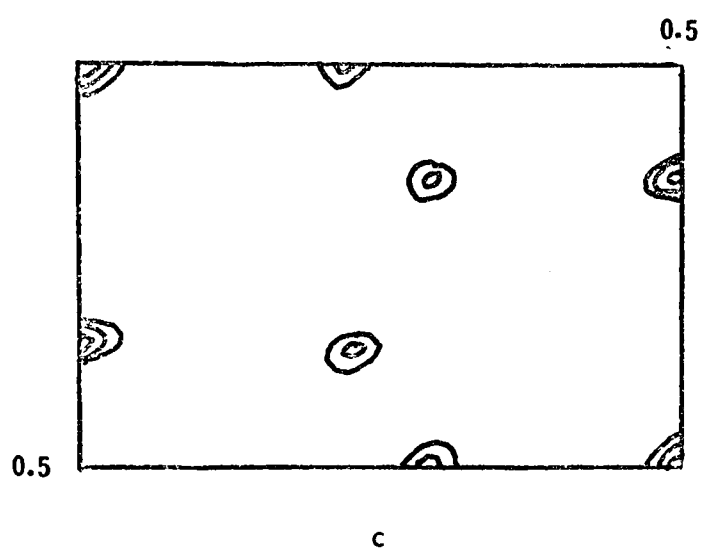
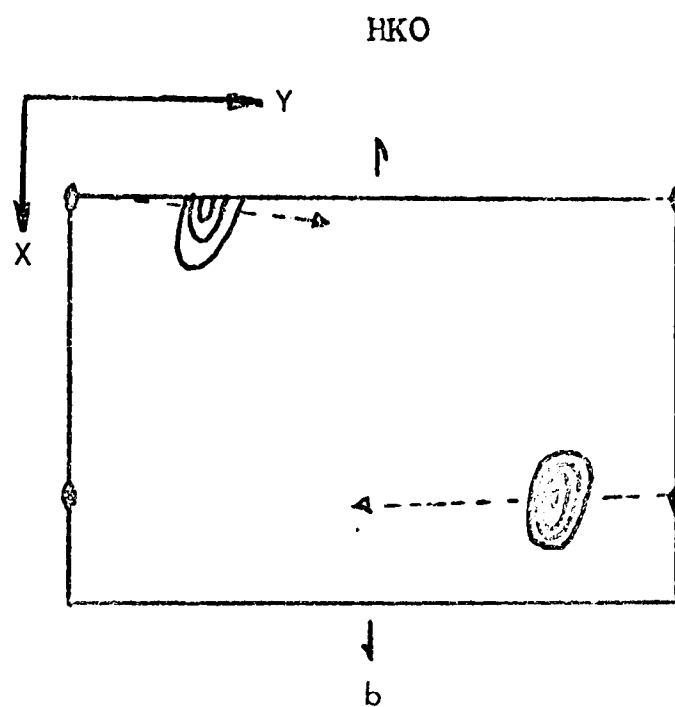
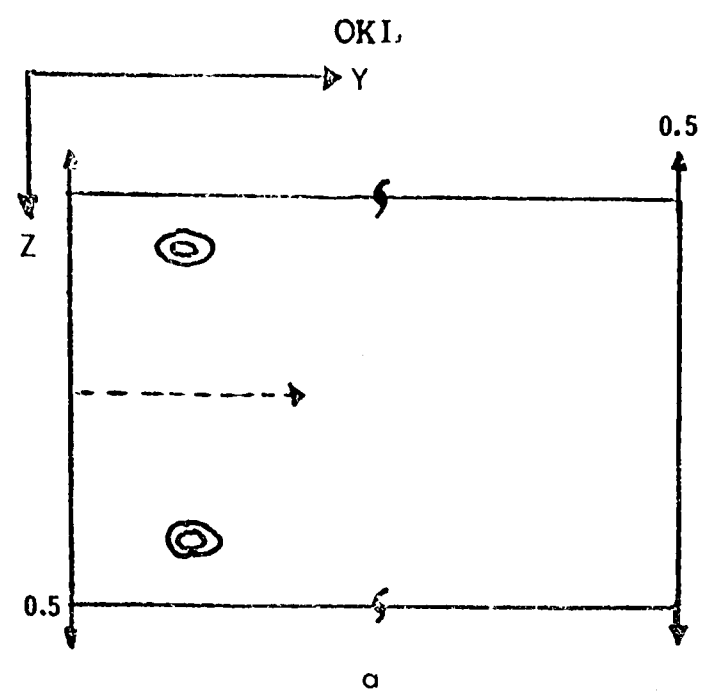


Fig. 14 Patterson functions of PAQ. a. OKL Fourier b. HKO Fourier c. OKL Patterson d. HKO Patterson e. even parity Patterson f. even parity Patterson g. odd parity Patterson h. odd parity Patterson.

terms. Since the Patterson function uses the squares of the difference terms ΔF as coefficients, it is heavily biased towards the function of a perfectly body-centred cell.

The inherent asymmetry of the sites, caused by the rotation of local axes away from the cell axes is masked, particularly in these centric projections. This effect is worsened in the OKL zone by the persistence of cell-centring in the c^* direction. The Patterson in fig.14c was solved in terms of two sites $(x, 0.110, 0.078)$ and $(x, 0.110, 0.425)$ shown in the difference Fourier projection 14a with an indication (dashed line) of the local two-fold axis.

Unfortunately the HKO Patterson 14d shows that the x co-ordinates of the sites are very similar and the sites are not resolved. The HKO difference Fourier 14b reflects this situation. It was not clear on which side of the origin, the site nearest to zero in x should be placed. The correct positioning of these sites is clearly important in describing the derivative, as the degree to which these sites differ in their x and y co-ordinates, determines the correct phases. One solution to the problem of centring lay in the collection of 3D data to a resolution of $5 \overset{0}{\text{\AA}}$ or greater, but this was not an acceptable method for derivative survey work.

Odd and even Pattersons.

Because of the dominating effect of the even ordered terms a novel form of the Patterson synthesis was attempted. The data was split into two parity groups, one for which $h + k + l = 2n + 1$ and the other for which $h + k + l = 2n$.

The Patterson function at the point (x, y, z) is obtained by summing over $\frac{1}{2}$ of reciprocal space :

$$F^2 \cos 2\pi (hx + ky + lz) \quad (4)$$

With the simplification of space group symmetry, using the equations in fig.13, expansions of equation (4) give

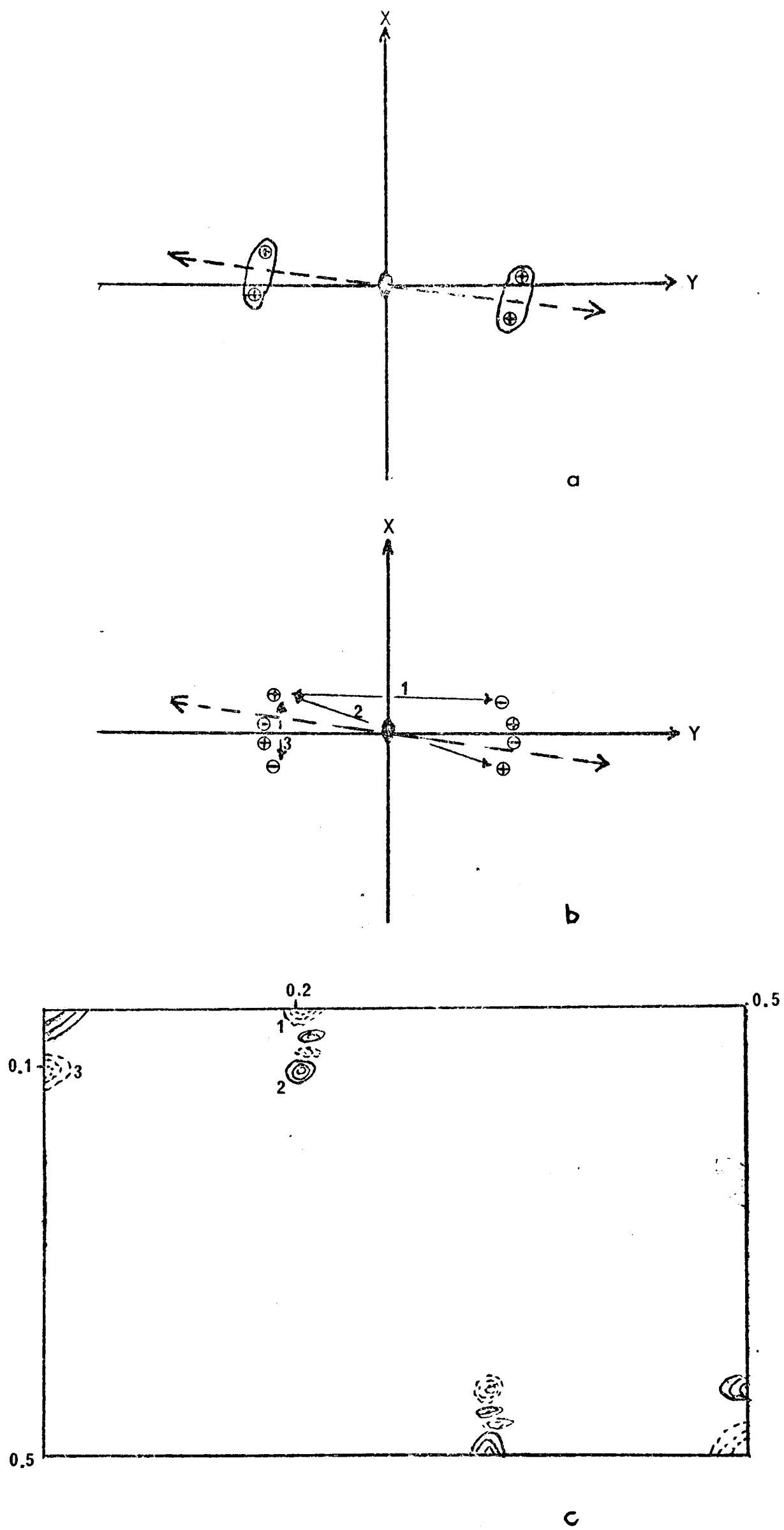


Fig. 15. Features in HKO odd parity synthesis.

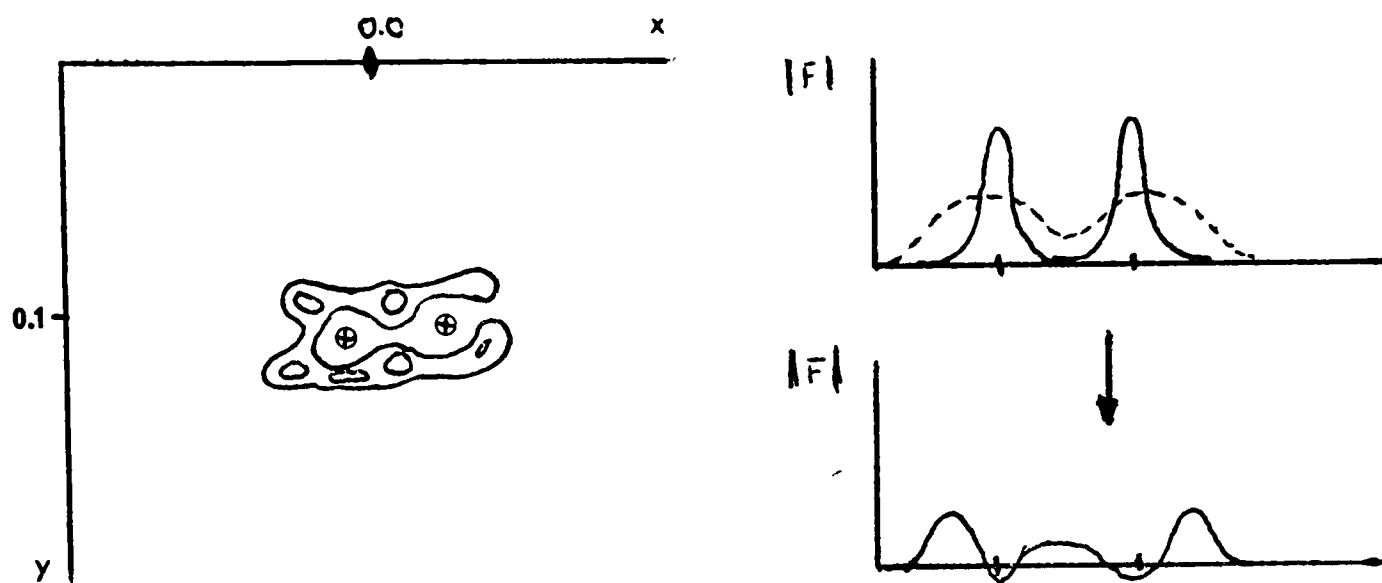
$$4. \Delta F^2 \cos 2\pi hx \cdot \cos 2\pi ky \cdot \cos 2\pi lz \quad \text{for } h+k = 2n \quad (5)$$

$$\text{and } -4. \Delta F^2 \sin 2\pi hx \cdot \sin 2\pi ky \cdot \cos 2\pi lz \quad \text{for } h+k = 2n+1 \quad (6)$$

Consideration of these equations shows that for planes of even parity $h + k + l = 2n$, the Patterson synthesis gives a perfectly centred solution, as the cosine function is symmetric with respect to the origin. For planes of odd parity $h + k + l = 2n + 1$ the Patterson function is anti-symmetric with respect to reflection through the origin. The Patterson synthesis (6) would also be expected to have negative peaks with respect to an inversion centre at $1/4, 1/4, 1/4$.

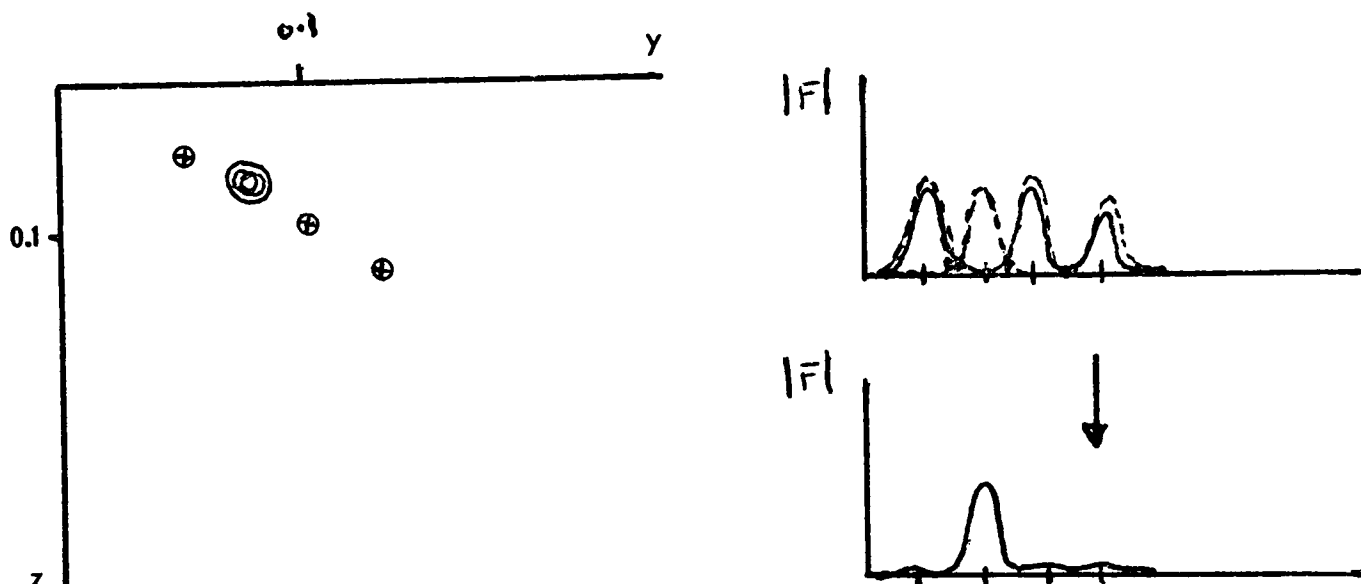
Figure 14e, f, g, h show the appearance of these odd and even-termed Patterson syntheses. The resemblance of the even functions (e,f) to those of the functions using all the terms (c, d) is obvious. The odd parity OKL Patterson is virtually featureless, as would be expected from the small size of the terms in ΔF^2 , however that of the HKO projection yields some useful information.

An explanation of the features of the odd-parity Patterson is best given by considering the Fourier synthesis corresponding to the odd terms only. Referring to fig.15a, a guess at the positions of the two mercury sites related by the local axis is shown as . A Fourier synthesis containing only terms for which $(h+k+l) = 2n+1$ might have an appearance as shown in 15b. There are positive peaks at the real sites, but also negative ones of equal height at positions related to the real peaks by reflection in the x and y axes. (The complete data would contain cosine terms to cancel these peaks). Thus 15c shows the Patterson that would be obtained by super-position of this vector set. This model bears a close resemblance to that actually observed (fig.14h). The best fit to the observed pattern of positive and negative peaks was obtained when site 1 has an x co-ordinate of less than zero and close to 0.98, and site 2 has an x co-ordinate of 0.055. These values were found to be very close to the



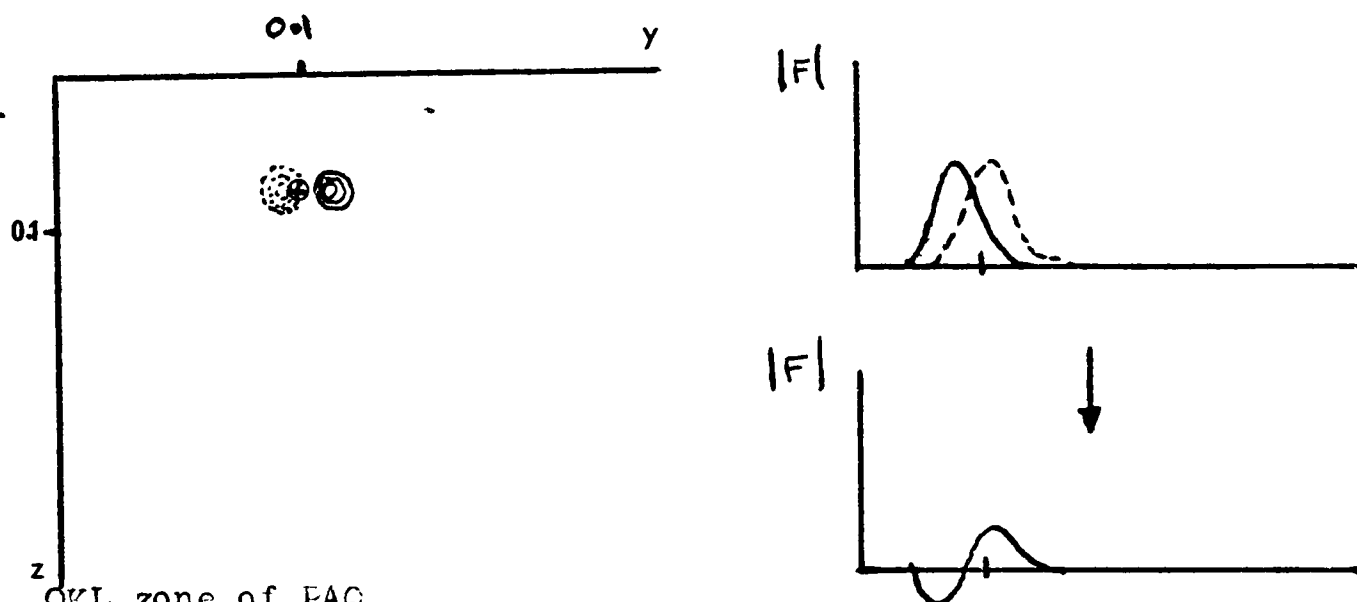
HKO zone of PAQ.

Effect of too small a temperature factor.



OKL zone of PTI.

The synthesis indicated that a fourth site should be added to the existing heavy atom solution.



OKL zone of PAQ.

The y-co-ordinate of site HGl is underestimated (should be 0.107).

Fig. 16 Three examples of the double difference synthesis observed during the centric data refinement at $3\text{\AA}$ resolution. The circles containing crosses represent the current heavy atom positions. On the right of each Fourier the top graph shows the one-dimensional synthesis of F_{Hobs} (broken line) and F_{Hcalc} (solid line). The lower graph shows the result of the subtraction ($F_{Hobs} - F_{Hcalc}$) which should be compared with the actual synthesis.

values obtained later from refinement of the 3D 2.5 Å data for this derivative (0.979 and 0.054).

Difference syntheses.

The use of difference Fourier syntheses in the description of heavy atom features within the unit cell is discussed by Phillips (1966). If coefficients ΔF are used in equation (2.2) where:

$$\Delta F = (k |F_{PH}| - |F_P|) \cdot e^{i\alpha_P} \quad (7)$$

peaks appear at the heavy atom positions with half their true height (Luzzati, 1953). In the case of centric reflections, F_{PH} and F_P are co-linear and provided the signs of F_P and F_{PH} are identical, the peaks in such a projection synthesis have their true height.

When a model is available for the heavy atom replacement, an estimate F_{calc} of the true contribution F_H may be made. The synthesis using coefficients $\Delta\Delta F$ in equation (2.2) where

$$\Delta\Delta F = (k |F_{PH}| - |(F_P + F_H)|) \cdot e^{i\alpha_{PHcalc}} \quad (8)$$

should reveal deficiencies in the current model, as when

$$|F_{PH}| = |F_P + F_H|$$

the synthesis should be zero at every point. Bloomer (1972) has analysed in detail the origin of features in such a synthesis, and has shown that it is most sensitive towards the derivative to native scale factor, and towards the occupancy of the heavy atom sites. Because of the low resolution of protein data, distortion and spurious features often appear in difference syntheses, due to ripple from Fourier series termination. The double difference synthesis is free from this effect and allows a closer analysis of features. Some of these features and their causes are shown in fig.16.

Initial refinement of heavy atom parameters.

The solution of a Patterson synthesis usually gives quite accurate positional parameters for a heavy atom. The site occupancy however, is harder to estimate and usually a refinement procedure

which minimizes the function

$$E = \sum_h (|F_{obs}^h| - |F_{calc}^h|)^2 \quad (9)$$

is employed. Where the data comes from a centro-symmetric zone, the function to be minimized is :

$$E = \sum_j (kF_{PH} + F_P - F_{calc})^2 \quad (10)$$

The calculated heavy atom contribution is expressed by

$$F_{calc} = \sum_j^N f_j \cdot Z_j \exp(-2B_j \sin^2 \theta / \lambda^2) \exp(2\pi i(hx_j + ky_j + lz_j)) \quad (11)$$

where f_j is the atomic scattering factor for atom j ; Z_j is the occupancy of the site j ; B_j is the temperature factor of the atom j ; N is the total number of atoms in the unit cell. The minimization can be carried out directly by a least squares procedure where shifts in the parameters k , x_j , y_j , z_j , B_j and Z_j are calculated from normal equations of the partial derivatives of (11) (Lundberg, 1965). For all the work on prealbumin the method of Hart (1961) was used.

From the initial values of x_j , y_j , z_j , Z_j , B_j and k , shifts were made by -2, -1, 1 and 2 times a specified unit shift. Error sums as in (10) were accumulated for each of these twenty combinations. The sign ambiguity in (10) was solved at each stage by choosing that combination giving the smallest value of E . At the end of a complete cycle of refinement, shifts were applied to the parameters in the direction most influential in reducing the size of the error sum E .

The advantages of such a search technique for refinement are that quite large errors in the initial values of the heavy atom parameters can still lead to convergence. Also the method is less likely than the least squares method to converge to local, rather than true minima, because of the artificial shifts that are supplied. The progress of the refinement is monitored by the centric reliability index R where

$$R = \frac{<|F_{calc} - k \cdot F_{PH} - F_P|>}{<|k \cdot F_{PH} - F_P|>} \quad (12)$$

Typical reported values of R for useful derivatives, range from 40 to 58% for data to 2.5 Å⁰ resolution.

Refinement of prealbumin derivatives in projection.

For refinement using Hart's method, the centre of symmetry must be at the origin of the cell. Then the three centric zones can be referred to the same origin for refinement and the expression for F_{calc} may be written :

$$F_{\text{calc}} = \sum_j^N f_j \cdot Z_j \exp(-2B_j \sin^2 \theta / \lambda^2) \cos 2\pi(hx_j + ky_j + lz_j) \quad (13)$$

Referring to fig.13 it will be seen that the origin of $P2_12_12$ is on the two-fold, but in the plane of the screw axes, $1/4$ of a cell translation in z. To relate the OKL and HOL projections to an origin in one corner of the cell, a factor of 0.25 was subtracted from x and y by the programme when F_{calc} was computed for these zones. This means that the phases for these reflections differ by $\pi/2$ from the phases related to the usual origin of $P2_12_12$. The phases of the axial reflections H00 and OKO were unchanged in these zones, because $L = 0$ and these were considered to belong to the HKO zone. At the completion of refinement a set of phases chosen to minimize expression (10), were output, with the current estimates of F_{calc} . Both self-sign and cross-sign maps were calculated with such phases. However, because the phases of reflections in the HOL and OKL zones had been retarded by $\pi/2$, the Fourier were carried out in $P222_1$ (No. 17) which has the symmetry of the shifted cell. In addition the signs of H00 and OKO reflections were changed if H or K (modulo (4)) was non-zero, to give a consistent sign set for the new space group.

Sets of 3 Å⁰ data from the centric OKL and HKO projections of prealbumin for five derivatives were already available (Swan, 1970). Details of the conditions under which these derivatives (PAQ, PTF, PTI, GUN, HISC) were obtained are given in table 4. Some refinement of

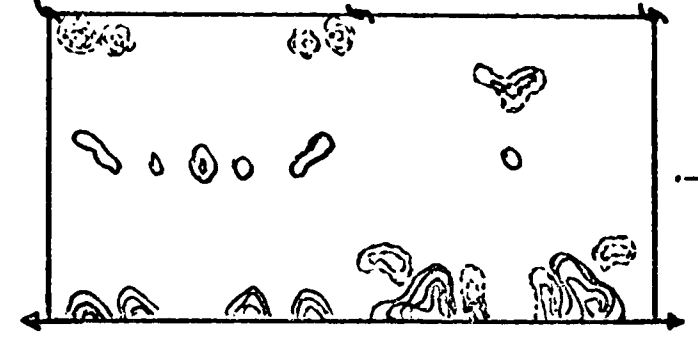
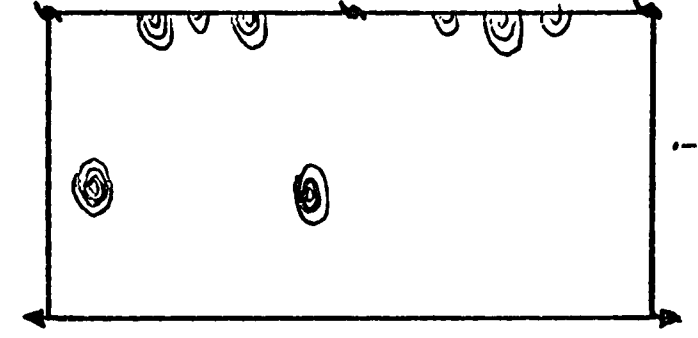
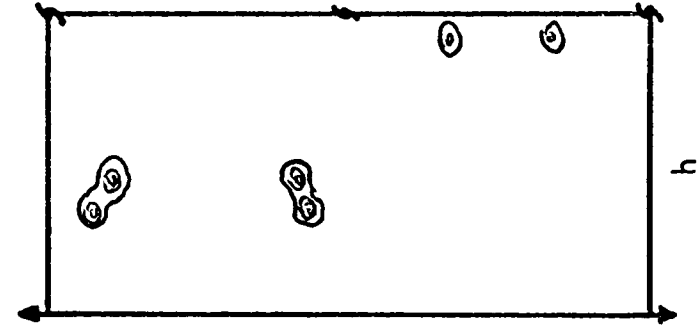
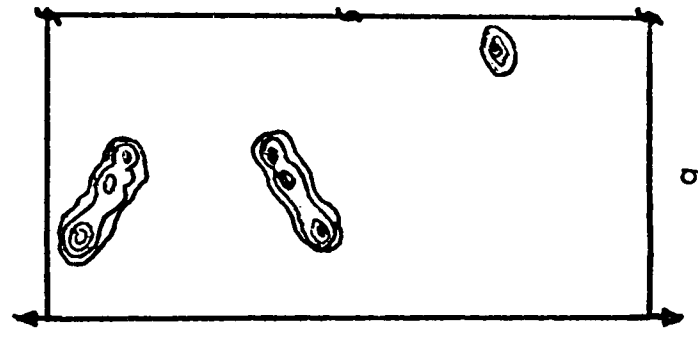
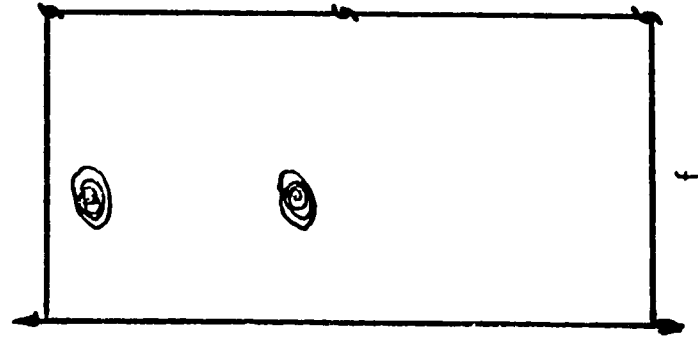
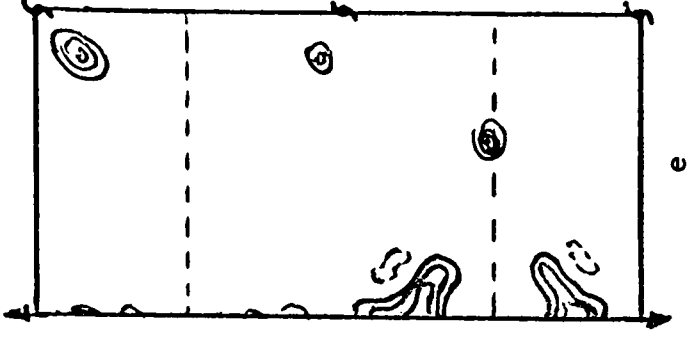
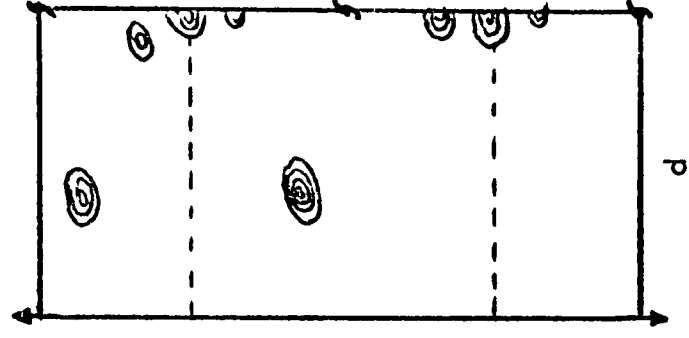
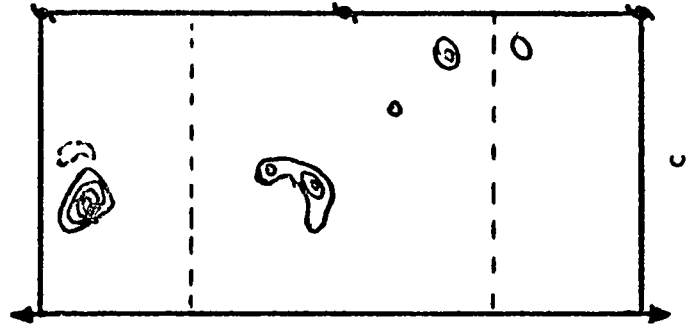
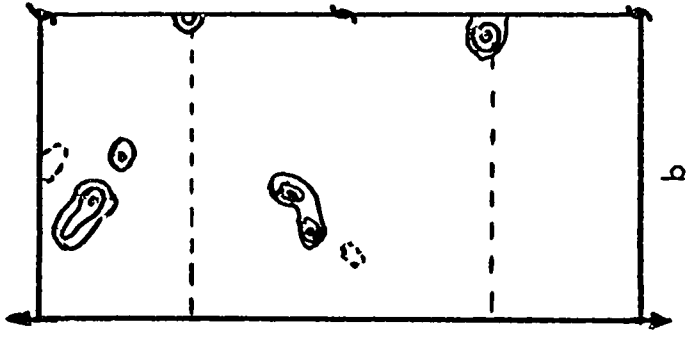
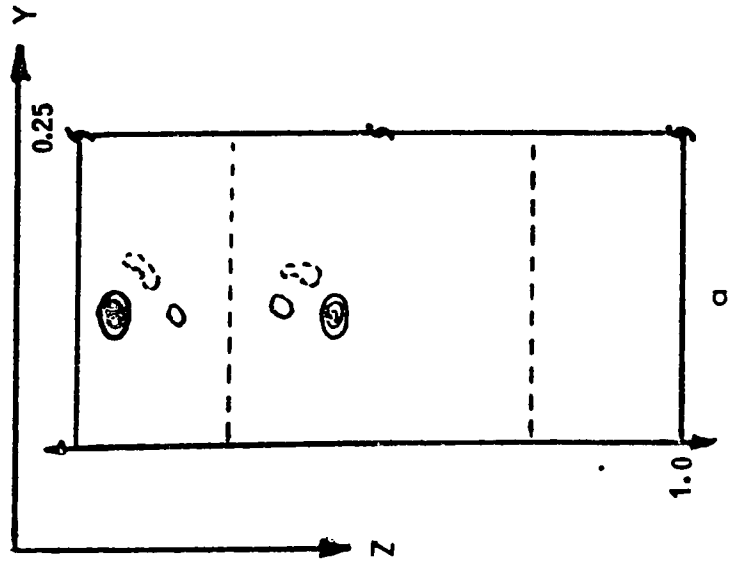


Fig. 17 OKL projections of PAQ,PTF,PTI,GUN,HISC. One asymmetric unit is shown.
a-e after refinement of all data. f-j after refinement of even terms
and exclusion of the HISC derivative from the phasing.

the sites in the OKL projection had been tried. What was needed was an estimate of the suitability of these derivatives for 3D high resolution work. Swan had encountered difficulties in these refinements because of the low values of the odd-ordered terms in the OKL projection. In the HKO projection, where the body-centring of the diffraction data was less pronounced, it appeared that the x-co-ordinates of all the sites were closely similar. The resulting Fourier overlap made interpretation difficult in this zone. The OKL difference Fourier projections at this stage are shown in fig. 17a,e.

Blow and Crick (1959) discuss in detail a method of using the observed isomorphous differences from centro-symmetric zones to obtain phase information. Appendix 2 shows how this method may be applied in a general phasing programme, where more than one derivative contributes this information. The result of this is to provide a 'best' phase angle and an associated weight 'm'. When such phases are used with the coefficients $m \cdot \Delta F$, a map is obtained with the least mean square error in electron density over the unit cell. The phases produced by the Hart programme only arise from the derivative which is being refined. When these phases are used with ΔF 's of that derivative or with an independent derivative, the corresponding syntheses are termed self-sign or cross-sign maps respectively.

The flow chart in fig.18a shows the process that was used to arrive at the difference maps in fig. 17 a,e. These maps have been produced with the observed ΔF 's appropriate to the particular derivative, together with jointly determined phases with a mean value of $m = 0.83$. To arrive at this stage, the projection Patterson function of each derivative has been solved, and independent refinement of the four site parameters for each observed site has

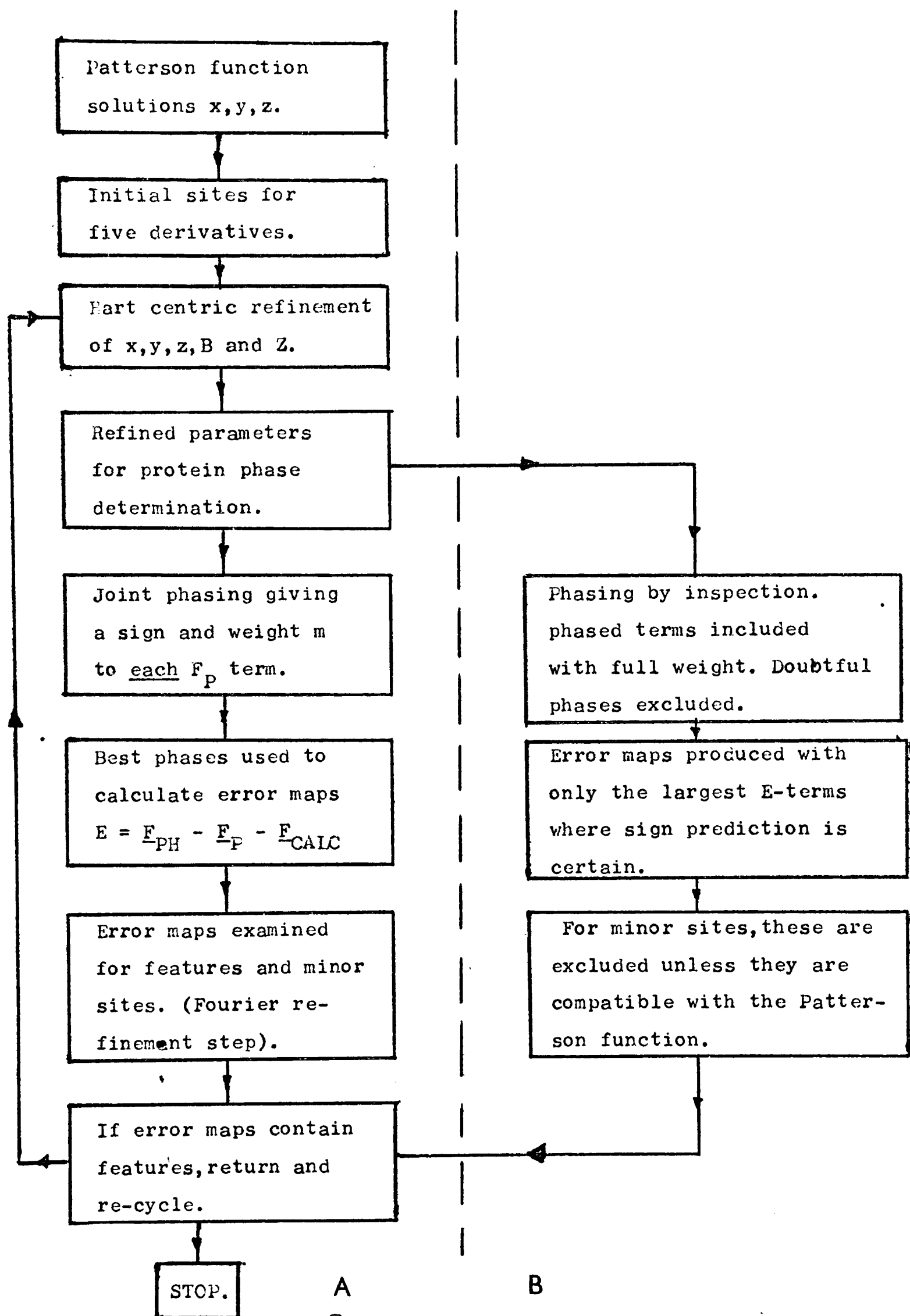


Fig. 18 Flow diagrams illustrating two approaches to centric phase determination. The recycling in both of these forms a feed-back loop for false information. However, this process is reduced to a minimum in scheme b.

given a starting set of heavy atom parameters. Best phases from the combined set of five derivatives have been applied to each in turn, to produce ΔF and $\Delta\Delta F$ Fourier projections using the methods described. Features appearing in these maps, such as minor sites in both syntheses or indications of incorrect occupancy in the $\Delta\Delta F$ synthesis have been taken account of in the successive cycle.

Thus each cycle is seen to include two refinement steps to improved parameters. 1) A Hart refinement and 2) a Fourier refinement.

Under such circumstances, one might anticipate that convergence would be quickly reached. The principal test that was applied for such convergence was the lack of feature in each $\Delta\Delta F$ synthesis. This expected convergence did not occur. The process of refinement started by Swan was continued for three such cycles, where apparently clear indications of additional sites and shifts eventually led to a reduction of the initial $\langle m \rangle$ from 0.83 to 0.71. At this point some of the sites included did not have plausible vectors in the projection Pattersons. Further, as seen in fig. 17a,e, the pseudo two-fold symmetry indicated by the dashed lines does not relate the sites as well as might have been expected from the diffraction patterns.

The E-value of the derivatives determines the weight each will have in the phasing step. At this point the E-values were fairly similar, with that of PAQ and GUN being the lowest. The phasing of the odd-ordered reflections for the OKL zone, instead of being random with low weight, appeared on examination to have an unexpectedly high value of $\langle m \rangle$. That this must be the case, is apparent from the observed break-down of local symmetry in figs.17a,e.

In order to assess the exact contribution of these reflections the data was split into two parity groups :

$$(k + 1) = 2n \text{ and } (k + 1) = 2n + 1.$$

Ten Patterson syntheses were produced in this way. The appearance

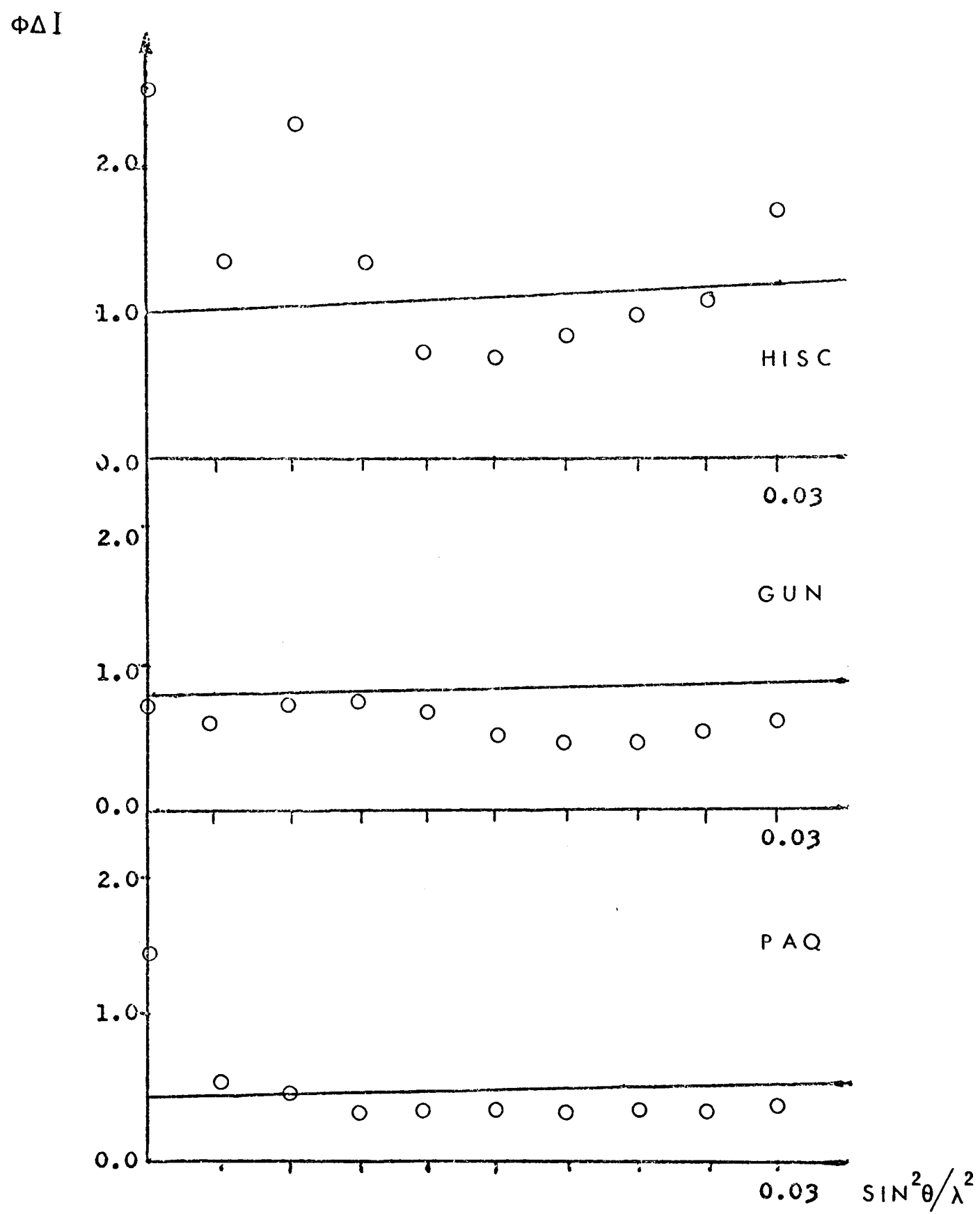


Fig. 19 Crick-Magdoff plots for PAQ, GUN AND HISC Centric data

and interpretation of that of PAQ has already been discussed. The odd parity Patterson functions of GUN, PAQ, PTI, PTF were essentially featureless, but that of HISC had higher features relative to the origin peak than that of the corresponding even-parity function. The origin peak of the odd-parity function was 25% that of the even. Such a breakdown in body-centring is possible only if two quite different environments exist, one of which will bind as HgI_4^{2-} ion while its local-symmetry related position will bind the ion only weakly, or not at all. The other possibility is that the introduction of the HgI_4^{2-} ion induces the breakdown in local symmetry, which implies that this derivative is non-isomorphous with the native protein.

Several tests of non-isomorphism have been proposed. That of Crick and Magdoff states that :

$$\Phi \Delta I \simeq M \cdot \left[\frac{N_H}{N_P} \right]^{1/2} \cdot \left[\frac{f_H}{f_P} \right] \quad (14)$$

where $\Phi \Delta I \simeq \frac{[(\overline{\Delta I})^2]^{1/2}}{\overline{I}_P}$ and M is 2 for centric data or $\sqrt{2}$ for acentric data, N_H and N_P are the number of heavy atom and protein atoms respectively, and f_H and f_P their atomic scattering factors. Clearly $\Phi \Delta I$ is a function of $\sin^2 \theta / \lambda^2$. If an estimate can be made of N_P and N_H , assuming Wilson statistics to be obeyed, plots of $\Phi \Delta I$ for the observed data and for the theoretical value obtained from equation (14) should be identical. Any deviation from the theoretical value may indicate non-isomorphism which is not represented in (14). Usually the observed values in these plots are far from ideal, as the assumption of Wilson statistics and estimates of N_H and N_P are both liable to error. However, the relative quality of derivatives of the same native protein may be better indicated by such plots. Figure 19 shows plots of $\Phi \Delta I$ for centric data for the derivatives PAQ, GUN and HISC. It can be easily seen that HISC deviates by a long way from the ideal. It was

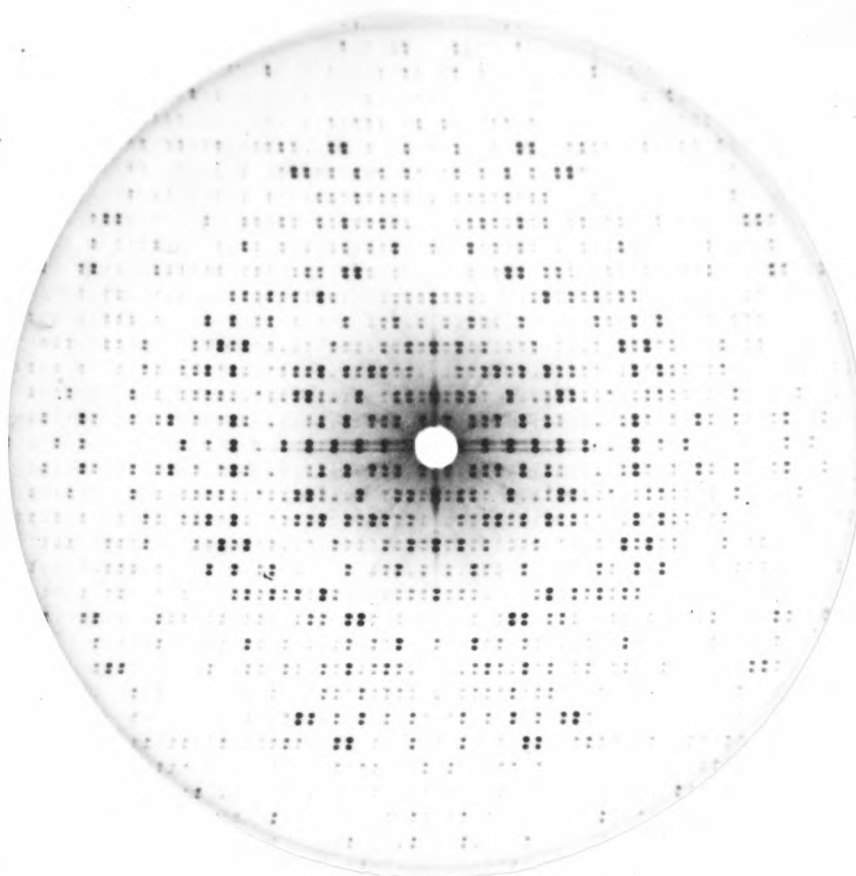


Fig. 20a The HKO zone of PAQ (upper pattern) superimposed upon that of the native. The pattern extends to a resolution of $2.5 \text{ \AA}$. There are no changes in cell dimensions. Characteristic, strong intensity changes may be seen in $0\ 14\ 0$, $0\ 8\ 2$ and $0\ 18\ 4$.

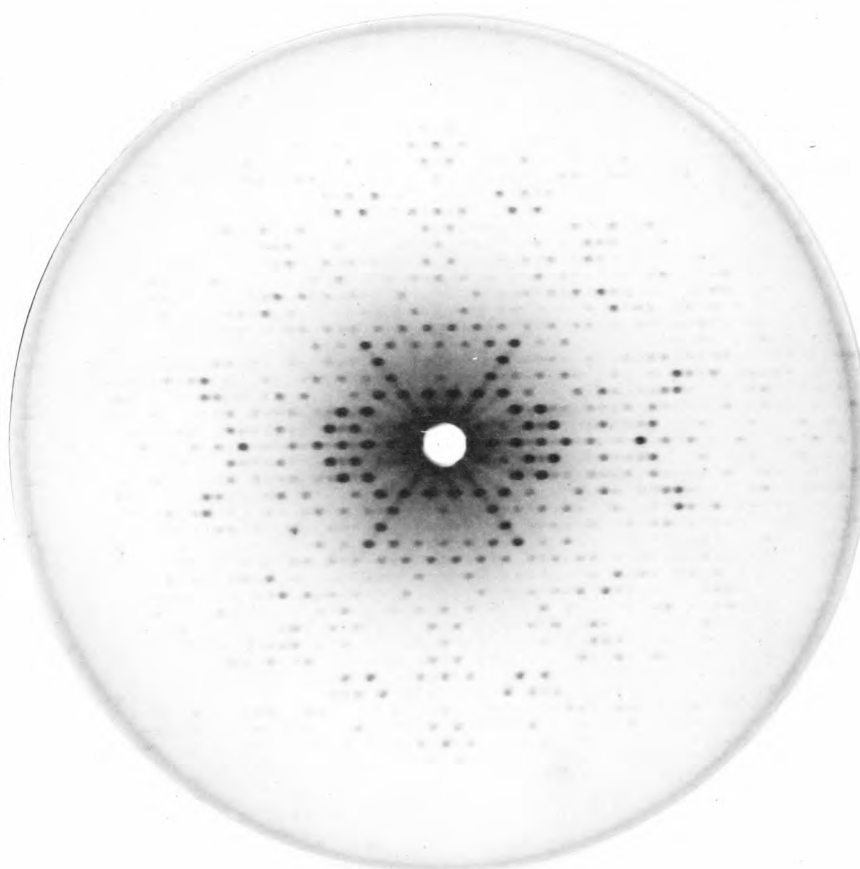


Fig. 20b The OKL zone of native prealbumin soaked in 10mM gold cyanide. Quite severe disordering is shown by the relatively high intensity of the inner part of the pattern. Little intensity can be seen at the resolution limit ($3\text{\AA}$). Low concentrations of platinum compounds and higher concentrations of uranium salts produced similar disorder.

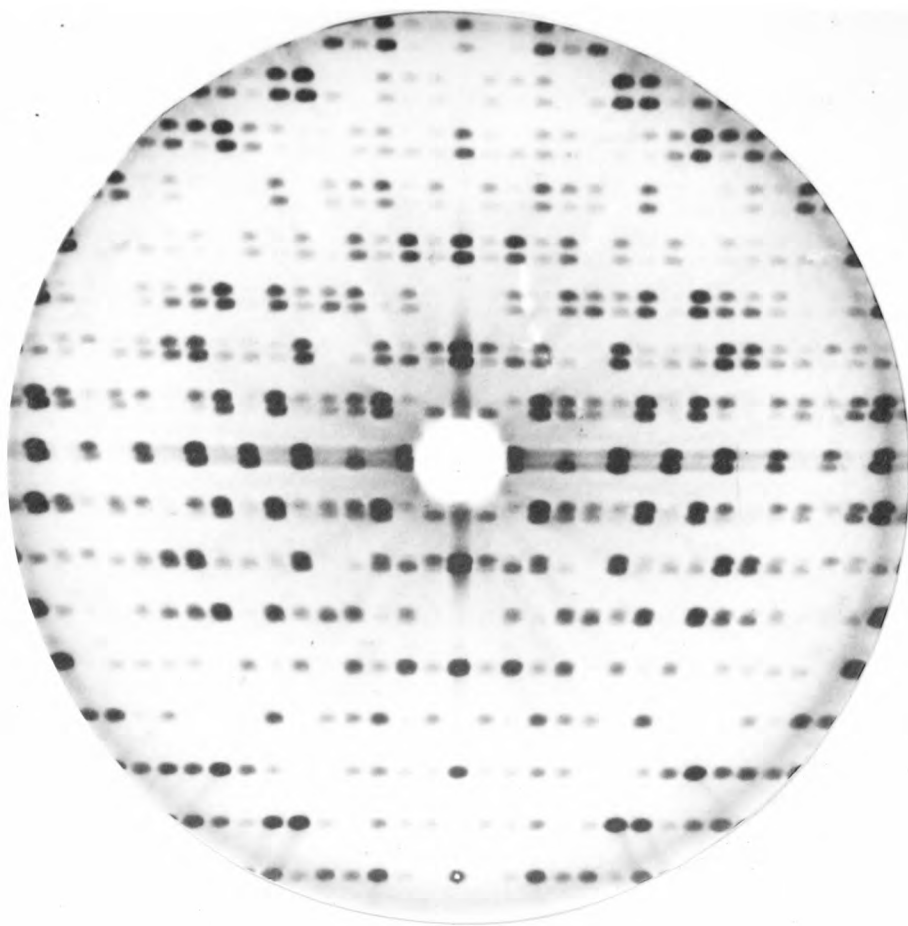


Fig.21a The HKO zone ($6 \text{ \AA}$) of a uranyl derivative of prealbumin produced by soaking the crystal in 5 mM $\text{UO}_2(\text{NO}_3)_2$ for 36 h. This picture is fairly typical of the pattern produced by uranyl compounds. The most striking intensity change is the axial reflection $0\ 4\ 0$, which vanishes in the uranyl derivative (upper pattern). This represents a change in structure factor of approximately 1200 electrons. Although there is a very large heavy atom contribution to this term, it does not account for all of this change.

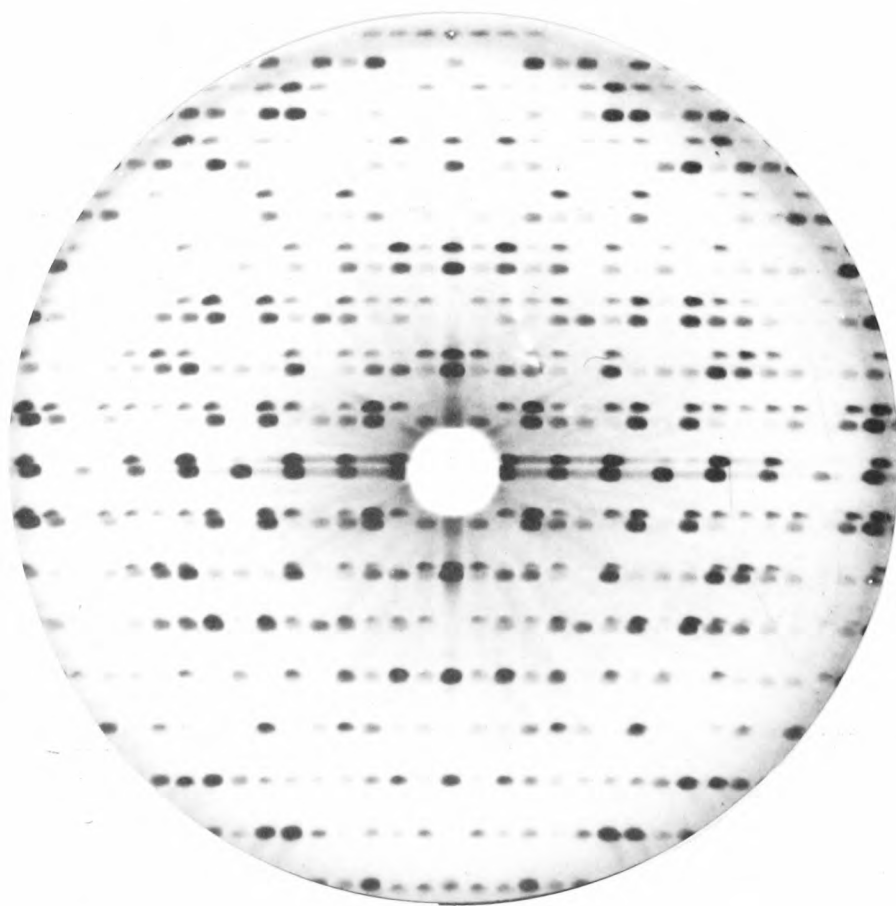


Fig.21b The HKO zone ($6 \text{ \AA}$) of a derivative made in the same way and under the same conditions as the one above, but using $\text{UO}_2(\text{CH}_3\text{COO})_2$. The substitution is very different. Here $0\ 4\ 0$ is only slightly reduced in intensity while $0\ 8\ 0$ has completely disappeared. It was very difficult to produce a reproducible uranyl derivative of prealbumin. The large change in a^* that is evident in both these patterns seems to be a characteristic of uranyl binding.

decided at this stage to remove this derivative from consideration.

One disadvantage with the automatic process of 'best phase' refinement is that it allows no easy comparison of the quality of contributing derivatives. If any derivative gives rise to features in a difference map that are wrongly interpreted as heavy atom sites, these sites tend to persist very strongly in subsequent Fouriers. This 'Fourier feedback' is discussed by Dickerson (1967) who shows the degree of persistence by using model derivatives. Because the E-value for HISC was not much greater than the mean value of the others, and as the occupancy of the observed sites was high, this derivative may have introduced error into the phase calculations. It was probably this derivative alone that had led to the marked breakdown in pseudo-symmetry of the OKL maps. As the OKL odd-parity Pattersons were virtually featureless for the remaining four derivatives, it was decided to dispense with the odd-ordered terms for this projection, so that subsequent work is described using terms for which $(k + l) = 2n$ only.

The new method of refinement which was adopted, is represented by the flow diagram in fig. 13b. The method described here was tedious, but was far more informative of the behaviour of individual derivatives than the previous automatic procedure. The even-parity OKL Pattersons were solved for the y and z parameters of the projection data. These solutions were essentially similar to the parameters obtained from the Pattersons of the full data sets, except that the y values of local two-fold related sites were taken to be identical. (The actual difference in y from refined $2.5 \overset{0}{\text{\AA}}$ data was 0.011 for the two PAQ sites.)

A cycle of refinement consisted of several cycles of Hart parameter-shift refinement for each of the four derivatives. After convergence, the phased structure factors $\underline{F}_p$, $\underline{F}_{PH}$, $\underline{F}_{calc}$ were output, together with the error term $\underline{F}_{PH} - \underline{F}_p - \underline{F}_{calc}$. An example

Compound.	Site no.	x	y	z	Z	B/ $\text{\AA}^2$
PAQ	1	0.0520	0.1080	0.0750	0.21	7.0
	2	0.9800	0.1150	0.4250	0.20	8.0
PTI	1	↑ all 0.05*	0.0626	0.0472	0.08	6.0
	2		0.0840	0.0666	0.09	6.0
	3		0.1090	0.0906	0.18	15.0
	4		0.1187	0.1250	0.11	14.0
	5	—	0.2365	0.7500	0.09	7.0
	6	↑ all 0.98*	0.0625	0.4528	0.09	6.0
	7		0.0840	0.4334	0.09	6.0
	8		0.1090	0.4094	0.18	15.0
	9		0.1187	0.3750	0.11	14.0
PTF	1	—	0.2302	0.6999	0.08	6.0
	2	—	0.0919	0.0659	0.18	6.0
	3	—	0.1027	0.0830	0.16	11.0
	4	—	0.0919	0.4341	0.18	6.0
	5	—	0.1027	0.4170	0.16	11.0
	6	—	0.2302	0.8100	0.08	6.0
GUN	1	0.0540	0.1080	0.0750	0.21	8.0
	2	0.9800	0.1170	0.4250	0.19	8.0
	3	0.1000	0.2363	0.6630	0.17	11.0
	4	0.0500	0.2415	0.7500	0.25	8.0
	5	0.7000	0.2429	0.8310	0.17	8.0

* These values are only approximate. There is such extensive overlap in the HKO zone of platinum derivatives that it is impossible to obtain precise values of the x-co-ordinates.

Table 5 Refined site parameters (position, occupancy and temperature factors) from the $3\text{\AA}$ centric refinements.

H	K	L	FP	FPH	FCALC	E	COMMENTS
0	1	3	-517	-481	113	-77	Non H-atom contribution
0	3	3	-602	-517	69	16	
0	4	3	-89	-100	-8	-3	Odd Ordered term
0	5	3	-235	-387	-137	-15	Large Fcalc contribution
0	6	3	-96	-89	8	-1	
0	7	3	-101	-140	-15	-24	
0	8	3	-79	-54	4	21	
0	9	3	-192	-24	136	32	
0	10	3	96	84	-10	-2	
0	11	3	413	338	-37	-38	
0	12	3	100	129	2	27	
0	13	3	-383	-503	-113	-7	
0	14	3	-209	-199	9	1	Breakdown of centring
0	15	3	-213	-119	76	18	
0	16	3	157	153	-7	3	
0	17	3	-237	-195	74	-32	
0	18	3	-360	-368	-6	-2	
0	19	3	-535	-599	-96	32	
0	20	3	-79	-79	-30	30	
0	22	3	-126	-94	1	31	
0	23	3	-73	30	95	8	Cross-over term
0	24	3	-25	-29	-10	6	
0	25	3	175	171	-8	4	
0	26	3	-256	-229	3	24	Limit of 3Å. Data set

a

DERIVATIVE				PTF			PTI			PAQ			GUN			JOINT
H	K	L	FP	FPH	FC	S	FPH	FC	S	FPH	FC	S	FPH	FC	S	PHASE
0	1	3	517	801	152	+				450	109	-	581	88	+	
0	3	3	602	585	27	-	592	48	-	516	78	-	486	99	-	-
0	5	3	235	423	-169	-	529	-176	-	386	-109	-	303	-73	-	-
0	7	3	101	134	109	+	125	55	+	139	4	+	158	-16	-	+
0	9	3	192	38	70	-	90	92	-	24	150	-	44	200	-	-
0	11	3	413	309	-154	+	334	-56	+	337	-11	+	350	70	-	+
0	13	3	383	268	57	-	304	-61	+	502	-97	-	414	-82	-	-
0	15	3	213	49	91	-	161	61	-	119	74	-	74	129	-	-
0	17	3	237	222	-117	+	242	40	+	195	77	-	120	187	-	-
0	19	3	535	600	12	+	521	-86	+	598	-96	-	455	-19	+	
0	21	3	79	58	87	-	58	29	-	79	-50	-	55	3	-	-
0	23	3	73	58	-74	+	60	43	-	30	80	-	55	154	-	-
0	25	3	175	264	-15	-				171	-12	+				

b

Table 6. Phased structure amplitudes from the Hart-refinement program (a)
Sign determination by four derivatives and the resulting joint
phase (b).

of this output is shown in Table 6 for the OK3 reflections taken from the OKL zone of PAQ. Several features are picked out below.

- 1) The small values of F_P , F_{PH} , and F_{calc} for the odd-ordered reflections (0 4 3), (0 6 3) etc.
- 2) The large changes ΔF in reflections (0 5 3), (0 9 3) and (0 13 3), each giving good indications that the protein phase is negative.
- 3) The large error term for reflection (0 1 3) where the F_{calc} is too large not to account for the observed ΔF . 0 1 3 is a low angle term and this effect is often observed. It may be due to changes in the electron density of the solvent between crystals of native protein and derivatives.
- 4) Terms (0 4 3), (0 7 3) and all the odd-parity terms are considered to be unphased, despite the sign indication on the term for F_P .

Variations of ± 15 in F_{calc} were within the experimental error of the determination.

- 5) Reflection (0 23 3) is a cross-over, because the F_{calc} is larger than the term $|F_P|$, while $|F_{PH}|$ is small.

The final parameters of the heavy atom sites which gave rise to the values of F_{calc} shown are listed below:

Parameter	Site 1	Site 2
Z	0.21	0.18
B	19.0	19.0
x	0.0	0.0
y	0.1070	0.1120
z	0.0722	0.4240

Centric R factor = 48%

E-value = 25

R.M.S. $|F_{calc}|$ = 50

The positional parameters for these two mercury sites are close

to those finally obtained in the high resolution 3D work. The centric reliability factor is average for data at this resolution. The R.M.S. value of F_{calc} is only twice that of the E-value, but these values include contributions from both odd and even ordered reflections. A similar refinement using only the even-ordered terms gives a slightly different E-value, but an R.M.S. F_{calc} of 95, because the weak reflections arising from cell-centring have been excluded. Thus it must be remembered, when comparing the R.M.S. F_{calc} with the E-value for prealbumin data at resolutions of less than 4 Å, that this ratio is depressed by the effect of the odd-ordered terms.

It is apparent from the listing in table 6a that terms such as (0 7 3) have not been satisfactorily phased because of the low heavy atom contribution to the derivative structure amplitude F_{PH} . This missing phase information is hopefully provided by an independent derivative, whose heavy atoms do contribute to this term. Previously, information from the other available derivatives had been combined automatically as described in Appendix 2. To compare the protein phase determination in a more interactive way, a programme was written to present the structure factor information from the Hart programme as shown in table 6b. The index of each reflection is followed by F_{P} and for each derivative in turn : F_{PH} , F_{calc} and S where S is the protein phase determined by that derivative. For example in the derivative PTF, for the reflection (0 5 3) there is a very large negative contribution which increases the magnitude of F_{P} by a factor of nearly 2. This derivative predicts a negative sign for F_{P} as shown in column S. The reflection (0 7 3), poorly phased by PAQ, is well phased by PTF and PTI, giving a confident prediction of a positive sign for the protein phase. 226 such terms were examined, and only those terms where either all derivatives predicted the same sign, or at least two with values of

$|F_{\text{calc}}|$ greater than 50, predicted the same sign, were assigned a phase. This phase is shown in the column headed 'Joint Phase'. The automatic procedure would have phased all the terms, but with varying weights, or figures of merit m .

This list of jointly-determined phases was used in a selective best-sign procedure to produce signed E-terms of the type shown in the last column in table 6b. Only E-values greater than 40 were used in these E-maps. For example, term (0 23 3) in PTF has an E-value of +59 and gives a protein phase opposite to that of the remaining derivatives. Assuming a correct protein F_p of -73 the term to be included in the E-map is +59.

The maps were produced in space group $P2_1 2_1 2_1$ with appropriate adjustment of the phases of the OKO terms as described earlier. The maximum number of terms used in any one map was 50. These maps gave very clear indications of deficiencies in the current heavy atom models. The effect of an incorrect temperature factor was immediately apparent as shown in fig. 16a. An impression was gained that these selective E-term syntheses were sharper and easier to interpret than the conventional $\Delta\Delta F$ synthesis, already described. Fig. 16 shows the appearance of an E-map for the derivative PTI, after 3 cycles of refinement as outlined. Sites have already been included at positions 1, 2, 3, 4, 5 and 6. After sites had been included at the positions of the positive peaks the subsequent cycle of refinement (cycle 4) gave an E-map for this derivative which was featureless. On a previous cycle, the end peaks of the quartet of peaks had been mistakenly displaced in y by 0.2 of a unit cell. Even though these false positions had been included in the phasing step, the resulting E-map showed holes surrounding the false positions and peaks at the correct sites.

After four cycles of such refinement, all the E-maps for these

derivatives were flat. At this stage, the best phases were used to produce the maps shown in fig. 17f,g, which are to be compared with the initial maps immediately above. Where the phase for a term was well determined, and a heavy atom contribution indicated a cross-over, these were included correctly in the lower set of maps. At the same time, the terms, corrected for cross-overs, were used to produce a set of OKL Pattersons and the features in all the maps checked against the vectors in these Pattersons. These were in good agreement with all the maps. A final refinement using all the terms, odd and even, produced the set of parameters shown in table 5

A programme which tested the sign agreement of each derivative against the joint set showed that PAQ and GUN agreed best at high values of $|\underline{F}_{\text{calc}}|$. At low or medium values of $|\underline{F}_{\text{calc}}|$ there was not much distinction between the derivatives. A test of the sign predictions of the current heavy atom model for HISC, which had been excluded from the phasing cycles, showed relatively poor sign agreement at high values of $|\underline{F}_{\text{calc}}|$, even though HISC contained the PAQ sites. The most damning appraisal of this derivative came from application of the best phases to the HISC differences, which produced the projection difference map shown in fig. 17g. The negative features in this map are as large as the positive ones.

The final parameter to be determined for the derivatives was the x co-ordinate of the sites. This was to be obtained from examination of the HKO projection data. At this it was decided to abandon the two derivatives PTI and PTF. Both these have multiple sites which overlap extensively in the HKO projection, which would make refinement difficult. The OKL work had suggested that these derivatives were not as isomorphous with the native as PAQ and GUN. Moreover, the addition of platinum to prealbumin crystals is always accompanied by some disordering, so that an artificial temperature factor has to be

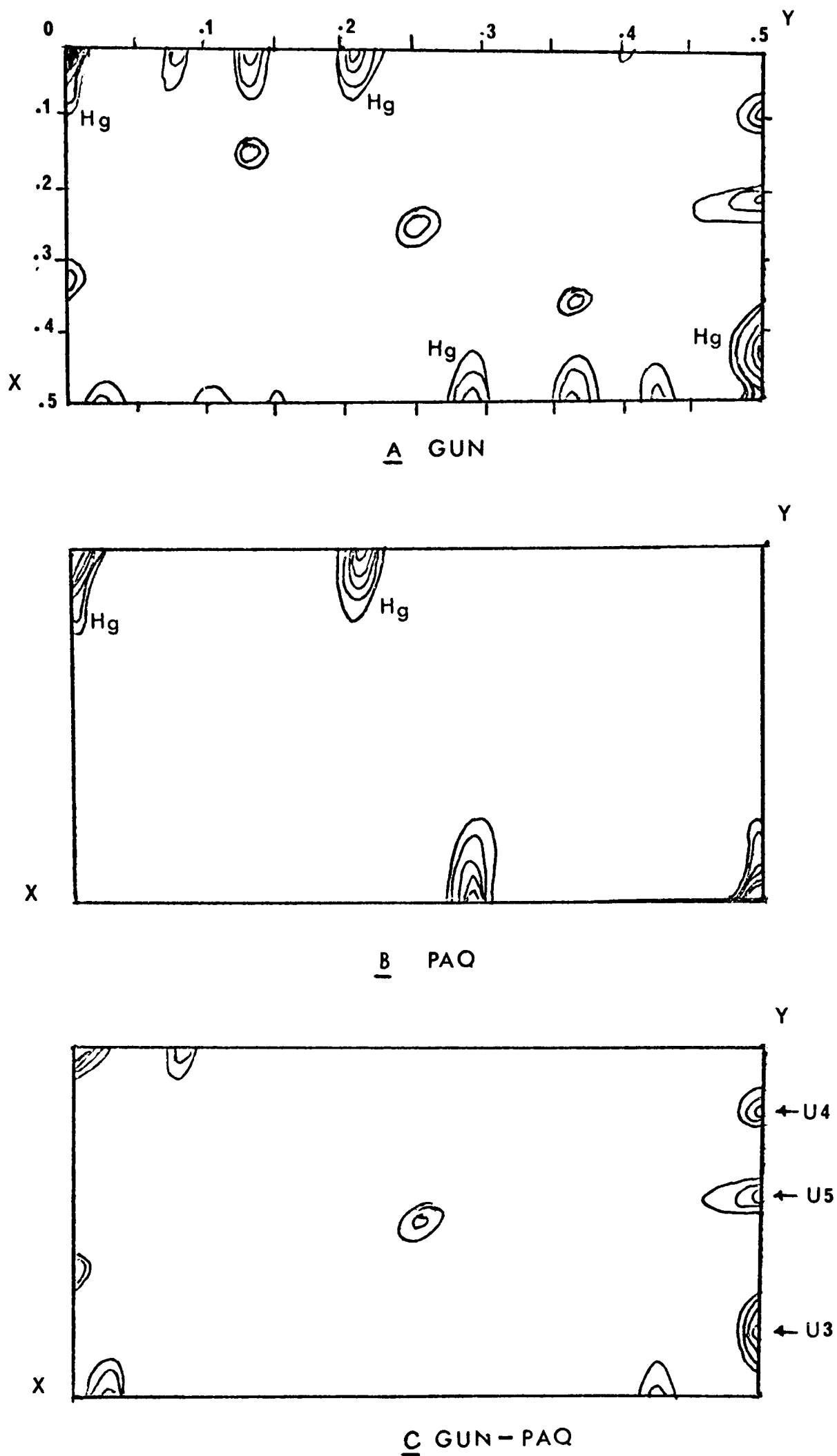


Fig. 22 The Pattersons of GUN, PAQ and (GUN - PAQ) in the HKO zone. The GUN Patterson is complicated by the cross-vectors between mercury and uranium atoms, although the mercury atom-vector set is easily seen (Hg). The Patterson in C established the identity of several of the peaks in A as mercury-uranium cross vectors and left the uranium self and cross-vector set. The identification of the three self-vectors (4,5,3) was made easier, and allowed an estimation of the uranyl site x-co-ordinates shown in Table 5.

applied to all platinum derivatives.

The photographic HKO data was on a scale 1.3 times that of the OKL data as judged from the common OKO row. Occupancies, temperature factors and the y parameters for the sites were thus carried over into the other zone. The symmetry of the Patterson leaves an uncertainty in the position of the atomic co-ordinates derived from it, to the extent of half a unit translation in the x and y directions. A set of co-ordinates was chosen for one PAQ site and cross-vectors determined the relative position of the second mercury site.

The GUN HKO projection Patterson was not so readily interpretable, however the PAQ self-vector set is readily distinguished. Making an assumption that the occupancy of these sites is identical in both crystals, a difference Patterson where PAQ is used as native should reveal the uranyl self-vector and cross-vector set alone. The data was combined in this way and the Pattersons produced are shown in fig. 22 . The positions of the mercury-uranium cross-vectors are indicated, as revealed by the difference Patterson of GUN-PAQ.

From these two solutions a final complete co-ordinate set for PAQ and GUN was produced and is shown in table 5.

Derivative trials.

It has been seen that the projection Pattersons of prealbumin derivatives are nearly always interpretable in terms of the sites seen in a difference Fourier. The OKL zone of prealbumin gives particularly good projection Pattersons, possibly because of the greater population of this zone, and the short a axis ($43 \text{ \AA}$). When photographic data was measured on a trial derivative, a projection difference map (to 3 or $6 \text{ \AA}$) was first calculated using the best phases obtained as described in the last section. Subsequently a projection Patterson was calculated and its features checked against the sites seen in the difference map.

In this manner, about thirty possible heavy atom derivatives were surveyed. The results of these binding studies are summarized in table 4. The source of this data is given as a code letter. P for projection, or 3D for diffractometer data. The number indicates the zone collected (1=OKL, 2=HKO and 3=HOL) . None of the compounds listed, however, yielded derivatives of a comparable quality to PAQ and GUN, as judged by the various criteria mentioned at the beginning of the chapter. It is unfortunate that these two derivatives have common mercury sites. Many attempts to prepare a single derivative of prealbumin using uranyl nitrate were frustrated by cracked or disordered crystals. An experiment was finally tried in which the phosphate buffer was omitted and this seemed to give crystals which were less disordered, and had higher uranium content. It was subsequently found that this was how the derivative GUN had been prepared. The single uranyl derivative is called PAUN and the double derivative, PAQUN, to distinguish them in later chapters.

These three derivatives were used for the high resolution work on prealbumin. They represent a situation which is far from ideal, because between them, there are 2 sets of common sites, which led to difficulties in refinement. Also the isomorphism of the uranyl compounds is significantly lower than that of the mercury derivative. At the present time, however, no other suitable derivatives are available.

Specific differences in prealbumin native crystals.

At the end of 1972 a 2.5 Å⁰ Fourier synthesis of the native protein was produced. This map was based on two derivatives PAQ and PAQUN. The map was largely uninterpretable. It is now realised that this was principally due to the native crystals used. These were from a batch crystallised by Behringwerke, Marburg am Lahn. The diffraction pattern of these appeared identical to that of small crystals (PAN) used in previous work. Mr.K.Snape, of this laboratory,

found that it was possible to remove mercury from the PAQ crystals using dithio-threitol as described in table 4. This gave a native crystal (PAQDTT) similar to PAN, but with certain specific intensity changes. The table below gives a comparison between the two native data sets when used with data from the PAQ crystals. These reflections have been selected because the E-values were high when PAN was the native data set used.

Crystal reflection			PAQ		PAQ - PAN					PAQ - PAQDTT				
H	K	L	F_{PH}	F_{calc}	F_P	$F_{PH} - F_P$	E	Sign		F_P	$F_{PH} - F_P$	E	Sign	
3	5	0	541	-123	365	+176	53	-		436	+105	18	-	
11	5	0	127	-9	200	-73	64	+		113	+14	5	-	
4	6	0	125	-46	335	-110	64	+		159	-34	12	+	
9	8	0	205	-17	117	+88	71	-		195	+10	7	-	
7	11	0	221	-25	335	-114	89	+		267	-46	21	+	
10	9	0	338	+1	400	-62	61	-		327	+11	11	+	

For a heavy atom solution giving the values shown under F_{calc} , a better fit to the model is obtained using data from PAQDTT as indicated by the size of the E term ($F_{calc} \pm F_{PH} \pm F_P$). It should be noted that although the $\langle E \rangle$ and centric R values are higher in the case of the (PAQ - PAN) data set, only a small fraction of the protein centric phases have changed. The only phase changes in the 6 OKL reflections above are not significant, because of the small values of F_{calc} . This means that the previously described $3 \overset{0}{\text{\AA}}$ refinement work is still valid. However, the acentric phases would be very seriously in error, since these are far more sensitive to changes in the relative magnitudes of $|F_P|$ and $|F_{PH}|$. This observation made it obligatory to recollect a high resolution native set using PAQDTT crystals. A further derivative PAUN was measured at the same time.

Some evidence of the nature of the difference between these two

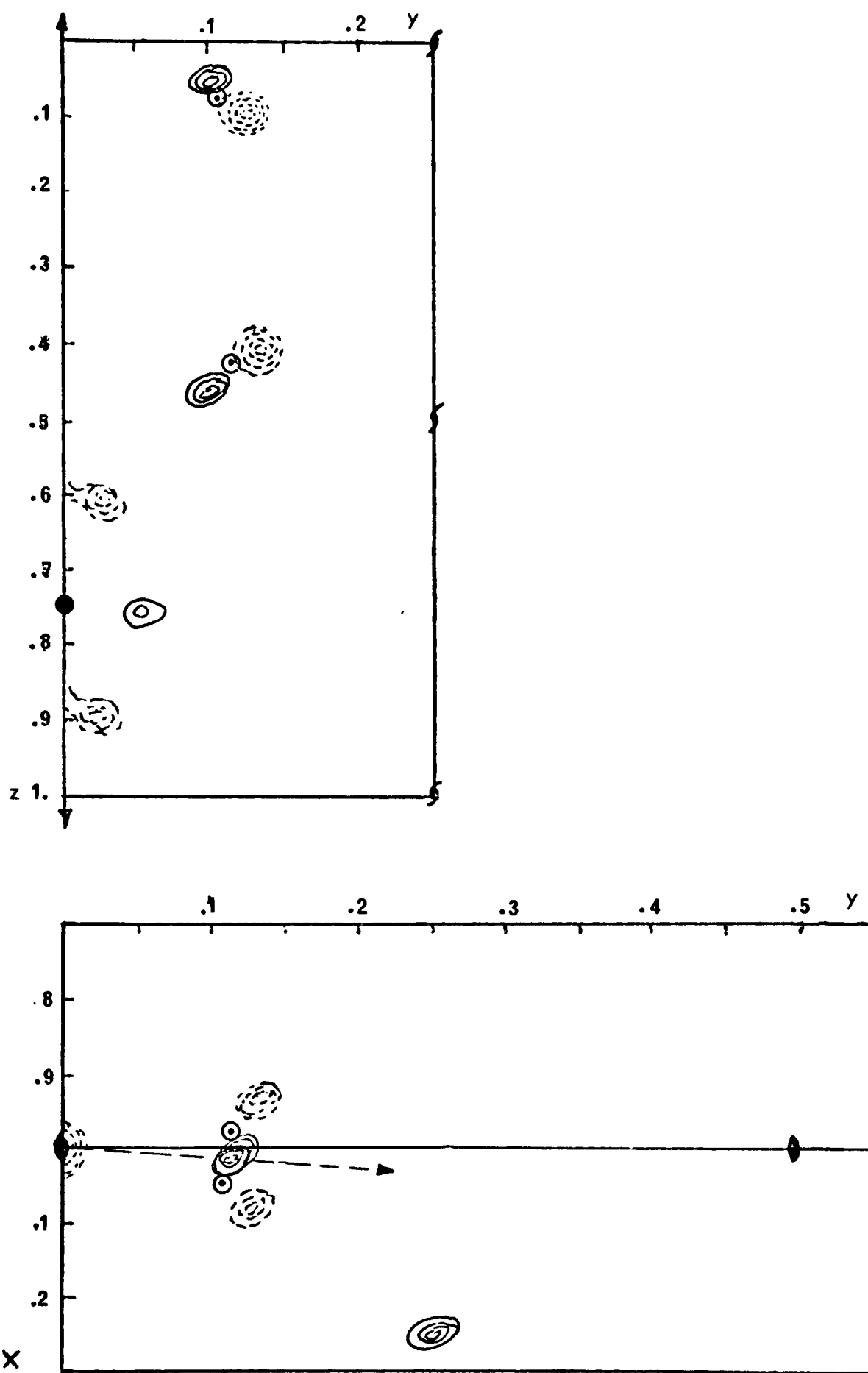


Fig. 23 The OKL(top) and HKO(bottom) difference projection
Fouriers of the two native crystal forms ; PAN and
PAQDTT. Although there are no changes in space group
or cell dimensions, as the Fouriers show, there are quite
specific and localised changes between the two forms.
The sense of the difference terms is such that negative
features reflect scattering matter present in PAN, but
not in PAQDTT (PAQDTT-PAN). The mercury sites are marked
by (⊙).

crystals can be presented. It is noticeable that reflections where a larger difference exists between the two native structure amplitudes also have large contributions from the mercury sites. This effect is less pronounced at higher angles. In general the MFCI is larger in the (PAQ - PAN) set than in the (PAQ - DTT) set, so that there are more contributions to the difference terms than from the mercury sites alone.

These observations suggest that there may be some local disturbance close to the sites of mercury substitution in PAN. To examine this possibility, two difference Fourier projections were produced using coefficients ($F_{\text{PAQDTT}} - F_{\text{PAN}}$) and phases from the best sign sets produced as described. Figure 23 shows these projections. The positions of the mercury sites in projection are marked $\odot$. It can be seen that there are deep holes near, but not at, the mercury sites as well as remote features along the two-fold axis, near to the molecular centre($\bullet$). The possibility that the features near the mercury sites are caused by Fourier feed-back can be discounted because this would cause the features to appear exactly at the mercury positions. A negative feature in such a difference Fourier means that there is less electron density in PAQDTT at that position than in PAN, and vice versa for a positive feature.

Several explanations may be advanced for this observation. It has been established that the sulphydryl groups of prealbumin are relatively inaccessible.

1) The local conformation of the protein may have altered to accommodate the mercury ion when the co-crystallised PAQ derivative was prepared. This altered conformation may then have been retained when the Hg^{2+} ion was removed to produce the PAQDTT crystal. A most likely candidate for such a shift would be the sulphydryl group itself. As sulphur is relatively electron-dense, it would be seen on such a projection map. That the sulphydryl group itself is not lost, may be

demonstrated by successfully re-diffusing mercury to give a derivative with an indential diffraction pattern to that of PAQ.

2) It may be possible that some of the sulphur atoms have been lost. The PAQ crystals have been used for many years due to lack of other suitable crystals. Over this period some of the sulphur atoms may have been lost to the mercury atoms as the sulphide. There would then be a lower occupancy at the cysteine sulphur positions, relative to that in the native PAN crystals. However, the appearance of positive features close to the holes in the projection Fouriers suggest that explanation 1. is more likely.

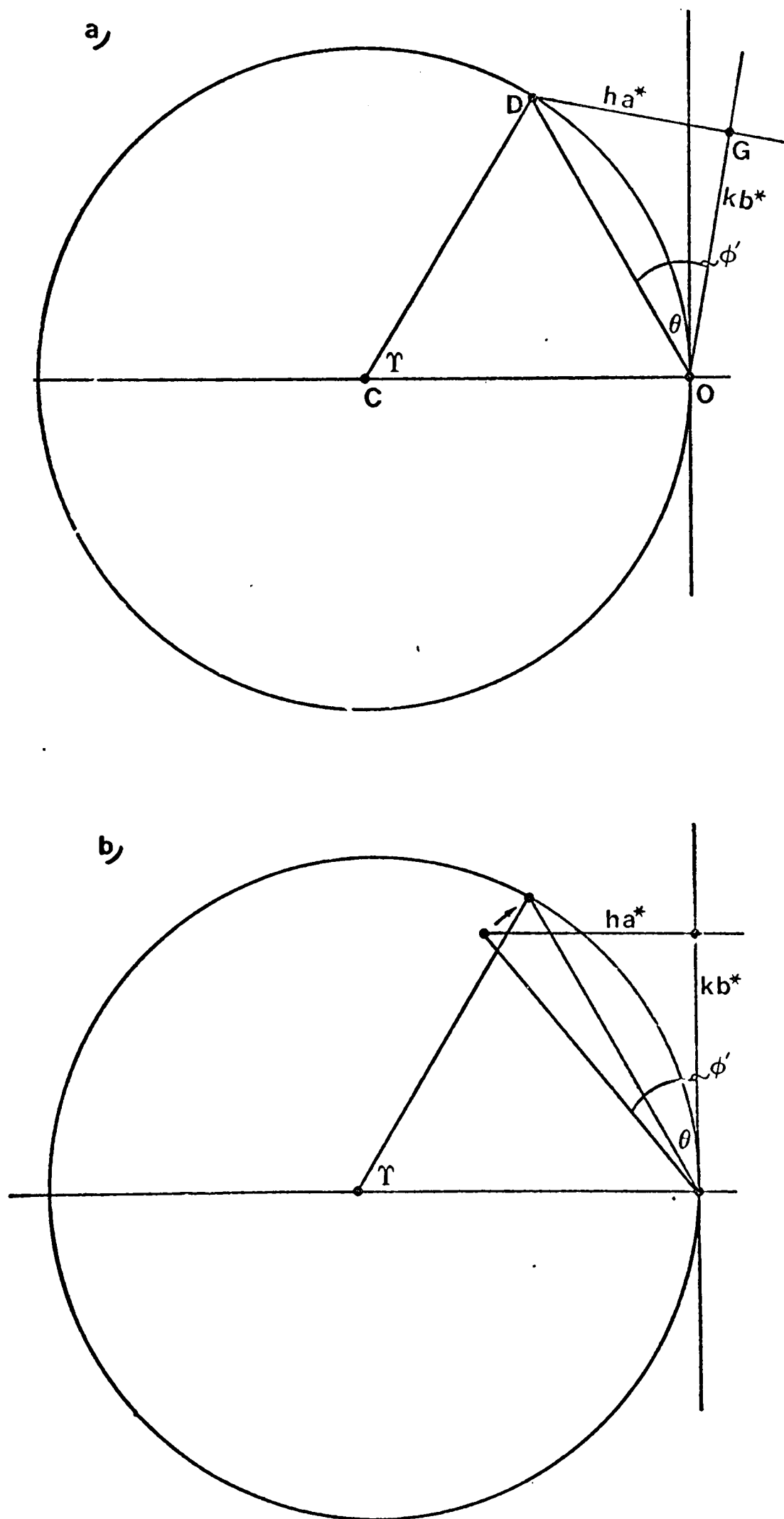


Fig 24

Geometry of the linear diffractometer.

a) Equatorial projection. The slides lie along DG, OG.

b) The corresponding Ewald construction.

CHAPTER 4.

Measurement and processing of 3D data from prealbumin crystals.

The measurement of a 3D data set for prealbumin to a resolution of $2.5 \text{ \AA}$ means the recording of nearly 9 000 unique reflections, or 18 000 Friedel-related reflections. At the time that this work was started, the two machines available were the Hilger and Watts four circle diffractometer and the Royal Institution Linear Diffractometer. The four circle machine (often abbreviated to 4 $\odot$) employs a single detector and achieves a speed of data collection of 40 to 80 reflections an hour. This would mean that a $2.5 \text{ \AA}$ prealbumin data set would take about 300 hours of continuous operation. The machine of choice was then the linear diffractometer, which can measure 250 reflections an hour.

The linear diffractometer has been described by Arndt and Phillips (1961). It has three orthogonal slide-arms, linked mechanically to a shaft on which a goniometer head may be mounted. The slide arms act as mechanical analogues of translations in reciprocal space. These can be set to automatically bring reciprocal lattice points of a mounted crystal into the reflecting conditions (fig.24). An oscillation mechanism then rotates the crystal about its mounting axis (Φ) so that the integrated intensity may be recorded. This machine is usually operated in the flat-cone setting (Buerger, 1942) where reflections symmetrically placed with respect to the flat-cone level diffract simultaneously. Phillips (1964) pointed out that with protein crystals, reflections in several levels on either side of the flat-cone level can be brought into the reflecting condition for a comparatively small rotation ($\delta\Phi$) of the mounting axis. This is due to the large unit cells of proteins, leading to a densely populated reciprocal space. The diffractometer in this laboratory is equipped to measure 5 reflections simultaneously. Angles of incidence

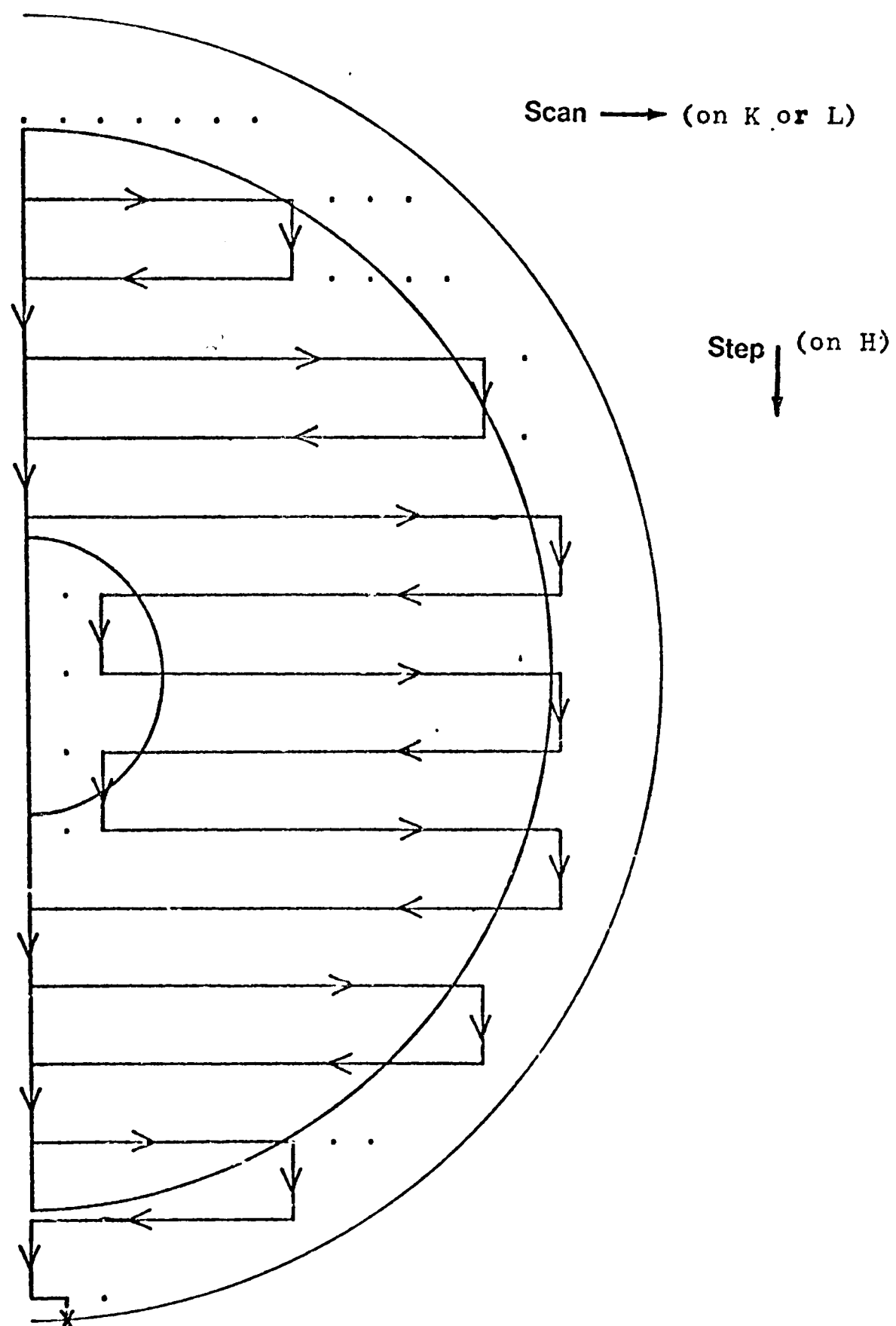


Fig. 25 Sequence of measurements on the linear diffractometer.

The trace is shown of a scan made through a single level in reciprocal space. The inner and outer resolution limits are set manually for each set of levels.

of up to 30° can be accommodated and the machine can access a hemisphere of reciprocal space defined by $\xi = \pm 1$ r.l.u. and $\zeta = \pm 1$ r.l.u. where ξ is the radial component and ζ is the vertical component in cylindrical polar co-ordinates.

In the flat-cone setting, however, there is a blind region with respect to measurement of radius $\xi_{\min}$ given by

$$\xi_{\min} = 1 - (1 - \zeta_f^2)^{1/2} \quad (1)$$

where ξ_f refers to the value of the flat-cone level being measured.

This is a theoretical maximum restriction, because the LP factor of reciprocal lattice points close to the mounting axis or the blind region is very large, as is also the divergence $\delta\gamma$ between the reflected beams for adjacent reflections, when ζ becomes large.

In practice, a maximum value of $\delta\phi$ and $\delta\gamma$ is adopted to suit the individual experimental conditions, and the area of reciprocal space to be measured is reduced accordingly. With the linear, these desired theoretical limits to the data cannot always be achieved. The linear sets itself with decreasing precision at low values of ξ , such that at $\xi = 0$ the crystal is free to swing through any angle. This is a mechanical limitation and not a property of the flat-cone geometry. With prealbumin data, this has led to an inner resolution cut-off of $3 \text{ \AA}$, corresponding to a $\xi_{\min}$ of 0.13 r.l.u. for each set of levels collected. The blind region of the flat-cone and that area of reciprocal space excluded by the criteria $\delta\phi \leq 0.4^\circ$, $\delta\gamma \leq 0.1^\circ$ both lie well within this cylindrical inner cut-off, as does an appreciable region of practically accessible reciprocal space. In order to collect these missing reflections, the crystal is usually mounted with a different principal axis co-incident with the ϕ axis of the machine. The previously inaccessible data now occurs at low values of ξ , but at acceptable values of ζ , and data already

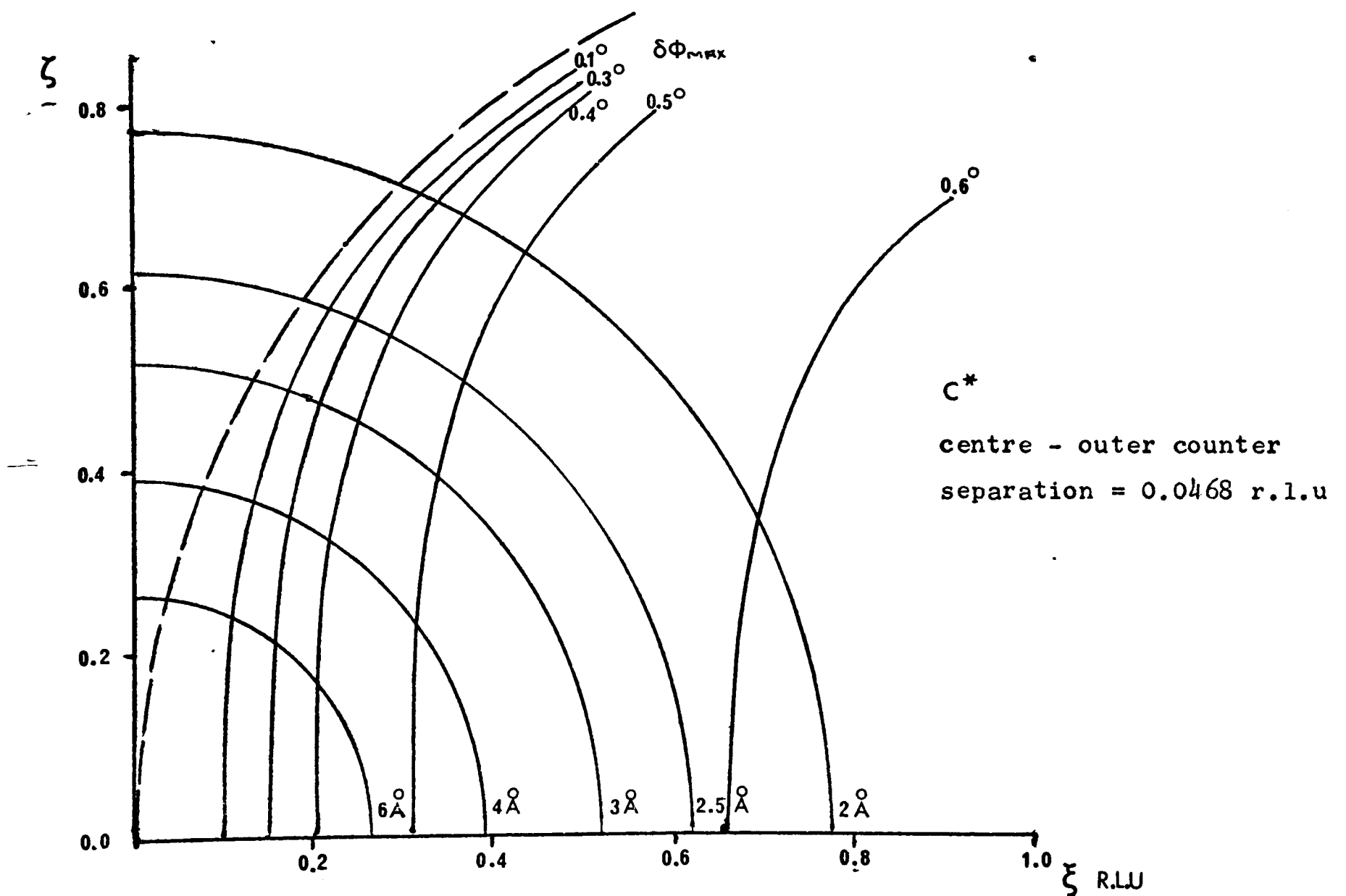
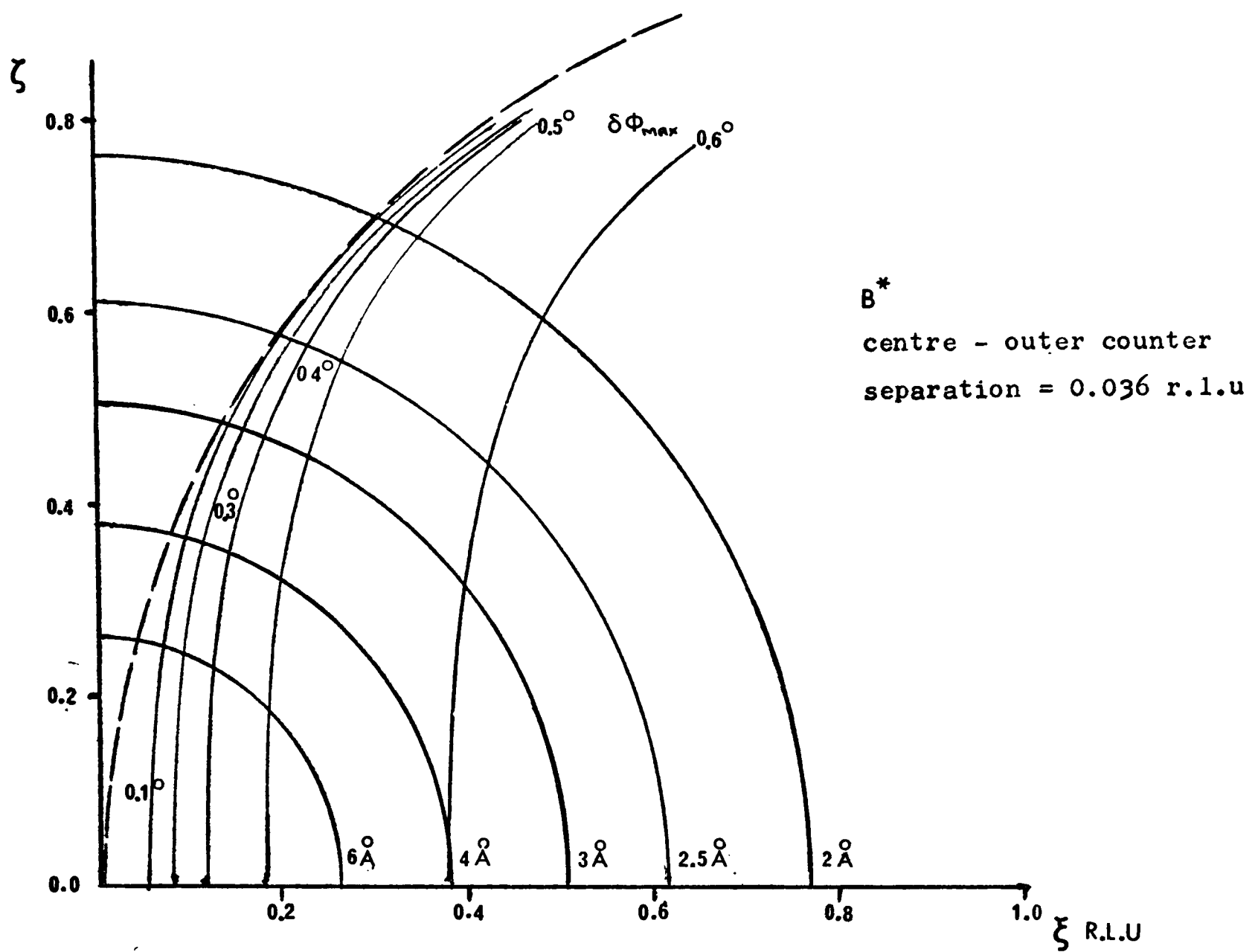


Fig. 26 Regions of reciprocal space accessible in the flat-cone setting for given values of $\delta\Phi$. The excluded regions lie to the left of the given $\delta\Phi$ contour. For PAUN a $\delta\Phi$ of 0.4° was used in the B* mounting.

measured now lies in the new blind region. There remains a quantity of low-angle reflections which must be collected using another diffraction geometry. A substantial quantity of overlapping reflections results from the two mountings. This is used to provide scale factors for relating data measured by different counters.

One set of prealbumin data has been collected on a Hilger-Watts diffractometer, modified to collect five quasi-simultaneous reflections in the flat-cone setting. This is a $4\odot$ machine provided with an extra axis ' Σ ' on which is mounted the detector array.

When this machine is operated in the flat-cone setting, all of the fore-going is applicable, except that nearly the entire sphere of reciprocal space is accessible and the machine can be set with equal precision right up to the blind region. The latter means that more care is required in the choice of scan widths to accommodate the larger values of $\delta\Phi$ between adjacent reflections, near the blind region. Had this machine been available from the outset, this would have been the diffractometer of choice. The linear is better suited to tetragonal or higher crystal systems. The method of data collection on the linear (fig. 25) means that many hours elapse between the recording of Friedel pairs, when radiation damage can reduce the desired agreement between these reflections. Because it is not convenient to record (hkl) and $(\bar{h}\bar{k}l)$ on the linear, a non-equivalent absorption correction factor has to be applied to the equivalents (hkl) and $(\bar{h}kl)$. The five circle does not have the above restrictions.

Figure 26 shows the regions of reciprocal space that are accessible in the flat-cone setting for a prealbumin crystal with the Φ -axis co-incident (a) with $b^*(35.6 \text{ \AA})$ and (b) with $c^*(65.8 \text{ \AA})$. Concentric circles indicate the area of reciprocal space within various resolution limits. A family of curves is shown centred on $\xi = 1$. These indicate the closeness of approach to the blind region

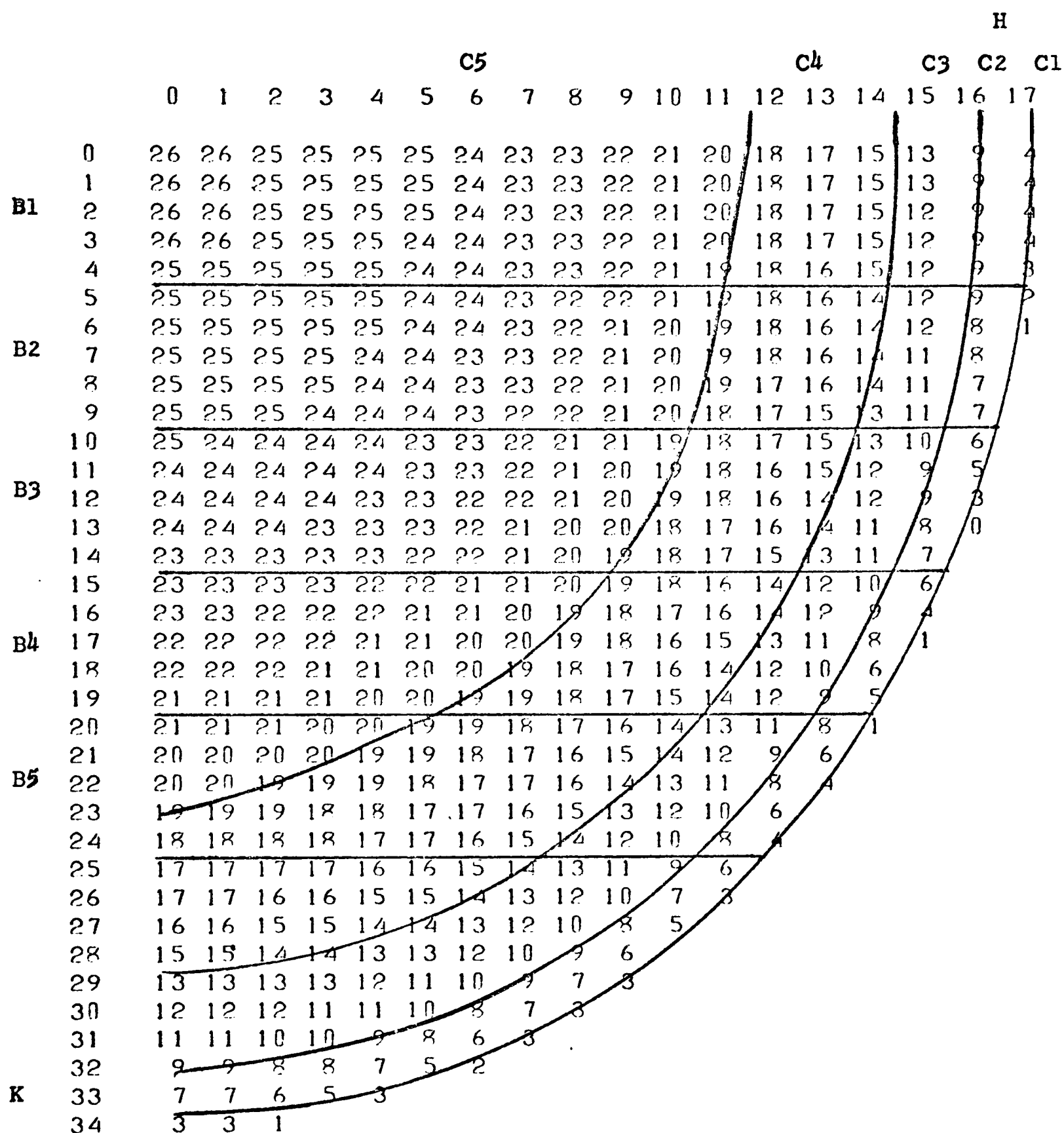


Fig. 27 The diagram, computer generated, shows the unique octant of of data for prealbumin (cell dimensions $a=43.3\text{\AA}$, $b=86.6\text{\AA}$ and $c=65.6\text{\AA}$.) The maximum value of L is given for each value of H and K to a resolution of $2.5\text{\AA}$. Also indicated are the sets of five levels that are measured simultaneously in the flat cone geometry. Groups of levels about b^* are shown as horizontal lines and groups about c^* as concentric circles. The diagram shows how a complete set of data may be collected from the b^* and c^* settings.

for various maximum values of $\delta\phi$. It can be seen that the longer b axis allows a closer approach to the blind region for a given maximum $\delta\phi$ value than does the c axis. Evidently the b^* mounting of a prealbumin crystal would allow the most data to be collected from one setting.

Figure 27 shows the unique reflections from a prealbumin crystal to $2.5 \text{ \AA}$ resolution. For a given value of H and K in the plane of the paper, the maximum value of L within the $2.5 \text{ \AA}$ spherical resolution limit is shown. Horizontal lines show the data quintuplets that can be collected in the b^* setting. Each group of 5 levels is termed B 1, B 2, etc. Concentric circles show the data quintuplets that are accessible in the c^* setting, termed C 1, C 2, etc. No indication of the blind regions are given. From this diagram it can be seen that B 1 to B 5 and C 1 to C 3, or B 1 to B 4 and C 1 to C 4 must be collected for a complete data set. Both these data collection schemes were used. In fact for the derivatives PAQ and PAQUN levels C 3 were not collected. The missing data for PAQ only were collected at a later stage.

On the five-circle machine, more than 80% of the data can be collected in the b^* setting with $\delta\phi$ and $\delta\tau$ tolerances of 0.4° and 0.1° between centre and outer reflections. In the case of the only set of data collected on this machine, the blind region was collected in the symmetrical A setting in a single counter mode. Reflections which had been collected in both modes were then used to scale the sets together.

Experimental procedure.

All the crystals were mounted in thin-walled glass capillaries of 1.5 mm internal diameter. The size of the crystal used was fairly constant, the greatest dimension (b^* direction) being about 1.6 mm which was usually reduced by removing the points of the crystal

as shown in Table 6. The other dimensions had maximum values of 1 mm and 0.4 mm.

Crystals significantly smaller than these, gave very feeble diffraction from 3 to 2.5° Å spacings. This is due to the large axial ratio, which means that the crystal volume rapidly becomes small as the b^* direction is reduced. Crystals of this size typically gave maximum count-rates of 1500 - 2000 counts/second from 2.8° Å spacings in the b^* setting. In the c^* setting, the intrinsic absorption is comparatively high and the maximum count-rates are half these. In most cases the maximum count-rates from derivative crystals were smaller than these values, since the addition of heavy atoms tends to increase the overall temperature factor of the protein.

In addition to this, the mosaic spread of the crystals is rather high. Unsplit peak widths measured on the three diffractometers in this laboratory vary from 0.5 to 0.8° . This is the angle through which the crystal must be rotated to record 90% of the peak. Peak broadening is also a function of heavy-metal replacement, which nearly always increases the peak width. Some loss of crystallinity may be due to the age of the crystals used (six years) although a more recently crystallised batch also had large peak-widths.

The c^* mounting is unstable, because the plate-like crystal is across the diameter the tube. Some crystal slip was noticed and corrected for while the c^* levels were being collected on the linear diffractometer. But this mounting led to great difficulties on the five-circle diffractometer, where the rapid movement of the circles usually disturbed the crystal alignment after about 200 measurements, so that re-alignments had to be constantly made. Unfortunately, the short a^* axis (43° Å) precluded data collection by the quasi-simultaneous flat-cone method.

These factors together make the extension of data collection

beyond $2.5 \text{ \AA}$ unlikely and the value of the current diffractometer data in the region $2.8 - 2.5 \text{ \AA}$, rather questionable. Photographic methods would undoubtedly produce better results in this region.

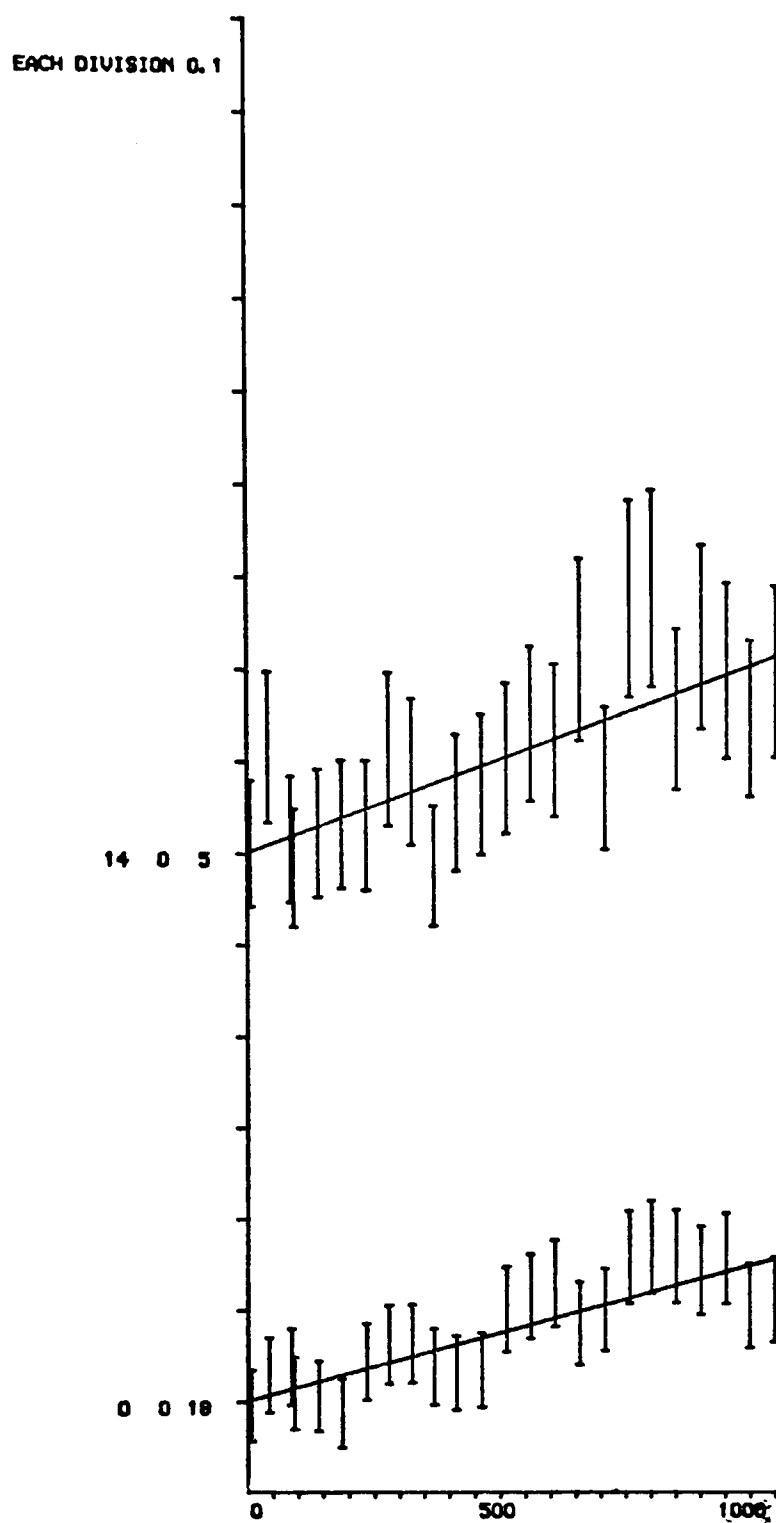
Alignment and measurement of cell dimensions.

All the diffractometers were operated with one reciprocal axis co-incident with the machine ϕ axis. As far as possible (by examination of crystal morphology), the remaining 2 axes of the crystal were constrained to be parallel to the orthogonal arcs of the goniometer head. In the equi-inclination setting, a low angle axial reflection was found and the mean value of θ , the Bragg angle, was determined. The arcs were set so that such axial reflecting planes were normal to the mounting axis. At this stage, the variation of intensity I_{hkl} of a strong axial reflection was recorded for values of ϕ at intervals of 10° or 15° . When normalized by a factor

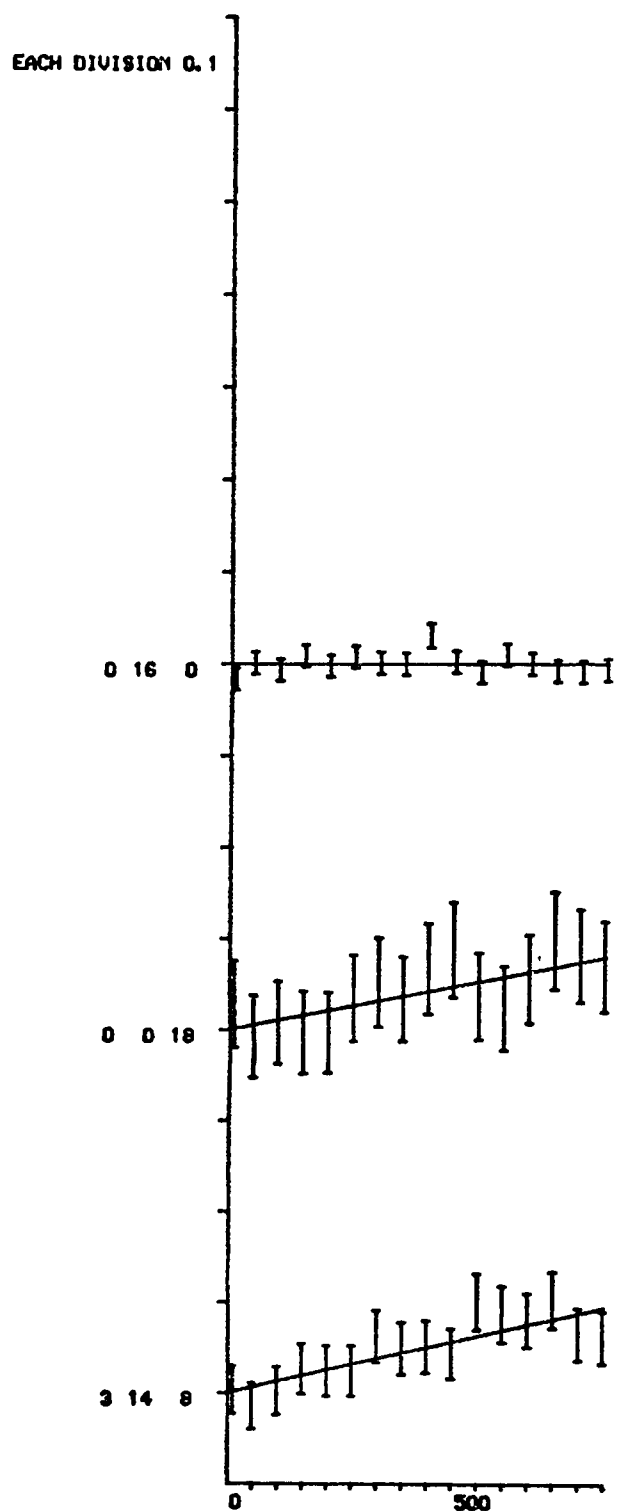
$$\frac{I(\phi)_{\max}}{I(\phi)}$$

these values of absorbance were used to calculate absorption corrections at the data reduction stage. In the B^* setting of prealbumin the absorption correction was usually 1.9 - 2.5 at the position of maximum absorption. The positions of the axes normal to the mounting axis were determined by scanning 2θ (the detector angle) and ϕ , close to the estimated positions of a strong reflection. The zero value of ϕ for the absorption curve was taken as that angle where the axis one cyclically removed from the mounting axis, was tangential to the Ewald sphere, while the other axis lies within the reflecting sphere.

The absorption curve obtained in this way was used as described by North et al., (1968). The ϕ values of the incident and diffracted beam are determined for each (hkl) and the absorption correction appropriate to it is taken to be the mean of the values of the coefficients $A(\phi)$ at these points. North et al., critically discussed limitations of this method. It can be seen that this



0 0 18 slope = 0.0015
 14 0 5 slope = 0.0021



0 0 18 slope = 0.0011
 3 14 8 slope = 0.0012

Fig. 28 Calculation of radiation damage correction factors. A measurement of a strong reflection is made after every 100 normal measurements. The plots above show the ratio of the initial measurement to each subsequent measurement of the reference reflection. 28A. shows the similar fall in intensities for 2 high angle reflections in the PAUN data set. 28B. indicates the angular dependence of such radiation damage. Reflection 0 16 0 is much lower than the others in its θ value.

correction does not take into account the variation of absorption with θ_{inc} . This can be corrected for, to some extent, by measuring a family of absorption curves from axial reflections lying in the particular planes of reciprocal space being collected. In the flat-cone setting, the incident beam makes a very steep angle with the mounting axis ϕ , and generally no axial reflections of a comparable value of θ_{inc} can be measured in the equi-inclination setting.

The values of the unit cell translations a^* , b^* , and c^* were determined on the linear diffractometer by determining the mean $\bar{\phi}$ values at (ϕ and $\phi + 180^\circ$) for high angle axial reflections on either side of the main beam. Then

$$n d^* = 2 \sin [(\bar{\phi}_{left} - \bar{\phi}_{right})/2] \quad (2)$$

where n is the order of the reflection and d^* is the appropriate unit cell constant. On the five-circle diffractometer these units were found by scanning 2θ to find the value of θ_{ref} , when the value of ω was set to this and ϕ was scanned for maximum counts. Given these settings a new value was found for 2θ and the process repeated until no further improvement could be made. The cell dimensions of all the crystals measured were very closely similar. The observed range for each reciprocal unit cell translation is shown below.

$$\begin{aligned} a^* & 0.0356 - 0.0357 \text{ r.l.u.} \\ b^* & 0.0179 - 0.0180 \text{ r.l.u.} \\ c^* & 0.0234 - 0.0236 \text{ r.l.u.} \end{aligned}$$

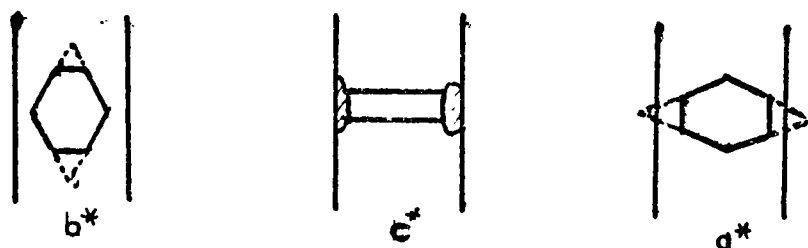
Radiation damage.

Prealbumin crystals are relatively stable towards radiation damage. The most severe damage occurred in crystals PAQDTT and PAUN, which amounted to 15 and 20% loss of initial intensity respectively, by the end of the run. These were the only crystals from which a complete

Setting parameters for the linear diffractometer.

Mounting axis.	Counter-crystal distance.		Aperture size.	Aperture spacing.	Absorption curve reflection
		mm.	mm.	mm.	
b*	0.0180	416.6	5	7.5	0 6 0 (1.2×10^6 c/s)
c*	0.0235	318.0	5	7.5	0 0 6 (8×10^4 c/s)

Crystal mountings for prealbumin.



Typical count - rates. c/s

0 6 0	1.2 x 10 ⁶
0 8 0	8.0 x 10 ⁵
0 16 0	4.0 x 10 ⁵
0 28 0	2.0 x 10 ³

Settings used for 3D data collection in the flat-cone geometry.

Name of level grouping.	Index of centre counter.	$\bar{\zeta}$	$\bar{\mu}^0$	Number of reflections.	Reference reflections
B1	2	0.036	2 4	3800	0 0 8-0 4 8
B2	7	0.126	7 15	3650	0 0 14-0 4 14
B3	12	0.216	12 28	3500	8 0 0-8 4 0
B4	17	0.306	17 49	3300	
B5	22	0.396	23 18	3200	
C1	2	0.047	2 40	4203	10 0 0-10 0 4
C2	7	0.164	9 45	3900	0 16 0-0 16 4
C3	12	0.282	16 27	3800	
C4	17	0.399	23 26	3000	

Setting parameters for the five-circle diff ractometer.

Mounting axis.	Counter-crystal distance.		Aperture size.	Aperture spacing.	Incident beam collimator.
b *	0.0180	558.6	5.5	9.5	1.4
c *	0.0234	425.0	5.5	9.5	1.4

Data shell	2.5 $\frac{\circ}{\text{\AA}}$ 2.8	2.8 $\frac{\circ}{\text{\AA}}$ 4.0	4.0 $\frac{\circ}{\text{\AA}}$ ∞
Maximum $\delta\phi$	0.4 $^\circ$	0.4 $^\circ$	0.4 $^\circ$
Maximum $\delta\gamma$	0.1 $^\circ$	0.1 $^\circ$	0.1 $^\circ$
Count time			
Peak -	50 secs	30 secs	30 secs
Background -	50 secs	30 secs	30 secs
Scan width	1.5 $^\circ$	1.5 $^\circ$	1.5 $^\circ$
Scan width at $\delta\phi_{\text{max}}$	3.0 $^\circ$	3.0 $^\circ$	3.0 $^\circ$

Table 6. Setting parameters for the linear and five-circle diffractometers.

2.5 Å⁰ data set was collected, all other sets used 2 or more crystals. In the case of PAQDTT, no damage correction was applied, as reference reflections checked after each set of levels were measured showed only a small fall-off during a run (3 500 reflections). The overall damage was accounted for when separate levels were scaled together at a later stage.

Data collection.

The instrument settings used for data collection from prealbumin crystals are summarised in table 6. Also shown are typical values of count-rates under the conditions given, and reference reflections used for alignment and checking the occupancy of heavy atom sites in derivative crystals. The collection of a complete data set for one derivative took between one and two weeks, including downtime and manual resetting of the crystal alignment.

In the case of PAUN, the data was collected in shells starting with the highest resolution data (2.5 - 2.9 Å⁰) and working inwards. Within each shell a radiation damage correction was applied as a function of elapsed data collection time. Figure 28 shows the decrease in intensity of reference reflections as a function of the number of reflections collected. However, the process of collecting in shells minimized the radiation damage, which is generally worst at high angles, within a given time.

No investigation of how to minimize the damage which does occur was carried out. This might involve choice of pH or salt conditions of the crystals. However, a cooled air device, which directs dry air at 0°C over the capillary containing the crystal, looks most hopeful. An HKO 2.3 Å⁰ precession picture of prealbumin taken at 0°C showed little change from the diffraction pattern normally observed at room temperature.

There were only small changes at low angle, which are to be expected

<u>No.</u>	<u>Quantity.</u>	<u>Expressions.</u>	
1.	Background-corrected count N_0	$N_0 = N - t_B/t_P(n_1+n_2)$	
2.	Standard deviation σ_{N_0}	$\sigma_{N_0} = \sqrt{N + t_B/t_P(n_1+n_2)}$	
3.	Background significance ratio R	$R = (n_1 - n_2) / \sqrt{(n_1 + n_2)}$	
	Setting:	<u>Flat-cone</u>	<u>Equi-inclination</u>
4.	Lorentz correction L	$L = 1/(\cos\mu \cdot \sin\Upsilon)$ $\sin\mu = -\zeta$ $\cos\Upsilon = \frac{2 - (\xi^2 + \zeta^2)}{2(1 - \zeta^2)^2}$	$L = 1/(\sin 2\theta)$
6.	Polarization correction P	$P = \frac{1}{2}(1 + \cos^2 2\theta)$	$P = \frac{1}{2}(1 + \cos^2 2\theta)$
7.	Absorption correction A(hkl)	$A(hkl) = (A(\phi_{inc}) + A(\phi_{ref}))/2$ $\phi_{inc} = (\phi_{hkl} - \epsilon_{hkl})$ $\phi_{ref} = (\phi_{hkl} + \delta_{hkl})$	
		<u>Flat - cone</u> $\tan\epsilon_{hkl} = (\xi^2 - \zeta^2) / (4\xi^2 - (\xi^2 + \zeta^2)^2)^{\frac{1}{2}}$ $\tan\delta_{hkl} = (\xi^2 - \zeta^2) / (4\xi^2 - (\xi^2 + \zeta^2)^2)^{\frac{1}{2}}$	
		<u>Equi-inclination</u> $\tan\epsilon_{hkl} = \tan\delta_{hkl} = \xi / (4 - (\xi^2 + \zeta^2))^{\frac{1}{2}}$	
8.	Structure amplitude and standard deviation F and σ_F	$F = 10 \cdot \sqrt{N_0^{corr}}$ $\sigma_F = 10 \cdot \sigma_{N_0} / \sqrt{N_0^{corr}}$	
9.	Form of data storage on magnetic tape (24 bit words)	Word.	Contents.
		1. hkl	packed 8-bit indices.
		2. I or F	intensity or amplitude
		3. σ_I or σ_F	standard deviation
		4. N	reflection counter

Table 7. Equations used in the primary data processing.

from ordering of the solvent molecules.

Stage 1. Primary processing of intensity data.

The reduction of photographic or counter data to give corrected intensities is expressed by the equation

$$I(hkl) = 1/(A.L.p.) * N_o(hkl) \quad (3)$$

where N_o is an optical density measurement or pulse count, corrected for background variation. L is the Lorentz factor, which corrects for the variation in time spent by lattice points in the reflecting condition; p is the polarization factor, which allows for the dependence of the scattered amplitude on the orientation of the electric vector of the X-ray beam. A is an absorption correction which corrects for the variation in pathlength through the crystal for individual diffracted rays.

Diffractometer data was processed by a programme written by Dr.A.C.T.North (1964). Modifications of this programme exist for flat-cone and equi-inclination geometries and several modes of data collection. Table (7) summarizes the most general forms of the expressions used in this programme to arrive at the corrected intensity I and its standard deviation σ_I .

The uncorrected intensity is derived from the background-corrected count N_o . On the linear diffractometer the counting sequence is fixed. A 15 second count is made with the crystal stationary on one side of the reflecting position. The crystal is then moved through the diffracting position over 30 seconds and a final stationary 15 second count is made of the background. Then

$t_B = t_p = 30$ in table 7 (1). On the four and five-circle diffractometers a wide scan is made through the peak and the counts k_i at any ordinate i are stored. The width of the peak is specified as t_p separate consecutive ordinates. The background width is then $t_B = \text{TOTAL} - t_p$. The position of the consecutive t_p peak ordinates

is chosen automatically to give the greatest value of N , the total peak count. The remaining ordinates are summed to give the background count ($n_1 + n_2$). The background-corrected count is then obtained as before from (1). This method has the advantage over the background-peak-background mode, of allowing for a small amount of crystal slippage during the run.

The standard deviation σ_{N_0} is derived in (2) in table 7 assuming Poisson statistics. The data is monitored for four classes of reflections.

<u>Class</u>	<u>Comment</u>
(1) $N_0 < \sigma_{N_0} \times \text{NSD}$	NSD was set to 1 for all prealbumin data.
(2) $N_0 < 1$	In this case N_0 is set to zero and σ_{N_0} to 1 to indicate a valid measurement.
(3) $3 < R < 10$	The value of R is calculated as shown in (3) in table 7 and indicates if the backgrounds are as close as might be expected from counting statistics.
(4) $R > 10$	

If condition (3) applied the background was taken as twice the value of the lower of n_1 or n_2 . If condition (4) applied the reflection was rejected. The Lorentz and polarization factors appropriate to the particular setting are applied as shown in 4 and 6 in table (7).

An absorption correction measured as described earlier was applied using the method of North and Phillips (1968). The forms of the correction appropriate to the settings used are summarized in (7) in table (7).

The programme finally writes either corrected intensities or structure factors derived as in (8) in table (7) to magnetic tape. These are stored sequentially as blocks of 200 words, where 10 words describe a reflection as outlined in table 7 (9). With the data processing programmes for the four- and five-circle machines a

Compound.	Crystal.	Mounting and levels	N	$N_0 < \sigma_{N_0}$ %	-N ₀ %	BG %	A _{max}
Native	PAN2	B* 0-24	15 949	17	7	0.7	2.1
	PAQDTT1	B* 0-14	10 204	15	6	0.2	2.0
	PAQDTT1	C* 0-19	14 000	20	11	0.2	1.9
	PAN3	C* 0-9	4 203	15	8	2.0	1.8(E)
	PAN4	C* 0-9	7 800	20	10	0.6	1.6
Mercury	PAQ2	B* 0-24	16 969	13	4	0.5	1.4
	PAQ3	C* 0-9	8 000	15	7	0.9	1.2
	PAQ7	C* 0-14	12 000	10	5	0.1	1.9
Uranium	PAQUN5	B* 0-24	15 458	19	7	0.5	2.6
+ Mercury	PAQUN6	C* 0-9	7 685	21	10	0.6	1.2
Uranium	PAUN1	B* 2.5-2.9 Å	5432	19	18	5.0	2.0
	PAUN1	B* 2.9-4.0 Å	6136	12	10	0.8	2.0
	PAUN1	B* 4.0-∞ Å	3745	6	5	0.4	2.0
	PAUN1	B* blind region	1141	6	4	0.4	2.0
	PAUN2	C* scaling data	2334	10	9	0.8	1.2
Mercury	PAQNEW1	B* 3.0-4.0 Å	5755	3	-	-	2.0
	PAQNEW2	B* 4.0-6.0 Å	4349	1	-	-	1.7
	PAQNEW2	B* 6.0-∞ Å	1357	6	-	-	2.0

N - total no. of reflections in a data set.
 $N_0 < \sigma_{N_0}$ - no. of reflections less than their standard deviation.
-N₀ - no. of reflections with negative intensity.
BG - no. of reflections with significantly unequal backgrounds.

Table 8. Statistical quantities from the prealbumin native and derivative data sets after the primary data processing stage. E = an estimated quantity.

Data for PAQNEW were collected on the 40 diffractometer using an ordinate analysis technique which always returns a positive or zero measurement of intensity. For this reason N- and BG are not quoted for this derivative.

reflection counter is included in the output, which identifies the sequence in which data has been measured. Then if it is desired to apply a radiation damage correction as a function of data collection time, this can be carried out routinely.

Table 8 summarizes the statistical observations from the most complete data sets measured on the linear and five-circle diffractometers. Any crystals which had more than 21% of their background corrected counts (N_0) less than the associated standard deviations were rejected. This happened with several of the double derivatives eg. (PAQUN) which tended to diffract very poorly. A general correlation was noted from the data, that the percentage of negative background-corrected counts (N_-) was about half those with N_0 less than σ_{N_0} . This observation was also used to achieve an acceptable data set. The native crystal PAN 3 was damaged before a proper absorption curve could be measured. In this case an 'empirical' absorption curve was obtained by comparing the intensities obtained from the zero-level set $C = 0$ with those measured from an HKO 2.8 Å precession photograph. The absorption factor varies much more slowly and evenly across the film in the precession geometry. The plot of $\bar{I}_{\text{counter}} / \bar{I}_{\text{photo}}$ in a series of overlapping ranges of ϕ where

$$\tan \phi = ha^* / kb^*$$

gave an absorption curve very similar to that usually obtained in the c^* setting. However no data from the crystal was subsequently used, due to an unacceptably high number of unequal background counts.

Stage 2. Collation of equivalent reflections.

All subsequent stages of data processing to a complete set of structure factors, refer to the flow chart shown in fig.29. The magnetic tape output from the primary data reduction programme contained a mixed sequence of reflections (HKL) and ($\bar{H}$ KL) from the linear or ($\bar{H}$ $\bar{K}$ L) from the five-circle diffractometer. These were

Stage NO.

Program

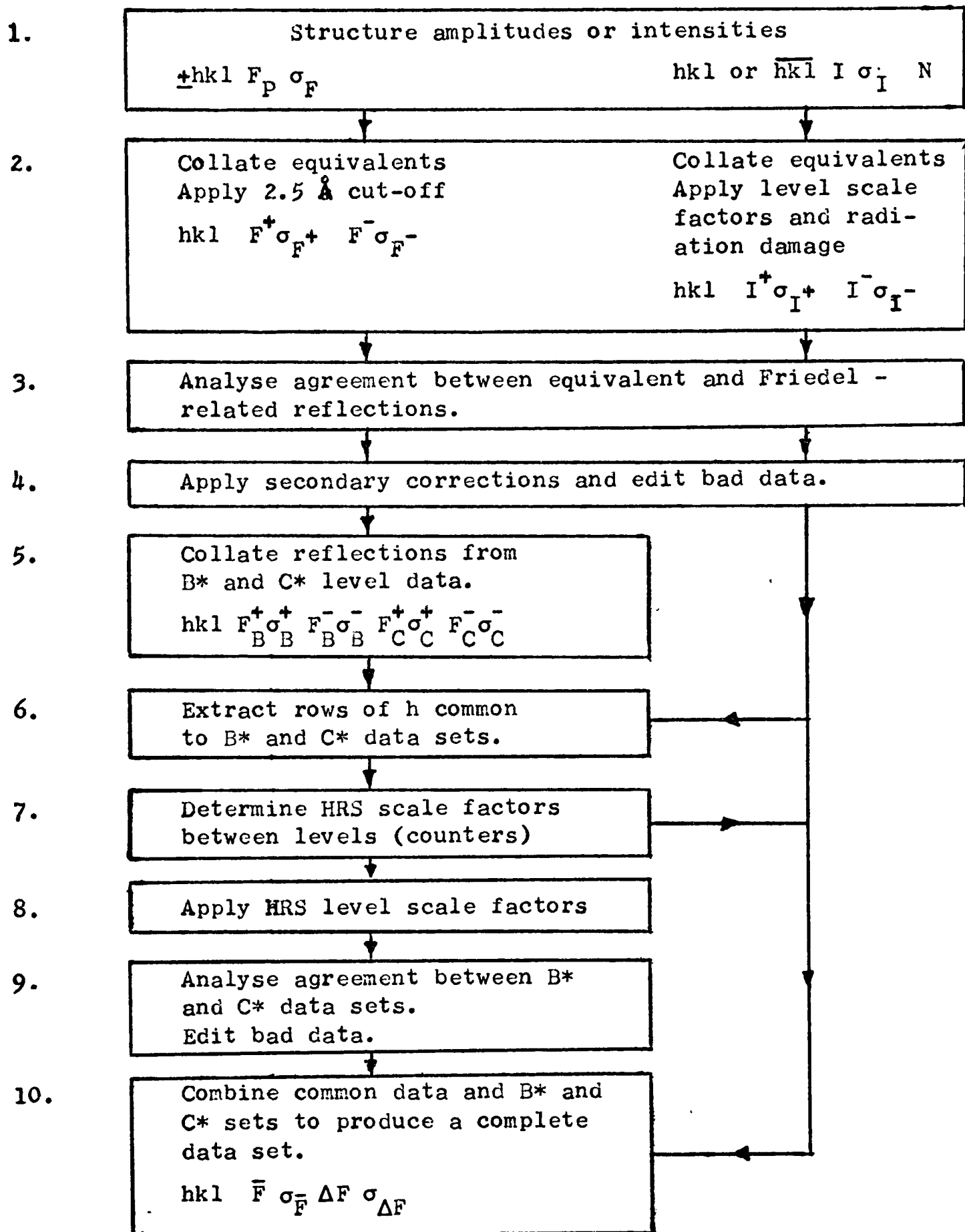


Fig. 29 Flow chart to indicate the sequence of secondary data processing to produce a complete set of data (the unique octant of orthorhombic reciprocal space plus the friedel related octant). The sequence is shown for both the 5-circle and linear diffractometer data. F^+, F^- are a Friedel pair and σ_+, σ_- their standard deviations B ,C implies data collected in either the B* or C* mounting on the diffractometers. HRS = Hamilton, Rollet and Sparks.

collated for the linear data and any that fell outside the $2.5 \text{ \AA}^0$ data limit were rejected. At the same time, a list of missing reflections was output. The five-circle data was treated more completely at this stage. A radiation damage correction was first applied, using the form

$$I^{\text{corr}} = I^{\text{obs}}(1 + N * \text{SLOPE}) \quad (4)$$

where SLOPE is determined for a given data shell as shown in fig.28 and N is the reflection sequence number. Four scale factors relating the counter data from counters 1, 2, 4 and 5 to the central counter 3 were applied. These were obtained by recording the same reflection in each of the 5 counters for about 20 strong reflections, measured in the flat-cone setting.

Counter-counter scale factors for intensities.

PAUN	Counter no.	Scale factor
b* mounting	1	1.25
	2	1.10
	3	1.00
	4	1.02
	5	1.10

Such scale factors are adequate on the grounds of counting statistics, but over the entire data set, they may exhibit angle and time dependence, because of secondary absorption effects and changes in machine settings. Because of the possibility of these effects, the level-level scale factors were checked by common reflections measured on a c* mounted crystal (PAUN 2).

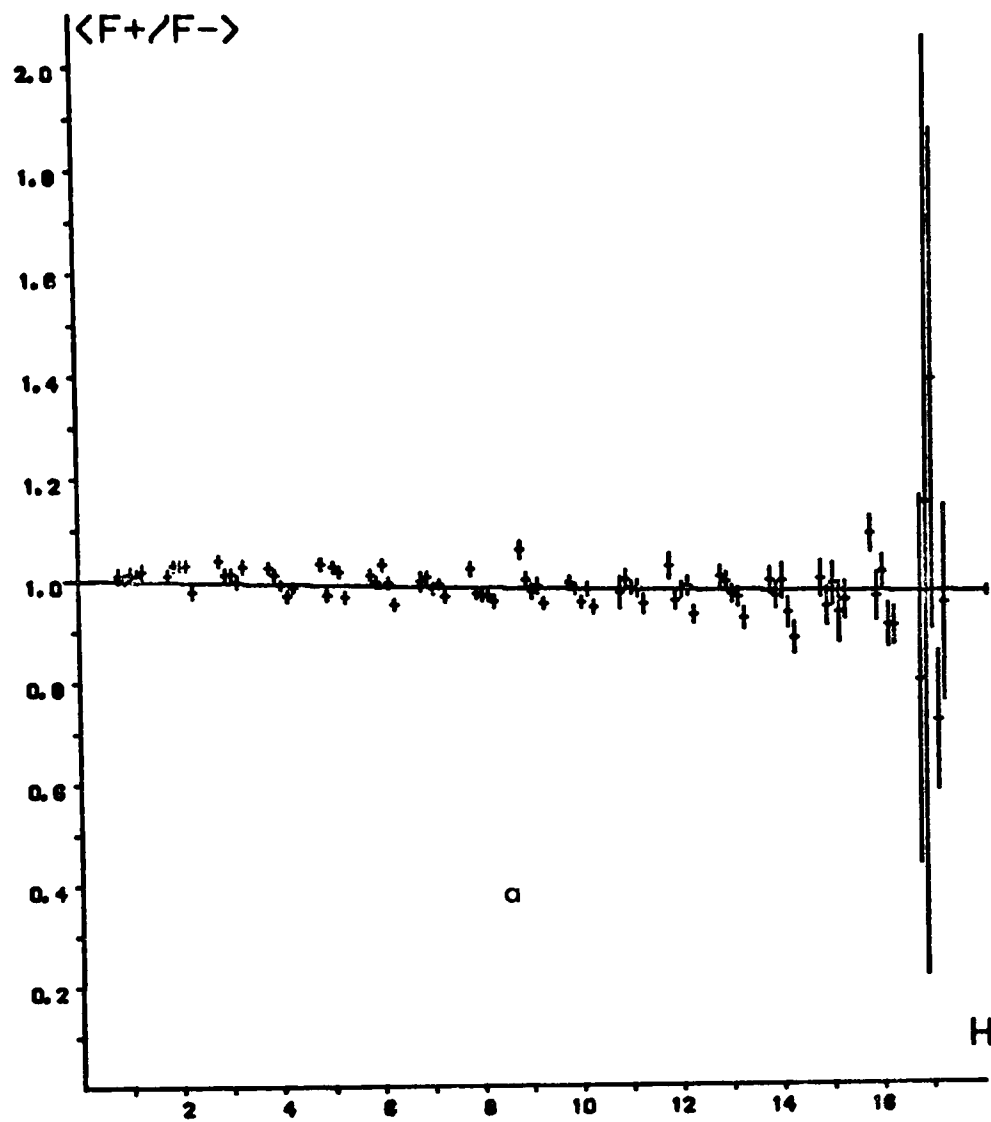
Stage 3. Analysis of agreement between Friedel-related reflections.

After the collation of Friedel pairs it is desirable to measure the extent of the agreement between these. This analysis is the basis of a thorough internal consistency check on the data. There are two reasons for this check. Firstly, the diffractometers have

P. A. 2.5Å. FRIEDEL PAIR ANALYSIS

LEVELS K=0, 1, 2, 3, 4 CRYSTAL PAQ2

RATIO 1.0000(0 0 0) -0.0015(008)H



P. A. 2.5Å. FRIEDEL PAIR ANALYSIS

LEVELS K=5, 6, 7, 8, 9 CRYSTAL PAQ2

RATIO 1.0000(0 0 0) -0.0010(011)H

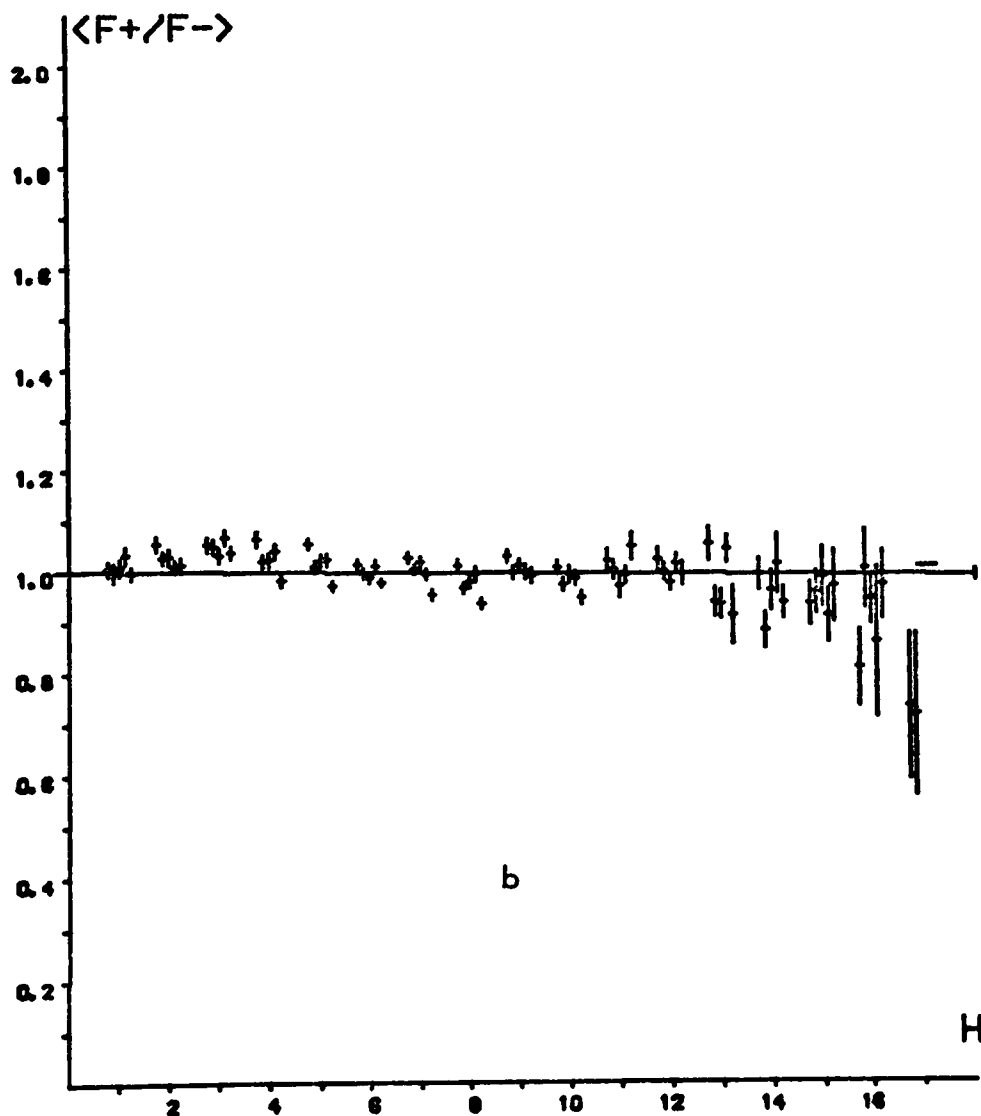


Fig. 30 The mean value of the ratio (F^+/F^-) plotted as a function of the index h for levels B1 and B2 of crystal PAQ2. The index k of the 5 counters increases by 1 from left to right.

been subject to a good deal of electrical noise, especially the linear and this check should reveal such noise. Secondly, it is desirable that the equivalence should be of as high a degree as possible for the significant measurement of anomalous scattering effects. It is clear that $\sum I(hkl) = \sum I(\bar{h}\bar{k}\bar{l})$, and that this equivalence should hold even if only a small number of reflections are included in the summation. This is due to the extremely small magnitude of the anomalous scattering, which causes a breakdown in the above equality for individual reflections. Accordingly, several forms of analysis were used which monitored the ratio $\langle I^+ \rangle / \langle I^- \rangle$ to look for systematic discrepancies in the equivalence.

Levels of reflections from the linear diffractometer containing either 700 reflections (B^*) or 900 reflections (C^*) were split into blocks of constant value of one of the stepping index (H) or the scanning index (K or L). Figure 30 shows two typical plots of the weighted mean of the ratio F^+/F^- as a function of the stepping index H for intensity data from a PAQ crystal. There are five points for each value of H, representing in order from left to right, levels 0.1.2.3.4. (fig.30a) and 5.6.7.8.9. (fig.30b). The vertical line through each point represents the standard deviation of $\langle F^+/F^- \rangle$. At high values of H this becomes very large as each value of F becomes small with respect to its own standard deviation. Individual points falling significantly off the least-squares line, such as those from counter 5 in (30b) for H = 4, 5, 7 and 8 led in several instances to the discovery of a faulty proportional counter. In these cases, the appropriate part of the run was rejected.

The reciprocal lattice rows $17\ 0\ L$ and $\bar{17}\ 0\ L$ were collected about 17 hours apart on the linear diffractometer. Any radiation damage or movement of the crystal will adversely affect the agreement of equivalent reflections in this mode of collection. More

significantly, the absorption correction $A(hkl)$ is not in general identical to $A(\bar{h}kl)$, as discussed by North et al., (1968) because of the asymmetric nature of the mounting of the protein crystal. These effects cause the plot of $\langle F^+/F^- \rangle$ to deviate from 1.0 at high values of H . The slopes of these lines were used to correct the values of F^+ and F^- so that the structure amplitudes in the more highly absorbed octant were increased (stage 4).

Intensity data was collected from the five-circle diffractometer in such a way that equivalent reflections were measured at time intervals of not greater than 20 minutes. Also the equivalents measured were hkl and $\bar{h}k\bar{l}$. As the data was collected in shells of decreasing resolution, the agreement of equivalents could not be examined in the same way as that of linear data, nor would such an analysis be appropriate. The programme written analysed the ratio $\langle I^+/I^- \rangle$ in blocks of reciprocal space of about 50 reflections each. This is similar to an idea suggested by Tsernoglou et al., (1972). The blocks were determined by taking the five adjacent reflections in the mounting axis direction for 3 orders of H between the resolution limits of the particular shell being measured.

The correction factor applied to the data was k such that

$$I^{+corr} = 1/\sqrt{k} \cdot I^+$$

$$I^{-corr} = \sqrt{k} \cdot I^-$$

where $k = \langle w \cdot I^+ / w \cdot I^- \rangle$ (5)

Mr. P. R. Evans has implemented a programme on the University 1906A computer, which analyses the agreement of I^+ and I^- in ranges of ϕ and ξ for specific groups of levels about the mounting axis. Because the primary absorption correction is applied as a function of ϕ and may be inadequate for reasons already discussed at high values of ξ or ξ , this would be the preferred method for any future work.

Table 9 summarizes the results of the agreement analyses. Factors

Compound.	Crystal.	Levels.	Noref.	N _B	R _B %	N _A	R _A %	$\overline{\Delta F}$	$\overline{F}$
Native.	1AQDTT1	C1	2015	14	12	-	-	22	173
		C2	1889	41	15	-	-	27	166
		C3	1580	40	16	-	-	26	153
		C4	1100	50	20	-	-	25	124
	PAQDTT1	B1	1362	51 *	7	-	-	13	176
		B2	1556	63 *	9	-	-	15	166
		B3	1414	41 *	9	-	-	14	157
Mercury	PAQ2	B1	1782	30	9	28	8	12	147
		B2	1669	26	10	26	8	12	141
		B3	1515	26	14	16	10	12	119
		B4	1295	12	12	11	10	11	110
		B5	979	10	11	12	10	11	100
	PAQ7	C1	2324	70	12	65	9	22	252
		C2	1923	80	15	65	11	30	252
Uranium + Mercury	PAQUN5	B1	1706	58	12	-	10	13	123
		B2	1610	60	13	-	11	14	122
		B3	1110	70	14	-	11	14	120
		B4	1181	32	14	-	14	14	98
		B5	952	32	17	-	17	14	81
	PAQUN6	C1	2204	52	16	35	13	12	93
		C2	1900	35	14	34	13	13	90
Uranium	2.5 - 2.9 Å.		5400	10	10	-	8	-	90
	2.9 - 4.0 Å.		6136	15	7	-	5	-	120
	4.0 - 40 Å.		3745	20	11	-	9	-	200
	Blind region.		1411	10	5	-	4	-	120
	PAUN2 C*		2334	20	6	-	5	-	100
Mercury	PAQNEW								
	3.0 - 4.0 Å.		5749	0	6.7	0	5.5	-	145
	4.0 - 6.0 Å.		4200	5	3.2	-	-	-	203
	6.0 - 40 Å.		1320	0	4.7	-	-	-	176

Table 9. For each crystal is shown the number of reflections measured in each set of five levels (B or C) or in the data shell between the resolution limits stated.

Also shown are the percentage agreement factors between Friedel-related reflections, calculated as explained in the text. N represents the number of Friedel pairs where the difference in amplitudes is greater than 3 or 2(*) times the associated standard deviation. Subscripts A and B refer to these quantities, where calculated, after and before any secondary absorption or scaling has been applied.

R_A and R_B are quoted for the quantity

$$\frac{\langle |F^+ - F^-| \rangle}{\langle F \rangle} \cdot 100 \%$$

before and after correction. These values refer to complete sets of five levels collected simultaneously in the flat-cone setting. N_A and N_B are the number of individual reflections for which

$$|F^+ - F^-| \geq 3\sigma_{\Delta F} \text{ or } 2\sigma_{\Delta F}$$

where $\Delta F = |F^+ - F^-|$

NOREF is the total number of reflections after rejection of bad data as defined above. $\langle F \rangle$ is the value of the mean crystal structure amplitude for the quintuplet levels.

Stage 5. Collation of equivalent reflections from two crystals.

Structure amplitudes from stage 4 of the data processing were collated to give a magnetic tape format as shown in fig. 29(5). F_B^+ , F_B^- is a Friedel pair measured from a b^* -mounted crystal and F_C^+ , F_C^- from a c^* -mounted crystal and $\sigma_{B,C}$ are the associated standard deviations. The magnetic tape contains the data sorted on L so that each file contains data of constant L, where L runs from 0 to 26.

The B^* and C^* data sets, besides possessing areas of reciprocal space not common to both, contained several thousand common reflections. These reflections were used to determine internal scale factors for each level of index K in the B^* set and L in the C^* set. The method described by Hamilton, Rollett and Sparks (1965) (HRS) was used. This method was applied using a Gauss-Seidel iteration to minimise the weighted sum for all observations of the residual :

$$R = (F_L^2 - G_L^2 \cdot \bar{F}_L^2)^2 \quad (6)$$

where F_L is the estimate of the value of a structure amplitude in a level L, $\bar{F}_L$ is the best estimate of F_L from all the observations and G_L is the reciprocal of the scale factor k_L to be applied to that individual observation. Clearly the final magnitude of R is dependent

upon the magnitude of the included observations as well as the extent of the agreement between common reflections. I have defined a reliability factor R_{HRS} where

$$R_{HRS} = R / \langle I \rangle \quad (7)$$

where $\langle I \rangle$ is the mean intensity of the common data. This is quoted in table 10 and should be used for comparison between data sets rather than R .

The Argus Ferranti computer necessitated the choice of an iterative, rather than full-matrix method of solving the normal equations, because of the limited store size. In addition, the number of observations was reduced by considering the reflections in groups with constant values of k and l . The values of F_L in equation (6) were then represented by rows of individual reflections summed over h , and common to both sets. This is probably inferior to the method as applied to individual reflections, because it assumes identical intensity distributions in the a^* direction. Such an assumption is not necessarily valid because the common data has been collected in different crystal settings with correspondingly different primary and secondary absorption corrections and anisotropic radiation damage effects.

The averaging process reduced the number of observations to between 200 and 300, from which 35 level scale factors were determined; the common rows contained an average of 11 reflections. Bloomer (1972) discusses in detail the problem of weighting the individual reflections in common rows and whether intensities or structure amplitudes should be used. In this work it was found empirically that the use of structure amplitudes, while giving a lower residual and a more closely scaled data set, gave very poor agreement between the largest values of F in the two data sets. This observation is closely analogous to the results of scale factor determination between derivative and native sets.

The scale factor between separate experimental observations of structure amplitudes (or difference structure amplitudes) is not constant, but varies with the magnitude of F . Figure 32 expresses this result in graphical form. This phenomenon is at its worst in protein crystallography, where the intensity distribution is very steep, so that high resolution data sets often contain a very large number of weak measurements of F . The use of counter data worsens this effect, which can be improved by the use of photographic data at the same resolution. Blow and Crick (1959) mention this effect, but do not develop it to a discussion of its effect on scale factors.

Because the largest amplitudes are most important in phase determination, all subsequent scaling including HRS was carried out with intensity data. A more sophisticated analysis might lead to a set of scale factors appropriate to ranges of F possibly where breaks in the Wilson plots of the data occur (see fig.12).

The common row sums F_L in expression (6) were weighted by the function :

$$w_{HRS} = \frac{1}{\sigma_B^2} + \frac{1}{\sigma_C^2} \quad (8)$$

where σ_B , σ_C are the standard deviations of the C^* and B^* row sums respectively. These were calculated as shown in Appendix 1. If both measurements of a Friedel pair were present in the B and C data sets, they were included in the common row averages as two independent observations. If the difference between two equivalents in a common row from either data set exceeded 4 standard deviations, that reflection was excluded from both row sums. The scaling usually converged after 5 cycles of calculation of the G_L . Table 10 lists the HRS scale factors for three of the prealbumin data sets measured.

The scale factors reflect expected trends. They are generally larger for the C^* mounting, because this data is weakest in most cases. They tend to increase with increasing index K or L. This

CRYSTAL PAIRS.				
Levels	Index	PAQDTT1	PAQ2 + 7	PAQUN5 + 6
B1	0	0.9492	0.9889	1.0012
	1	0.9693	0.9649	0.9970
	2	0.9868	0.9565	1.0415
	3	0.9749	0.9093	1.0359
	4	0.9403	0.8807	0.9450
B2	5	0.9491	0.9801	0.9251
	6	0.9745	0.9687	0.9668
	7	1.0020	0.9819	1.0106
	8	1.0064	0.9502	0.9923
	9	0.9749	0.9147	0.9575
B3	10	0.9889	1.2125	0.9529
	11	1.0446	1.2636	0.9415
	12	1.0705	1.2602	0.9653
	13	1.0529	1.1354	0.9993
	14	1.0674	1.1190	0.9146
B4	15	-	1.2000	1.0549
	16	-	1.2280	1.0830
	17	-	1.2209	1.1261
	18	-	1.1162	1.1241
	19	-	1.1509	1.0352
B5	20	-	1.0863	1.1254
	21	-	1.1058	1.1461
	22	-	1.0676	1.2045
	23	-	1.0533	1.2357
	24	-	1.0373	1.2081
C1	0	1.0387	0.6323	1.3874
	1	1.0533	0.6465	1.3819
	2	1.0715	0.6524	1.3968
	3	1.0803	0.6485	1.3861
	4	1.0460	0.6418	1.2985
C2	5	1.0509	0.5853	1.2843
	6	1.0643	0.6191	1.3116
	7	1.1220	0.6331	1.3496
	8	1.1062	0.6252	1.3196
	9	1.0303	0.6282	1.2596
C3	10	1.0274	-	-
	11	1.0364	-	-
	12	1.0649	-	-
	13	1.0525	-	-
	14	1.0454	-	-
C4	15	0.9892	-	-
	16	0.9898	-	-
	17	1.0174	-	-
	18	1.0071	-	-
	19	0.9368	-	-
F		169	143	120
R		1423	2343	905
R(HRS)		0.5	1.1	0.7
N(common)		3074	2911	2793
N> σ		80	163	37
N _R		32	20	6
$\Delta F/F$ %		13	20	6

Table 10. H.R.S. scale factors and agreement of common reflections for three derivatives.

reflects the effect of radiation damage with time, since the collection of data started with the lowest indices of K or L. The mean values of the structure amplitudes are similar for each set of data, although the uranyl data is weakest at high diffraction angles. It should be noted that the lowest value of R_{HRS} was obtained from the PAQDTT 1 data where both B and C data sets were measured from the same crystal. The worst value of R_{HRS} came from the pair of crystals PAQ2 and PAQ3, (not shown). The data for PAQNEW, which extended to $3 \text{ \AA}$ only, was collected in three shells on the four-circle diffractometer and these were scaled together using the HRS method on common reflections between shells.

Stage 8. Application of HRS scale factors.

The scale factors listed in table 10 were applied to the appropriate levels of the data output from stage 5.

Stage 9. Analysis of the extent of agreement between reflections common to the two data sets.

The same programme calculated the values of

$$\Delta F_{CB} = \overline{F}_C - \overline{F}_B \quad (9)$$

$$\overline{F}_B, \overline{F}_C = \langle F_B^+, F_B^- \rangle, \langle F_C^+, F_C^- \rangle \quad (10)$$

$$\overline{F}^-, \overline{F}^+ = \langle F_C^+, F_B^+ \rangle, \langle F_C^-, F_B^- \rangle \quad (11)$$

$$\overline{F} = \langle \overline{F}^+, \overline{F}^- \rangle \quad (12)$$

$$\Delta F = \overline{F}^+ - \overline{F}^- \quad (13)$$

as shown in more detail in Appendix 1. The discrepancy index (9) was calculated for reflections common to both sets of data and if this value exceeded four times the associated standard deviation, the reflection was monitored. The number of these reflections for each data set is listed in table 10 as $N > \sigma$. A very large number of poor agreements (364) exist between the B^* and C^* sets from crystals PAQ2 and PAQ3. Two possible explanations for this were examined.

- (1). Error in data processing of one of the data sets.
- (2). Difference in mercury occupancy between the data sets.

To check possibility (1) the original intensities were examined. The two worst agreements between the B^* and C^* data sets were from reflections $4\ 18\ 0$ and $4\ 16\ 0$. Two measurements of these reflections (F^+ and F^-) existed in each of the separate data B^* and C^* data sets and were in agreement to within 2 and 3 standard deviations respectively.

The ratios I_{B^*}/I_{C^*} of the original intensities of groups of reflections such as $4\ 13\ 0$, $4\ 18\ 2$, $4\ 18\ 3$ and $4\ 13\ 4$ were calculated. These reflections occur very close together in reciprocal space and may confidently be expected to suffer similar corrections with respect to absorption, Lorentz and polarization factors. In each case, these ratios were similar, except for the reflections with high orders of disagreement between the two data sets, eliminating the possibility of processing error.

Some of the large disagreements occur in the centric data. To check possibility (2) the bad agreements were listed and compared with previous refinements at $3\ \overset{0}{\text{\AA}}$ of the PAQ derivative. A reflection was considered to be sensitive to replacement by mercury if a large value of F_{calc} was recorded for that reflection. Although $4\ 13\ 0$ changes by nearly 50% on saturation of the mercury sites of prealbumin, no very clear correlation between heavy-atom sensitivity of reflections and discrepancy between the b^* and c^* data sets could be found. When a series of PAQ crystals were examined on the diffractometer, some variability in these sensitive reflections was noticed. Since the PAQ2 data set was the most extensive (more than 75% of the data), it was decided to find a PAQ crystal that gave the most clearly similar intensities to that of PAQ2. The result of that search is the PAQ7 set, which, as may be seen from table 9, is an improvement but still contains a significant number of discrepancies. It has subsequently

emerged that a better mercury derivative can be obtained by rediffusion of mercuric acetate in to PAQ-DTT crystals (see chapter 3).

Apart from these specific differences in the PAQ crystals, the agreement between the other data sets was acceptable. At this stage a number of reflections from each data set were removed. Stage 3 had allowed an internal consistency check of equivalent reflections. Where no Friedel equivalent previously existed for a reflection, no test could be made of its significance. The presence of common reflections from B^* and C^* sets allowed some of these reflections to be checked. Several of these gave rise to discrepancies between the sets that could mainly be ascribed to faulty proportional counters on the linear diffractometer. Such noise pulses usually simultaneously affected all five scalers in the machine. It was then found that a discrepancy between equivalents existed for five adjacent reflections of constant H and K or H and L for C^* and B^* data respectively. This observation was often of use in deciding to omit data at this stage. The number of reflections rejected from the data at this stage shown in table 9 as N_R . Also shown are the mean values of

$$\frac{\Delta F}{\bar{F}} = \frac{|\bar{F}_B - \bar{F}_C|}{\bar{F}} \cdot 100\%$$

and the number of reflections common to both sets of data (N_{common})

Stage 10. Combination to complete data set.

After editing, the data were combined to produce a complete data set in the form shown in fig. 29(10). The values of $\bar{F}$ and ΔF were determined as indicated by equations (10-13) and appendix 2. Stages 6 and 7 were repeated to check that the HRS rescale factors were unity after editing. In no case were the scale factors for any data set significantly different from unity to warrant re-scaling. The four data sets used in the subsequent high resolution work are listed below with the number of unique data determined for each.

Stage No.

Programme (cont. from p. 63)

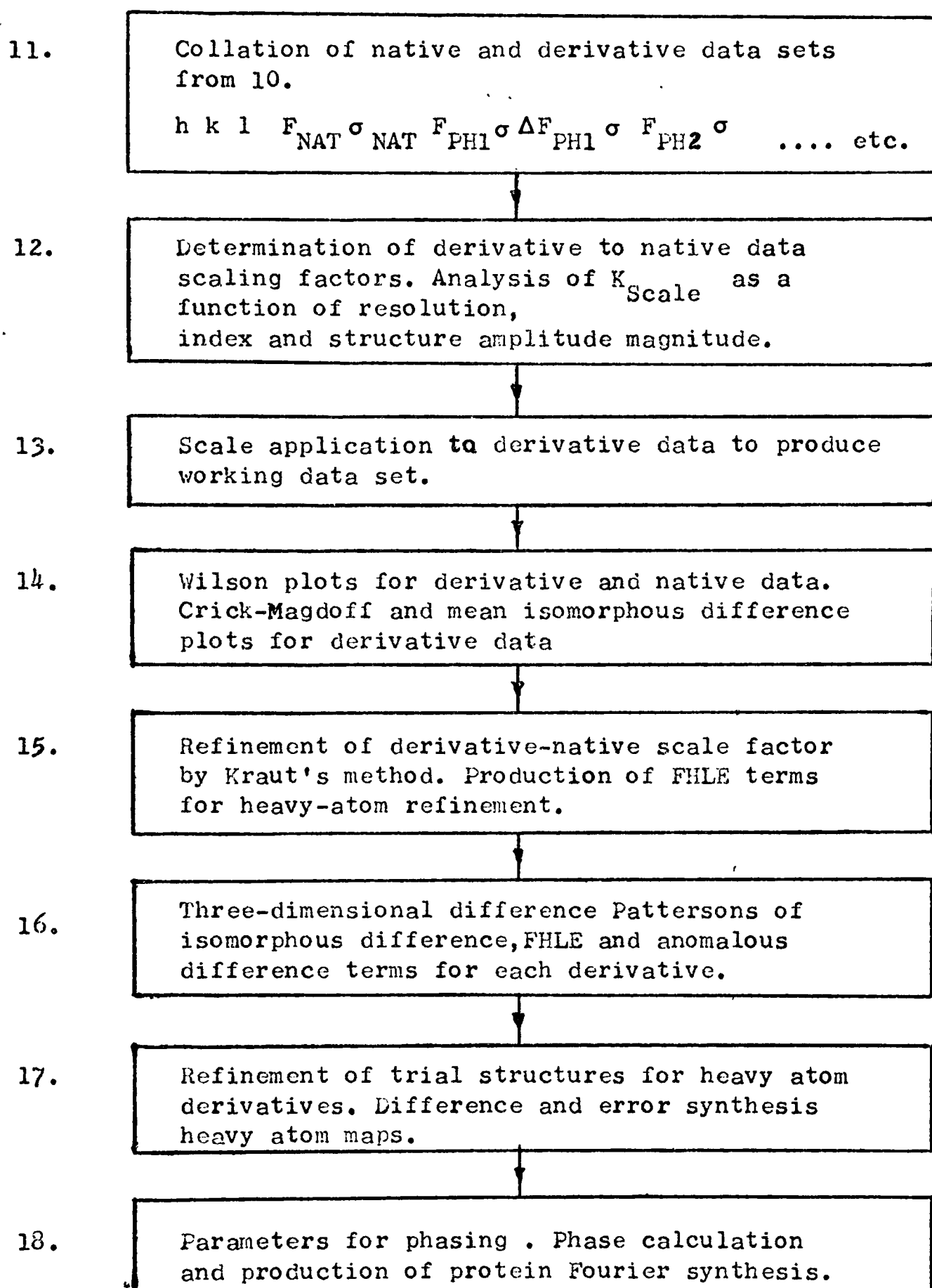


Fig. 31 Flow chart to indicate the final stages in the data processing to the production of the protein Fourier synthesis. Stages 15 to 18 are covered in chapter 5. The data processing in these last stages has been carried out on the Ferranti Argus computer using programmes written by the author for P2₁2₁2₁. The latest work in these final stages has been carried out on the 1906a computer using more general programmes.

Compound	Abbreviation	Number of reflections	$\bar{F}$
Native	PAQDTT	8789	170
Mercury	PAQ	8466	143
Uranium	PAUN	8000	95
Uranium + mercury	PAQUN	8045	120
Mercury	PAQNEW (to 3 Å)	5700	220

Stage 11. Collation of data sets.

Separate data sets from stage 10 of the data reduction were combined to give a magnetic tape of the form shown in fig. 31(11). The organisation of this data is again in files of constant L. When the University 1906A computer service became available, a copy of the data was transferred to this machine. The organisation of this data differed by the inclusion of the term $4 \sin^2 \theta / \lambda^2$ after the indices. The value of this quantity was slightly different for each derivative but that of the native protein was used to express this value in subsequent analyses.

Stage 12.. Primary scaling of derivative and native data.

The combined derivative and native data from stage 11 was used to calculate least squares regression lines of

$$K_I = \left[\frac{\sum I_{\text{NATIVE}}}{\sum I_{\text{DERIVATIVE}}} \right]_Q \quad (14)$$

where Q represents values of h, k, l, $(\bar{F}/20)$ and $(500 \cdot \sin^2 \theta)$.

These lines define a scale factor k_I such that $k_I = A \cdot Q + B$

where A and B are coefficients determined from the regression analysis.

Initial values of these coefficients are given in table 11. The

graphs in fig. 32 show the variation of k_Q of PAQ, PAUN and PAQUN as

a function of various Q. Also calculated were 3D scaling factors which

examined the value of equation (14) as a simultaneous function of ϕ , ζ

and ξ . This was intended to allow for possible systematic error

Scaling function.	Coefficient.	PAQDTT	PAQ	PAQUN	PAUN	PAQNEW
Q						
Index of h	A	-	-0.013 (3)	-0.0002 (50)	-0.023 (10)	-0.009 (10)
	B		1.660 (20)	2.128 (30)	3.32 (6)	2.045 (5)
Index of k	A		-0.003 (2)	0.017 (3)	0.054 (7)	0.004 (6)
	B		1.593 (20)	1.929 (30)	2.653 (70)	1.944 (12)
Index of l	A	-	-0.008 (3)	-0.0269 (5)	0.087 (9)	-0.026 (2)
	B	-	1.632 (30)	2.352 (50)	2.61 (7)	2.150 (6)
500.sin ² θ (S)	A	-	-3.36 (40)	-1.95 (101)	15.4 (18)	-3.800 (20)
	B	-	1.78 (3)	2.26 (7)	2.34 (3)	2.220 (3)
EXP(A.S + B) temperature factor	A	-	-2.176 (24)	-1.0 (5)	4.7 (6)	-1.86 (11)
	B	-	0.59 (2)	0.83 (3)	0.89 (4)	0.80 (2)
Overall scale factor	K	-	1.567	2.119	3.220	2.000
Mean intensity	I	41802	26674	18859	13521	21323

Table 11. Derivative to native scale factors as a function of various Q such as index or resolution.

The coefficients A and B are the slope and intercept, respectively of the function :

$$K_I = A.Q + B, \text{ where } K_I \text{ is a scale factor to be applied to the derivative intensity}$$

in order to place this on the same relative scale as the native intensities. The figures shown in parentheses are the standard deviations of the quantites (slope, etc.) shown.

In the final scaling, only temperature factors were used as scaling parameters.

introduced into the sets of structure amplitudes by the primary absorption correction. It can be seen that the greatest variation of scale factor against index is shown by PAUN. Fig. 32 shows the 3D-scale analysis for this derivative. A systematic difference appeared as a function of Φ . This data was the only set collected on the 5-circle diffractometer, but apart from this observation, no easy way to account for the Φ variation of scale factor can be proposed. A re-examination of the absorption correction for PAUN showed this to have been correctly applied. A Φ -dependent scale factor was applied to the PAUN data.

It will be noticed that although the PAQDTT (native) data has the largest mean intensity, the relative temperature factor of PAQ is negative. It is usually found that the derivatives have higher temperature factors than the native compound. However, in this case the native compound has been derived from the PAQ crystals by chemical treatment with di-thiothreitol. This treatment has evidently caused an increase in the temperature factor of PAQDTT relative to the untreated PAQ crystals.

The line of regression corresponding to temperature factor in table 11(5) was applied to the three derivative data sets. The re-scale factors were then examined by repeating the analysis. No significant departure from unity was observed in the plots of equation (14) for PAQ and PAQUN, except at low angles ($d > 20 \overset{0}{\text{\AA}}$). Examination of the data in this region showed several very large changes (1, 1, 0) etc., between the derivative and native sets. This effect is probably caused by salt or solvent changes between the crystals used. A Φ -dependent scale factor was then applied to the PAUN data, as already described. A final scale factor analysis showed no further trends of scale variation with the variable quantities shown in table 11.

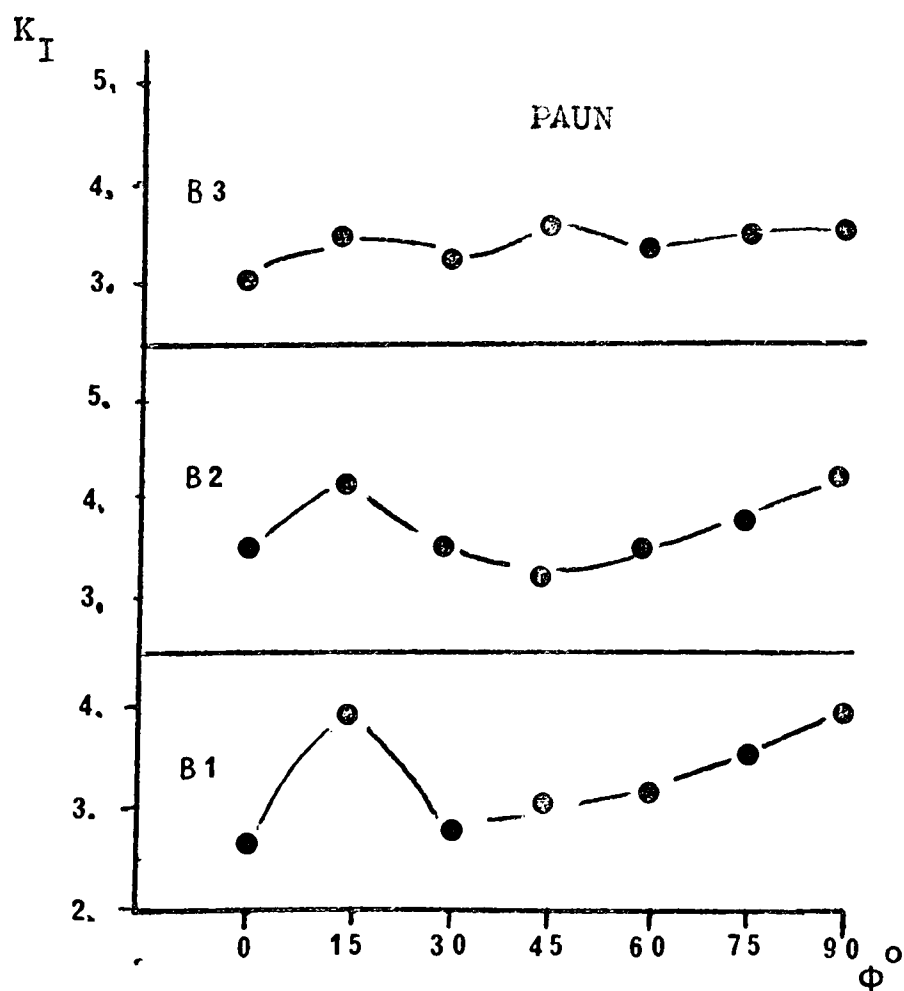
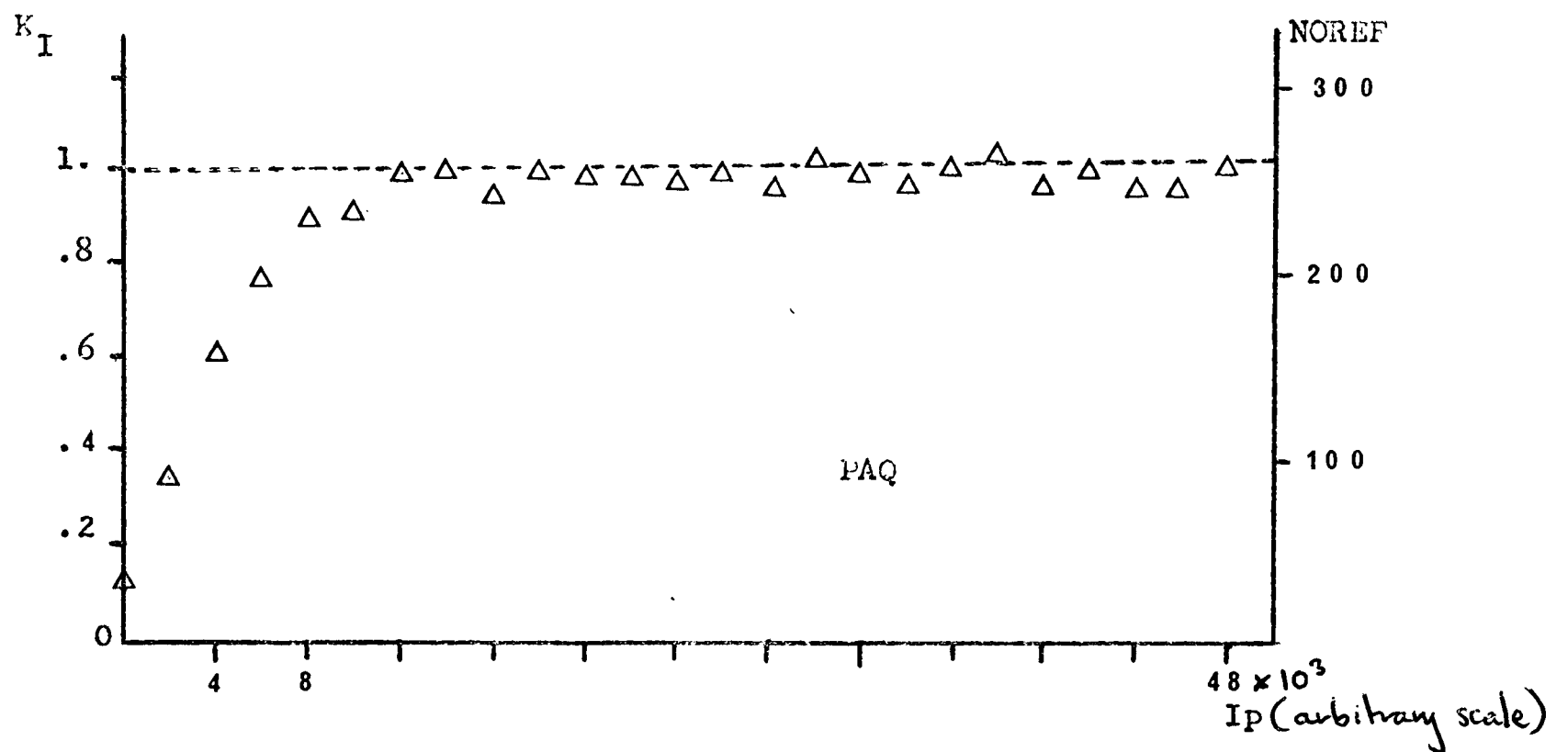
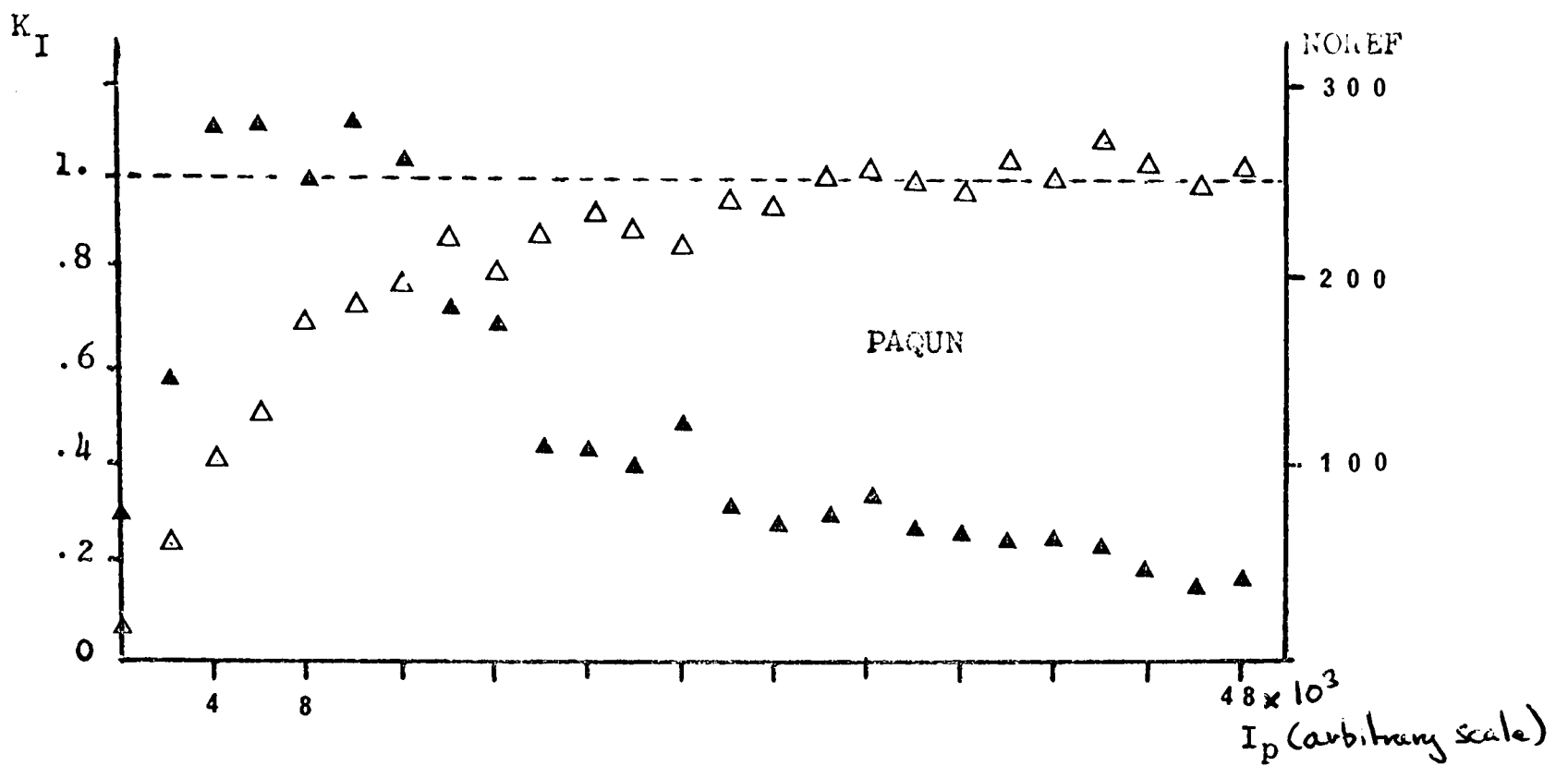


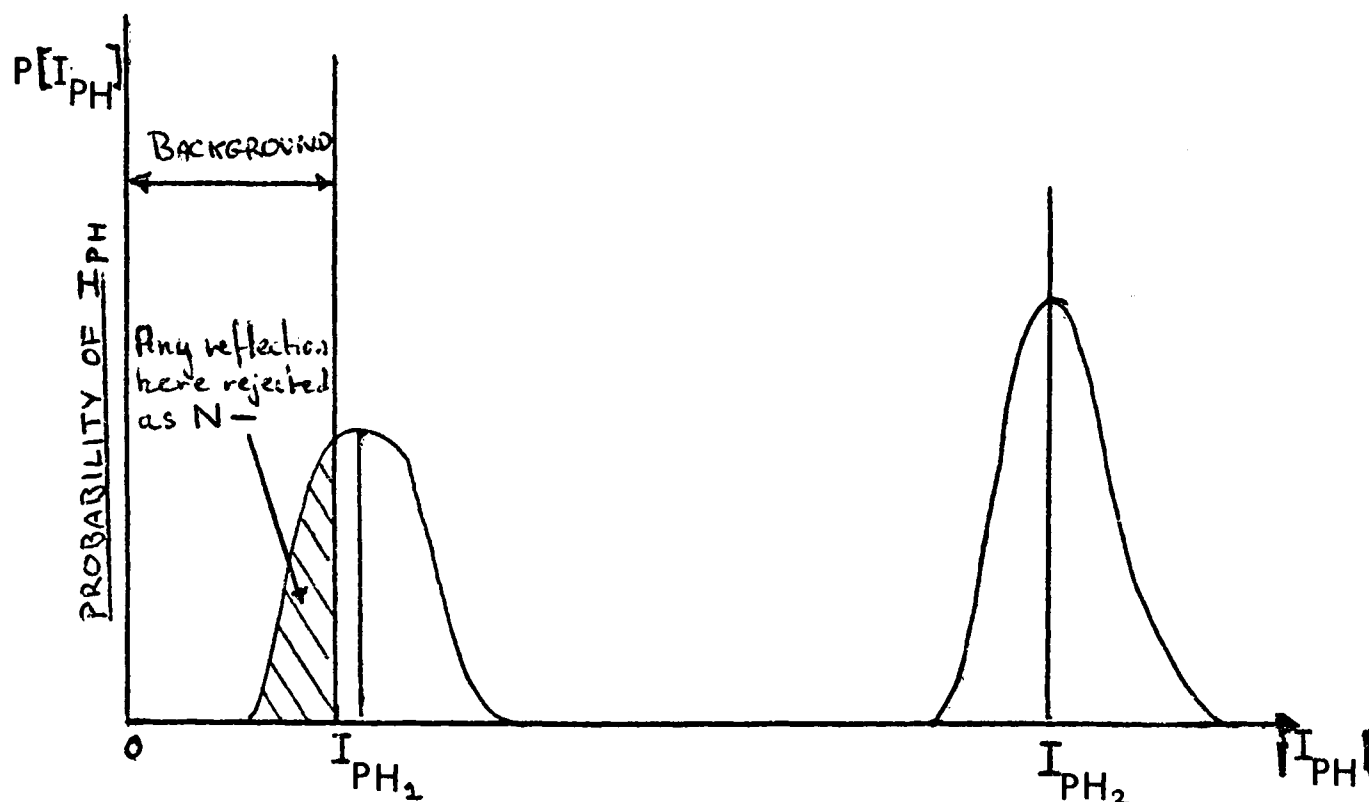
Fig. 32 Plots of derivative to native scale factor as a function of I_p (above) and ϕ & ζ (left). The much weaker PAQUN data shows a nearly continuous dependance on I_p . Filled triangles represent the number of derivative and native pairs contributing to the scale factors shown as open triangles.

Variation of derivative to native scale factor with $|F|$.

After scaling the derivative data to the native set, the residual variation of K_I with $|F_P|$ was analysed. These plots are shown for PAQ and PAQUN in fig. 32.

The plots are very non-linear at low to medium values of $|F_P|$. It appears that the scale factor of $|F_{PH}|$ is too high at these values of $|F_P|$, but is correct at medium to large values of $|F_P|$.

Two factors may explain this observation. The diagram below shows the probability distribution of $|I_{PH}|$ for two reflections, one very close to the background region and one very much above it.



The probability distributions are skew, because the counting statistics most closely follows the Poisson distribution where σ_I is a function of I such that $\sigma_I \propto \sqrt{I}$. Clearly, the probability distributions of reflections very near the background level of I_{obs} are very skewed indeed, since any observation falling below this level has been rejected (N- in table 3). For a given value of I_P in fig. 32, the corresponding value of I_{PH} is likely to be the same or higher because of the loss of weak data as N-. In this way the scale factor of I_{PH} is seen to be very large at low values of I_P , because a Gaussian distribution of intensities has been assumed in the analysis of the averages. However in the

analysis of PAUN scale factors k_I seems to be continuously dependent upon I_p . It is not too easy to see the reason for this, but the derivative to native scale factor was greatest for this derivative, and the analysis may reflect this fact. Previous workers have reported that difficulty was experienced with the refinement of derivative to native scale factors at a later stage in the crystal structure analysis (Bloomer, 1972), (Banyard, 1973). It is to be expected that as the data is initially constrained to obey equation (14), the introduction of further scattering $\underline{F}_H$ into the equation :

$$k \cdot \underline{F}_{PH} = \underline{F}_P + \underline{F}_H \quad (17)$$

would cause the value of k to be greater than unity by a factor that represents the mean value of

$$\frac{|\underline{F}_P| + |\underline{F}_H| \cdot \cos \Phi}{|\underline{F}_{PH}|} \quad (18)$$

where Φ is the angle between the vectors $\underline{F}_P$ and $\underline{F}_H$. However, as these workers report, the value of k defined by equation (17) frequently drops significantly below unity when allowed to refine. This effect has even been observed for centric data, where moduli replace the vectors in (17) and k should certainly be greater than 1.0. I propose as part explanation for this effect that it is a result due to scale factor dependence upon $|\underline{F}_{obs}|$. All the refinement procedures which are reported to have displayed this effect have used the measured amplitudes $|\underline{F}_{obs}|$, whilst the initial scaling took place using $|\underline{F}_{obs}|^2$. This scaling procedure is biased towards large values of $|\underline{F}_{obs}|$ which are in the minority (see fig. 32). An experiment with prealbumin data showed that the value of K_I determined for each derivative was significantly larger than the value of K_F , determined using structure amplitudes instead of intensities. Presumably the additional scattering introduced by $\underline{F}_H$ in 17 is not sufficient to offset the effect of the numerous weak terms whose K_I after scaling

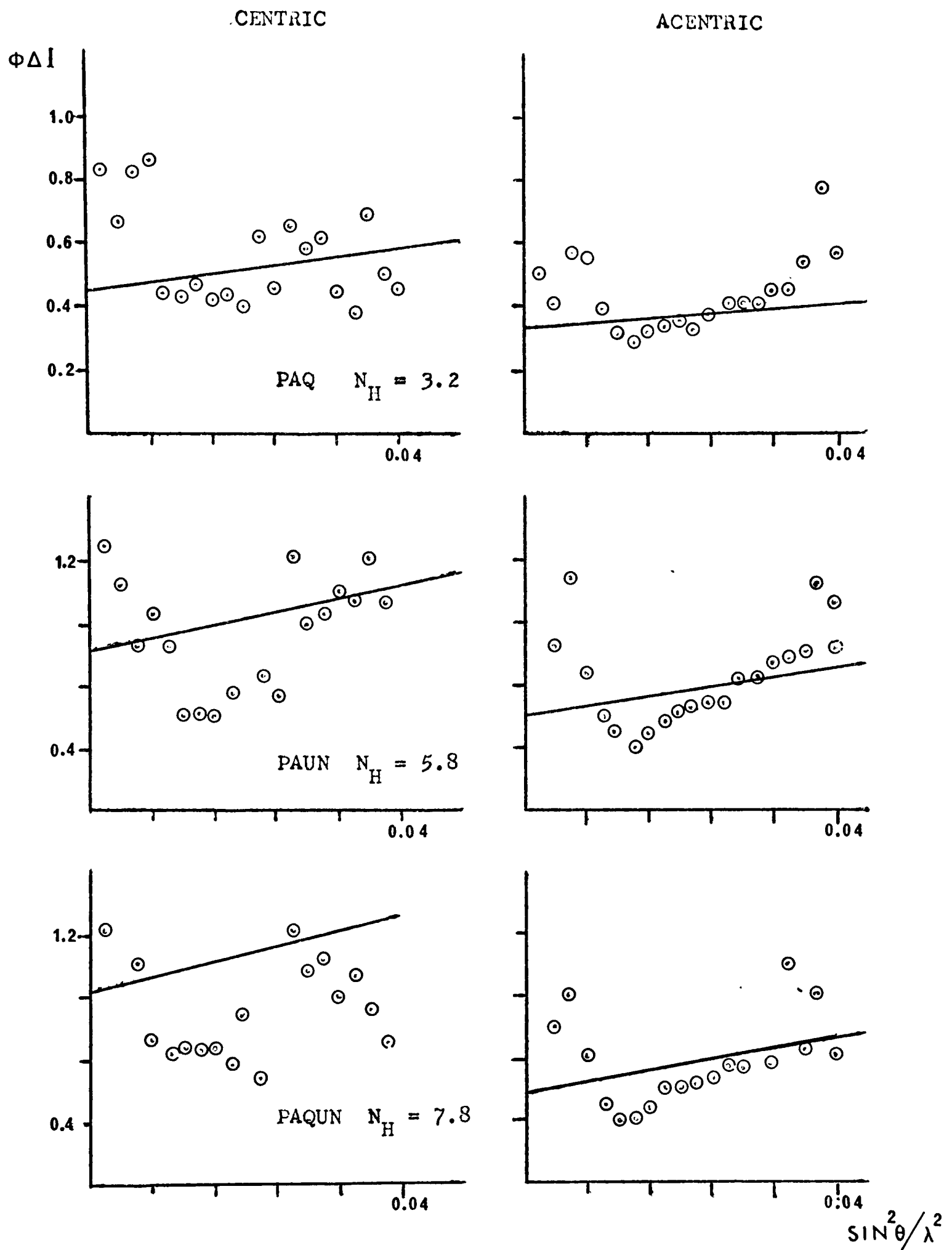


Fig. 33 Crick-Magdoff plots for the three derivatives used in the high resolution work. Observed (⊙) and calculated (—) values of equation 3.14 are shown for acentric and centric data. $N_p = 7714$ and N_H is as shown. The greater scatter of the points in the centric plots presumably reflects the smaller number of terms.

is less than unity. A second factor proposed to explain this effect is discussed briefly in the next chapter on refinement methods.

Isomorphous quality of the derivatives.

The most sensitive test of derivative quality is an ability to refine to a low value of R , the crystallographic reliability index, and a high ratio of RMS ($F_{\text{calc}}/E_{\text{ISO}}$) and these results are discussed in the chapter. However when the data sets have been placed on a common scale, several indicators of derivative quality can be examined. The native protein absolute scale and temperature factors were estimated at this point as shown in fig. 12. These allows some absolute estimates of the various parameters in terms of electrons to be made. Virtually all of these derived values are based on the assumption of Wilson statistics (Wilson, 1942) so that

$$\langle F_{\text{PH}}^2 \rangle = \langle F_{\text{P}}^2 \rangle + \langle F_{\text{H}}^2 \rangle \quad (19)$$

As stated in chapter 2, the values obtained from this result are probably poor absolute estimates, but may present a good method of comparing the quality of derivatives. The isomorphism test of Crick and Magdoff has already been discussed in chapter 3, so these plots for the derivative data to 2.5° Å resolution are presented in fig. 33 without further explanation. The values for N_{H} in equation (3.14) were obtained from the final refined values of the heavy atom site occupancies.

Anomalous scattering.

The derivatives used, especially the uranyl compounds, contain atoms which scatter X-rays anomalously, so that in general

$F_{\text{hkl}}^+ \neq F_{\text{hkl}}^-$ (Bijvoet, 1954). When anomalous dispersion occurs the vector $\underline{F}_{\text{H}}$ is composed of two sets of components from each contributing heavy atom. Then

$$f_j = f_{0j} + \Delta f_j + if_j'' \quad (20)$$

where Δf and f'' are the real and imaginary dispersion corrections,

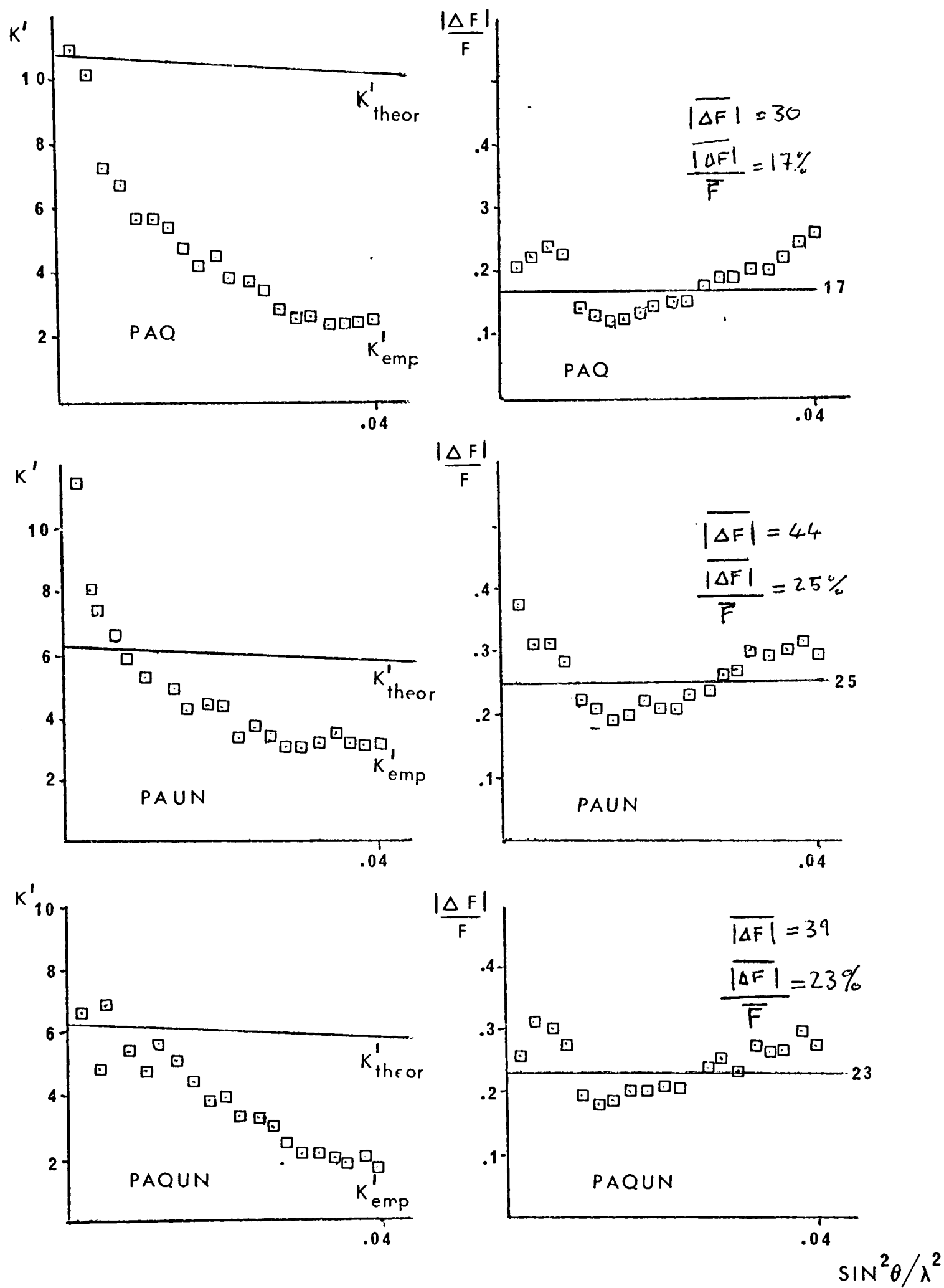


Fig 34. Plots of empirical and theoretical k' and mean fractional change in protein structure amplitude for three derivatives as a function of resolution. Also shown are the mean values of the isomorphous difference ΔF and the percentage mean fractional difference $\Delta F / \bar{F}_p * 100$.

and f_{0j} is the normal value of the scattering factor when anomalous scattering does not occur. Δf_j and f_j'' depend upon λ and $\sin\theta$, as well as the nature of the anomalous scatter. They are less sensitive functions of $\sin\theta$ than f_{0j} because the contributory electrons (the K-electrons) are concentrated in a small volume near the atomic nucleus. This means that the ratio

$$k'_{\text{theor.}} = f' / f'' = \frac{f_{0j} + \Delta f_j}{f_j''} \quad (21)$$

termed the real to imaginary scattering ratio decreases very slightly with increasing $\sin\theta$.

Matthews (1966a) shows that an estimate of this ratio can be made by the approximation :

$$k'_{\text{emp}} = f' / f'' = 2. \langle |F_{\text{PH}} - F_{\text{P}}| \rangle / \langle |F_{\text{PH}}^+ - F_{\text{PH}}^-| \rangle \quad (22)$$

where the numerator is simply the mean value of the isomorphous difference and the denominator is the mean difference between Friedel-related reflections. The value of expression (22) is shown for each derivative, as a function of $\sin^2\theta / \lambda^2$ in fig. 34. It can be seen that the experimental points are poor approximations to the theoretical value of k' especially at high angles. This effect is general experience. Dodson and Vijayan (1971) point out that the isomorphous differences, being larger, are expected to have smaller percentage errors than the anomalous differences and that the error in the measurement of the latter is likely to make $\langle |F_{\text{PH}}^+ - F_{\text{PH}}^-| \rangle$ larger than is the case, particularly at high angles.

Mean fractional changes in intensity due to introduced heavy atom.

In chapter 3 it was mentioned that a mean fractional change in intensity of 15% was sought for adequate isomorphous replacement. In fig. 34 are set out plots of the function :

$$\frac{\sum_h |F_{\text{PH}} - F_{\text{P}}|}{\sum_h F_{\text{P}}} \quad (23)$$

for ranges of $\sin^2 \theta / \lambda^2$. To some extent this value is similar to Crick-Magdoff plots, in that with increased resolution, one would expect a smaller value of expression (23), because of the monotonic fall in the heavy atom form factors. This reduction occurs at first in each derivative, but then the value rises, indicating perhaps, a loss of isomorphism at that level of resolution.

Summary of chapter 4.

Several general points can be made from the experience gained from the collection and processing of data from prealbumin crystals.

It is necessary to determine the maximum resolution of the data to be collected with some care. It appears that the upper resolution limit for prealbumin might be more reasonable at $3 \overset{\circ}{\text{\AA}}$ for diffractometer data, than $2.5 \overset{\circ}{\text{\AA}}$. This observation is made from an examination of the various plots in chapter 4, which indicate that the quality of the data beyond $3 \overset{\circ}{\text{\AA}}$ is dubious. The principal advantage of limiting the data collection to $3 \overset{\circ}{\text{\AA}}$ resolution is that the data collection time would be cut by half and this would allow twice as much time to collect data within the more realistic resolution limits.

It is also desirable to collect all the data with the same statistical precision. Because of the rapid decrease of the intensity of diffracted beam, with increasing angle, I would suggest that data be collected in shells encompassing a moderate resolution range, where counting times are proportionally increased with increasing resolution. If the highest resolution shell is collected first, and then shells of decreasing resolution, this minimizes the effect of radiation damage. Further, the application of an absorption correction for a spherical shell of data is perhaps least liable to error from θ -dependence of absorption factors. An empirical curve is produced for an axial reflection within each shell.

These techniques are impossible or highly inconvenient on the

linear diffractometer, but some were tried with PAUN on the five-circle machine. Although this derivative is the weakest data set, it has the best value of k'_{emp} and gives peaks at the heavy atom sites in an anomalous difference Patterson, while the other derivatives do not. The generally poor quality of the measured anomalous data would be improved by repeated measurements of Friedel pairs. This is usually difficult on a large protein because of the prohibitively long collection times. In such cases it might be possible to collect a complete native set and restrict the measurement of derivative data to reflections where I_{NATIVE} is greater than a prescribed value, which may be related to the particular counting statistics, beyond a resolution of $3 \text{ \AA}$. This is particularly relevant to prealbumin, where there are a rather small proportion of reflections beyond $3 \text{ \AA}$ that can be measured with good counting statistics with the data collection times used.

The analysis of the variation of scale factors with F_{obs} suggests that every possible step should be taken to ensure both low background and equal counting precision for all the data. To improve the derivative to native scale factor for low values of F_{p} it may be useful to put into the native protein data set the unobserved reflections as half the minimum observable intensity, as is customary in small molecule work.

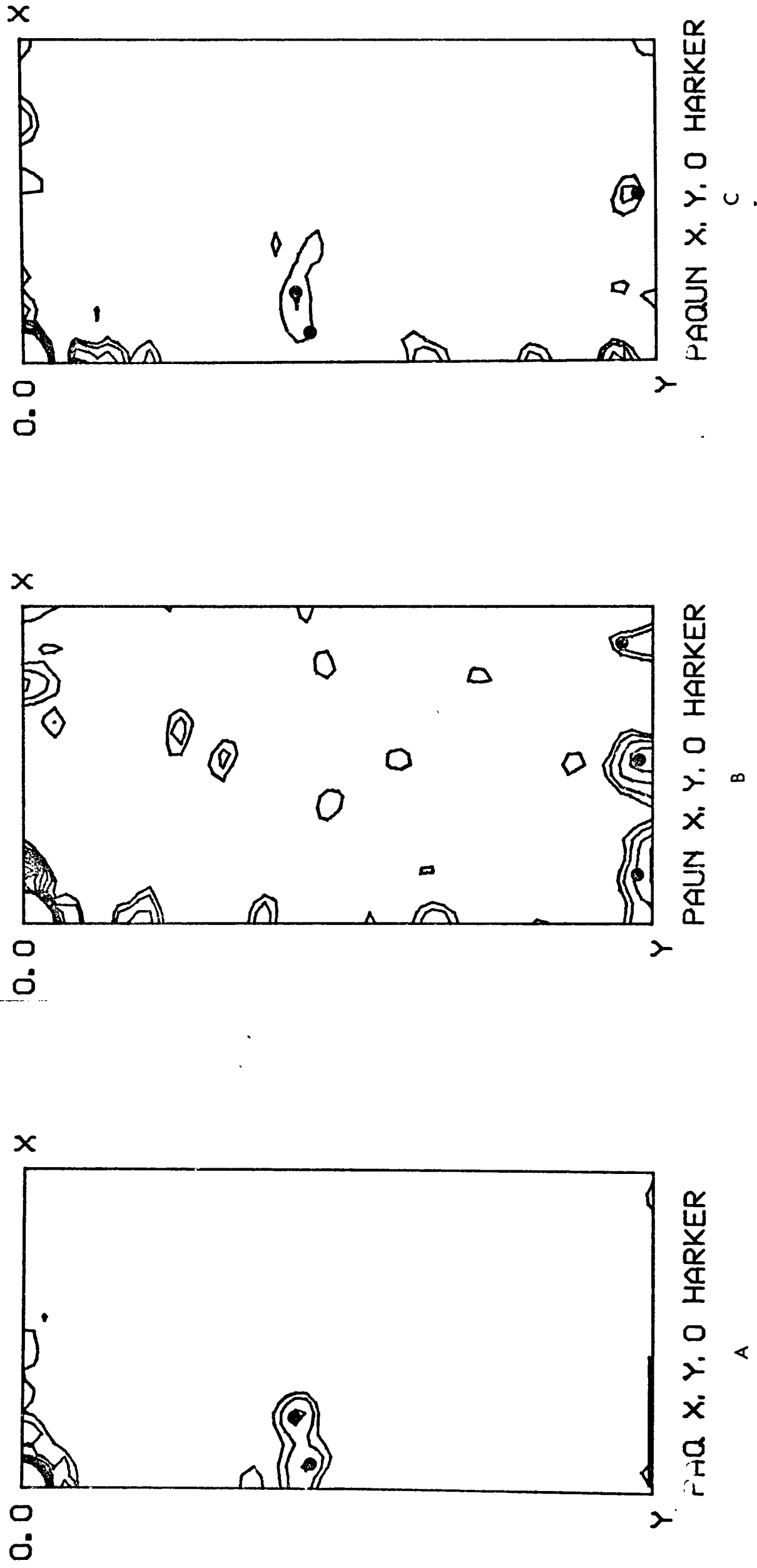
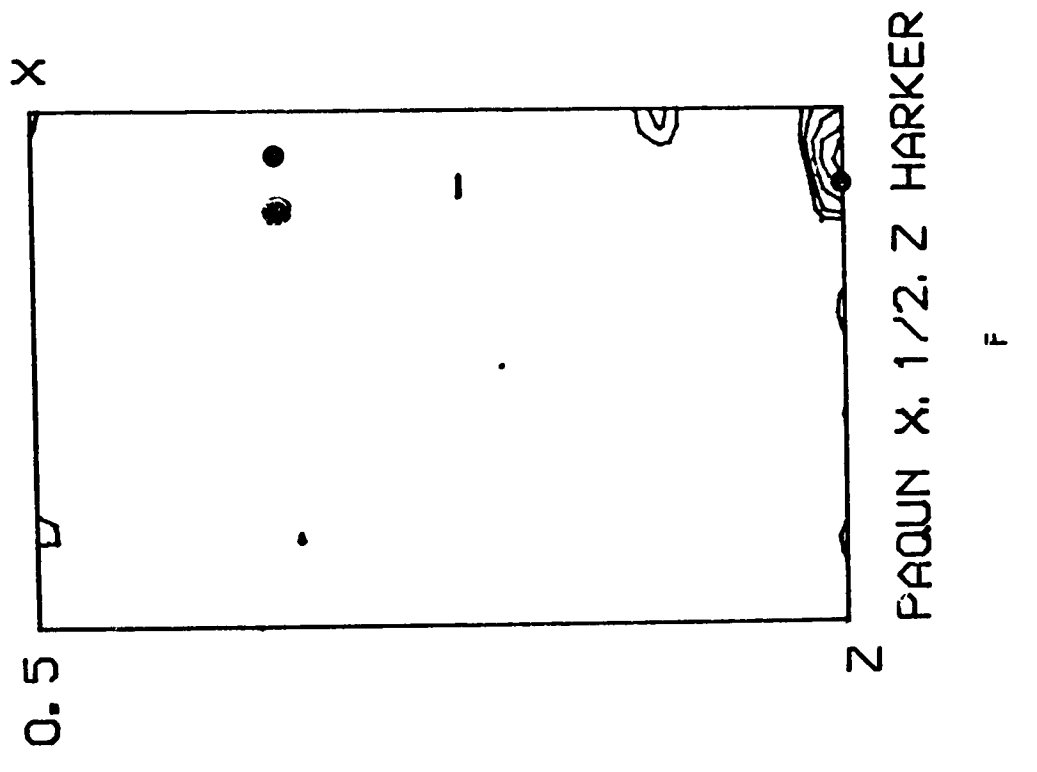
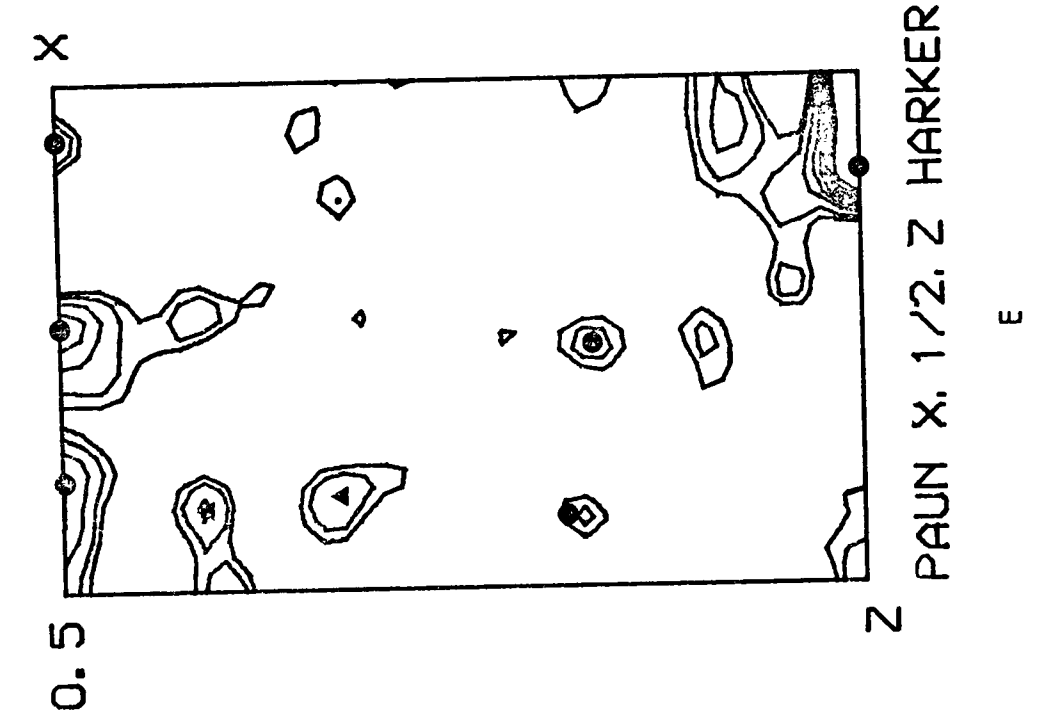
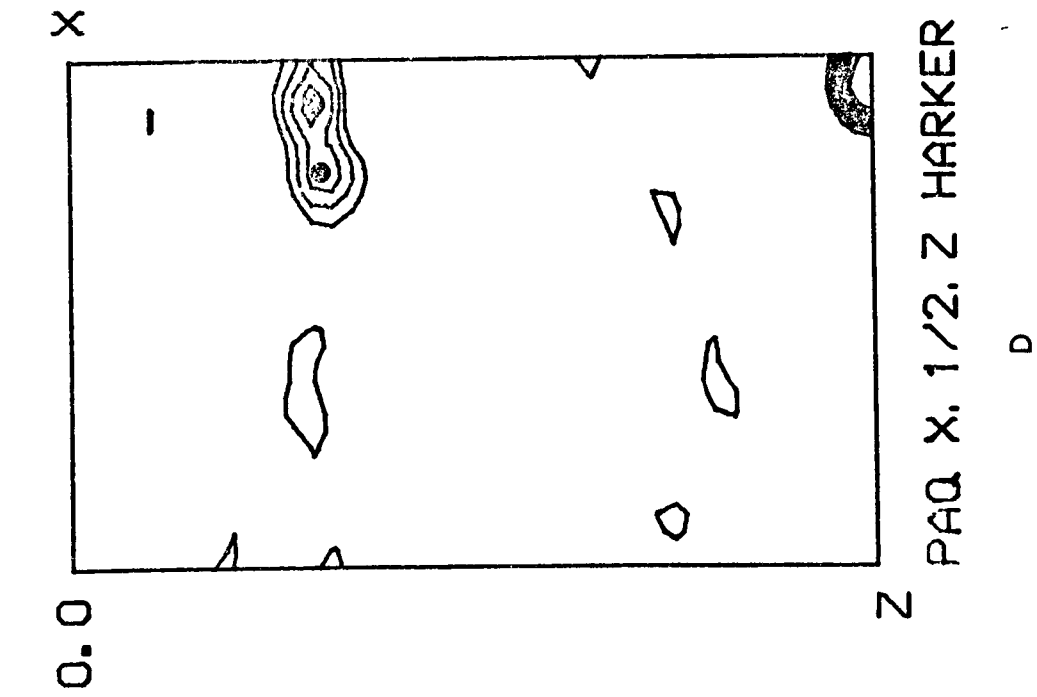
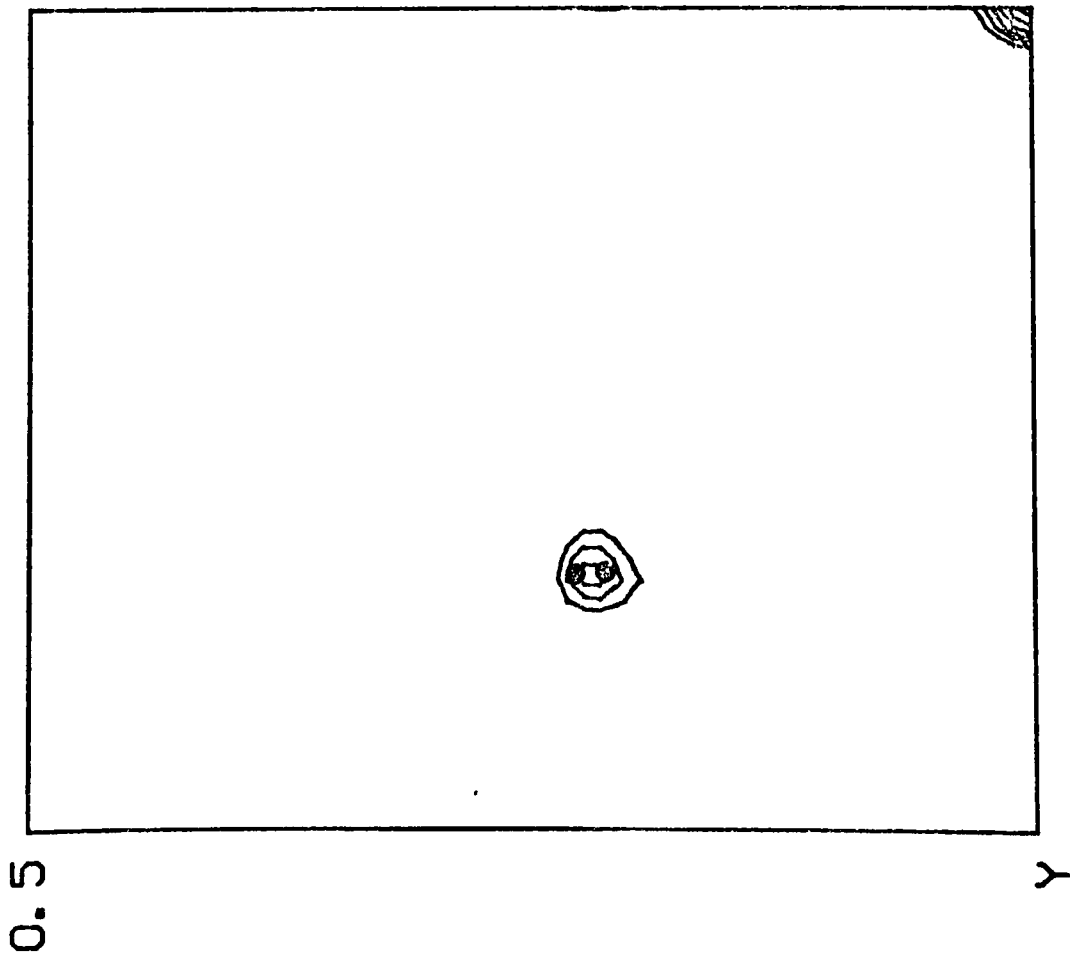


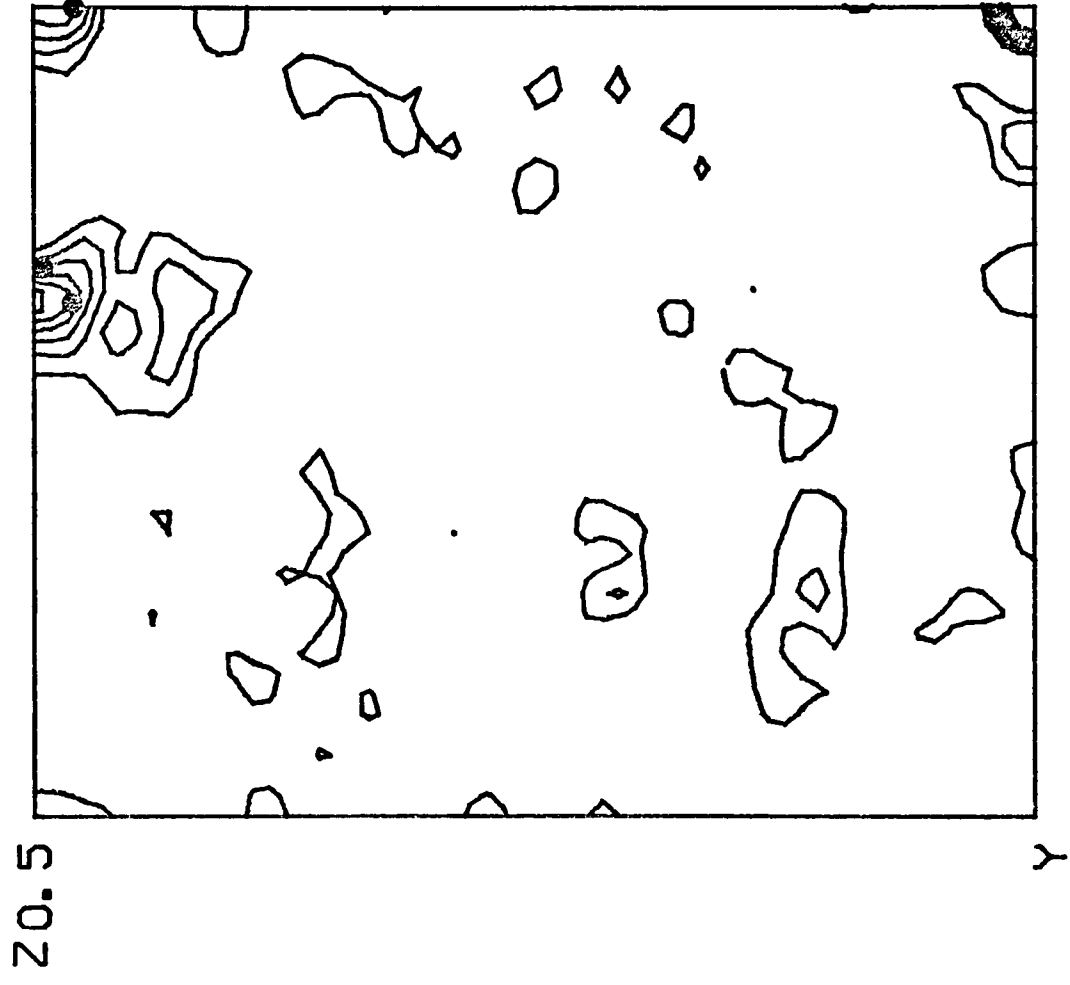
Fig 35 Harker sections XY0 (a to c), X_2^1Z (d to f) and XYZ (g to i) of the 3D Pattersons of PAQ, PAUN and PAQUN. also shown are the general sections X0Z (j to l).





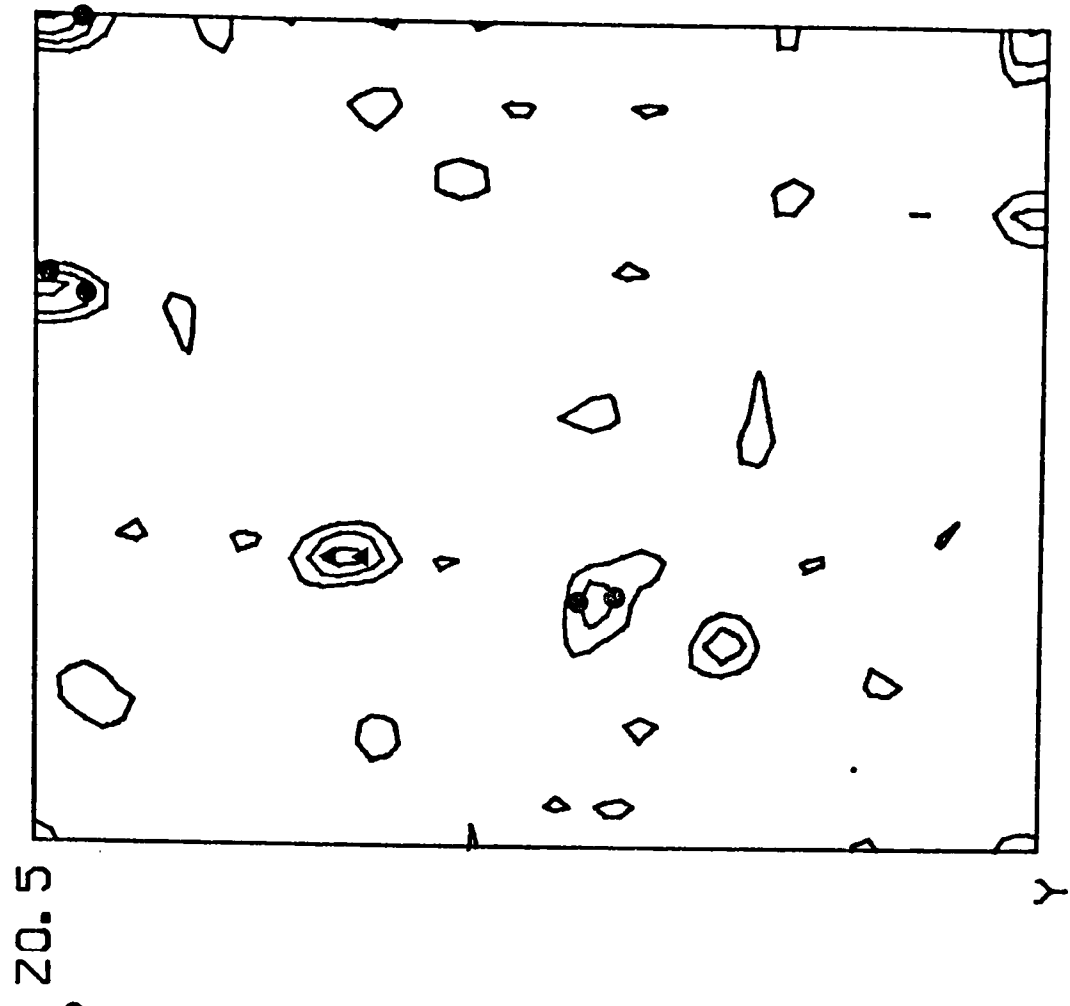
PAQ 1/2. Y. Z HARKER SECTION.

G



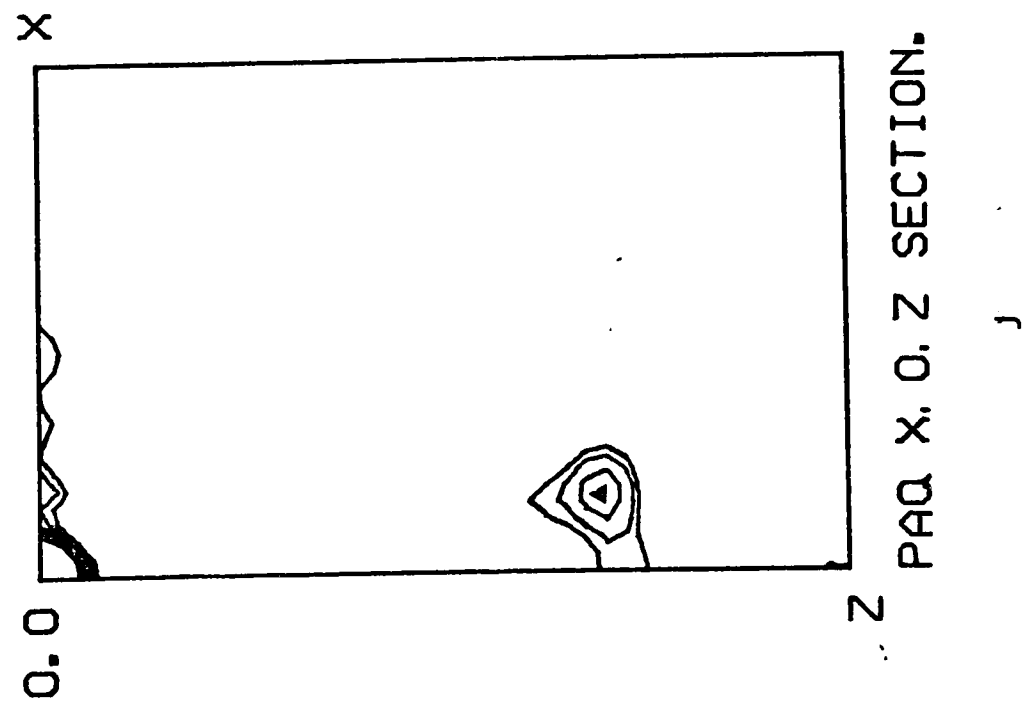
PAUN 1/2. Y. Z HARKER SECTION.

H

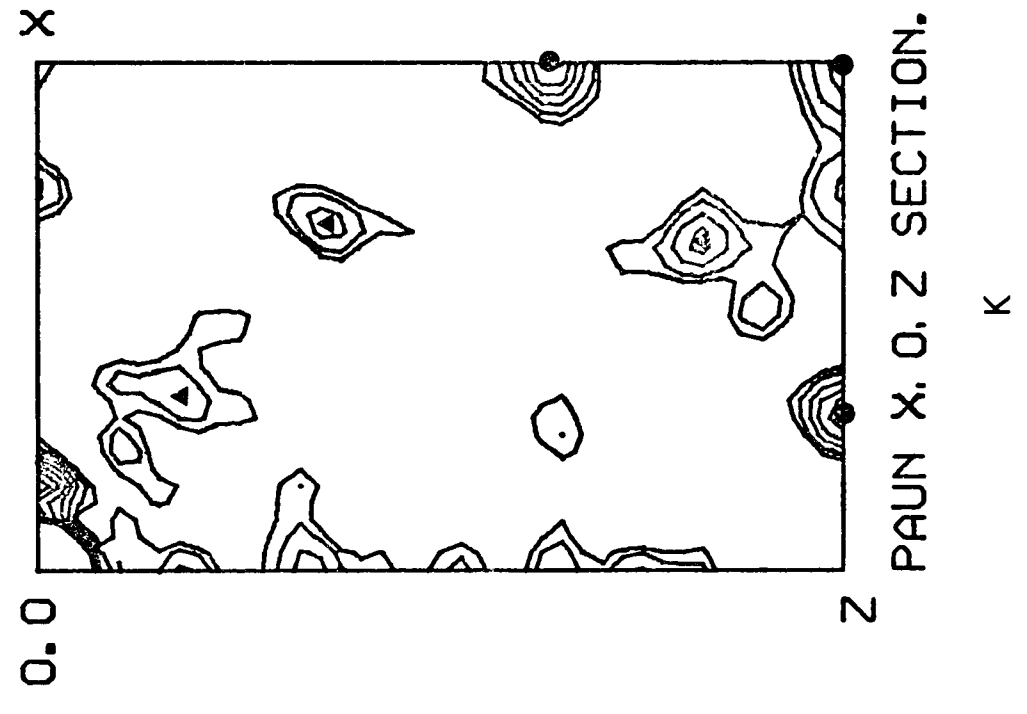


PAQUN 1/2. Y. Z HARKER SECTION

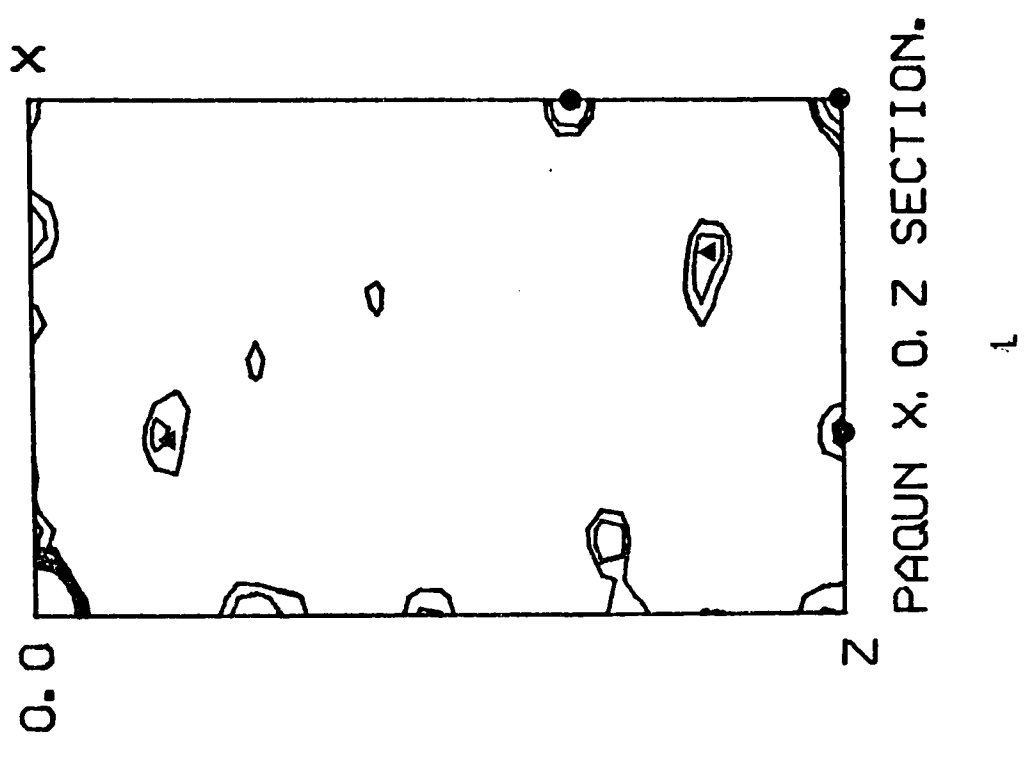
I



J



K



L

CHAPTER 5.

Solution of heavy atom derivative Pattersons, refinement and phasing.

The structure amplitudes for the three derivatives, PAQ, PAUN and PAQUN were scaled to the native set at stage 14 of the last chapter. This data was used to calculate three 3D difference Pattersons using equation 3.1 where $I(hkl)$ is replaced by $(F_{PH} - F_P)^2$. The Patterson synthesis is very sensitive to changes in the derivative-native scale factor. Several Pattersons were calculated with scale factors varying from 1.0 to 1.05, after the data had been initially scaled so that $\sum_h F_{PH}^2 = \sum_h F_P^2$. The effect of increasing the scale factor caused the lowest features to change rapidly, while the principal peaks increased in area with increasing scale factor. None of the main features disappeared, but previously resolved peaks coalesced in some cases. This suggests that care in scaling is necessary to avoid loss of resolution.

The Harker sections $(X, Y, 0)$, $(1/2, Y, Z)$ and $(X, 1/2, Y)$ and the general section $(X, 0, Z)$ from the 3-dimensional difference Pattersons are reproduced in fig. 35(a to f). A solid circle represents a self-vector and a solid triangle, a cross-vector. The vectors marked on these Pattersons are those obtained from the final sets of heavy atom co-ordinates (table 13) used in phase determination for the protein Fourier. For clarity, only the principal vectors are shown in the two complex uranyl derivatives.

The Patterson of PAQ was very encouraging. For the first time, the two mercury sites in the asymmetric unit were resolved in the x-direction (a and d). These sites have nearly identical y co-ordinates, so that the section $(X, 0, z)(j)$ contains the cross-vector. This gives the separation in x as 0.075 (3.25 Å). It is clear from this that one site must have an x co-ordinate less than zero, as indicated by the odd parity Patterson synthesis, discussed in chapter 3. The Patterson was

solved in terms of the two sites:

	X	Y	Z
EG1	0.054	0.103	0.075
HG2	0.980	0.115	0.424

The Patterson functions of the two uranyl derivatives are very much more complicated. The principal features of these Pattersons corresponded with known vectors in previous $6 \text{ \AA}$ syntheses and $3 \text{ \AA}$ projections. The mercury self-vector set can just be distinguished in the double derivative Patterson of PAQUN, if a and g (PAQ) are compared with c and i (PAQUN). The strongest evidence for the presence of the mercury sites in PAQUN is given by the cross-vector between these and the main uranyl site (two close triangles in i). The cross-vector appears in the Harker section (1/2, Y, Z) because the mercury sites are half a cell translation in x from the uranyl site. There is no evidence for the mercury sites in the PAUN derivative Patterson. This shows that the mercury atoms were successfully removed when the derivative was prepared (see table 4). A comparison of the two uranyl Pattersons reveals a similar vector set, but different site occupancies. The X, 0, Y sections k and l show this best. The (X, 0, Z) section is highly populated, because all the uranyl sites have a similar y co-ordinate (0.25). Phases derived from the mercury sites were used to produce difference Fourier syntheses of the uranyl derivatives. The sites found in these Fouriers were then checked against the vectors in the Pattersons. Table 13 shows the sites corresponding to peaks in the Pattersons.

These sites formed a basis for the refinement of the three derivatives at $2.5 \text{ \AA}$ resolution.

Refinement of heavy atom site parameters.

The refinement of heavy atom parameters remains a highly time-consuming step in protein crystal analysis. The problem central to

this stage, is the determination of a suitable F_{obs} term for the expression:

$$E = \sum_h w_h (|F_{obs}| - |F_{calc}|)^2 \quad (1)$$

which every refinement procedure minimizes. The centrosymmetric case, discussed in chapter 3, is the easiest one to cope with, since

$$|F_{obs}| = |F_{PH}| \pm |F_P| \quad (2)$$

This method has been used with success, as the sole refinement procedure prior to calculation of a 3D protein Fourier (Blake et al., 1962). In the case of prealbumin, however, only 15% of the data to 2.5 Å resolution is centric and these zones, as discussed in chapter 3, may not be fully representative of the 3 D data.

For acentric data, there is an unknown relationship between the observed difference ΔF in structure amplitudes where:

$$\Delta F = |F_{PH}| - |F_P| \quad (3)$$

and the true isomorphous difference F_H shown in the Argand diagram in fig. 5.

Rossman (1960) suggested that the magnitude of F_H be approximated by (3) and the expression:

$$E = \sum_h w_h \left[(|F_{PH}| - |F_P|)^2 - |F_H|^2 \right]^2 \quad (4)$$

be minimized, where the choice of weights as

$$w_h = (|F_{PH}| - |F_P|)^2 \quad (5)$$

gives most importance to terms where the vectors F_{PH} and F_P are nearly co-linear (large values of ΔF). Although this method has been used with some success (Cullis et al., 1961). It was decided not to attempt it with prealbumin, since the weighting function (5) inevitably favours the even-ordered terms in a pseudo-centred cell.

The procedure of Dickerson (1961) is to obtain an estimate of F_{PH} from the current values of the protein phases and to minimize the expression:

$$E = \sum_h w_h (|F_{PH}| - |F_P \cdot \exp i\alpha_p + F_{calc}|)^2 \quad (6)$$

This so-called 'phase refinement' has the disadvantage of requiring a

starting set of protein phases.

The final method reported was first suggested by Kartha (1965) and has been developed by workers such as Dodson and Vijayan (1971). This technique uses the measurement of the anomalous dispersion, together with the isomorphous difference to give an estimate of the true isomorphous difference $|F_H|$. In principle this method yields a complete set of F_{obs} terms for equation (1), since although no anomalous dispersion effects can be measured in centric zones, equation (2) provides an estimate of F_{obs} in this case. The estimate of F_H is calculated from:

$$F_H^2 = F_{PH}^2 + F_P^2 \pm 2F_{PH}F_P \left[1 - (k' [F_{PH}^+ - F_{PH}^-] / 2.F_P)^2 \right]^{\frac{1}{2}} \quad (7)$$

where k' is the ratio of the real to the imaginary components of the contribution from the anomalous scatterer. The positive sign in expression (7) is relevant when $|\alpha_P - \alpha_{PH}| > 90^\circ$ and the negative sign, when $|\alpha_P - \alpha_{PH}| < 90^\circ$. Since the protein structure amplitudes $|F_P|$ and $|F_{PH}|$ are in general much larger than the isomorphous difference $|F_H|$, the lower estimate of expression (7) is likely to be the right one. The derivation of a suitable set of F_{obs} from (7) then involves a rejection of terms where the upper estimate of $|F_H|$ could conceivably be the correct one. This refinement is referred to as ' F_{HLE} refinement' in later discussion.

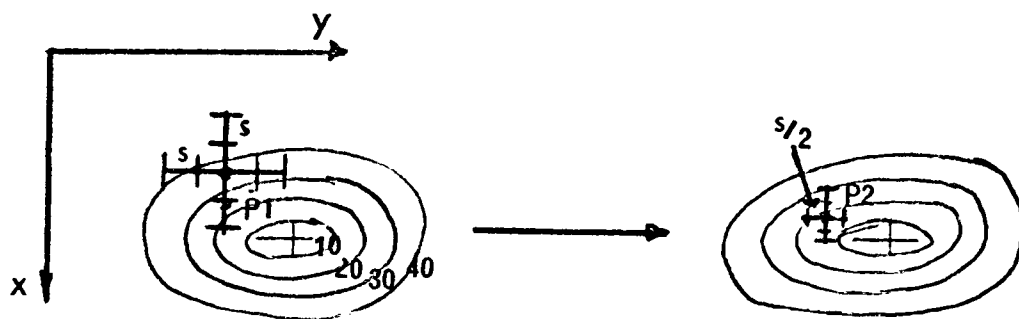
Fourier syntheses.

At intervals in the refinement procedures described, the difference and residual Fouriers, using coefficients expressed in equations 3.7 and 3.3, were calculated. This also constituted a step in the refinement process as the residual Fourier synthesis indicated gross shifts in occupancy or site co-ordinates. It is also the only method of introducing minor sites into the refinement procedure. These maps were produced for the asymmetric unit $x: 0-0.5, y: 0-1.0, z: 0-0.5$ at Fourier grid spacings of $1/60, 1/120$ and $1/30$ giving a separation

of approximately $0.3 \text{ \AA}$ between each grid point. After the positions of the main sites had been identified small areas of Fourier space containing these sites were often calculated, to save computing time.

Centric zone refinement.

Until the 1906A computer became available all refinement was carried out in two dimensions using the Hart-type refinement programme already described. Data from the 3 centric zones was refined together so that simultaneous adjustment of x, y and z co-ordinates was possible. The cycle time for this process was very slow. For 1 500 reflections, approximately half a minute was required per site for each parameter varied. This method has the additional disadvantage of very slow convergence. This is illustrated by the diagram below:



The refinement is shown for the co-ordinates X and Y. The contours represent a minimum of the error sum E plotted around the true values of x and y in two-dimensional space. The initial values of x and y are quite close to this minimum, so that after calculating shifts of $-2s$, $-1s$, $+1s$, $+2s$, the new co-ordinates are found at P_2 . However, the size of the shifts is now reduced as in b. If the minimum is not found on the second or third cycle, many cycles give only a small progression towards the minimum.

All the derivative to native scale factors, refined by this programme to a value a little less than unity. As mentioned in chapter 4, this is partly due to the dependence of scale factor upon $|F_{\text{obs}}|$. However, the refined scale factor was also dependent on the

derivative quality. Those derivatives with the largest E-values, where E is the R.M.S. value of (1), consistently gave the lowest scale factors. In the case where many of the native structure amplitudes are weak, it is most likely that the corresponding derivative amplitudes F_{PH} are as weak or larger, weaker measurements having been rejected through poor counting statistics. If this is so, the reduction of $\langle |F_{PH}| \rangle$ obtained by lowering the scale factor decreases the error sum E, particularly when the calculated heavy atom structure factors do not agree very well with the observed differences. Some evidence for this supposition was obtained when the refinement was carried out using only reflections for which $|F_P|$ was greater than 50. In these cases, the scale factors did not drop below unity on being allowed to refine. Such cut-offs and weighting functions, similar to that used by Lundberg (1965) are difficult to apply to prealbumin data in two-dimensional refinement. Not only do they all reduce the number of observations, but they discriminate against the odd-ordered terms. For these reasons other types of refinement were tried to check the validity of the results of the Hart-refinement.

Phase refinement.

At this stage the refined site parameters for the derivative PAQ were used, together with the measurements of the anomalous dispersion to produce a set of 3D phases. The self-phase map of PAQ showed, as expected, two peaks in the asymmetric unit, with no indication of minor sites. Allowing for slight variations in the relative scale of the Fourier grid spacings, these sites appeared elliptical, with the major axes in the y-direction. This elongation of the mercury sites had been noticed in previous $3 \text{ \AA}^0$ OKL work (see fig. 17). In the $3 \text{ \AA}^0$ HKO projection the two sites overlapped extensively, but now that the 3D synthesis separated the sites, it could be seen that the ellipsoid was oblate, the axes in the x and y directions being greatest.

Two models were tried for these sites. At first, two atoms were placed within the electron density, separated by a distance of $1 \text{ \AA}$ in y. At this separation, it is impossible that both sites can be simultaneously occupied. This means that the model represents a statistical distribution of occupancy, within the elliptical envelope observed in the Fourier synthesis. This model caused a significant drop in the centric R factor. An attempt was made to model the elliptical sites with an anisotropic temperature factor. It was decided to abandon centric refinement, as it is doubtful whether meaningful thermal shape parameters can be derived from zone data.

In order to determine the F_{obs} values appropriate to the acentric terms, the phase refinement method of Dickerson was used, the expression (6) being minimised. This required a starting set of protein phases. It was undesirable to use the phases derived from the PAQ isomorphous and anomalous differences. Not only are the PAQ anomalous measurements poor (see fig. 34) but the phases will be biased towards the starting model. This bias is very much enhanced with acentric data. Because α_P is not restricted to α_H and $\pi + \alpha_H$ as in the centric zones, the vector triangle is always closed in favour of the lowest ϵ value. The phases derived in such a situation always carry information about the heavy atom site. If these phases are used with random numbers, as difference structure factors, the heavy atom site is usually easily visible.

To provide better phases for refinement, difference maps of PAUN and PAQUN were calculated with the PAQ phases. In both of these derivatives the PAQ sites were observed, even though the Patterson of PAUN shows no evidence of these. Two of the principal uranyl sites were well-defined and spherically symmetric. In PAUN, the third site was rather extended in the x direction, but in PAQUN this site continued with undiminished density through 16 Fourier sections - a

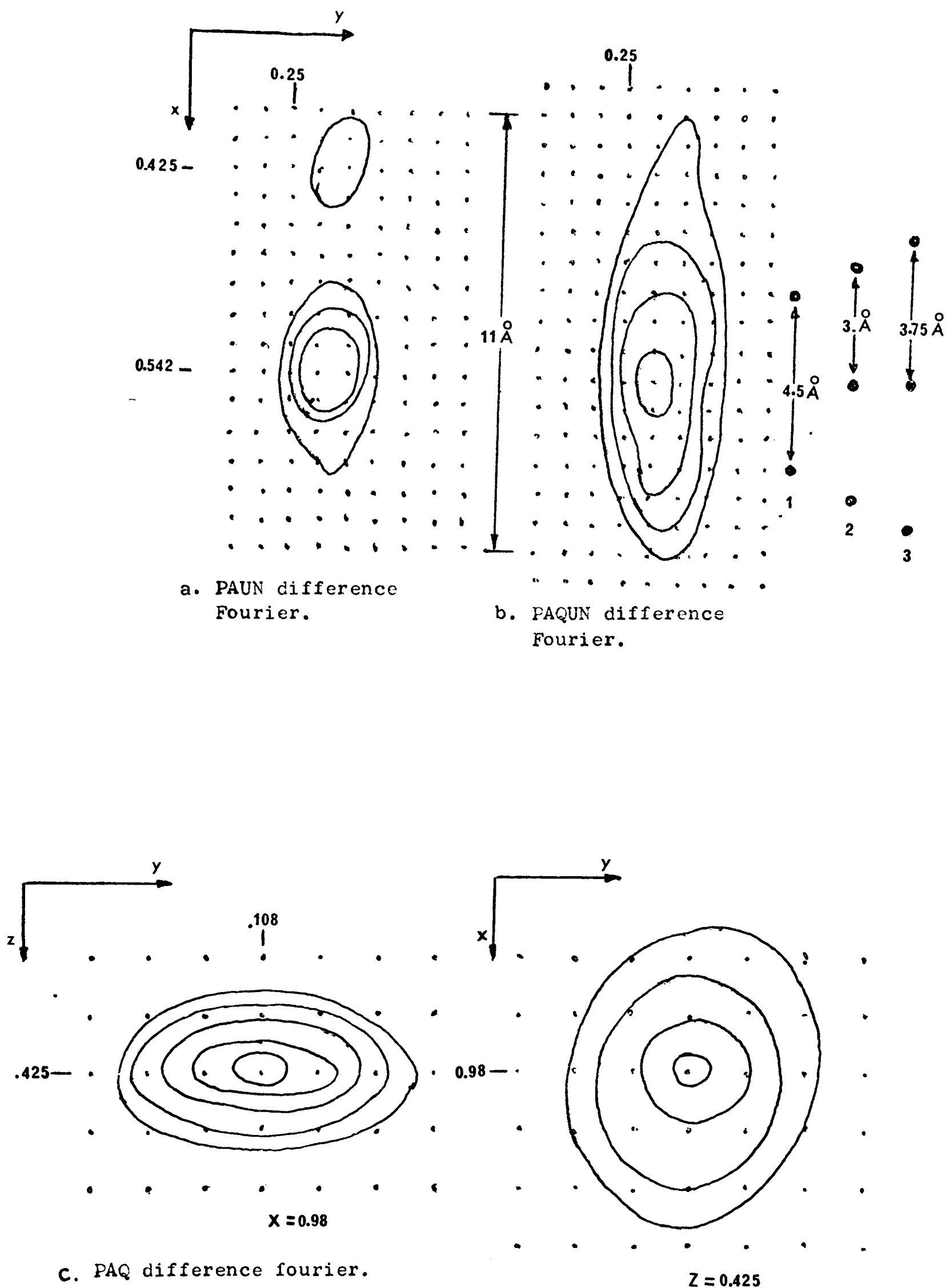


Fig.36 Sections of the difference Fourier of PAQ, PAQUN and PAUN at 2.5 Å resolution. The grid spacings are approximately 0.3 Å. The sites indicated to the right of (b) represent attempts made to model the extended density. The two cross-sections of the PAQ site HG1 show that the site is oblate.

distance of $11.2 \text{ \AA}$. Figure 36 shows the appearance of the heavy atom sites at $2.5 \text{ \AA}$ resolution. At this stage of the refinement, there was no indication of a waist in the density, which might have indicated where several close sites could be accommodated. To the right of the Fourier section (b) are 3 models used for this density. (The model 3 was the one finally adopted). This behaviour of the uranyl derivative might be explained as a distribution of sites, as with PAQ. Alternatively there may be a series of close carboxyl groups on the protein at that point. It is a possibility that the presence of phosphate buffer is responsible for the association of several uranyl ions. This site is less extended in PAUN (a) and the main difference in the conditions of preparation of the two uranyl derivatives was lack of phosphate in PAUN.

At the first stage in the refinement, single uranium atoms were placed at the centres of gravity of these complex sites. The derivatives PAQ and PAUN were used to produce a further set of phases, against which the PAQUN sites seen in the difference Fourier were refined.

The phase refinement was carried out using the programmes CRYLSQ and ORFLS from the X-ray 70 package of Oak Ridge. Both programmes allow a partial structure factor to be read as A and B, where

$A = F \cos \alpha$, $B = F \sin \alpha$. In the protein work, the partial structure factor read was A_p and B_p , calculated from the observed native structure amplitude and the current estimate of the phase angle. These quantities were then added to the components of F_{calc} to give a F_{PHcalc} , which was compared with F_{PHobs} in the refinement. Usually three or four cycles of least-squares shifts caused convergence. The resulting new site parameters were then used in a further phase run to obtain an improved set of phases, when the process was repeated.

Blow (1973) has pointed out that this method can be extremely

slow to converge to a final parameter set. As a result of his experience with α -chymotrypsin he adopts a method of excluding one derivative from the phasing step, and using the phases from (N-1) derivatives to determine the parameters for this. On the next phasing cycle, this derivative is included, but another is excluded and treated like the first. This technique, he reports, promoted rapid convergence.

The same poor convergence was evident in the prealbumin phase refinement. Because the mercury derivative appears more isomorphous with the native than the uranyl compounds, the effect of the phase refinement caused a gradual increase in occupancy of the mercury sites, until this was greater than the final occupancy obtained from the centric refinement.

A trial refinement alternately excluding each derivative also failed to converge satisfactorily. The prealbumin situation is complicated by each derivative possessing common sites. This makes the alternate exclusion of sites impossible, since it was thought that a minimum of two derivatives was required for meaningful phasing.

Some attempt to achieve convergence was made by using derivative phases with figures of merit of 0.8 or greater. However, this method was very slow to converge, because these phases naturally favoured the heavy atom model which produced them.

Some useful information was obtained from the phase refinement. At a later stage, the mean figure of merit of the phases had improved, and a difference map of PAQUN now indicated structure in the extended uranyl site, and the three-atom model (fig. 36) was adopted for this site. Also the refined parameters of PAUN allowed a cross-phase difference Fourier of PAQ to be made (fig. 36). This was the first time that this derivative had been phased independently. The map showed just two sites with the expected co-ordinates of x, y and z, and the elliptical appearance seen in the self-phase maps. The cross-phase synthesis did

not show any evidence for minor sites in PAQ.

Temperature factors.

The simple form of the isotropic temperature factor B that had been previously used in the centric refinement is included in the structure factor calculation as

$$T = \exp[-2B/\sin^2\theta / \lambda^2] \quad (8)$$

and $B = 8\pi^2 \bar{u}^2 / A^2$ where $\bar{u}^2$ is the mean square displacement of the atom perpendicular to the reflecting planes. The least squares programmes follow the suggestion of Cruickshank (1956) and give temperature factors in terms of U's where $B = 78.977 \cdot U$.

If an atom vibrates anisotropically, the expression for its effect on scattering is given as

$$T = \exp[-(h^2 \beta_{11} + k^2 \beta_{22} + l^2 \beta_{33} + 2hk \beta_{12} + 2kl \beta_{23} + 2hl \beta_{13})] \quad (9)$$

where h, k and l are the Miller indices and β_{ij} the coefficients of the parametric equation of an ellipsoid, where β_{ii} represent the major axes. These coefficients are dependent upon the unit cell translations, which complicates a direct comparison of the amount of vibration in any given direction. The corresponding temperature factor expression in terms of U_{ij} is

$$T = \exp[-2\pi^2(h^2 a^{*2} U_{11} + k^2 b^{*2} U_{22} + l^2 c^{*2} U_{33} + 2hka^*b^* U_{13} + 2klb^*c^* U_{23} + 2hla^*c^* U_{13})] \quad (10)$$

The refined values of the U_{ij} for the two mercury sites are given below:

SITE	U_{11}	U_{22}	U_{33}	U_{12}	U_{23}	U_{13}
HG1	0.319	0.329	0.209	-0.015	-0.0023	0.0137
HG2	0.343	0.421	0.209	-0.023	-0.0151	0.033

These values are in excellent agreement with the shapes observed in the cross-phase difference synthesis of PAQ (fig. 36c). The temperature factor is similar in the x and y axial directions and reduced in the z direction, giving an oblate ellipsoid. The cross-terms are very small, showing that the major axes of the ellipse are parallel with the cell

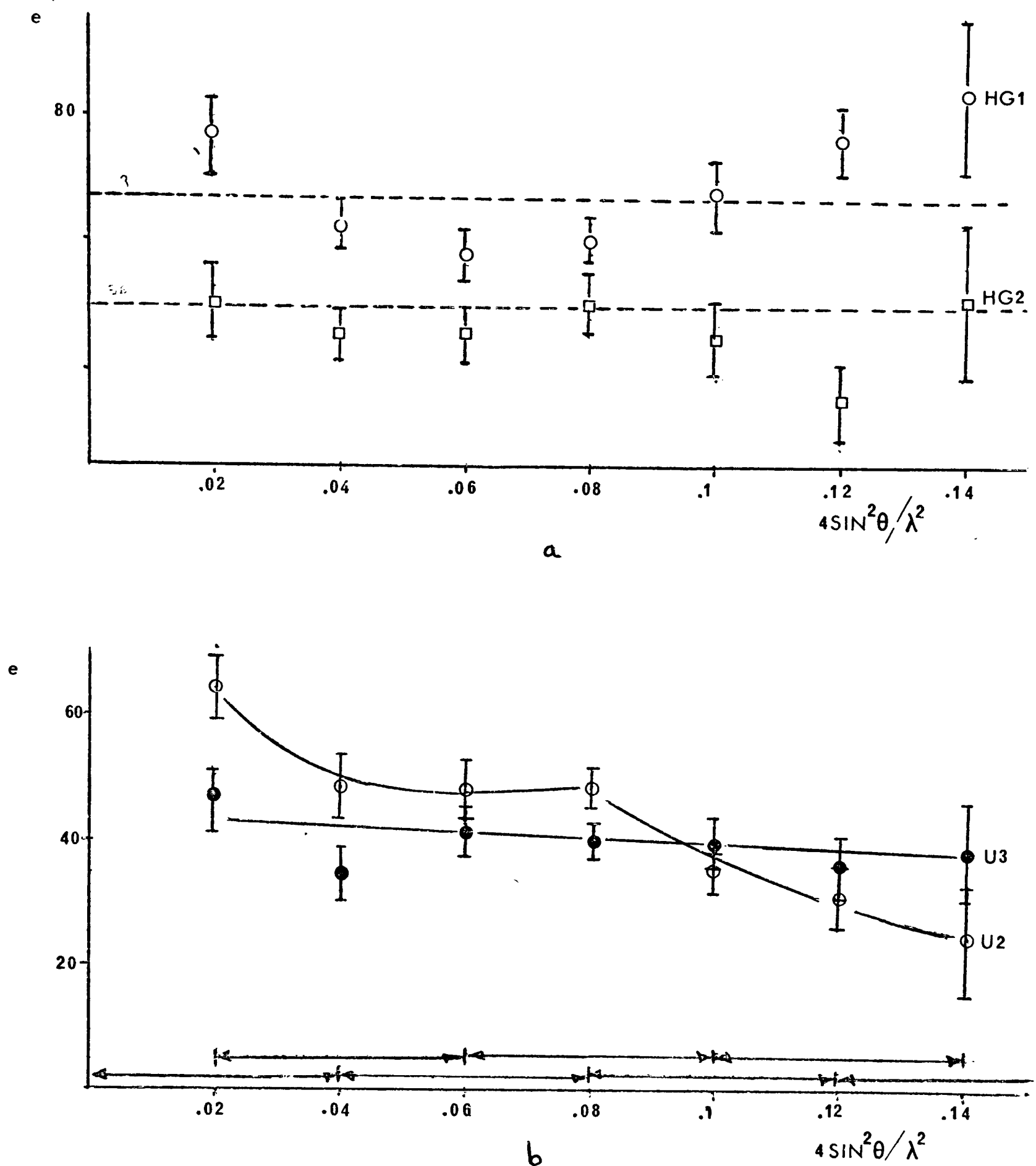


Fig. 37 Plots of heavy atom site occupancy for a) the two mercury sites in PAQ and b) two uranyl sites in PAUN, as a function of resolution. The vertical scale is approximately in electrons. The arrows in b) indicate the overlapping ranges used in the refinement. Very similar plots of occupancy were obtained for the same sites in PAQ and PAQNEW, so that in the latter these plots were used to derive the empirical form factors shown as HGA and HGB in table 13. An increased temperature factor was applied to site U2, as indicated by the above plot.

axes, as is observed. The anisotropic refinement of the FAQ sites behaved surprisingly well, although this is by no means the first time anisotropic parameters have been used with protein data (see Cullis et al., 1961 and Liljas et al., 1972). Temperature factor refinement of protein data is an acknowledged difficulty, probably because the curve of the partial differential of the structure factor expression (3.11) with respect to temperature factor, is almost flat. This is certainly the case at $6 \text{ \AA}$ resolution where a B of 12 to 15 has been used without refinement.

The occupancy factor has a much more well-defined minimum, and so some workers adopt the strategy of refining this factor in shells of increasing resolution. Wyckoff (1970) reports using this method to check that an estimated temperature factor was reasonable. The plot of the occupancy of each site against resolution should be flat if the temperature factors have been correctly determined. This technique has been used with the derivatives of prealbumin, after refinement of site and occupancy factors and discovery of minor sites. Figure 37 shows the results of this refinement on the three derivatives of prealbumin. The general result of this technique was that the most highly occupied sites from initial refinement had the lowest temperature factors and that minor sites generally disappeared altogether near the resolution limit of the data. A correction factor $\pm \Delta B$ was applied to the site temperature factors after the shell refinement where ΔB was the slope of a plot of the logarithm of the occupancy against $\sin^2 \theta / \lambda^2$. The shells were chosen to overlap each other, to give an adequate number of experimental points, while keeping the number of observations within a shell to a maximum. It is possible to use such an analysis as the basis for empirical form-factor curves for each site (Margaret Adams, 1963), but this is probably a method of absorbing the effects of non-isomorphism and

experimental error and may not necessarily lead to more accurate protein phase angles.

As has been stated, the process of phase refinement did not appear to work satisfactorily in the case of the prealbumin derivatives because it was not possible to refine the derivatives individually. The F_{HLE} refinement, however, does not have this limitation.

F_{HLE} refinement.

The estimation of F_H using this technique places a heavy reliance on the accurate measurement of the anomalous scattering. In general the effect of experimental error causes an over-estimation of the anomalous difference $\Delta F = F_{PH}^+ - F_{PH}^-$. If the theoretical value of k' (k'_{theor}) is used in expression (7), Dodson and Vijayan (1971) have shown that the F_{HLE} terms and consequently the site occupation factors, are systematically over-estimated. However, if the empirical value of k' (k'_{emp}) obtained from the Matthews expression (4.22) is used, then the refined occupancies are close to their true values. This idea is similar to Matthews' original weighting factor w (Matthews, 1966). The effect of both is to weight down the anomalous contribution to F_{HLE} . It is interesting to note that the estimate of F_{HLE} is independent of the sign of ΔF_{anom} , since this term is squared in the Matthews' expression. The estimates of F_H derived from phase-refinement, are however, very dependent on the sign of ΔF_{anom} , and less upon the magnitude of this quantity. In this way, the two methods are complementary.

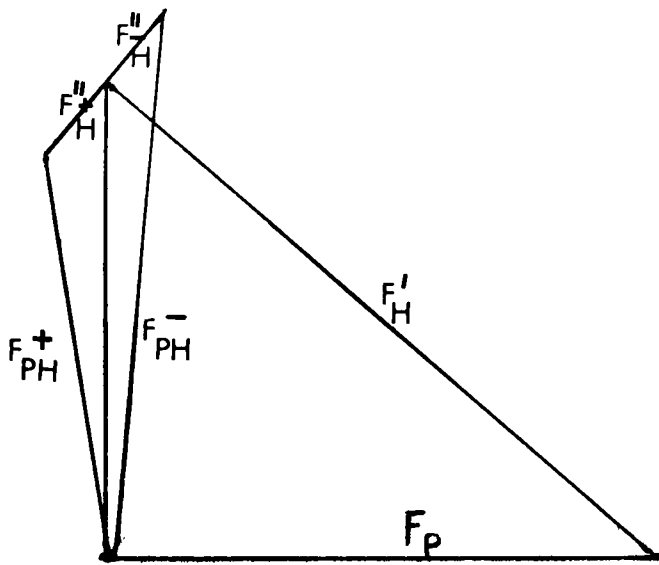
The k'_{emp} curves shown in fig. 34 were used in the calculation of F_{HLE} terms. Because the terms are presented as $|F_{obs}|$ to the refinement programme, no refinement of the derivative to native scale factor can be carried out. This scale factor was calculated prior to the computation of the F_{HLE} terms by a method suggested by Kraut (1962). The basis of this technique is that the origin peaks

$P(0, 0, 0)$ of two different Pattersons constructed from (i) $(F_{PH} - F_P)^2$ and (ii) $(F_{PH}^2 - F_P^2)$ should be equal if F_{PH} is properly scaled to F_P . The scale factor k is adjusted until $\sum_{hkl} F_H^2 = \sum_{hkl} (kF_{PH}^2 - F_P^2)$. The estimate F_H can be obtained from the Matthews' expression.

Since F_{HLE} is dependent upon the scale factor, the scaling was applied iteratively as suggested by Arnone et al., (1971). Scale factors of 1.03 and 1.05 were obtained for the mercury and uranyl derivatives respectively. Two species of atom exist in the PAQUN derivative, but in these computations the anomalous dispersion from this compound has been treated as if from a single species of atom. At the stage of phase calculation, these sites were assigned their proper anomalous form factors.

Equation (7) was used to calculate the two estimates of $|F_H|$, corresponding to the two choices of sign. In most cases F_{HUE} (positive sign choice in (7)) was greater than either F_P or F_{PH} . Selection criteria for F_{HLE} terms were as follows:

- 1) F_{HUE} must be greater than the maximum possible value of F_H . This excludes the possibility that F_{HUE} may be the correct estimate of F_H . For proteins with no centrosymmetric zones, this value must be estimated from the $F(0\ 0\ 0)$ term of the difference structure amplitudes. In prealbumin the largest centric difference term was used as an estimation of this quantity. Values of 400 and 600 electrons were used for the mercury and uranyl derivatives respectively.
- 2) F_{HLE} must be less than or equal to the maximum value of F_H . This avoids large over-estimates of F_H from acentric terms, where a spuriously large anomalous difference exists.
- 3) If the expression $(k' [F_{PH}^+ - F_{PH}^-] / 2.F_P)$ (11) is greater than one, a negative root results. Such terms were rejected. However if the angle $|\alpha_P - \alpha_{PH}|$ is nearly 90° , the effect of the anomalous scattering upon F_{PH}^+ and F_{PH}^- will be at a maximum (see diagram overleaf).



If expression (11) was between 1.0 and 1.2 and $|F_P|$ was small compared with $|F_{H \max}|$ the estimate of F_H was found from

$$F_H^2 = (F_{PH}^+ + F_{PH}^-)^2 \quad (12)$$

Using the criteria above, 3 200, 3 500 and 5 400 estimates of F_{HLE} were calculated from the PAQ, FAQUN and PAUN data sets respectively. The increased number of reflections retained from the PAUN data, perhaps reflects the better quality of the anomalous scattering from this derivative. The re-calculation of k'_{emp} from the F_{HLE} terms of the derivatives gave curves which were closer approximations to the plot of k'_{theor} , presumably because spuriously large values of ΔF_{anom} and ΔF_{iso} had been excluded. Some typical values of F_{HLE} and F_{HUE} from FAQ are given below:

H	K	L	F_P	F_{PH}	ΔF_{anom}	F_{HLE}	F_{HUE}
0	6	2	1093	1079	0	14	2172
2	6	2	840	739	14	109	1578
3	6	2	256	323	1	67	578
5	6	2	149	252	5	105	400
6	6	2	344	234	-5	110	577
7	6	2	345	416	-4	71	760
8	6	2	137	181	-1	44	317

The validity of the F_{HLE} terms was checked by Patterson syntheses. Only in the case of PAUN were the Patterson vectors increased in height relative to the isomorphous difference Patterson vectors. However, the background was significantly lower in these syntheses, particularly

Compound.	Refinement. scheme.	Site.	Parameters.				B _{ij}				B _{ij}			
			x	y	z	Z	B ₁₁	B ₂₂	B ₃₃	B ₁₂	B ₁₃	B ₂₃	B ₁₂	B ₁₃
Mercury PAQ	C		0.0529	0.1075	0.0780	0.3600	0.0043	0.0014	0.0012	-0.0013	-0.0005	-0.0000		
	F	HG1	0.0543	0.1068	0.0785	0.4367	0.0042	0.0013	0.0015	0.0001	-0.0003	0.0000		
	P		0.0544	0.1071	0.0776	0.4090	0.0070	0.0013	0.0007	-0.0001	-0.0002	-0.0001		
	C		0.9779	0.1147	0.4256	0.3370	0.0076	0.0016	0.0011	-0.0009	-0.0013	0.0004		
	F	HG2	0.9796	0.1153	0.4245	0.3511	0.0044	0.0016	0.0009	-0.0005	-0.0004	-0.0003		
	P		0.9811	0.1147	0.4250	0.3686	0.0073	0.0014	0.0004	0.0003	0.0009	-0.0001		
Mercury PAQNEW	F	HG1	0.0540	0.1070	0.0790	1.0000	20.000							
	F	HG2	0.9792	0.1150	0.4270	1.0000	20.000							
	C		0.3675	0.2455	0.1674	0.4300	29.000							
Uranium PAUN	F	U1	0.3689	0.2457	0.1636	0.5000	16.000							
	P		0.3697	0.2465	0.1636	0.4642	34.000							
	C		0.7170	0.2559	0.3406	0.2669	29.000							
	F	U2	0.7231	0.2632	0.3446	0.3056	36.000							
	P		0.7233	0.2583	0.3426	0.2175	33.000							
	C		0.9960	0.0326	0.0528	0.1000	29.000							
	F	U3	0.9069	0.0288	0.0499	0.2356	33.400							
	P		0.9359	0.0291	0.0418	0.1358	40.000							
	C		0.0262	0.0315	0.4494	0.1000	29.000							
	F	U4	0.0264	0.0311	0.4443	0.1661	29.000							
	P		0.0257	0.0323	0.4431	0.1126	40.000							
	C		0.5436	0.2596	0.2504	0.3100	29.000							
	F	U5	0.5431	0.2633	0.2500	0.3090	30.700							
	P		0.5415	0.2621	0.2502	0.2967	31.000							
	C		0.4874	0.2648	0.2458	0.1630	30.000							
Uranium + Mercury. PAQUN	F	U1	0.4885	0.2659	0.2461	0.1970	20.000							
	C		0.5845	0.2719	0.2564	0.1130	20.100							
	F	U2	0.5841	0.2724	0.2527	0.1759	30.000							
	C		0.5296	0.2556	0.2593	0.1100	21.000							
	F	U3	0.5315	0.2565	0.2548	0.1580	21.000							

Uranium +	C	0.3660	0.2420	0.1625	0.2062	30.000
Mercury.	F	0.3712	0.2427	0.1644	0.2573	30.000
PAQUN.	P	0.3637	0.2446	0.1649	0.1661	33.700
	C	0.7170	0.2498	0.3387	0.1190	25.000
	F	0.7143	0.2502	0.3340	0.1250	26.000
	P	0.7170	0.2522	0.3327	0.0920	30.000
	C	0.5000	0.2100	0.2400	0.1100	30.000
	F	0.4968	0.2143	0.2416	0.1040	35.000
	P	0.4915	0.2163	0.2430	0.0327	30.000
	C	0.4500	0.2450	0.2000	0.1000	20.000
	F	0.4515	0.2441	0.2000	0.1475	16.000
	P	0.4494	0.2456	0.1981	0.0817	30.000

Form Factors.									
Atom.	Type.	0.0	0.05	0.10	0.15	0.20	0.25	0.30	$\sin\theta/\lambda$ value.
Hg	Real	80.0	78.5	75.0	71.0	66.5	62.5	58.5	
U	Real	92.0	91.0	87.0	82.5	77.5	73.0	68.5	
HgA*	Real	42.5	37.0	26.5	22.5	26.0	-	-	* = empirically determined from occupancy refinement in shells of different resolution.
HgB*	Real	22.5	22.5	18.0	17.0	12.0	-	-	
Hg	Imag.	8.3	8.3	8.3	8.3	8.3	8.3	8.3	P = phase refinement.
U	Imag.	16.5	16.5	16.5	16.5	16.5	16.5	16.5	C = centric refinement.
									F = FHLE refinement.

Derivative to native refined scale factors.

Derivative. Scale factor.

PAQ	1.032
PAQNEW	1.025
PAQUN	1.053
PAQUN	1.057

Table 13. Refined site parameters for the derivatives and the form and scale factors used in the refinement.

in that of PAQNEW.

Refinements similar to those discussed in the section on phase refinement were carried out with the F_{HLE} terms. The R-factors were lower than those obtained from the centric refinements, but this is to be expected since the 'best' terms have been selected. Anisotropic temperature factors were refined and shell refinements on each derivative were carried out to establish the correctness of temperature factor assignments.

The F_{HLE} refinements were easier to handle and much more conclusive than the phase refinements, undoubtedly because the derivatives could be independently refined. Table 13 compares the final refined parameters for three derivatives from the three refinement procedures used. For each site in the derivatives the heavy atom parameters x , y , z , Z and B or B_{ij} are given from the final cycle of centric (c), F_{HLE} (F) and phase (P) refinement. There is no equivalent to sites U1, U2, and U3 from the phase refinement of PAQUN because these were modelled by one, anisotropic site at (0.525, 0.258, 0.250), close to the current site 'U3'. It was encouraging to see that the positional parameters from each method are in agreement within experimental error. Since the $|F_{obs}|$ differs considerably in each case, as do the regions and extent of the data refined, it is probably safe to say that these values are insensitive to the method of refinement. In the case of a single-site derivative, the phase angle α_H is independent of the site occupancy. Once the direction of $\underline{F}_H$ has been established, the protein phase angle, determined by equation (3.2) is also insensitive to small errors in the magnitude of $\underline{F}_H$. This ideal case suggests that convergence to a final set of phases would be quickly reached.

When, as is more usual, several sites of comparable occupancy make up a derivative, the phase of $\underline{F}_H$ becomes increasingly dependent

Derivative.	Refinement type.	R	E	F_H	No. of Fobs.	selection criteria.		
						F_{HLE} max.	F_{HLE} min.	FIGM min.
FAQ	C	0.52	26.6	53	1209	-	-	-
	F	0.46	30.1	56	3200	200	200	-
	P	0.14	31.6	55	1367	-	-	0.8
PAUN	C	0.65	54.0	70	1109	-	-	-
	F	0.60	59.0	75	5400	300	300	-
	P	0.22	48.1	65	1294	-	-	0.8
FAQUN	C	0.61	47.0	59	990	-	-	-
	F	0.55	48.0	60	3500	300	300	-
	P	0.19	39.0	50	1350	-	-	0.8
PAQNEW	F	0.44	27.0	50	4092	150	150	-

Table 12 Overall refinement statistics. In the columns headed selection criteria, FIGM = figure of merit, F_{HLE} = lower estimate of F_H and (-) means that a particular selection parameter did not apply.

Derivative.	Type.	Lack of closure in intervals of $0.004 \text{ s}^2/\lambda^2$									
		.000	.004	.008	.012	.016	.020	.024	.028	.032	.036
FAQ	E	60.	50.	40.	40.	35.	35.	35.	30.	30.	30.
	E'	15.	15.	15.	20.	20.	30.	35.	35.	36.	37.
PAUN	E	90.	50.	46.	50.	38.	44.	39.	37.	36.	36.
	E'	25.	30.	30.	30.	31.	32.	34.	36.	36.	37.
PAQUN	E	70.	60.	50.	40.	35.	35.	35.	34.	32.	32.
	E'	30.	30.	30.	30.	32.	35.	36.	38.	38.	39.
PAQNEW	E	40.	40.	30.	30.	30.	28.	28.	27.		
	E'	10.	10.	10.	10.	10.	10.	10.	10.		

Table 14 E-values for the four derivatives used in phase calculations. The last column (0.036) includes the terms out to the resolution limit of the data (0.04). The last two entries in the columns 0.032 and 0.036 are not known for PAQNEW since this data was only collected to a resolution of $3 \text{ \AA}$.

upon the relative occupancies of the sites. As table 13 shows, there is appreciable variation in this parameter, as determined by each of the three methods. This parameter should also be compared with the associated site temperature factor as at high resolution, there is a correlation of 60% between them.

As suggested by Dodson and Vijayan (1971) the site occupancy determined by F_{HLE} refinement is high by comparison with that determined by the other methods. Presumably because of an over-estimation of the anomalous contribution at high angles. However, the F_{HLE} occupancies are closer to the centric values than those determined by phase refinement. Since the F_{obs} from the centric zones are likely to be nearest the true value of F_H , this suggests that the phase refinement occupancies are spuriously low. This is most likely due to the dominating effect of PAQ upon the phase determination, since the mercury site occupancies gradually increased during phase refinement. The temperature factors refined poorly, with the exception of the anisotropic parameters of the PAQ sites, where the three methods are in surprising agreement. The temperature factors from the F_{HLE} refinement were the only ones checked by refinement in shells of $\sin^2 \theta / \lambda^2$ and so are probably more appropriate than the others shown. The statistics in table 12 give some indication of the relative derivative quality. Each method was selective in the terms used as can be seen by the number of reflections included in the refinement.

All of these selection methods discriminate against the smallest values of ΔF_{iso} , ΔF_{anom} or F_P . To check that there was no bias towards the even-ordered difference terms, at high resolution, a plot of $\langle |\Delta I| \rangle$ was made against resolution (where $\Delta I = I_{PH} - I_P$). Beyond a resolution of $5 \overset{0}{\text{\AA}}$ the mean values of $|\Delta I|$ for even and odd-ordered terms become identical. Since the data contain a large number of terms beyond $5 \overset{0}{\text{\AA}}$, the loss of odd-ordered reflections at resolutions

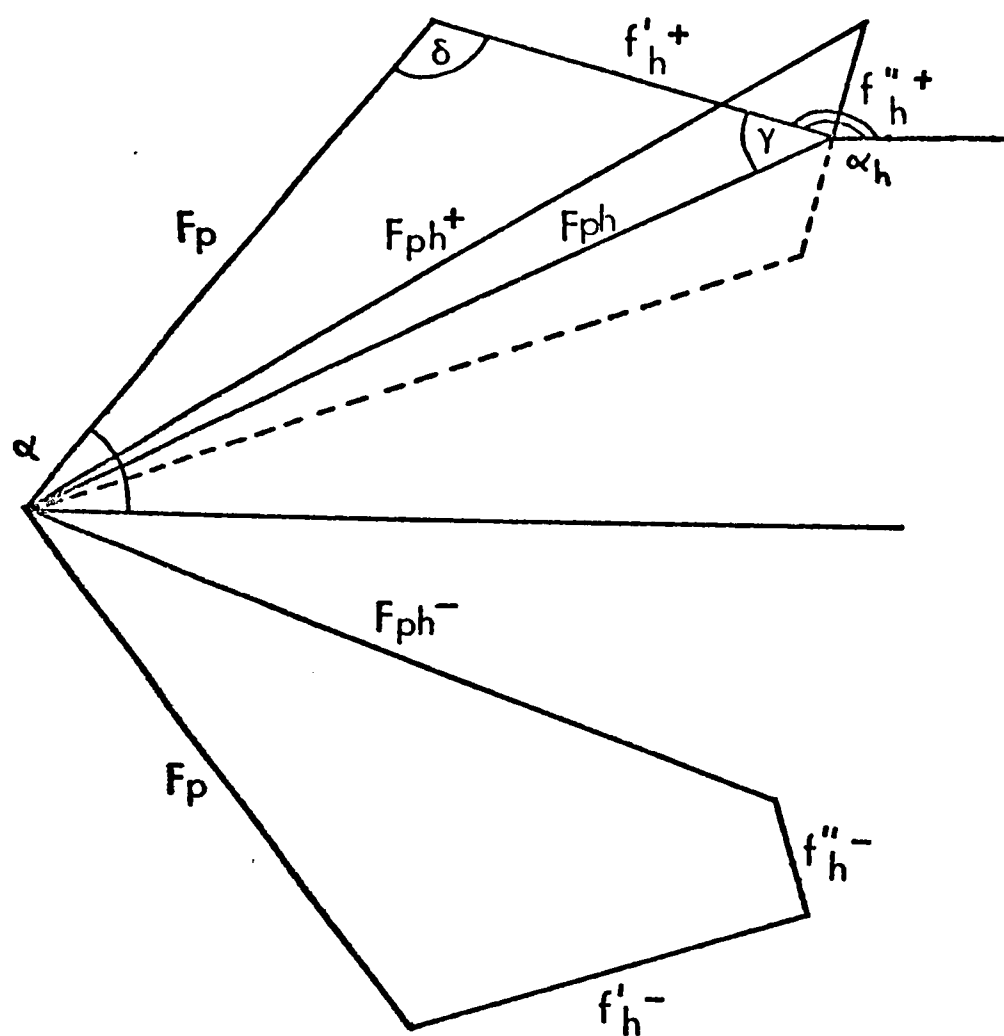
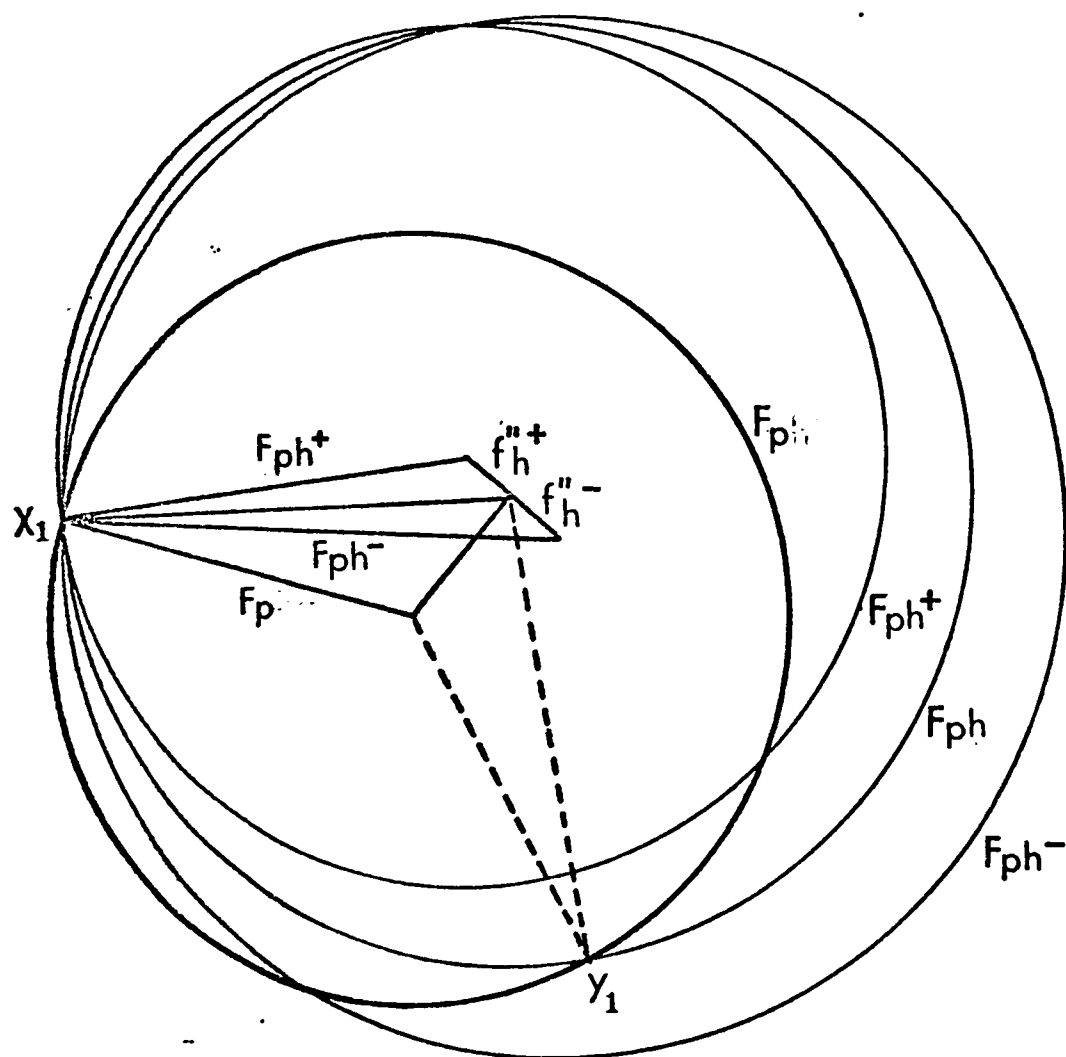


Fig. 38 (a) The effect of anomalous scattering on a Bijvoet pair of reflections.

(b) Harker Construction applied to a single isomorphous derivative with anomalous scattering.



$$F_{ph}^+ \equiv F_{ph_{hkl}} \neq F_{ph}^- \equiv F_{ph_{\bar{h}\bar{k}\bar{l}}}$$

less than this should be of minor importance.

Phase determination.

When a suitable model for the heavy atom substitution has been found, a vector $\underline{F}_H$ may be calculated as the heavy atom contribution to each derivative structure factor $\underline{F}_{PH}$. The magnitude of the $\underline{F}_{PH}$ calculated from the native protein structure factor and the vector $\underline{F}_H$ may be represented as

$$F_{PH}^2 \text{ calc} = F_P^2 + F_H^2 + 2F_P F_H \cos(\alpha_H - \alpha_P) \quad (13)$$

In general this value will not agree with the observed F_{PH} because of errors in the values of $|F_P|$, $|F_{PH}|$ and $\underline{F}_H$. The vectors shown by the Argand diagram in fig. 38a usually fail to form a closed triangle by a lack of closure error $\epsilon(\alpha)$ where

$$\epsilon(\alpha) = |F_{PH}^{\text{obs}}| - |F_{PH}^{\text{calc}}| \quad (14)$$

Blow and Crick (1959) assumed that the error (14) obeyed the Gaussian distribution and treated the quantity $|F_{PH}^{\text{obs}}|$ as being the sole source of this error. Their now familiar expression for the probability of the phase angle α being correct is represented by the normal error curve :

$$P(\alpha) = N \exp \left[-\frac{\epsilon(\alpha)^2}{2 E^2} \right] \quad (15)$$

This function is generally bi-modal, because there are usually two values of α_P for which equation (14) is satisfied. These two values are shown by the two vector triangles in fig. 38b. In fact, the ambiguity is four-fold, because there are two possible values of $\underline{F}_H$, depending upon a choice of x, y, z or $\overline{x}, \overline{y}, \overline{z}$ for the heavy atom solution.

A further derivative, with a common origin to the first and a different value of $\underline{F}_H$ resolves the first of these ambiguities by choosing the correct solution of (14). This is shown graphically by common intersections of F_{PH} radii upon the native protein circle in fig. 38b. An obvious way to combine the effects of these derivatives is to multiply the probability distributions, since the information

is independent. The exponent of (15) then is the sum of the individual lack of closure ratios $\epsilon_j^2 / 2 \cdot \bar{\epsilon}_j^2$ for each derivative j . North (1965) and Matthews (1966) have shown how the anomalous scattering measurements F_{PH}^+ and F_{PH}^- may be incorporated into (15) to resolve the phase ambiguity. Figure 38a and b shows the Harker construction relevant to a Friedel pair. The combination of the isomorphous and the anomalous information in (b) resolves the phase ambiguity, by choosing the intersection X_1 . However, it is necessary to establish the correctness of the choice of F_{PH}^+ and F_{PH}^- . It can be seen from (b) that reversing the choice of these selects the solution at Y_1 . Previous workers, using the combination of anomalous and isomorphous data from a single derivative have calculated protein phases for both choices of F_{PH}^+ and F_{PH}^- and used the phases which gave a physically meaningful protein Fourier. If two isomorphous derivatives exist, one or both with anomalous scattering measurements, it should be possible to establish the correct choice of F_{PH}^+ and F_{PH}^- for a given set of heavy atom positions. The isomorphous data establishes the intersection at Y_1 or X_1 and the F_{PH}^+ and F_{PH}^- can be chosen also to intersect at X_1 .

The choice of protein phase angle α_p is made by computing the value of (15) at intervals of α from 0° to 360° . Blow and Crick (1959) define a most probable phase α_{mp} where the probability distribution $P(\alpha)$ is at a maximum.

The value of α_p always used for protein maps is the best phase α_{best} , which is represented by the centroid of the total probability distribution. This distribution is normalized to unity by the factor N in (15), so that the length of the vector to the centroid lies between zero and one. This magnitude is known as the figure of merit m , and is zero if the distribution is perfectly bi-modal. When the coefficient

$$m \cdot F_p \exp(i\alpha) \quad (16)$$

is used in the Fourier synthesis, a 'best Fourier' results, with the

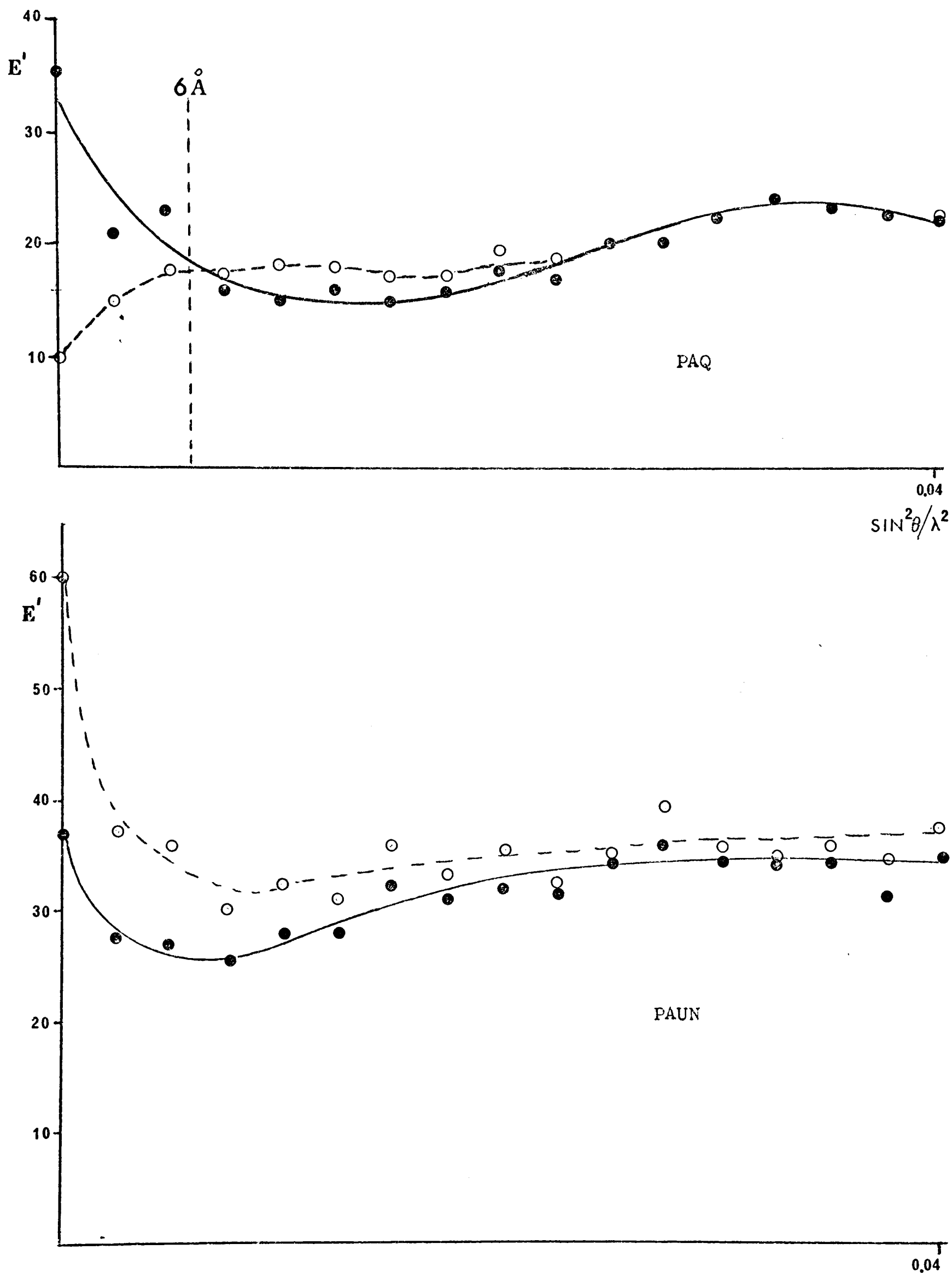


Fig. 39 Plots of the anomalous lack of closure index E' as a function of resolution for the derivatives PAQ and PAUN. The curves established the correct sense of inclusion of the anomalous data and led to the discovery that the 6 Å data for PAQ had been collected with respect to a left-handed co-ordinate frame. The broken line corresponds to a choice of F_{PH}^+ as $F_{PH}(hkl)$ and F_{PH}^- as $F_{PH}(\bar{h}k1)$. The solid line corresponds to the reverse,

least mean squares error in electron density over the unit cell.

Determination of F_{PH}^+ and F_{PH}^-

When several isomorphous derivatives are available, the standard method of determining which Friedel measurement corresponds to F_{PH}^+ or F_{PH}^- is to include the Friedels in each of these two possible senses in separate phase runs, and choose which sense gives the highest mean figure of merit $\langle m \rangle$ (see for example Watson, 1970). The change in this index is usually very small and a better indicator is the value E' where

$$E' = (\overline{\varepsilon'^2})^{1/2} \quad (17)$$

$$\text{and} \quad \varepsilon' = \Delta F_{obs}^{anom} - \Delta F_{calc}^{anom} \quad (18)$$

The value of E' can be calculated from (18) during a complete phase determination. The values of E' and E were calculated for the two derivatives PAQ and PAUN for each of the two choices of F_{PH}^+ and F_{PH}^- . As the phase programme is supplied with the anomalous difference ΔF_{obs}^{anom} , the two choices are $\pm \Delta F_{obs}^{anom}$. In this case $+\Delta F_{obs}^{anom}$ corresponded to $F_{PH}(hkl) - F_{PH}(\bar{h}kl)$. These values are shown in fig. 39 as a function of resolution.

Although the mean figure of merit is only marginally affected by the change in sign of ΔF_{obs}^{anom} , the value of E' for PAUN is lowest when $F_{PH}(\bar{h}kl)$ corresponds to F_{PH}^+ . The plots of E' for PAQ were ambiguous. The lowest values of E' below a resolution of approximately $6 \text{ \AA}$ indicated that F_{PH}^+ was best represented by $F_{PH}(hkl)$ and above this resolution by $F_{PH}(\bar{h}kl)$. The low resolution data of Swan (1970) had been incorporated into this derivative and a check on the primary data showed that this had been collected with respect to a left-handed set of axes on the four-circle diffractometer. There is a fixed data collection sequence on this machine, with the Miller index l moving fastest, k next and h slowest. If it is desired to alter this sequence, the indices may be permuted, by supplying appropriate

cell dimensions and matrix-defining reflections. A non-cyclic index permutation (khl) permits data to be collected in a left-handed co-ordinate system. This is what happened to the data of Swan. When the signs of the $\Delta F_{\text{obs}}^{\text{anom}}$ for PAQ data below $6 \text{ \AA}^0$ were reversed, a phasing run gave a consistent indication of F_{PH}^+ as $F_{\text{PH}}(\bar{h}kl)$ from both derivatives.

Two experiments were carried out as a final verification of this choice of F_{PH}^+ . The first is not suggested in the literature. It has been shown that it is difficult to assign F_{PH}^+ and F_{PH}^- correctly, without either an independent phase set or a second isomorphous derivative. Even if a second derivative is available the difference in phase angle obtained by the two choices of F_{PH}^+ is usually small, particularly if the anomalous data is poor. To avoid this bias towards the isomorphous phasing, the single derivative phases of PAUN, calculated with the two choices of F_{PH}^+ , were used with the PAQ difference terms to give a cross-phase difference synthesis. The two sets of phases had a mean difference of 50° . Both phase sets gave peaks at the expected PAQ positions, but the phase set corresponding to a choice of F_{PH}^+ as $F_{\text{PH}}(\bar{h}kl)$ gave a 50% increase in peak height at these positions.

Kraut (1968) suggested a 'Bijvoet-difference synthesis' where the observed anomalous differences are used together with the current phase information. The synthesis using phases appropriate to the correct enantiomorph should give positive peaks at the locii of the heavy atoms, whilst that corresponding to the incorrect enantiomorph, should give holes at the centrosymmetrically related sites. The phases appropriate to the anomalous terms $\Delta F_{\text{obs}}^{\text{anom}}$ of PAQ were derived by adding $\pi/2$ to the joint phases produced by PAQ and PAUN with $F_{\text{PH}}^+ \equiv F_{\text{PH}}(\bar{h}kl)$. A Bijvoet Fourier synthesis using these phases gave strong positive peaks at the PAQ sites, indicating the correct choice of enantiomorph.

Determination of E-values.

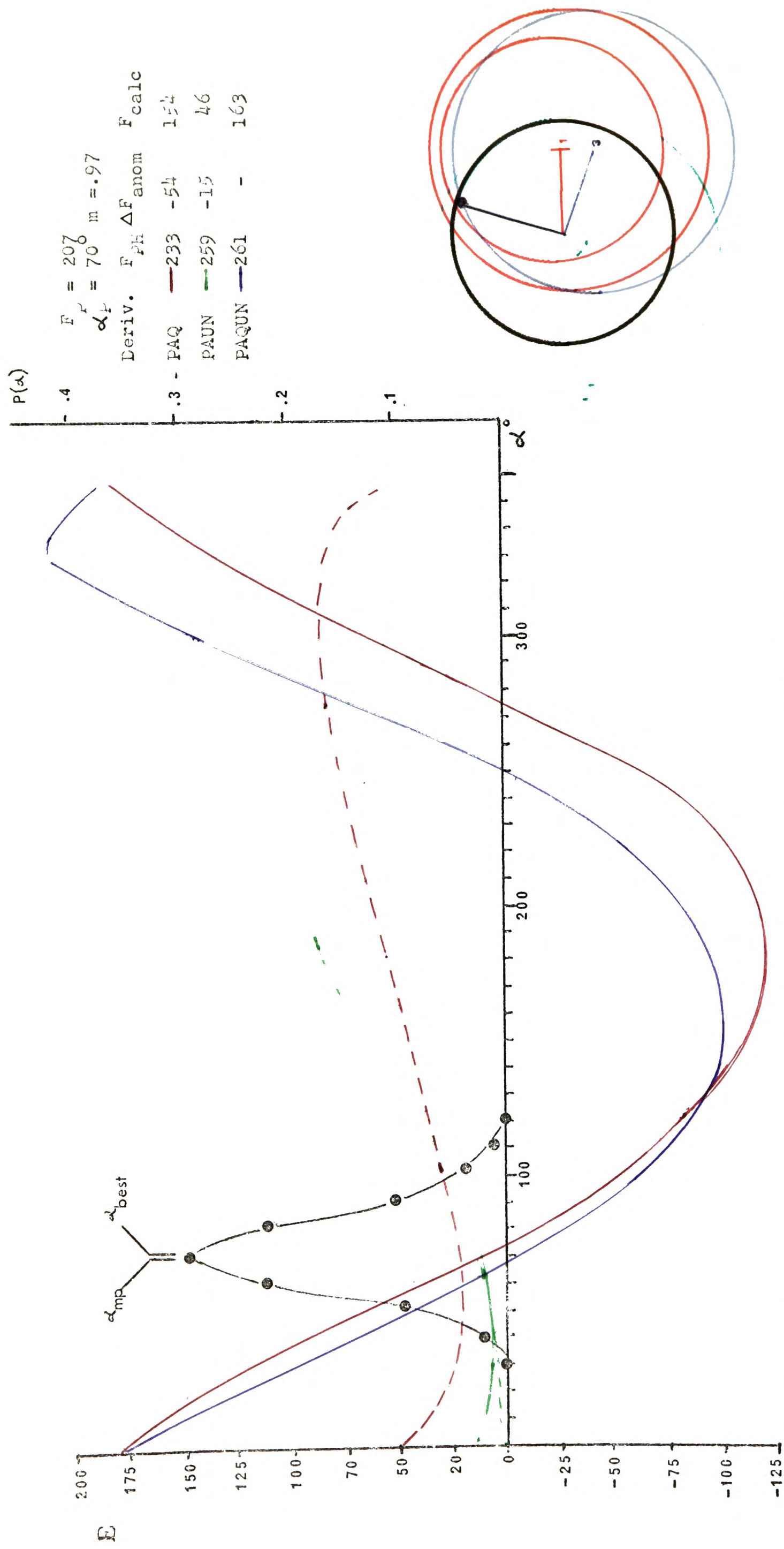
The relative weight that each derivative has in determining the protein phase α_p is determined by the values of E and E' in (15). The value of E is easy to estimate for centric reflections, where E is the RMS value of expression (1). Because the protein phase is restricted to two possibilities, this estimate is likely to be close to the correct value. If the centric data are representative of the whole data set, then this value can be used as an estimate of E for the acentric data. As indicated by the section on phase refinement, the value of E obtained from the protein phases is likely to be systematically underestimated, particularly for the derivative with the most weight in the phase calculation.

The value of E depends more upon lack of isomorphism than upon errors due to counting statistics, so that the error estimate E is not reduced substantially for reflections measured more than once. All reflections were given the same estimate of E. However, E and E' are dependent upon resolution, as indicated by the results in fig. 40. The phase programme written allowed E and E' to be specified as functions of $\sin^2 \theta / \lambda^2$.

The values of E and E' used in the phase determination were determined from expression (1) where $|F_{obs}|$ was $|F_{HLE}|$. This value was slightly lower than the centric estimate of E (because some of the worst individual reflections had been discarded in the F_{HLE} calculation), but greater than those values calculated using the protein phases as in phase refinement. An estimate of E' was obtained for each derivative from protein phases calculated from the F_{HLE} E values and an initial estimate of E' as E/3 as suggested by North (1965). The values of E and E' used in phasing are reproduced in table 14.

Phase calculation.

Sets of protein phases were produced for each individual derivative and combinations of each. The phase probability expression (15) was



Lack of closure error and phase probability plots for reflection 5 2 1 .

evaluated at intervals of 10° round the phase circle using the heavy atom site parameters shown under heading 'F' in table 13 and using form factors given in table 13.

Harker constructions and probability curves are shown for reflection 5 2 1. The black circle represents the protein phase circle. From its centre are drawn vectors of magnitude $|\underline{F}_H|$, (the phase of $\underline{F}_H$ has been changed by π to allow it to be added to the bottom of the vector $\underline{F}_P$). Derivative 1 (red) is PAQ, 2(green) is PAUN and 3(blue) is PAQUN. At the end of the vectors the vector $\underline{F}''$ is added and circles of radii $|\underline{F}_{PH}^+|$ and $|\underline{F}_{PH}^-|$ are drawn for each derivative, with that of $|\underline{F}_{PH}^+|$ centred on the most clockwise ends of the anomalous vectors. The best and the most probable protein phases are represented by dark solid and dashed lines respectively.

The corresponding phase probability diagram is shown for reflection 5 2 1. The lack of closure error ξ_j is shown for the isomorphous (solid line) and anomalous (dashed) contributions at phase angles from 0° to 360° . The normalised protein phase probability curve (black) is also shown. These results, produced in a tabular form, were useful for examining the contribution of individual derivatives to the phasing. It was apparent, after examining a series of these that the derivatives PAUN and PAQUN were very much inferior to PAQ. Plots of $\langle \underline{F}_H^{calc} \rangle$, $\langle E \rangle$ and $\langle m \rangle$ from the joint phase determination are shown in fig. 40.

Fourier syntheses.

The joint phases were used to produce a protein Fourier synthesis. The Fourier grid intervals used were $1/60$, $1/120$, $1/80$ along a, b and c to give spacings of approximately $0.8 \text{ \AA}$ between adjacent Fourier points. Map scales were based upon unit cell dimensions 43.33 , 86.05 and $65.63 \text{ \AA}$ for a, b and c respectively.

A map on a scale of $0.5 \text{ cm/\AA}$ was produced first, with x and z

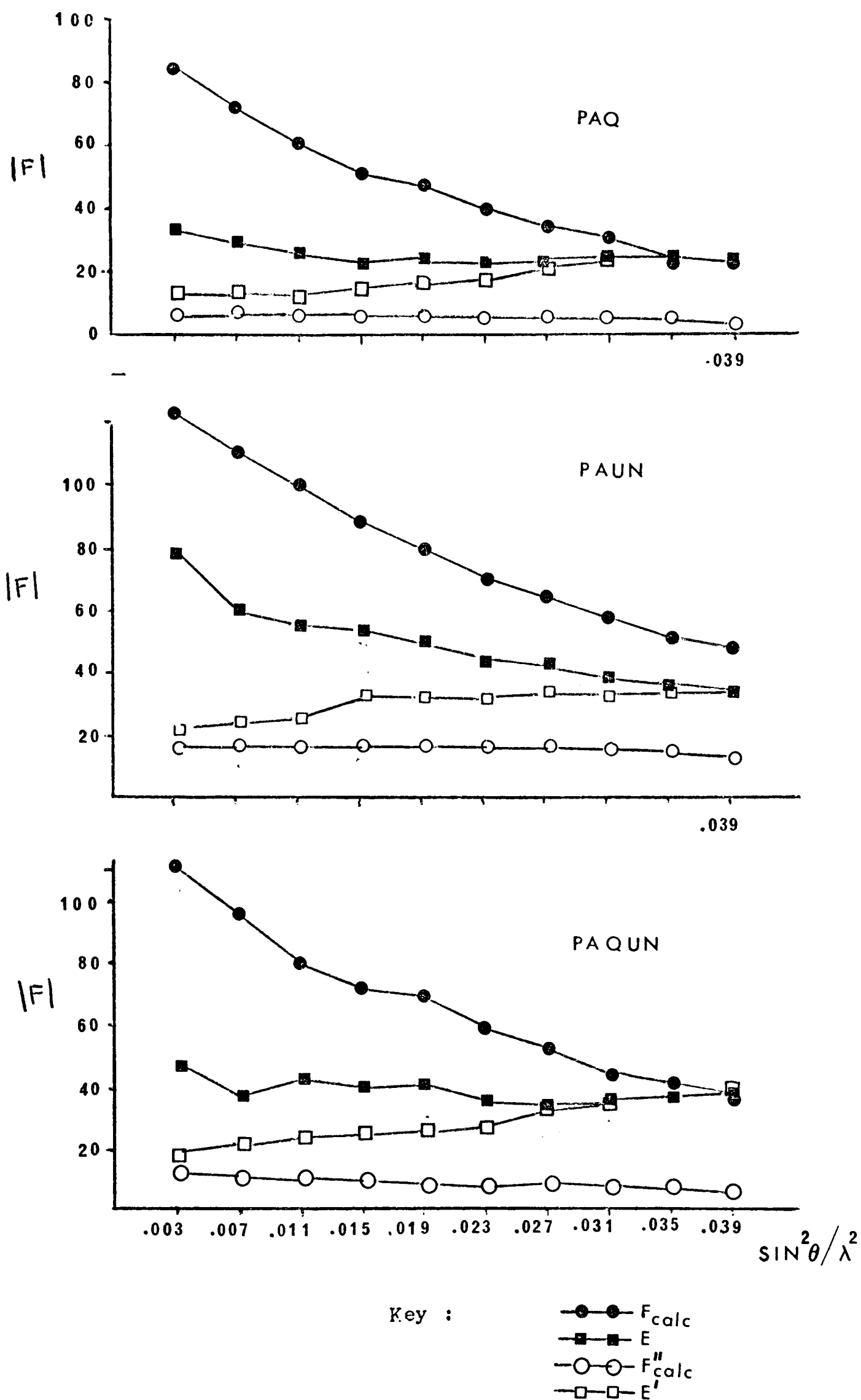


Fig. 40 Plots of heavy atom values of F_{calc} , E (isomorphous), E' (anomalous) and F'_{calc} using the final phase set. The E values are lower than those used to produce the phases and are probably under-estimated as discussed in the text. The vertical scale should be multiplied by a factor of two to give approximate magnitudes in electrons.

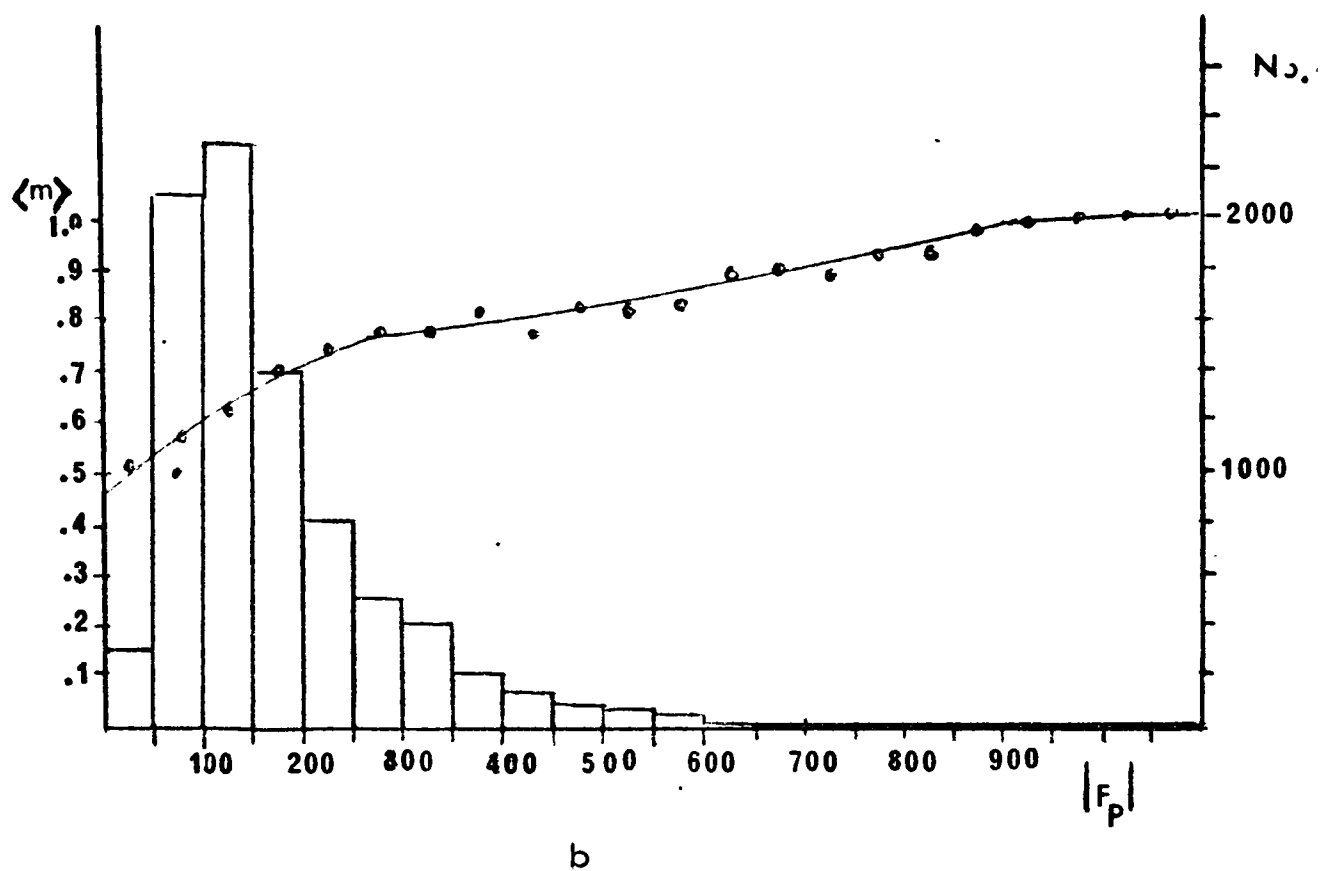
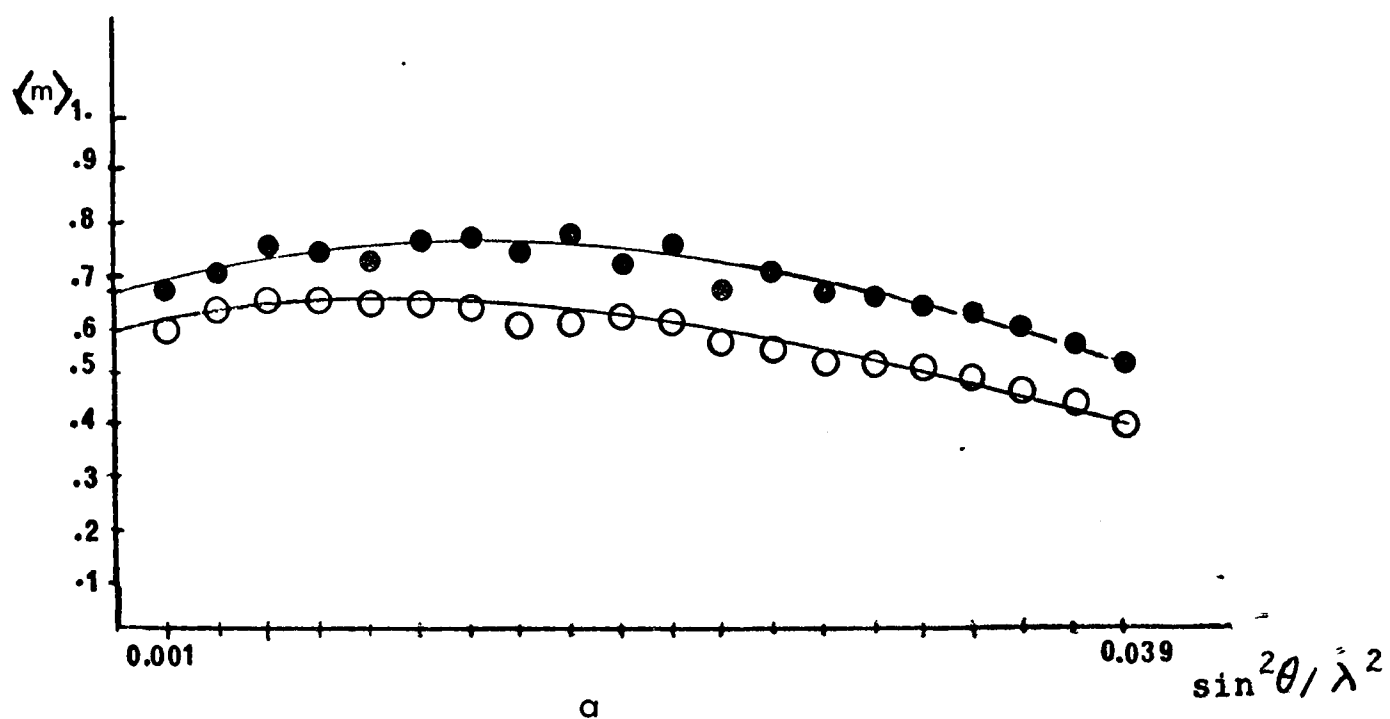


Fig. 40 Plots of mean figure of merit as a function a) of resolution and b) of $|F_p|$. The histogram in b) shows the frequency distribution of $|F_p|$ in intervals of 50.

All reflections with a F_p of 600 or greater have a mean figure of merit of 0.35 to 1.0. The slight fall in mean figure of merit at low angles in a) is due to the absence of phase information from PAQUN, which was not measured below a resolution of 7 Å.

repeat distances of 21.65 and 32.94 cm and an inter-sheet separation of 0.4 cm. Large scale maps of $2 \text{ cm}/\text{\AA}^0$ for model building using a Richard's optical comparator were produced with x and z repeat distances of 86.66 cm and 131.36 cm and an inter-sheet separation of 1.434 cm.

All the maps were contoured at equal intervals of 25 units starting from 25 units above the zero level. The Fourier output was multiplied by a factor of 1/750, so that given an absolute scale factor of 2, these units correspond to intervals of $0.3 \text{ e.\AA}^{-3}$. The Fourier syntheses sampled the cell from

$$x = 0 \text{ to } 1.0$$

$$y = 0.5 \text{ to } 1.0$$

$$z = 0.75 \text{ to } 1.75$$

The prealbumin tetramer is centred at $(\frac{1}{2}, \frac{1}{2}, \frac{1}{4})$. Contoured sections were hung in the optical comparator with the crystallographic two-fold axis (z) vertical and the x direction in the horizontal plane of the mirror.

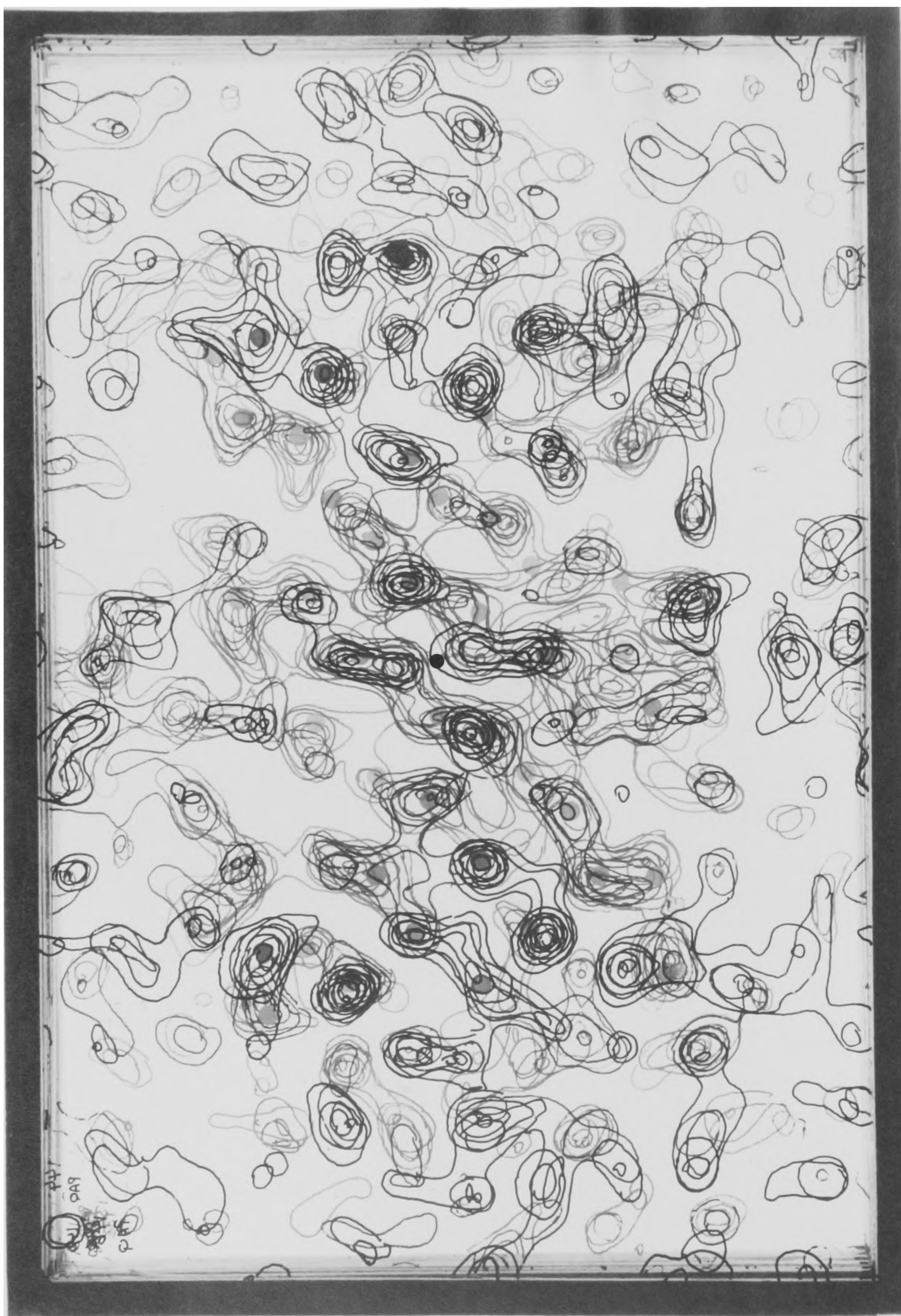


Fig. 41 View through 6 sections of the protein Fourier synthesis (0.5cm to 1Å). The long axis is z and the short one x, the local axis y' emerges from the centre of the map.

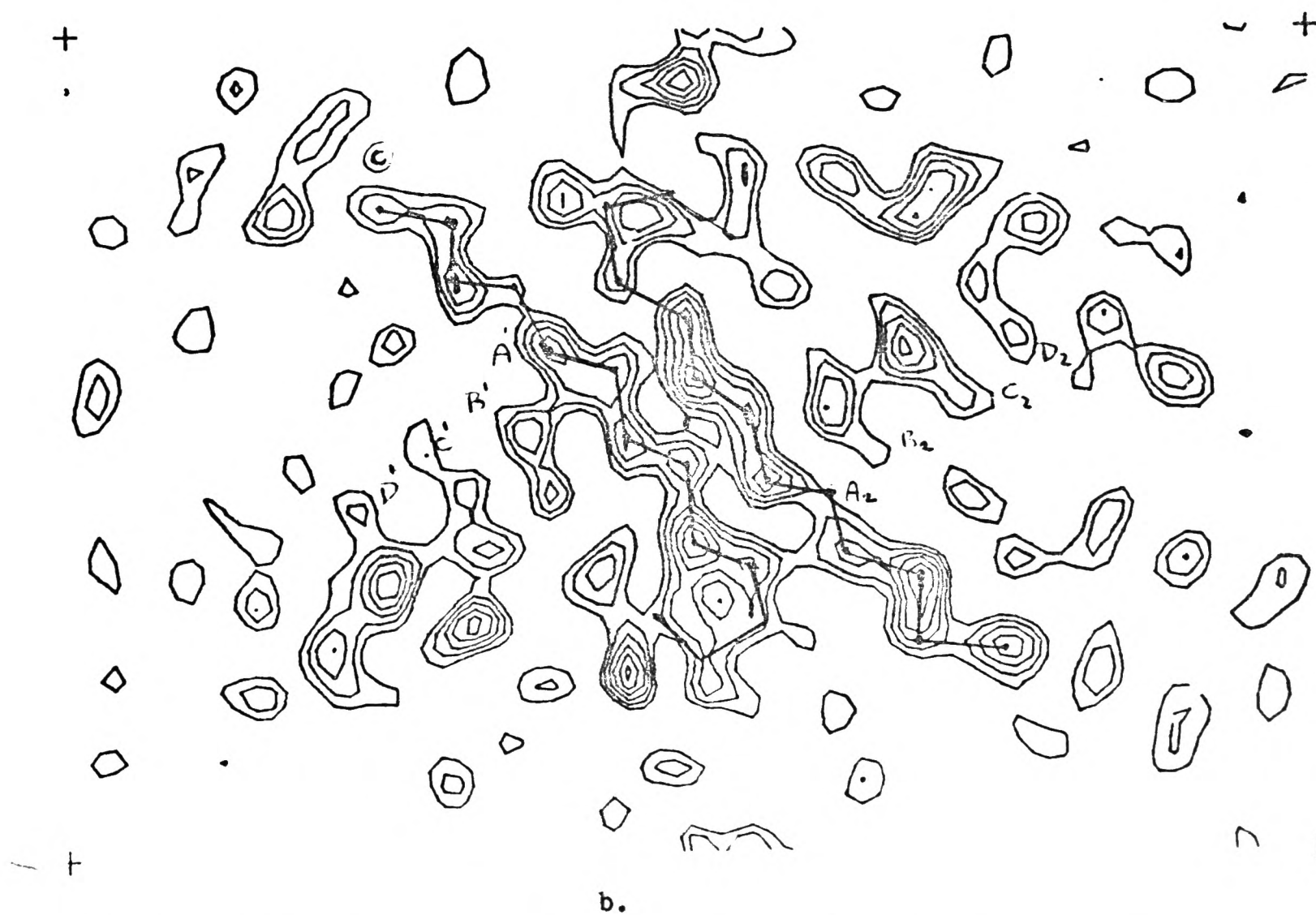
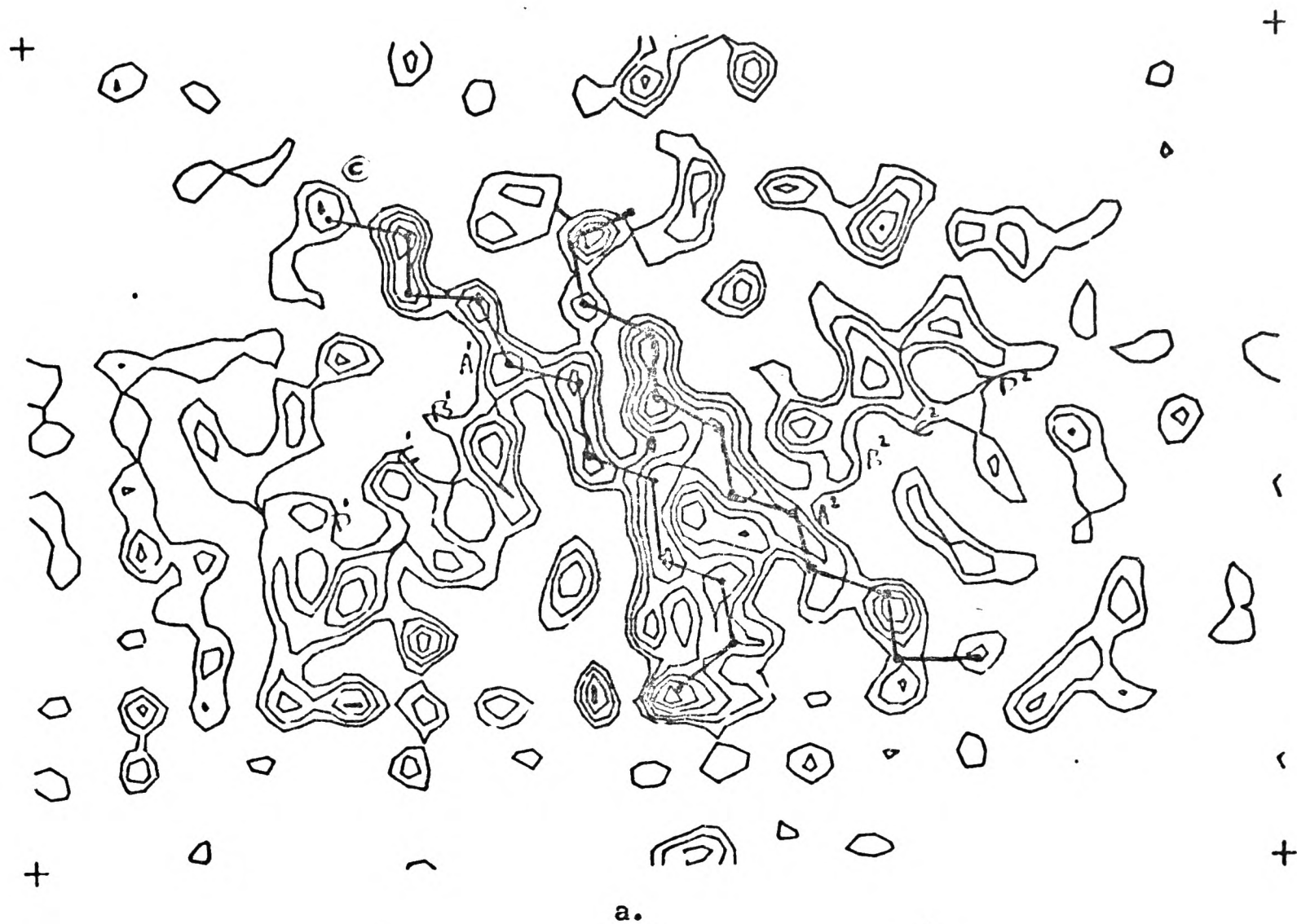


Fig. 42 Fourier section $y=17/30$ a) from 3 derivative phases.
 b) from single derivative phases (PAQNEW).

CHAPTER 6

The main chain conformation of prealbumin and the α -carbon positions.

In the following discussion the two protein subunits related by local symmetry are termed a_1 and a_2 and their exact symmetry equivalents are a'_1 and a'_2 . The prealbumin tetramer is then represented by $a_1 a_2 a'_1 a'_2$

The 2.5 Å synthesis revealed the subunit boundaries for the first time. The clearest intramolecular boundary is that between $a_1 a_2$ and $a'_1 a'_2$ which also contains the crystallographic two-fold axis. The gross location of the two dimers $a_1 a_2$ and $a'_1 a'_2$ is on either side of the slot in Fig. 1. The high resolution map shows that the low density bridges seen at the top and bottom of this slot at 6 Å, are loops of chain which cross the vertical planes (x,z) containing the two-fold axis z, make a close contact with the opposite dimer and return to the parent subunit. These four loops form the closest approach of $a_1 a_2$ to $a'_1 a'_2$. When these loops are included, the entire dimer $a_1 a_2$ lies within the real space defined by;

$$0.88 < x < 1.12 \quad 0.458 < y < 1.0 \quad 0.75 < z < 1.75$$

and the complete tetramer within;

$$0.88 < x < 1.12 \quad 0.0 < y < 1.0 \quad 0.75 < z < 1.75$$

In this discussion a_1 is closest to 1.75 in z.

The boundary between a_1 and a_2 is very much more intimate, and was not fully defined until the main chains within the $a_1 a_2$ dimer had been traced. The association of a_1 with a_2 is very close around the local symmetry axis y' which relates these subunits (horizontal axis in Fig. 1). Fig. 41 shows a view through 6 superposed sections of the high resolution synthesis, starting at the section $y = 130/240$. The faint contours are farthest away in y. A trace of the local axis y' is indicated in the centre of the diagram. The density seen on either side of y' has been interpreted as two extended anti-parallel chains, one contributed by each subunit. Small opaque discs mark the approximate positions of α -carbon atoms, placed 4 Å apart in the Fourier map.

The distance between these chains at the point closest to the local axis y' is $4.5\text{--}5\text{ \AA}$ and this spacing continues for $14\text{ \AA}$ in either direction.

It was concluded that these chains form hydrogen bonds to give two strands of anti-parallel β -pleated sheet. The amino acids in this secondary structure project alternately above and below the plane of the sheet and this regular alternation is clearly visible in the prealbumin map. The Fourier in fig. 41 shows that three further strands of extended chain make up a β -sheet containing 8 strands, with four contributed by each subunit. This sheet, together with the loops of chain mentioned, forms the $\alpha_1\alpha_2 - \alpha'_1\alpha'_2$ interaction and the two sheets lie within the dense walls of the low resolution synthesis.

Fig. 42 shows the single Fourier section $y = 17/30$, (a) from phases corresponding to 3 derivatives and (b) from the single isomorphous plus anomalous derivative PAQNEW. Very careful measurements of the anomalous scattering have been made in this derivative and it gave a cleaner Fourier in many regions (like the one shown) than the 3-derivative solution. The chains in different subunits are shown lettered A_1, A_2, B_1, B_2 etc. and the projection of the line joining the α -carbons is shown on β -strands A_1 and A_2 . Although the chain A runs for virtually its entire length in the plane of section $y = 17/30$, the strands B, C and D progressively twist away from this plane, so that the two sheets D_1 and D_2 make an angle of 90° to each other. This twisted β -sheet secondary structure has been observed in many proteins, but the best comparisons are provided by 2-zinc insulin (Adams et al., 1969) and Concanavalin A (Hardman and Ainsworth, 1973).

In the insulin dimer, the B chains form an anti-parallel β -sheet around one of the non-crystallographic two-fold axes. In Concanavalin A, the monomers interact to form two opposed dimers around three orthogonal crystallographic two-fold axes. Each dimer contributes 6 anti-parallel strands to form a 12-stranded, twisted β -sheet, rather similar to the 8-stranded sheet observed in prealbumin.



Fig. 43 The image of the partly-assembled β -sheet region (A,B,C) superimposed upon the Fourier density in the Richard's optical comparator. The β -strands make an angle of approximately 45° with the vertical axis(z).

The 'sense' of the twist of the prealbumin sheet is similar to that seen by other workers (Chothia, 1973). Looking down any strand, there is a right-handed twist of the peptide planes. In prealbumin the angle between successive peptide planes is not in general constant, but in strands B to D there is an average twist of 15° to 20° to the right between adjacent planes. It has been suggested that this specific twist reduces the non-bonded interaction of adjacent carbonyl and side-chain grouping, lowering the free-energy of the sheet (Chothia, 1973).

The appearance of the Fourier (Figs 41 & 42) in the sheet region is very similar to the published Fouriers of Carbonic anhydrase and Concanavalin A. The zig-zag appearance of main chain density in the plane of the sheet is interpreted by these authors to arise from the regular alternation of carbonyl oxygens. In this region of the prealbumin Fourier the carbonyl groups were at their clearest, perhaps because of the rigidity of the secondary structure.

After the identification of these features from the small scale Fourier of the protein, a map on the 2 cm to $1 \text{ \AA}$ scale of the Cambridge Repetition Engineers skeleton atomic models was produced and model-building started using a Richard's optical comparator. Only those sections containing the $\alpha_1\alpha_2$ dimer (the asymmetric unit) were traced. Fig. 43 shows the image of protein chains $B_1A_1A_2B_2C_2D_2$ superimposed upon the Fourier map. In this view the crystallographic z axis is vertical and the x axis is in the horizontal plane of the mirror. Density corresponding to α_2 is uppermost. Model-building commenced with the residues closest to the local axis y'. These are internal to the dimer and very clearly defined, so that given the image of the carbonyl group it was possible to define the direction of the chain. Initially, all the strands in the 8-membered sheet were built anti-parallel. A tight, "hairpin" loop joins A to B and A_1 and A_2 are anti-parallel because the axis y' passes between them.

At first the main chain was built into the density with α -carbons at

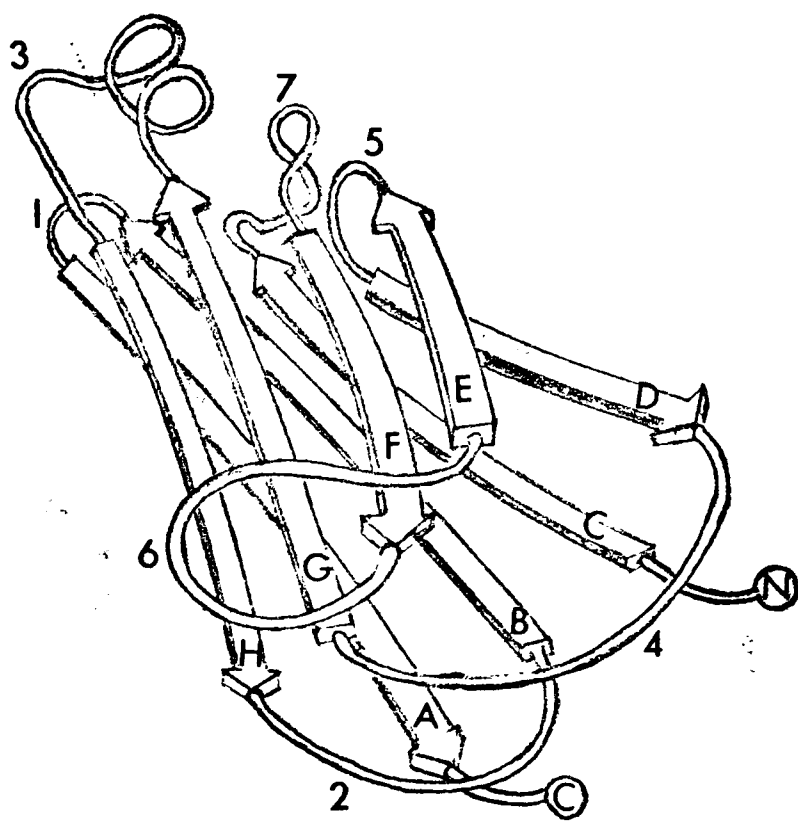


Fig. 44 A schematic drawing of the prealbumin monomer, showing the main features of the secondary structure.

the base of obvious side-chain density, or at bends in the density. It soon became apparent that nearly all the main chain in this protein is fully extended. Where the β -sheet runs perpendicular to the plane of the mirror, the pleat is very clear. The density runs normal to the mirror for 3 to 3.5 $\text{\AA}$, then in the plane of the mirror for 3.5 $\text{\AA}$.

Four other strands of chain in an extended β -structure form a sheet behind the first, and spaced 10 to 11 $\text{\AA}$ from it. A schematic drawing of the complete main chain is shown in Fig. 44 and a stereo pair in fig. 45a. These two diagrams should be used in following the discussion.

The two chains 'H' in the second sheet region E-H do not appear to form hydrogen bonds close to the local axis y' as do the A² chains in the first sheet, as the H chains are separated by 6 to 7 $\text{\AA}$ at this point. This separation is, however, reduced to the normal hydrogen-bonding distance of 5 $\text{\AA}$ where they run anti-parallel at a point 10 $\text{\AA}$ distant from y' .

The H-chains are hydrogen-bonded to the sheet composed of strands E, F and G. E and F clearly form a regular set of hydrogen bonds and the resulting pleat is easily seen, but F passes over G at a rather steep angle of 40° and may form fewer hydrogen bonds to it.

The chains H and G are fairly clearly connected by a loop (3), part of which forms a right-handed helix of 2 turns close to the β -strand G.

In the 6 $\text{\AA}$ structure this feature is defined as a loop, one limb of which is composed of a short rod, reminiscent of α -structure.

The remaining four loops which join the 8 strands of β -structure are less well defined, particularly loop (6) which joins the strands E and F. This loop extends well clear of the β -sheet formed by strands E-H, into a region of low contrast, which appears to be solvent between adjacent molecules. A continuous, but weak, density exists for this loop in the 3-derivative map.

In Fig. 44 the secondary structure is shown. β -structure by flat arrows which point towards the C-terminus and indicate the twist of the sheets. The seven joining loops are shown as close to the observed positions

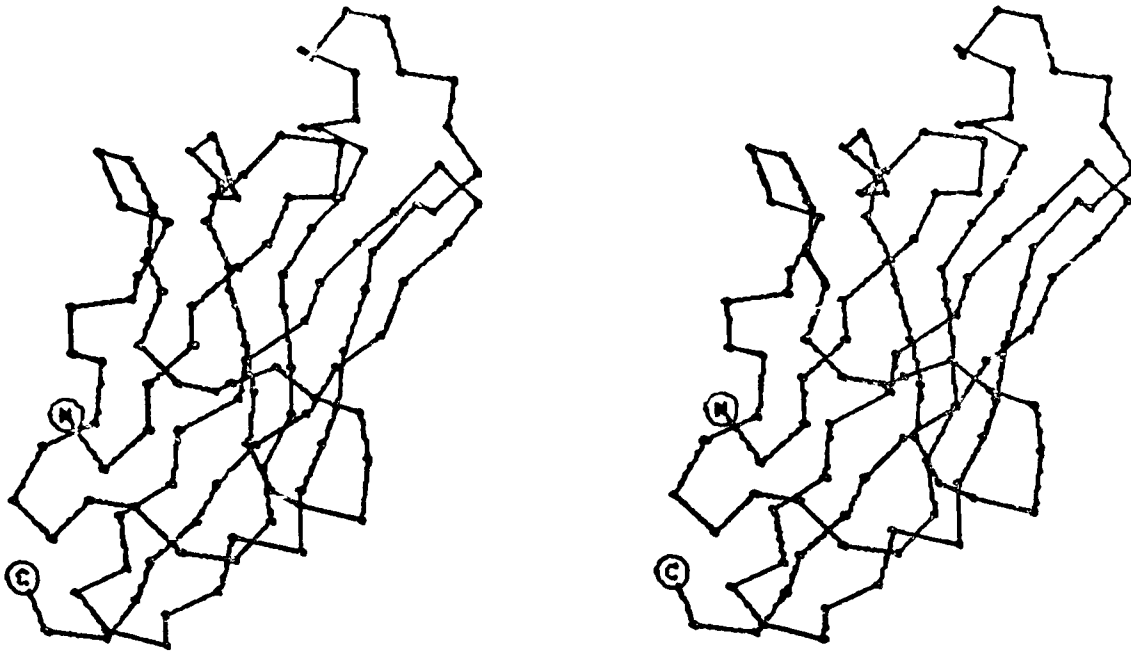


Fig.45a The prealbumin monomer α_1 . The C and N termini are indicated by the letters C and N.

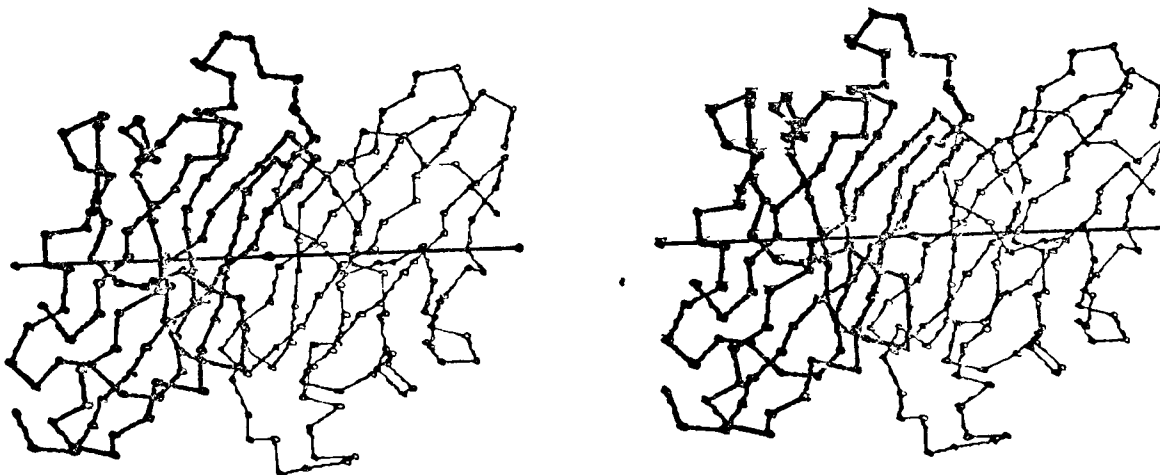


Fig.45b Stereo pair of the prealbumin dimer $\alpha_1\alpha_2$ showing the subunits not related by crystallographic symmetry. The two-fold axis z is shown with an indication of the local 222 point.

as possible, allowing for perspective and the short helix in loop 3 is indicated. Fig 45a shows a stereo of the monomer. In this diagram the edges of β -strands A and H are responsible for the $\alpha_1\alpha_2$ interaction in the dimer shown in Fig. 45b. Loop (7), the 'tetramer loop' is involved with the underside of strands A to D in the $\alpha_1\alpha_2 - \alpha'_1\alpha'_2$ interactions in the tetramer shown in Figs 46a and 46b.

The monomer represented in Fig. 45a has dimensions, measured between extremes of the polypeptide backbone, of $43 \text{ \AA}$ (in the long direction of the β -strands), $25 \text{ \AA}$ (perpendicular to the β -strands in the plane of the sheets) and $25 \text{ \AA}$ (normal to the β -sheets in the direction of y').

The $\alpha_1\alpha_2\alpha'_1\alpha'_2$ tetramer is then a globular unit of $44 \times 55 \times 50 \text{ \AA}$ in x, y and z .

The inter-sheet space in fig. 45b between ABCD and EFGH is a region of extremely high contrast in the Fourier maps, and appears to be entirely hydrophobic. This presumably causes the flattened appearance of the monomer, as a more globular arrangement would leave an internal cavity. Some of the clearest side chain density projects into this space from the sheets and the seven joining loops, again suggesting that the interior is inaccessible to solvent. The only "entrance" to this core is found between loops (4) and (6), where loop (6) turns away from the protein surface formed by chains E-H. However, this space appears to be partly filled by large side-chains from loops (4) and (6) and chains D and C.

The second monomer α_2 closes the hydrophobic core by hydrogen-bonding to chain A and part of chain H. It seems that the structural unit of prealbumin is the dimer, as the association of the $\alpha_1\alpha_2$ monomers is so close that the monomer is unlikely to have any separate existence, apart from the denatured state.

It is interesting to note that no two of the seven joining loops are similar, either in conformation or length, and apart from the helix-containing loop (3), appear to have no ordered structure. Assuming a total of 130 amino acids in each monomer, the estimated percentages of secondary

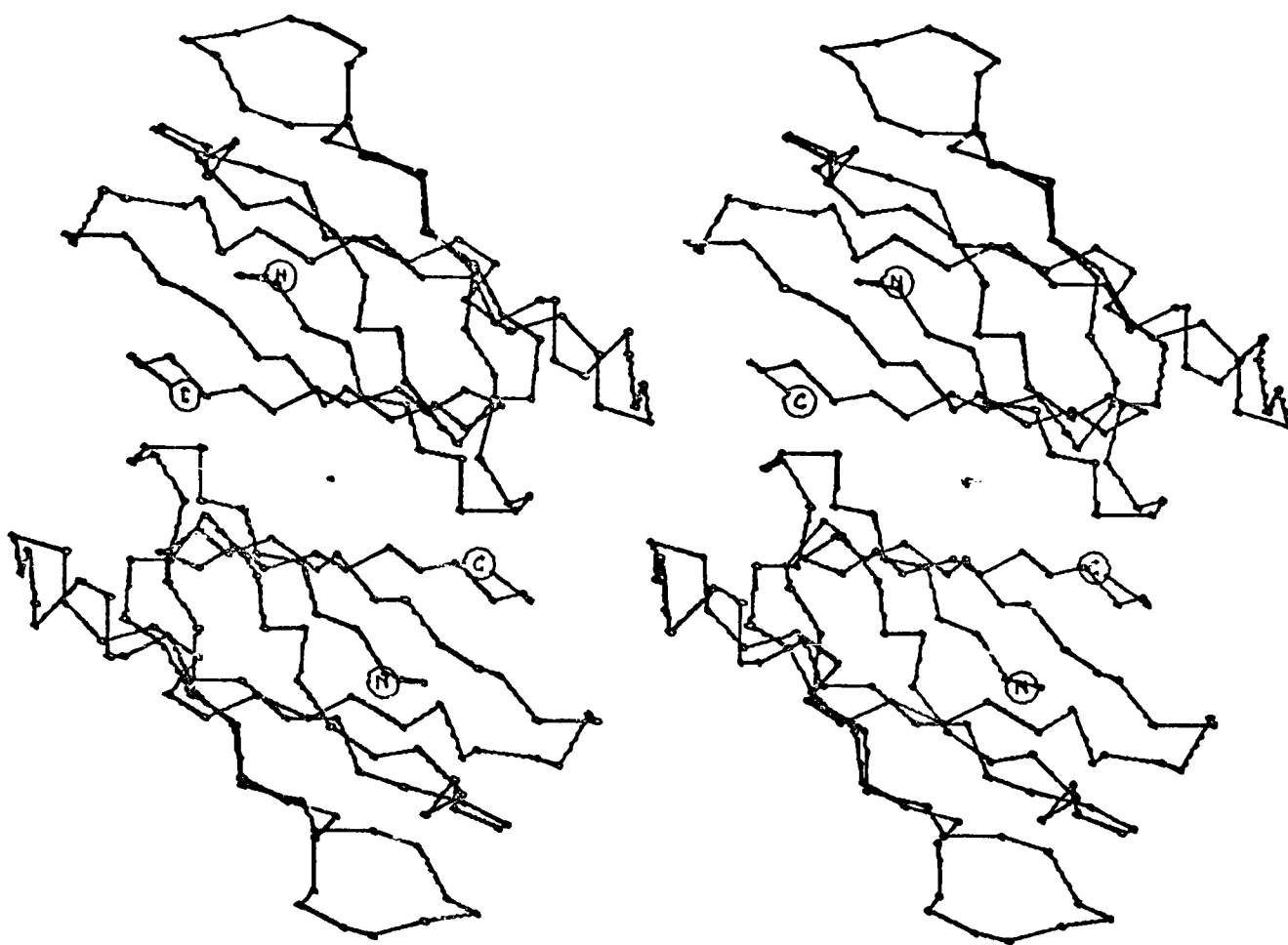


Fig. 46a The dimer $a_1 a_1'$ seen down the crystallographic two-fold axis
The interactions between the crystal symmetry related subunits
seem tenuous by comparison with the local symmetry-related
ones.

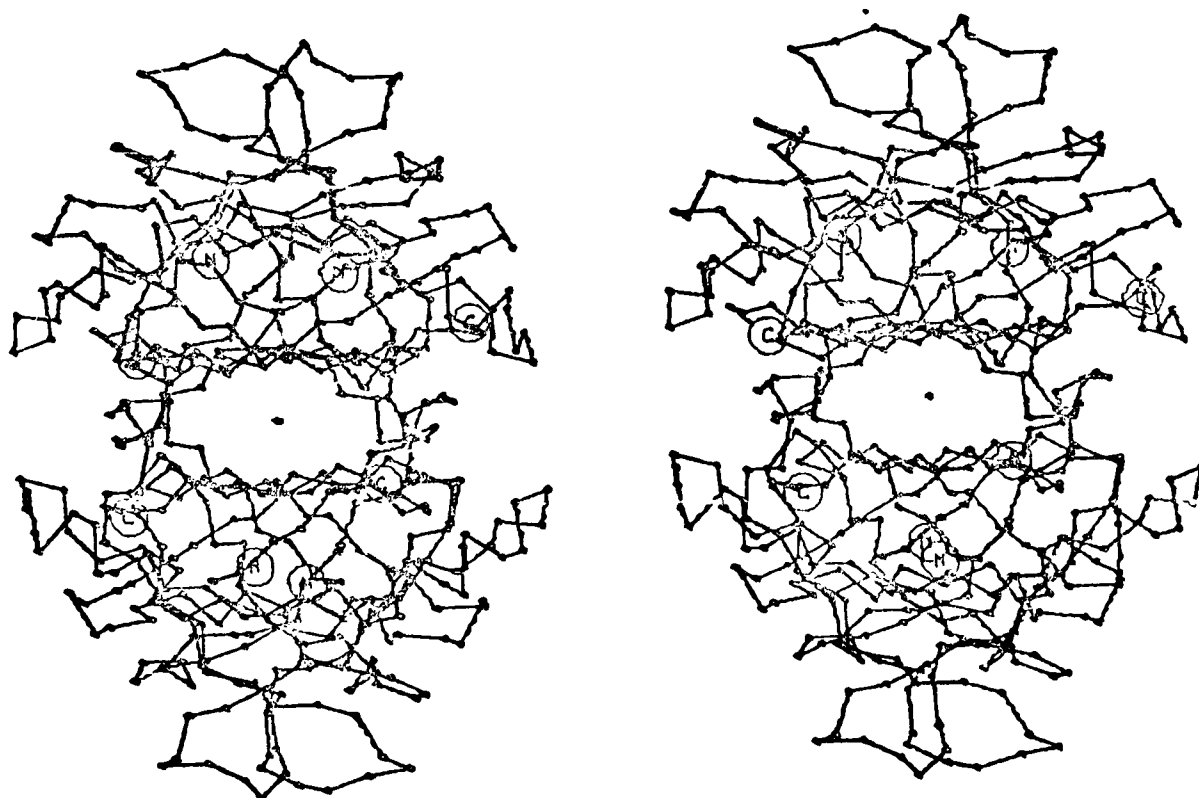


Fig. 46b The view down the crystallographic two-fold axis of the
complete prealbumin molecule. The diagram was generated by
rotating the $a_1 a_2$ dimer by 180 degrees about z.

structure are set out below.

Structure.	No. of residues.	X-ray results %	C.D results %
α -	10	8	—
β -	65	50	50
random coil	55	42	50

The loops (5) and (7) appear to contain 3 and 7 residues respectively and since 4 residues in loop (7) appear to make a close contact with the opposite monomer α'_1 , the β -strands C,D,E and F are drawn closely together in this region. The helix in loop 3 lies with its axis parallel to the closest β -strands (B and F) and with the rest of the loop, closes the interior of the monomer against the solvent. The opposite end of the monomer, formed by loops (2, 4 and 6) and the C and N termini, is less tightly drawn together. The internal side chains at this end are correspondingly larger than at the helix end. It appears, both from the appearance of the density and the regions of sequence fitted to the map, that the aromatic side chains phenylalanine, tyrosine, histidine etc. cluster at the larger end of the monomer, while the aliphatics segregate to the helix end of the monomer.

Sheet regions in prealbumin.

It is apparent that the β -sheet regions are responsible for the unusual stability of the protein, discussed in chapter 1. Fig. 43a shows that the extensive hydrogen bonding in the sheet essentially makes subunits α_1 and α_2 into one compact unit. There is some evidence that internal hydrogen bonds are formed, for example, between the two internal residues closest to the local axis y' on A_1 and A_2 . This density is currently fitted by tyrosine.

The hydrogen bonded sheet A-D is protected by association with the second dimer shown in Fig. 48b. Although this association probably further stabilises the protein, it is evident that solvent can penetrate

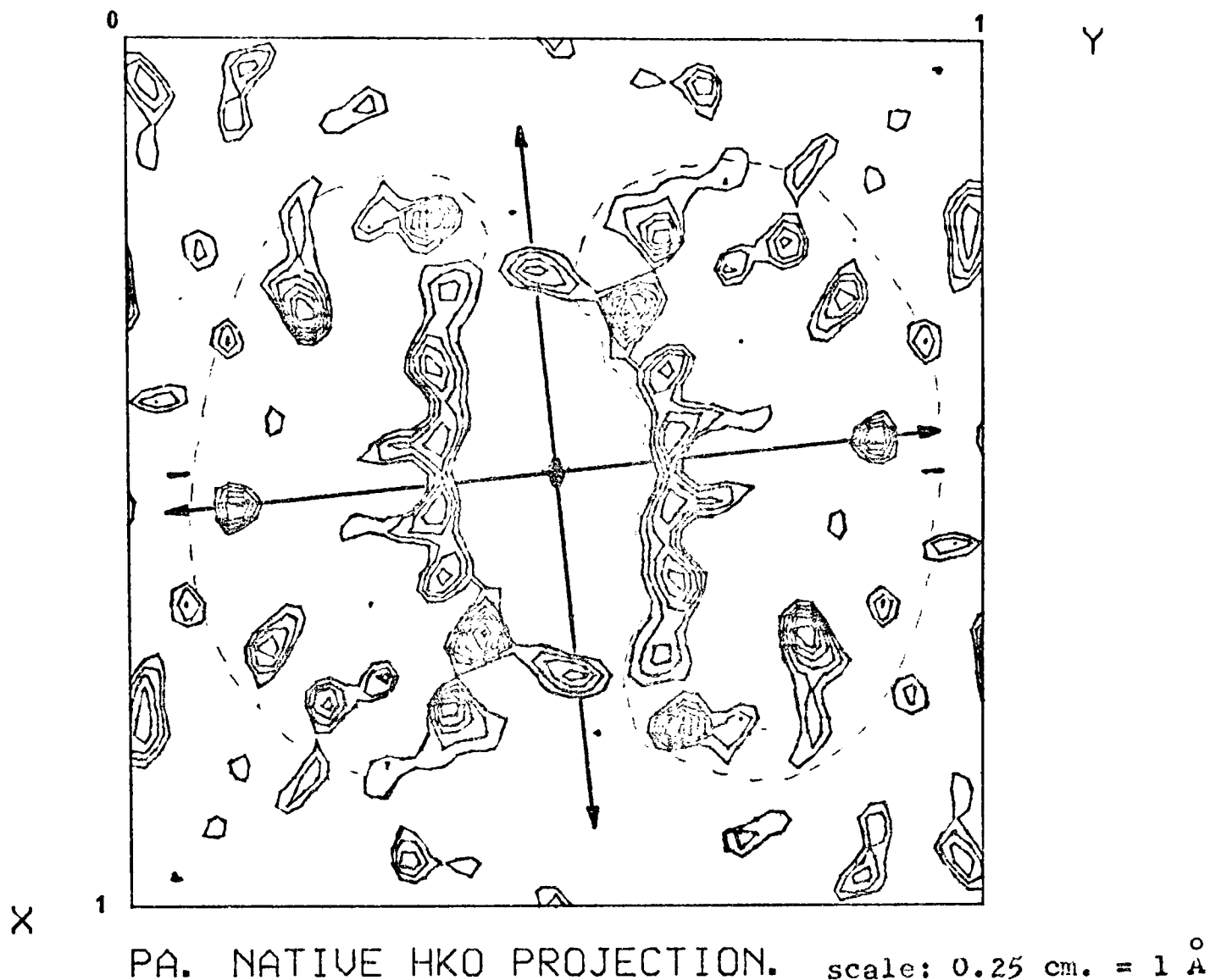


Fig.47 A projection fourier of the native protein at a resolution of $2.5 \text{ \AA}$. The phases are those taken from the final set of 3D phases from three derivatives.

It is possible, in retrospect, to identify the gross features of the protein structure from the centrosymmetric projections like the HKO zone shown here. In this projection the region of very low density around the crystallographic 2-fold axis (centre) corresponds to the slot between the two dimers. To left and right of this are dense regions where the pairs of monomers interact to form a continuous β -pleated sheet. Beyond this the density again falls, where the 3D synthesis indicates a hydrophobic space filled with side-chains. The local 2-fold axes have been placed on the projection from an inspection of the density. These make an angle of $6^\circ \pm 0.5^\circ$ with the cell axes, in good agreement with previous work.

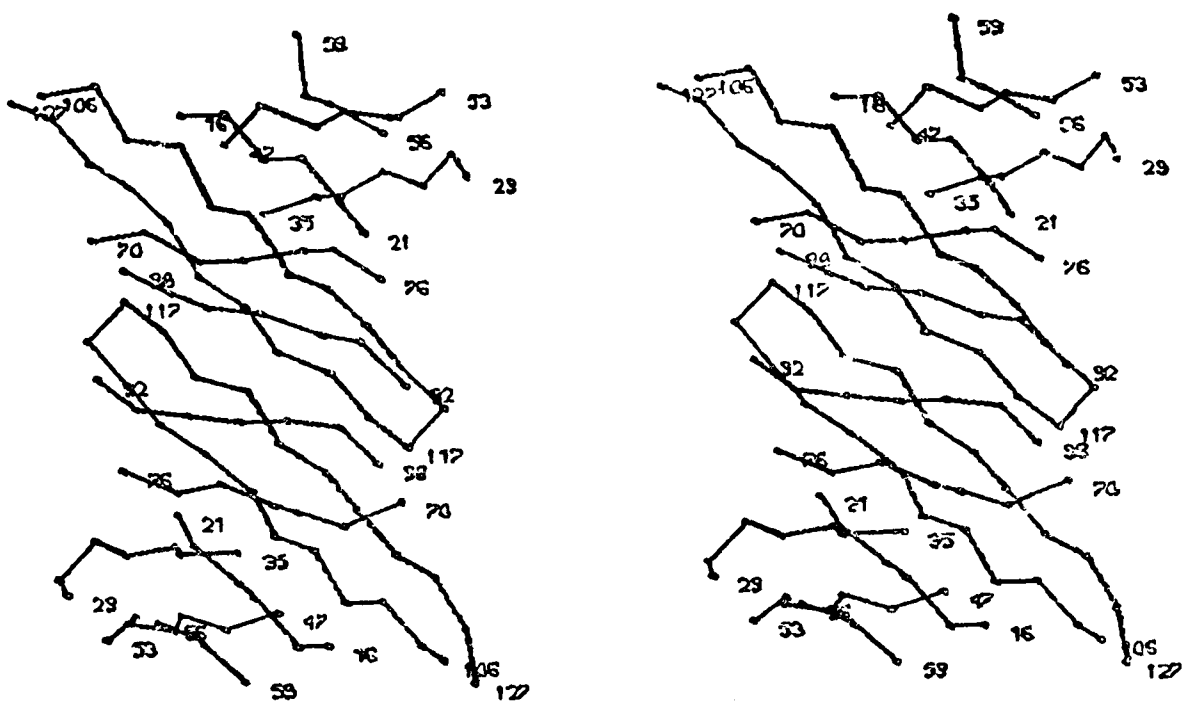
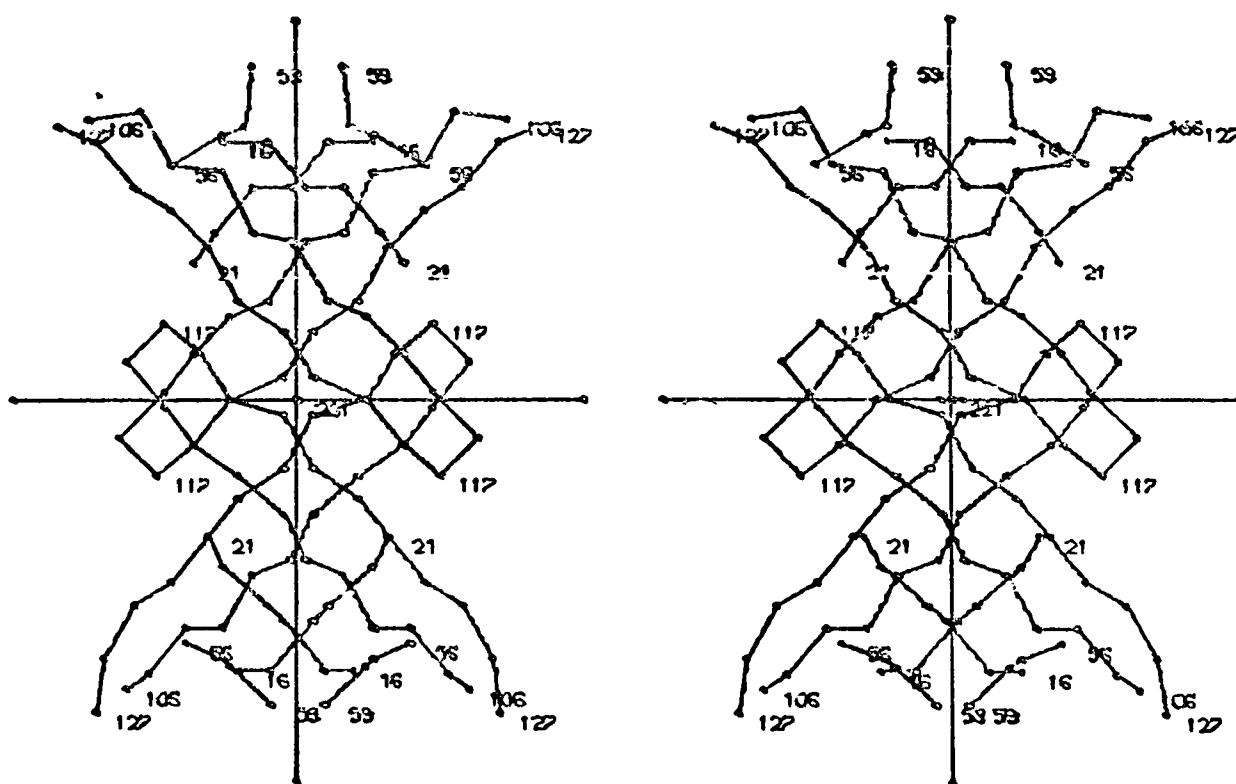


Fig. 48 The β -sheet regions in the $\alpha_1\alpha_2$ dimer. Only the most obviously hydrogen-bonded regions have been included and a significant proportion of the residues in β -conformations but with no obvious hydrogen bonds may have been omitted. The residue numbers refer to the key in Fig 50.



The β -sheet composed of the strands A,B,C and D has been isolated to indicate the interactions with the second dimer β -sheet across the crystallographic two-fold axis(vertical). The local axis x' is also shown (horizontal). This orientation corresponds to that in the Richard's optical comparator.

the funnel-shaped cavity between the sheets. The side-chains which project into the inter-sheet space near the molecular centre are all small. They appear, both from the map and from sequence data, to be either serine or threonine. As these residues are not sufficiently long to hydrogen-bond directly across the cavity, it may be that solvent molecules play an important part in the dimerisation.

A close parallel exists between the association of prealbumin and that seen in the C-domains of a λ -type Bence-Jones protein (Schiffer et al., 1973). In both these proteins, there are four regions of roughly parallel β -sheet separated by distances of $10 \text{ \AA}$, with two layers in each dimer. However, in this protein, the inter-sheet space appears to be filled with large side chains so that the solvent is excluded. Chapter 7 shows that the inter-sheet cavity in prealbumin forms the T3-binding sites, so the presence of solvent between the sheets is of functional importance in prealbumin.

It can be seen from fig. 44, that the only parallel strand in the β -structure is C ; the N-terminus. Also the association of α_1 and α_2 places the two C-terminal strands at the centre of the molecule. This suggests that the protein is very resistant to exopeptidase activity. The tight subunit association seems to rule out the speculated co-operative effects of prealbumin, that were originally advanced to explain the unique T3 binding site. The latest chemical evidence is in accord with this observation. Jaarsveld, et al., (1973) has shown that the RBP sites are independent and Pages et al., (1973) has indicated that the 2 binding sites for T4 analogues are independent.

It is interesting to note that the dense β -structure nearest the z axis can be picked out in a projection Fourier of the protein-containing only 300 terms, (fig. 47)

Association within the crystal

The approximately globular shape of the molecule leaves a considerable

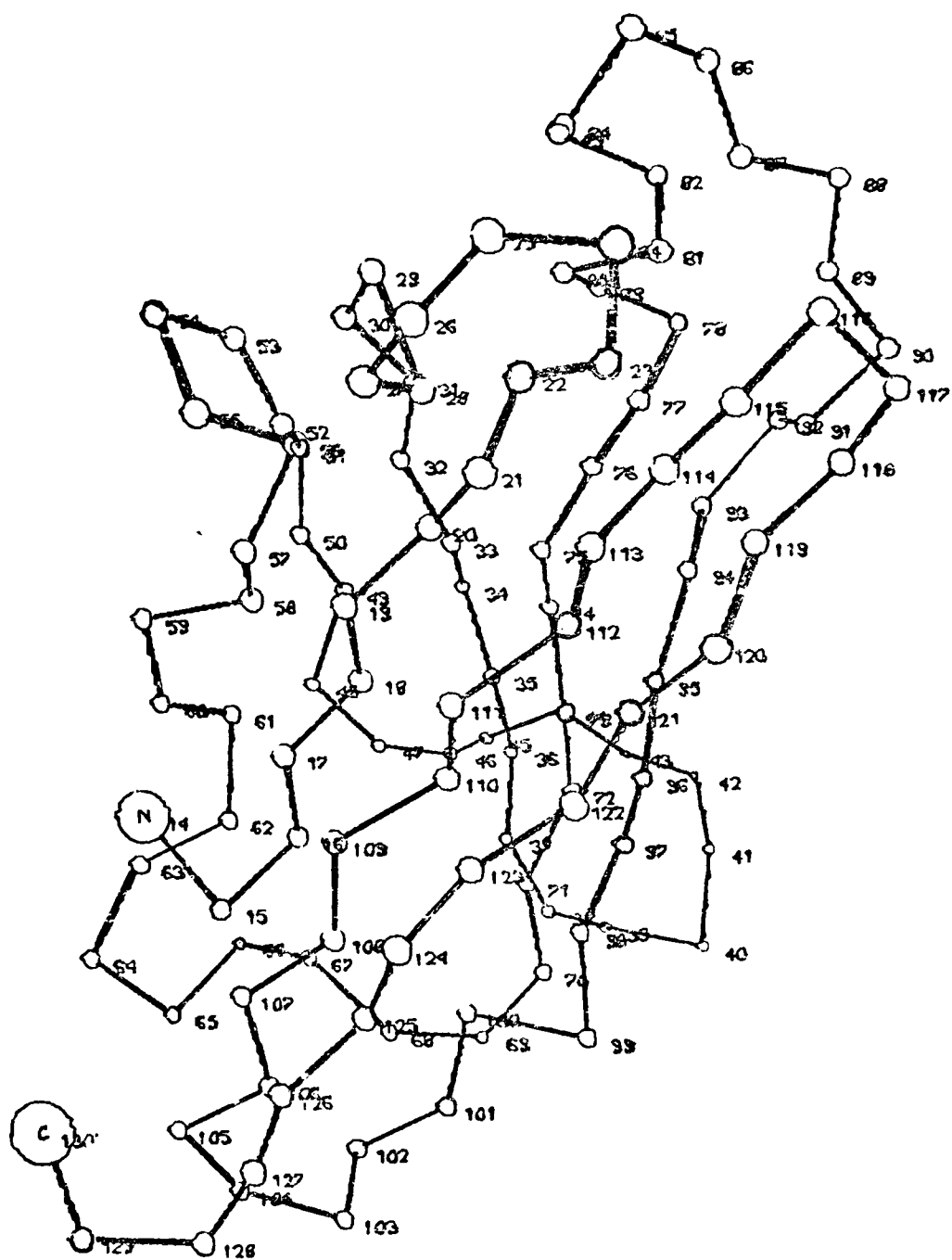


Fig. 49 This perspective diagram of the α_1 monomer shows the location of residues referred to in the list in Fig. 50.

volume in the crystal to be occupied by the solvent. However, several close-contacts with neighbouring molecules can be observed. The helix-loop (3) stands clear of the protein surface, giving the appearance of a 'handle'. The rotation of the tetramer away from the x and y axes of the cell moves the end of loop (3) sufficiently far away from the x,z plane for the loop (3) of the adjacent molecule to fit closely beside it.

The loop (7) has strong density corresponding to a side chain, which appears to point to a dense feature on the two-fold axis through $z=1.75$. This axis does not contain the molecular centre, so that the two-fold related molecule also has the same side chain, pointing towards the peak on the axis. The peak can only be an ion and the side chain density is fitted well by an arginine residue, which occurs in sequence at that point.

The greatest separation between molecules occurs in the y direction, where the two loops (6) project from each dimer. There is a region of very low contrast beyond these loops, up to the end of the unit cell at $y = 1$

Preliminary fit of the tryptic peptide sequence data.

A complete amino acid sequence for prealbumin has yet to be published. However, Dr G. Gonzalez and Dr R. Offord have sequenced four peptides (see fig. 3). Work on the sequence is also being carried out by Dr. S. Morgan and Canfield in the U.S.A. One difficulty of the present data concerns the number of residues in the protein. At present 126 have been accounted for in the x-ray map. It is hoped, with data from the two groups and the x-ray map, to arrive at a complete sequence for this protein.

The presence of two subunits within the asymmetric unit, allows two different views of the side-chain density corresponding to a single amino-acid. It was at first thought that this would be an interpretative aid, but the two subunits were eventually independently traced. This was mainly due to the difficulty of mentally 'rotating' the density about the

RESIDUE No.	RESIDUE.	SECONDARY
	TYPE.	STRUCTURE.

N-terminus.

14	cys - mercury binding	124	val
15	pro site.	125	val
16	leu	126	glu*
17	met	127	asx
18	val	128	
19	lys	129	
20	val	130	C-terminus.
21	leu		
22	asp		
23	ala		
24	val --		
25	arg*		
26	gly		
27	ser		
28	pro		
29			
30			
37			
46	pro		
47			
53			
56			
58			
70			
76	asp - uranyl binding site		
77	thr		
78	lys*		
79	ser		
80	tyr*		
81	trp*		
82	lys*		
83			
84			
85			
90			
99			
106			
107			
108	tyr*		
115			
116			
117	tyr		
118	ser		
119	tyr*		
120	ser*		
121	thr*		
122	thr*		
123	thr*		

A

Fig. 50 Residue location table. Those residues marked as * are good fits to the electron density.

local symmetry axis, in order to compare it. As the stereo drawings of the $\alpha_1\alpha_2$ dimer indicate, the main chain conformation of the subunits is closely similar, allowing for errors in placing the α -carbon atoms. This is clearly the right stage for an averaging procedure, since the independent chain tracing has not revealed any serious difference in the appearance or conformation of the subunits, and the averaging, if done accurately, should enhance the interpretation of the map. Until this is done, no complete assignment of the residues will be attempted. Accordingly, although over 50% of the residues have been tentatively assigned, only the clearest residues are included in Fig. 50. The angle made by the local axes with the cell axes has already been defined by the heavy-atom positions and rotation of the Fourier density is made easier by the co-incidence of one local axis with the z-axis of the unit-cell.

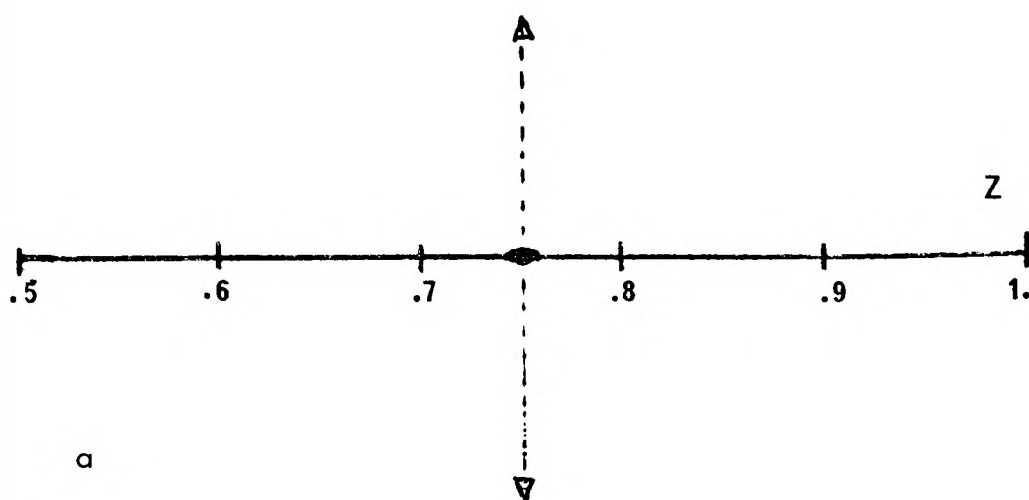
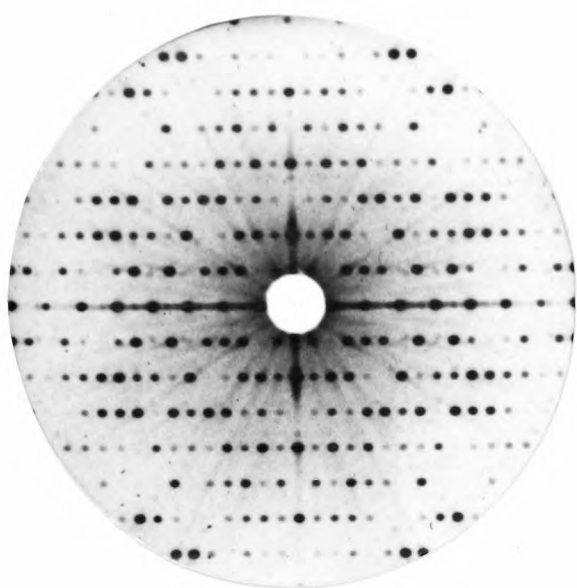
Fig. 50 lists the residue numbers and shows the location of the elements of secondary structure. The residue numbers do not necessarily co-incide with any partial, published sequence, but correspond to the numbers in stereo diagrams in this thesis. The first fit of sequence to the map was obtained at the C-terminus and this explains the lettering and numbering order of the β -strands and the loops. Although the exact number of residues is not definite, the C-terminal residue was numbered 130. The N-terminal strand has been traced as far as residue 14 which is the cysteine to which the mercury in PAQ is bound. Beyond this residue, the density is very weak, and although the chain has been traced right to the N-terminus in one subunit, this is not included in the stereo pairs of the monomer. The N-terminal residues beyond cysteine are rather polar and this nona-peptide may be partly solvated. It is interesting to note that the N-terminus may not contribute very much to the overall structure. This is because the surface 'skin' of chains is completed by strand C, and the proline residue before the cysteine appears to turn the N-terminus away from the surface of the monomer, breaking the run of sheet C.

CHAPTER 7.

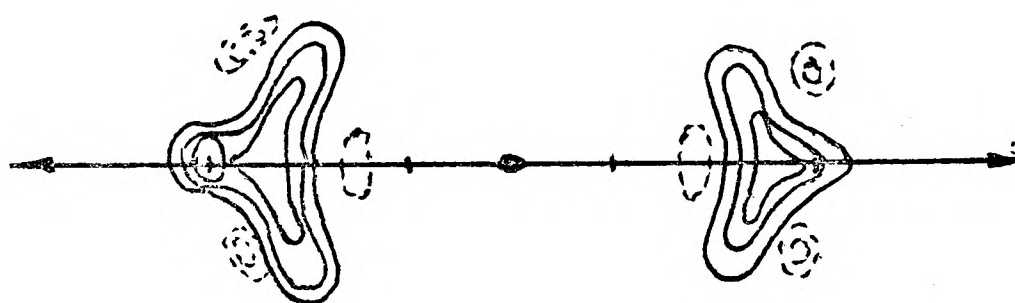
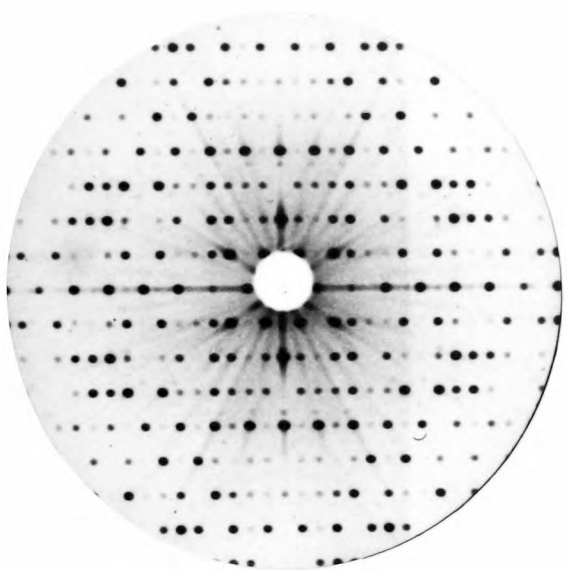
The interaction of T_4 and T_3 with prealbumin.

Although a widespread literature exists on the binding of the hormones L-3,5,3',5'-tetraiodothyronine (T_4) and L-3,5,3'-triiodothyronine (T_3) to prealbumin, until recently all attempts to observe binding to the crystalline protein have failed. Conventional diffusion experiments, where crystals have been allowed to soak for several weeks in 55% saturated ammonium sulphate solutions of T_4 or T_3 at pH values between 6 and 7, failed to produce observable changes in the native diffraction pattern. An experiment in which prealbumin crystals were cross-linked with 2% glutaraldehyde solution, allowed the binding of T_4 to be investigated under salt-free conditions at pH 11, where T_4 is completely soluble (Hedges, 1972). No binding was observed. Gonzalez (1972) was able to obtain prealbumin crystals from a solution containing T_4 , but their diffraction patterns were indistinguishable from that of the native protein.

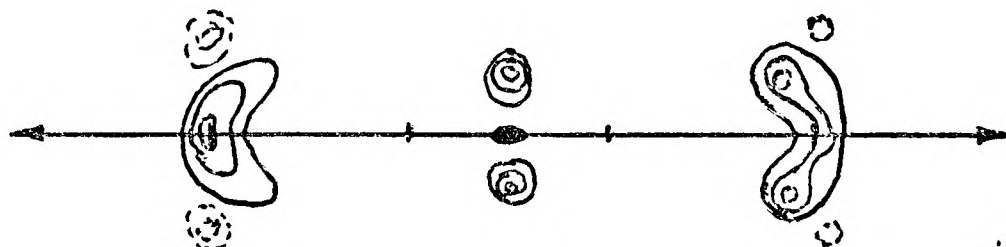
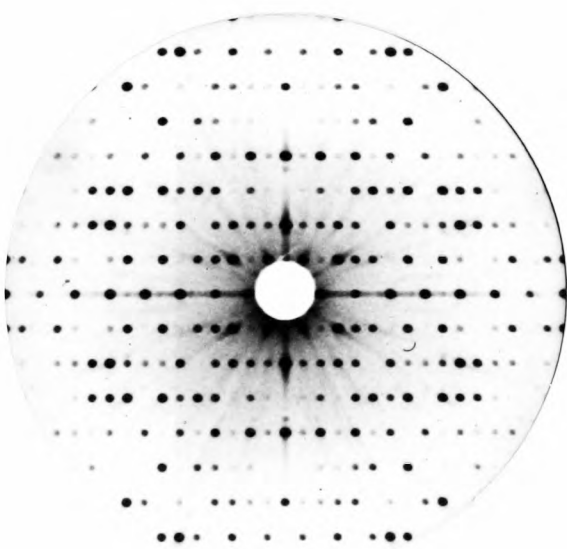
Recently both T_3 and T_4 have been bound to crystalline prealbumin by the author. The successful binding was achieved by soaking prealbumin crystals for between one and two months in 60% saturated ammonium sulphate solution buffered by tris hydrochloride buffer to either pH 7.4 or 8.6. The sodium salt of T_3 and the free acid of T_4 were used in these experiments. The hormones were dissolved at pH 9 in distilled water and after standing, the solutions were made 60% saturated with respect to ammonium sulphate and the pH adjusted to 7.4 or 8.6 with ammonium hydroxide or hydrochloric acid. The precipitated hormones were not filtered from the supernatant. Prealbumin crystals were placed in the solutions and the solutions were stored in darkness. Over the period of the experiment, crystals were removed for examination of the HKO zones of the diffraction pattern. At least one month of contact with the hormone suspension was necessary to produce observable changes in the pattern. The centre portion of



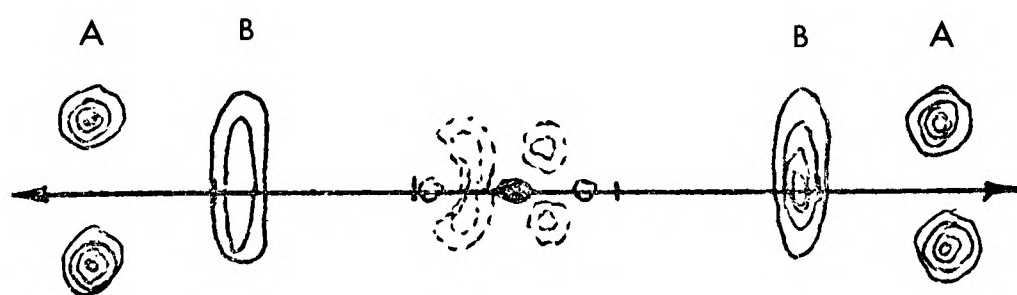
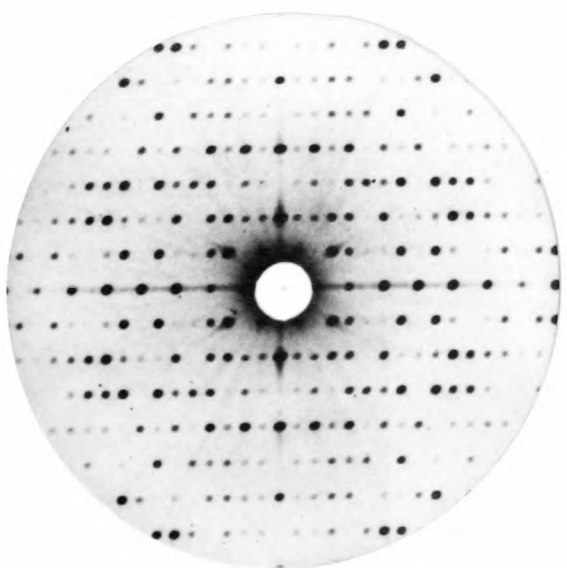
a



b



c



d

Fig. 51. Diffraction patterns of the RNO zone and part of the OKL difference projection syntheses of a) IAQTT b) iodoacetic acid c) gold cyanide d) T3.

an 18° precession photograph of the HKO zone of prealbumin- T_3 is reproduced in fig. 51 d, and should be compared with that of the native in fig. 51 a. As a point of interest the diffraction patterns (HKO) and projection difference Fourier (OKL) of iodo-acetic acid (b) and gold cyanide (c) are also shown. It can be seen that these

ions bind in the same general position as T_3 (d). The HKO diffraction pattern is shown because this is a useful diagnostic of binding on the crystallographic c axis. The characteristic intensity changes are 1 3 0, which with 2 0 0 forms a hexagon of equally intense reflections round the beam-stop, and 2 4 0 which increases dramatically when substitution occurs on the axis.

As the binding constants of the hormones are high ($1.3 \times 10^8 \text{ M}^{-1}$ for T_4 and $1.2 \times 10^7 \text{ M}^{-1}$ for T_3) in physiological media, it is not easy to understand why binding is so slow in the crystal. It may be that at the pH employed, the concentrations of T_4 and T_3 in 60% saturated ammonium sulphate solution are negligibly small, as these compounds are very poorly water soluble at neutral pH. It is also possible that the sulphate ion has an inhibitory effect on the binding, since in the native protein map there is significant density in a position subsequently found to be the hormone binding site.

Figure 51 d shows the $3 \text{ \AA}$ OKL difference Fourier projection of T_3 . Two distinct sets of peaks appear and are reproduced to form pairs by the two-fold axis indicated. The projection map of T_4 (not shown) is very similar, except that the peaks (A) are higher and the feature (B) is resolved into two discrete peaks. The HKO projection shows that these features are centred on the crystallographic two-fold axis, and in this zone, they form a 'cross' of density with the arms at 45° to the x and y axes. The photographic work indicates two binding sites for both T_3 and T_4 . These are on either side of the molecular centre and are related by local symmetry. The pairs of

features A are $2\sqrt{3}$ Å apart.

The most obvious explanation of the observed peaks is that they represent the iodine atoms (three in the case of T_3 and four for T_4). In projection it would be extremely unlikely that any of the light atoms in these hormones would make an observable contribution to the electron density. The position of the ~~two~~ binding sites on the crystallographic two-fold axis presents a problem of interpretation. Neither T_4 nor T_3 possess two-fold symmetry. However, no change in space group is apparent from the diffraction pattern. Two possible explanations exist for this apparent violation of symmetry. First, if the interaction of the hormones with the protein is through groups on both which do have exact two-fold symmetry. This might be the phenolic hydroxyl group, which could interact with 2-fold symmetry-related residues on the protein. This type of binding has been observed by Perutz (1970), where the asymmetric 2,3-diphosphoglycerate molecule is bound on a crystallographic diad axis, by virtue of its two phosphate groups.

Secondly, two ~~orientations~~ of the hormones may actually be present in the crystal, so that what is ^{seen, is} a statistical average of these. There are unlikely to be more than two ~~orientations~~, because of the small number of discrete peaks observed in projection syntheses. It is not possible for two hormones to be present in the binding site at the same time, as measurements on the high resolution model of the protein show that there is insufficient room. This means that the two-fold symmetry is broken by the hormone, and a random choice of association with either of the two-fold related subunits is made.

In order to define the binding site more precisely, data to $\sqrt{3}$ Å resolution were collected from the crystal that gave the $\sqrt{3}$ Å projection Fourier in fig. 51d. 1 400 reflections representing 700 Friedel pairs were collected with a count time of 100 seconds for each reflection. The R factor between Friedel-related reflections was 4%. The 3L difference

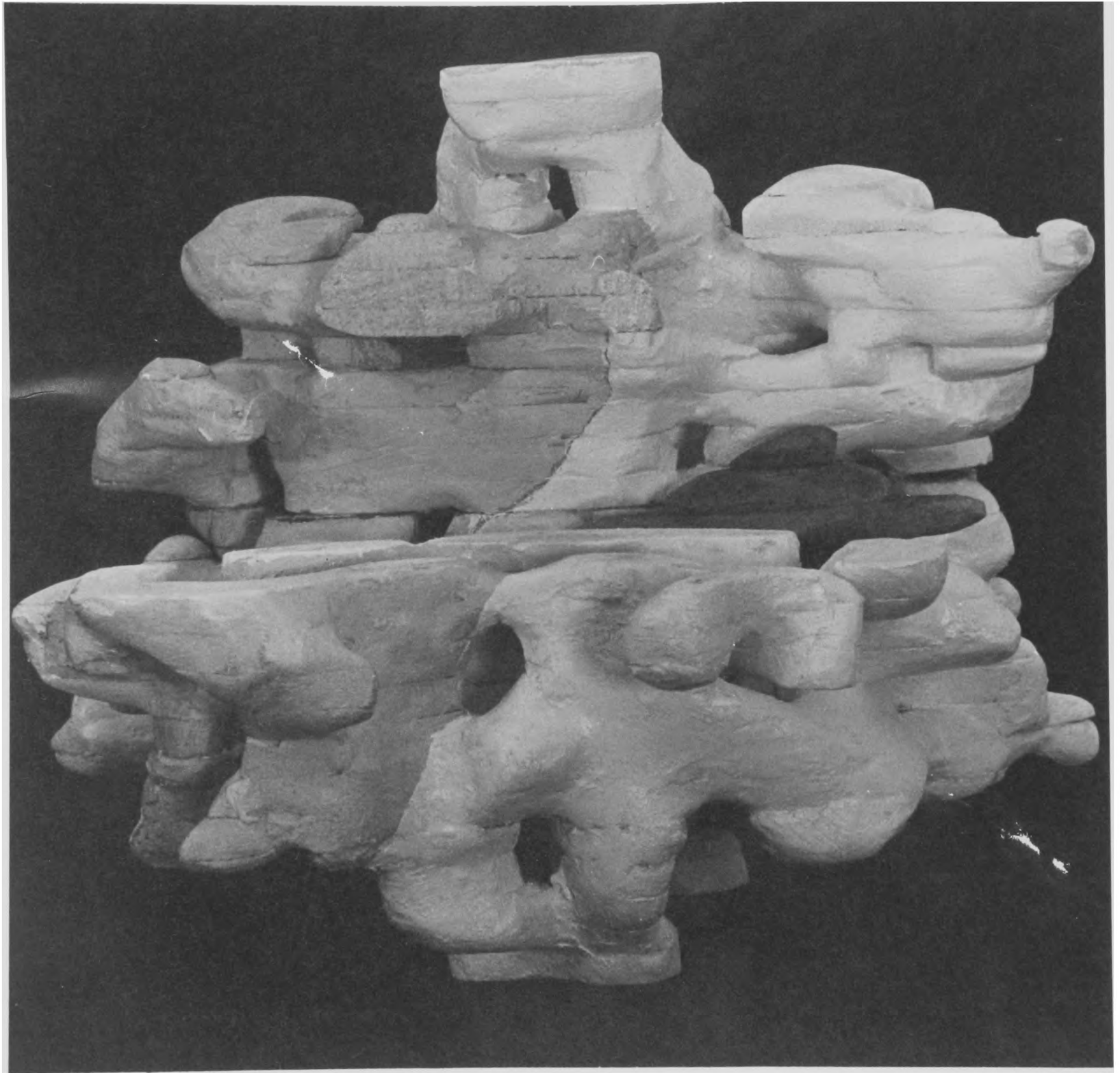


Fig. 52a. The density corresponding to T3 is shown in position in the 6 Å model of prealbumin. The top part of the model has been removed. The white and dark parts of the model indicate the approximate location of individual subunits. The density at the second T3 site has been omitted.

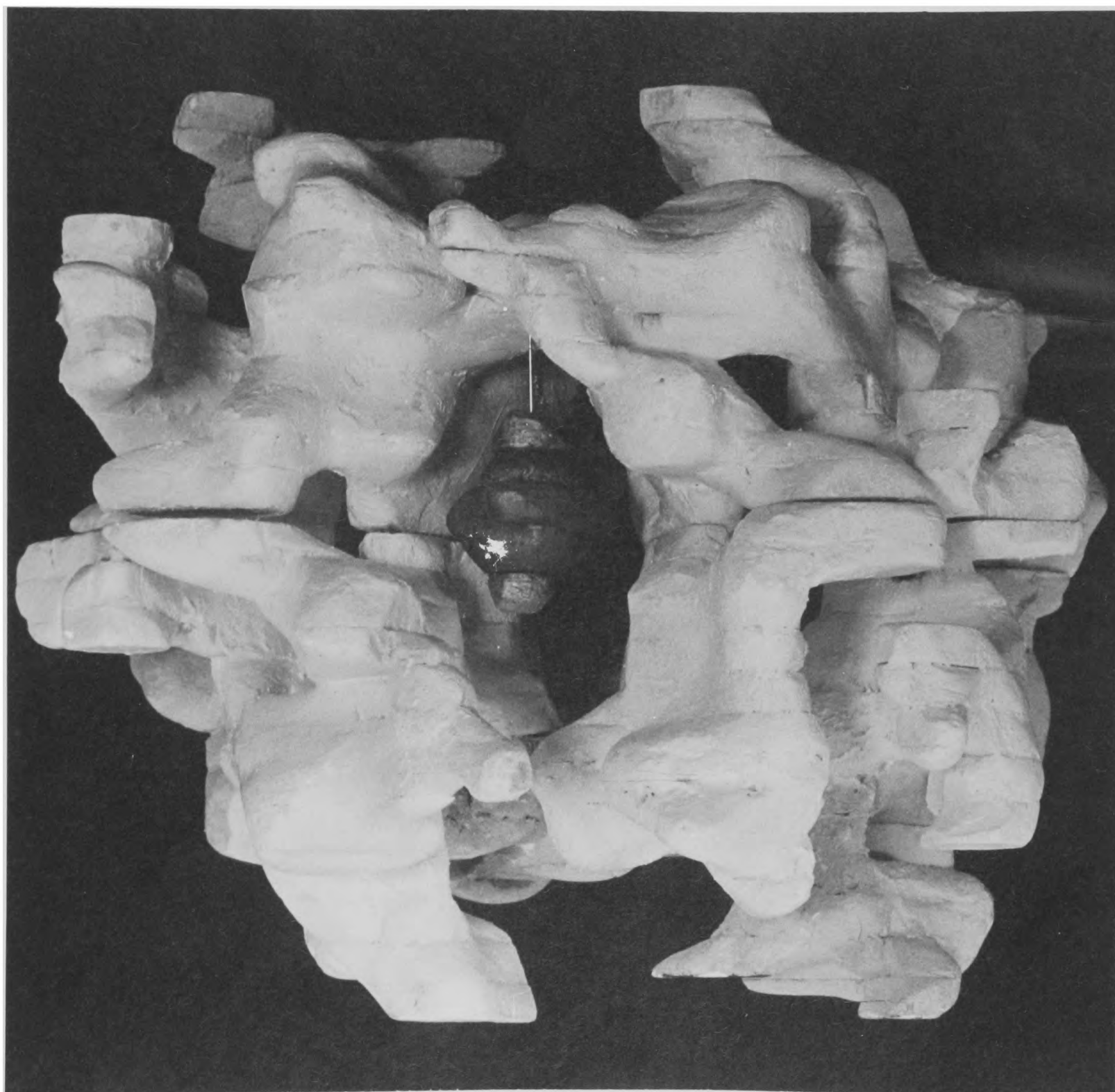


Fig 52b The complete tetramer seen down the crystallographic two-fold axis with the T3 model in position.

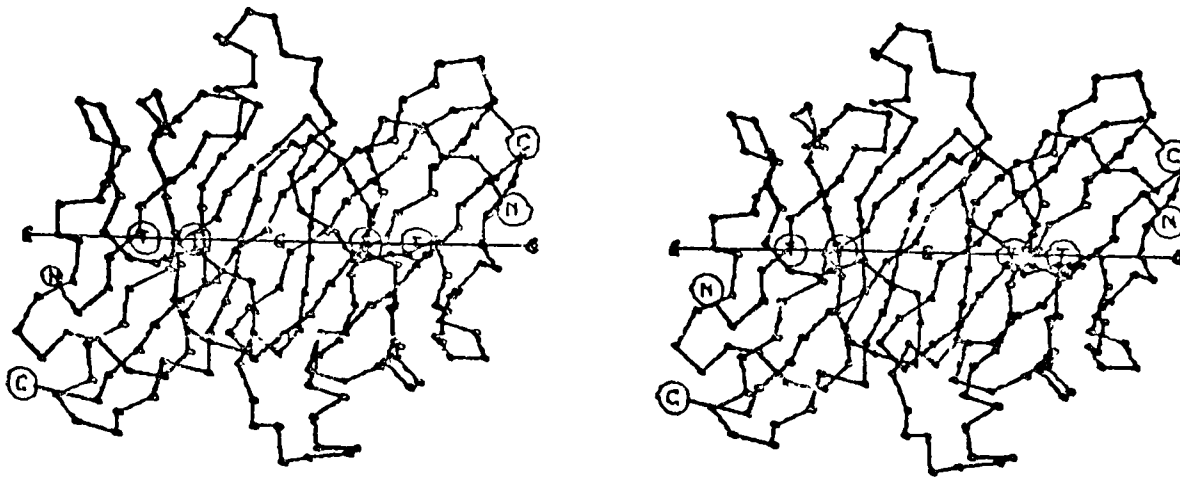


Fig. 53a Stereo pair of the $\alpha_1\alpha_2$ dimer of prealbumin showing the location of the T3 binding sites.

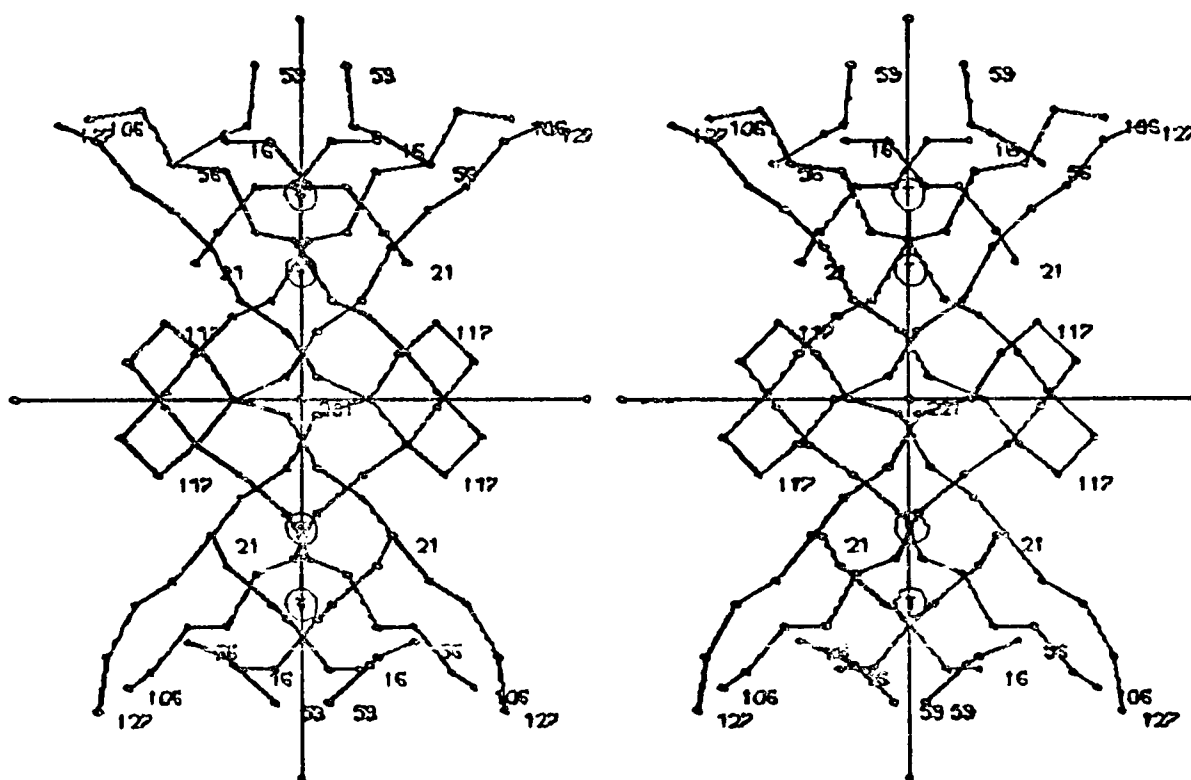
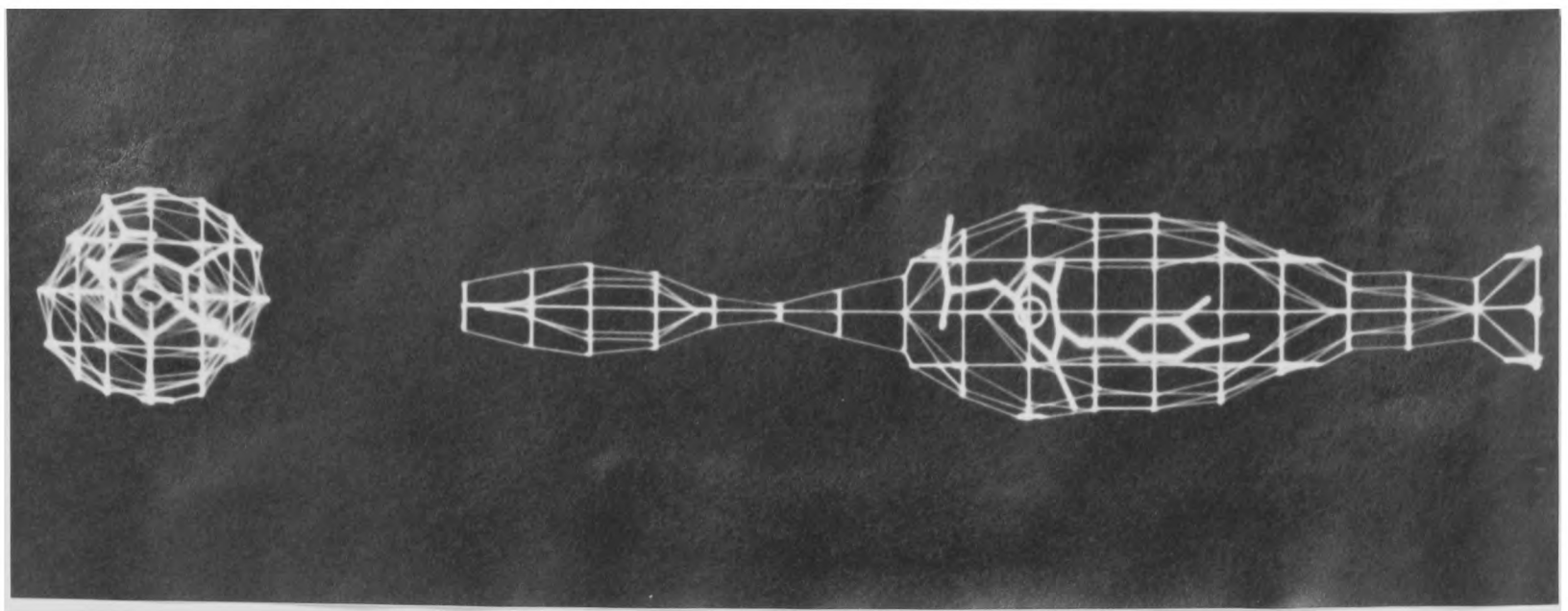


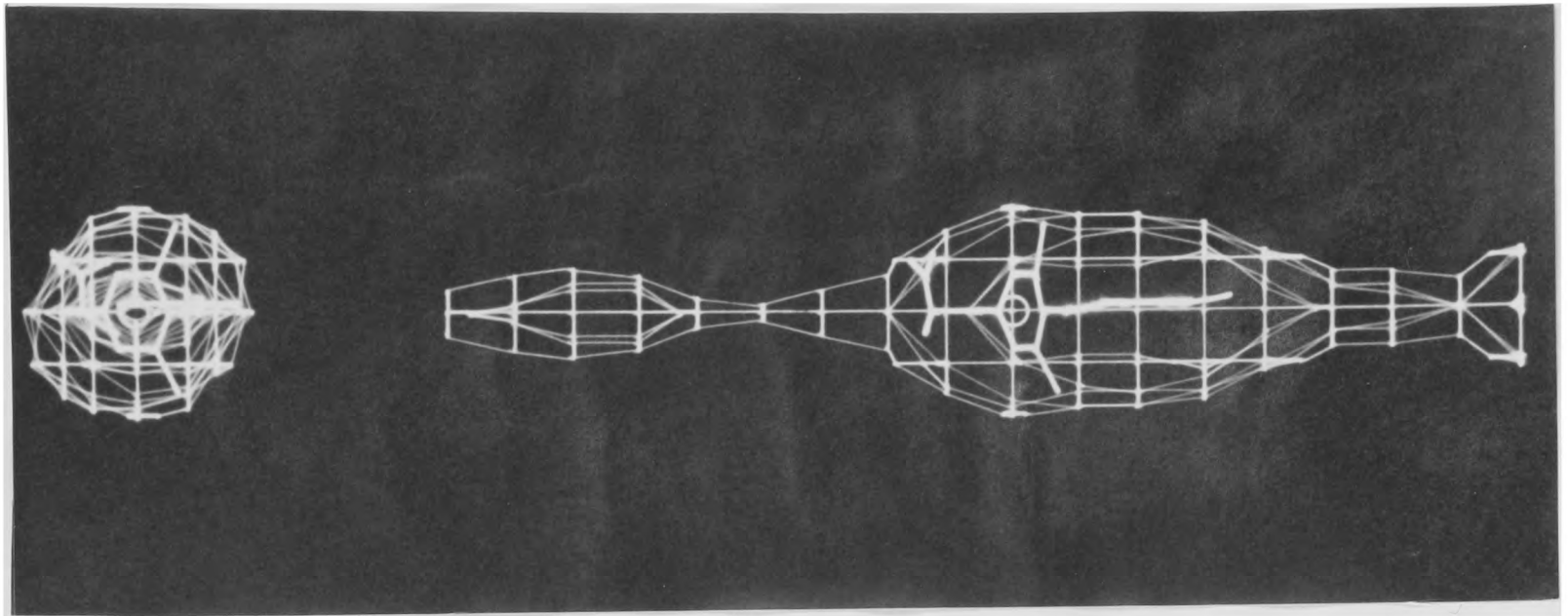
Fig. 53b The ABCD β -sheet regions of the tetramer showing the location of the T3 binding sites.

electron density map of T_3 was calculated on the same scale as the protein Fourier and a small balsa wood model was made of the T_3 -density. This is shown in position in the low resolution native model in fig.52 (a and b). Only one site is shown occupied although similar density is seen in the Fourier just to the left of the centre of fig.52 a. It is clear that the hormone is buried deeply in the channel formed by the β -pleated sheets of opposite subunits ($\alpha_1\alpha'_1$). The difference density is continuous from one site to the other at the lowest contour level, but the most significant density comes to within $3 - 4 \text{ \AA}$ of the molecular centre. The approximate length of the density is $14 \text{ \AA}$, which corresponds closely to the longest dimension of the hormone ($12 \text{ \AA}$). This would seem to eliminate the possibility of a contaminant or breakdown product of the original solution being bound.

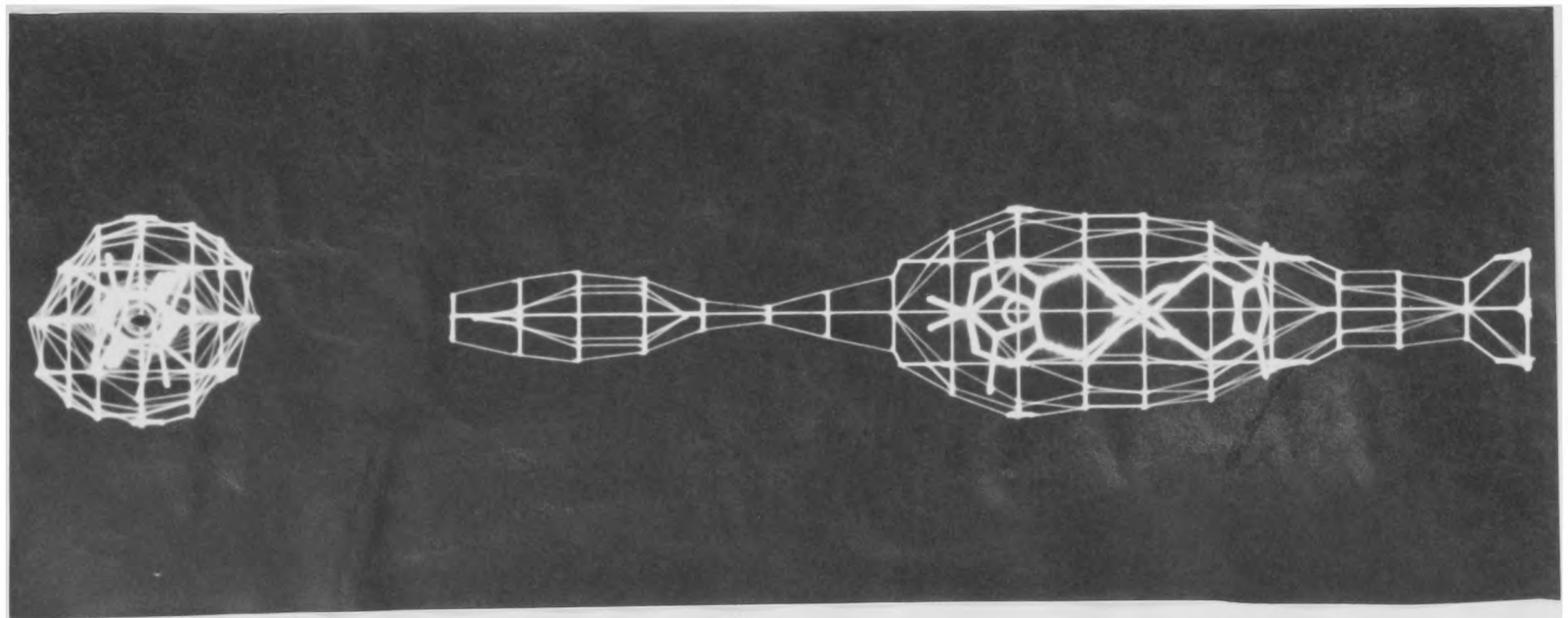
The hormone interacts most closely with the two protein subunits related by crystallographic symmetry (shown in white in fig.52) and not with those related by local symmetry, as it is bound entirely within the dimer formed by these two. When the $6 \text{ \AA}$ difference density was superimposed upon the corresponding sections of the high resolution protein Fourier, it appeared to make a close contact with the β -pleated sheets in the $\alpha_1\alpha'_1$ dimer. Apart from the immediate vicinity of the two-fold axis, very little density is present in the difference map and there are almost no negative contours, suggesting that the binding has no appreciable effect upon the protein conformation. Figure 53a shows the positions of the two T_3 binding sites with respect to the two subunits ($\alpha_1\alpha_2$) related by local symmetry. The circles marked 'T' represent an estimation of the positions of the two benzene rings of T_3 , placed either side of the centre of gravity of the T_3 difference density. The crystallographic two-fold axis is shown, with an indication of the $2 \times 2 \times 2$ symmetry point. Figure 53b shows the same sites between the sheet regions (A-B-C-D) of the complete tetramer.



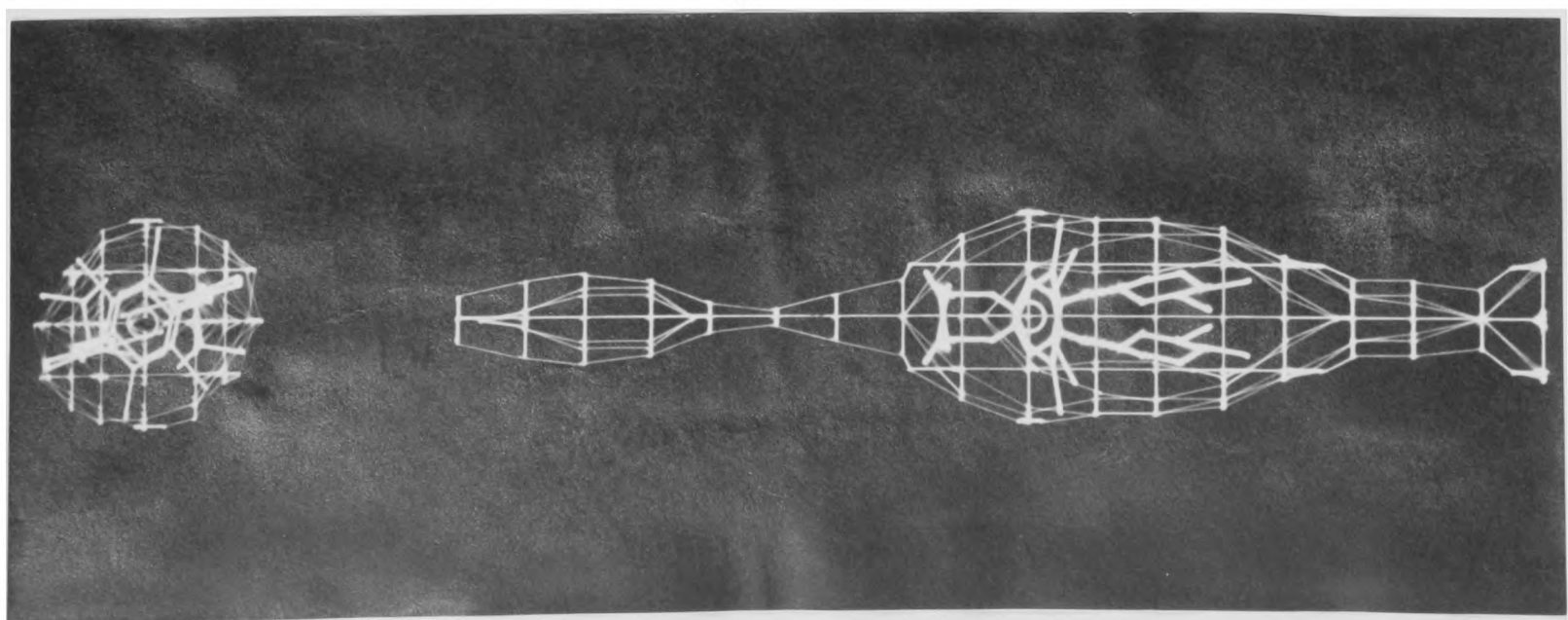
a



b



c



d

Fig. 54. Fit of T3 model to the $6 \text{ \AA}^0$ difference density.

Figure 54 (a to d) shows a series of views of an atomic representation of the T_3 molecule superimposed upon the $6 \text{ \AA}^0$ difference electron density. The left hand view is seen down the crystallographic two-fold axis and the right hand view, looking along the y axis in the plane of x and z. The co-ordinates for the T_3 molecule were calculated from the description of the crystal structure of T_3 of Camerman and Camerman (1971). The electron density is contoured at $0.3 \text{ e \AA}^{-3}$ using a cathode-ray display and a model-building system developed by Dr. C.D. Barry in this laboratory. These pictures do not represent a serious attempt to fit the hormone precisely, but to investigate the possibility of a unique orientation. It can be seen from (a) and (b) that a single orientation of T_3 is unlikely, since any orientation of the molecule leaves considerable density empty. Figures c and d show T_3 molecules related by 180° rotation about the crystallographic two-fold axis. Here the fit is improved, particularly in (d).

The $6 \text{ \AA}^0$ difference density of T_3 , together with the $3 \text{ \AA}^0$ projection data suggest that two orientations of T_3 may be present within different molecules in the same crystal. It is impossible to obtain a precise fit at $6 \text{ \AA}^0$, but the appearance of resolved peaks in the $3 \text{ \AA}^0$ projection syntheses suggest that it may be possible to define the precise positions of T_3 and T_4 using 3D data to this resolution.

The binding of thyroxine analogues.

Four groups of workers have investigated the binding of thyroxine analogues to prealbumin (Ross and Tapley, 1966; Hershman, 1963; Woebber and Ingbar, 1964; Pages et al, 1973). Ross and Tapley tested 49 single and double-ring systems. They obtained binding with several of these analogues and concluded that the di-iodo substituted phenolic ring, with the hydroxyl in the 4' and the two iodines in the 3' and 5' positions, was the important grouping for maximum binding. The second (α) ring was not essential for binding. Hershman, Woebber and Pages also

managed to bind single ring systems to prealbumin. The recent data of Pages et al. establishes that 4-hydroxy 3,5-di-iodobenzaldehyde, 3-(4-hydroxy 3,5-di-iodophenyl)propionic acid and 4-hydroxy 3,5-di-iodocinnamic acid bind to prealbumin, having association constants similar to that of T_4 . Careful measurements of the stoichiometry of these analogues gave a 1:1 association of protein with T_4 , T_3 and the propionic and cinnamic acid derivatives, but a 1:2 association for the di-iodobenzaldehyde compound. From these observations the authors speculate that there are two binding sites for the β -ring, but that long substituents of this ring (e.g. in T_3 , T_4 and the propionic and cinnamic acid derivatives) inhibit the occupation of the second site.

It is evident that there are two possible senses in which the T_3 molecule can bind to the protein. The projection and $6 \text{ \AA}$ maps show that the long axis ($12 \text{ \AA}$) is approximately co-incident with the crystallographic two-fold axis. Either the α or the β ring can be exposed to the solvent at the opening of the channel. Either orientation would place charged groups within the protein.

To attempt to establish this orientation, the analogue 4-cyano-2-iodo-6-nitrophenol (nitroxynil) was successfully bound to the protein. A prealbumin crystal was soaked for 24 hours in a 60% saturated ammonium sulphate solution of this compound, buffered to pH 8.0 with tris-HCl. The $3 \text{ \AA}$ OKL difference electron density showed only the set of peaks 'A' in fig. 51. These features lie within the ellipsoid of density defined by the $6 \text{ \AA}$ difference synthesis of T_3 , but at the end farthest from the protein molecular centre. If this analogue does indeed bind in the same position as the β -ring of T_3 , then this means that the α -ring lies deepest within the protein. This orientation agrees with the evidence presented by Pages et al., that long substituents of the β -ring interfere with binding at the second site. In order for this inhibition to occur, the substituents must point towards the second site,

(and hence towards the protein molecular centre). Pages et al., suggest that inhibition occurs by spacial interference with the second site.

The electron density map indicates that actual overlap of the sites cannot take place. A possible explanation of the inhibition is that either the amino-acid side chains near the centre are disturbed and transmit this information to the second site, or simply that simultaneous occupation of the two sites would bring similarly charged groups (e.g. CO_2^- in the acid derivatives) into close proximity. Both sites are seen to be occupied in the $6 \text{ \AA}^0$ difference map of T_3 , however the occupancy of these sites is very low and the crystallographic data is not incompatible with a 1:1 binding stoichiometry.

A 2 cm to 1 angstrom model of T_4 placed approximately in the position indicated by the difference maps brings the 4'-hydroxyl within $2 - 3 \text{ \AA}^0$ of lysine 15 on opposite subunits. These appear to be the closest residues to the β -ring of T_4 .

The binding of an aspirin analogue and sodium barbiturate.

The inhibition of binding of T_3 and T_4 to prealbumin by sodium barbiturate (veronal) and salicylate has been mentioned in chapter 1. Prealbumin crystals were soaked in 20 mM solutions of sodium barbiturate, sodium salicylate and ethyl mercury thio-salicylate. $3 \text{ \AA}^0$ OKL difference projections of these compounds were obtained. Sodium barbiturate exerted a marked effect on the diffraction pattern and a peak was obtained in the region occupied by the hormones T_3 and T_4 . Very little change in the native intensities were observed with salicylate but the aspirin analogue, ethyl mercury thio-salicylate, gave very large intensity changes and a very large peak was observed in the region of the T_4 binding site. No substitution of the sulphhydryl sites was observed. Evidently the inhibition of thyroid hormone binding by barbiturate and possibly by salicylate is through competitive binding of these inhibitors at the hormone binding sites.

Summary of Chapter 7

It has been demonstrated that the thyroid hormones T3 and T4 can potentially occupy two sites per molecule of prealbumin. These sites are deep within the cleft formed by opposed pairs of dimers. The two sites do not appear to overlap. It cannot at present be determined whether these sites are simultaneously or statistically occupied by the thyronines. The two sites are related by local symmetry and need not be identical, however it seems that the individual sites must contain two orientations of T3 or T4, since they lie on a crystal diad.

The dimensions of the difference electron density representing T3 shows that the long axis of the molecule is aligned with the crystal diad (along the slot). Evidence has been presented that the phenolic β - ring of T3 lies farthest from the molecular centre.

The slow interaction of T3 with the crystalline protein cannot be explained in the light of the very high association constant, but may result from an unfavourable energy barrier of diffusion into the crystal.

APPENDIX 1. Statistical expressions used in the data processing.

- 1) Standard deviation of a difference measurement ΔF .

$$\Delta F = F_1 - F_2$$

$$\sigma_{\Delta F} = \sqrt{(\sigma_{F_1}^2 + \sigma_{F_2}^2)}$$

- 2) Weighted mean and associated standard deviation of repeated measurements of F .

$$\bar{F} = \frac{(1/\sigma_{F_1}^2 \cdot F_1 + 1/\sigma_{F_2}^2 \cdot F_2 + \dots + 1/\sigma_{F_N}^2 \cdot F_N)}{(1/\sigma_{F_1}^2 + 1/\sigma_{F_2}^2 + \dots + 1/\sigma_{F_N}^2)}$$

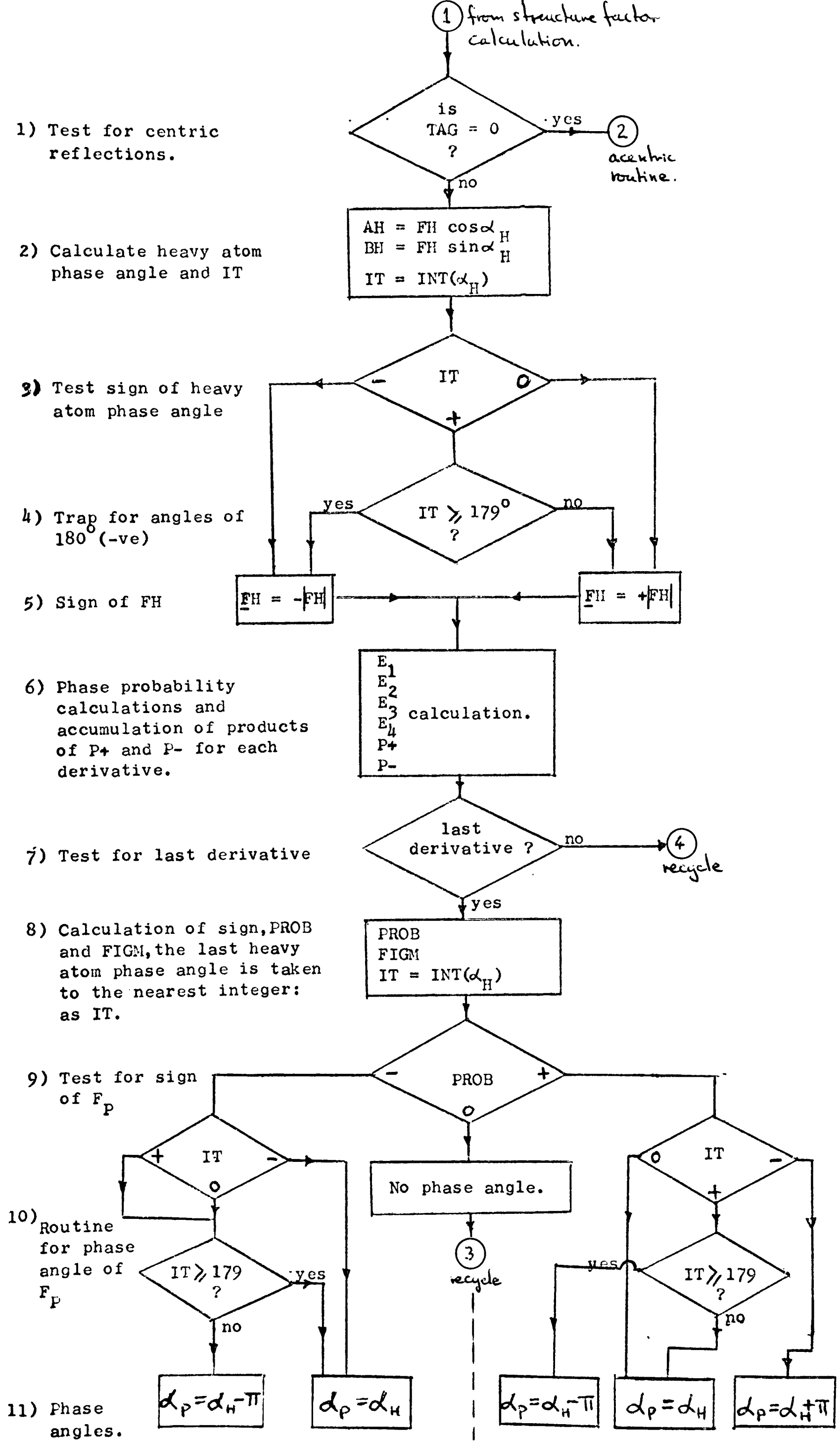
$$\sigma_{\bar{F}} = \sqrt{1/(1/\sigma_{F_1}^2 + 1/\sigma_{F_2}^2 + \dots + 1/\sigma_{F_N}^2)}$$

- 3) Crystallographic weight w .

$$w = 1/\sigma^2$$

- 4) Significance test for difference ΔF (NSD = number of standard deviations)

$$\Delta F / \sigma_{\Delta F} = \text{NSD}$$



APPENDIX 2. An algorithm for centric phase determination.

Blow and Crick (1959) give expressions for the probability of the two choices of sign of F_P . Difficulty was experienced when incorporating their expressions in a general phase programme. What is required after exiting from the sign determination, is the phase angle α_P in degrees. This is commonly some integer ^{sub} multiple of π for centric phases. The algorithm shown as a flow diagram opposite was used in the general phase programme. Essentially, the routine works by defining the sign of F_H where

$$\begin{aligned} \text{this is +ve if } 0^\circ > \alpha_H > 180^\circ \\ \text{or -ve if } 0^\circ > \alpha_H > -180^\circ \end{aligned}$$

The phase angle α_H of the last derivative heavy atom F_H is stored, and after the sign of the protein F_P has been determined, the routine works out whether α_P is equal to α_H , $\alpha_H - \pi$ or $\alpha_H + \pi$

Notes.

$$E_1 = F_{H\text{CALC}} - |F_{PH}| + |F_P|$$

$$E_2 = F_{H\text{CALC}} + |F_{PH}| + |F_P|$$

$$E_3 = F_{H\text{CALC}} - |F_{PH}| - |F_P|$$

$$E_4 = F_{H\text{CALC}} + |F_{PH}| - |F_P|$$

$$P+ = \exp(-E_1^2/2 \cdot E^2) + \exp(-E_2^2/2 \cdot E^2)$$

$$P- = \exp(-E_3^2/2 \cdot E^2) + \exp(-E_4^2/2 \cdot E^2)$$

$$PP = \prod_j P_j^+, \quad PM = \prod_j P_j^-$$

$$PROB = (PP - PM)/(PP + PM)$$

$$m = \text{ABS}(PROB)$$

E = RMS isomorphous lack of closure error.

PREALBUMIN α -CARBON CO-ORDINATES.

These co-ordinates are given in Ångstroms relative to an origin at the molecular centre and corresponding to the real cell co-ordinates : (0.5 0.5 0.25). The system of axes is right-handed. Residue numbers refer to the molecular drawing in fig. 49. The co-ordinates are given to the nearest 0.5 Å. The two monomers in the crystallographic asymmetric unit are given in the order α_1, α_2 in the listing. The complete molecule is specified by rotating the co-ordinates given by an angle of 180° about the z-axis through (0 0 0).

Residue No.	a_1			a_2		
	x	y	z	x	y	z
14	-4.0	-12.0	-22.5	3.5	-13.5	23.0
15	-7.5	-13.0	-20.0	6.5	-13.5	20.5
16	-5.0	-12.5	-17.0	4.0	-13.0	17.5
17	-2.0	-10.0	-17.0	1.5	-10.0	18.0
18	0.5	-9.0	-14.0	-1.0	-9.0	14.5
19	3.0	-6.0	-14.0	-2.0	-5.5	13.5
20	5.5	-6.0	-11.0	-5.0	-5.0	11.0
21	7.0	-3.0	-9.0	-6.0	-2.0	9.0
22	10.0	-2.5	-7.5	-8.5	-1.5	6.5
23	10.0	0.0	-4.5	-8.5	2.0	4.0
24	13.5	1.5	-4.0	-12.5	2.0	3.0
25	14.0	1.0	-8.0	-13.5	1.0	6.5
26	11.5	1.0	-10.5	-12.0	1.5	10.0
27	10.0	-2.0	-12.5	-10.0	-1.5	12.0
28	10.0	-5.5	-11.0	-11.0	-5.5	11.5
29	14.0	-5.5	-12.5	-13.5	-5.5	14.0
30	13.0	-9.0	-14.0	-14.0	-9.5	13.0
31	11.0	-12.0	-12.0	-11.5	-11.0	10.5
32	8.5	-15.0	-13.0	-9.5	-14.5	11.5
33	5.5	-17.0	-11.5	-6.2	-16.5	11.0
34	4.0	-20.5	-11.5	-6.0	-20.5	11.5
35	0.5	-20.0	-10.5	-2.0	-21.5	11.5
36	-2.5	-22.0	-10.0	-1.0	-23.5	8.5
37	-6.0	-23.5	-10.5	3.0	-23.5	9.0
38	-9.0	-24.0	-9.0	6.0	-24.5	7.0
39	-10.0	-27.5	-7.0	8.0	-28.0	6.0
40	-11.0	-28.0	-3.0	10.0	-29.0	3.0
41	-7.0	-28.0	-2.5	6.5	-29.5	1.0
42	-4.0	-29.5	-3.0	3.0	-30.5	2.0
43	-3.0	-31.0	-6.0	1.0	-30.0	5.5
44	-1.0	-29.0	-8.5	-2.0	-28.5	7.5
45	-2.0	-27.0	-11.5	-1.0	-27.5	11.0
46	-2.5	-23.0	-12.5	-1.0	-24.0	12.5
47	-2.0	-20.5	-15.0	0.5	-22.5	15.5
48	0.5	-20.5	-17.5	-3.0	-21.5	16.5
49	4.0	-19.5	-16.0	-6.0	-20.0	15.5
50	6.0	-16.0	-17.0	-6.5	-16.5	16.5
51	9.0	-14.0	-16.5	-9.5	-14.0	16.0
52	9.5	-10.0	-16.5	-9.0	-10.0	15.0
53	12.5	-9.5	-18.0	-11.0	-7.0	17.0
54	13.0	-6.0	-20.0	-11.0	-5.5	20.5
55	9.5	-4.5	-18.5	-8.5	-2.5	19.5
56	8.5	-6.5	-15.5	-7.5	-3.5	16.0
57	5.0	-7.0	-17.5	-5.0	-6.5	17.0
58	3.5	-10.5	-18.0	-4.5	-10.0	17.5
59	3.0	-11.0	-22.0	-1.5	-10.0	20.0
60	0.0	-13.5	-22.0	-2.0	-13.5	21.5
61	-0.5	-16.5	-20.0	0.0	-16.5	20.0
62	-4.5	-17.5	-20.5	3.0	-18.5	20.5
63	-6.0	-17.5	-24.0	6.0	-18.0	23.0
64	-9.5	-17.5	-26.0	9.0	-21.0	24.0
65	-12.0	-20.0	-23.5	6.5	-22.5	21.0
66	-9.5	-21.5	-21.0	8.5	-20.0	19.0
67	-10.0	-19.5	-18.0	8.5	-22.0	16.0
68	-13.0	-19.5	-15.0	11.5	-23.2	14.0
69	-13.5	-21.0	-11.5	12.0	-23.0	10.0
70	-11.0	-20.0	-9.0	9.0	-21.5	8.5
71	-7.5	-19.5	-9.5	5.0	-20.5	10.0
72	-4.0	-18.5	-7.5	2.0	-19.5	9.0
73	-1.0	-17.0	-7.5	0.5	-16.0	8.5
74	3.0	-18.0	-8.0	-3.0	-16.0	7.0
75	5.0	-15.0	-8.0	-6.0	-13.5	7.5

76	8.0	-15.0	-6.0	-9.5	-13.5	6.0
77	10.0	-12.0	-4.0	-11.0	-10.5	4.0
78	13.0	-14.0	-2.5	-14.0	-12.0	3.0
79	14.5	-15.0	-5.5	-15.0	-12.0	6.5
80	14.5	-11.0	-6.5	-15.0	-8.0	6.5
81	15.0	-10.0	-3.0	-17.0	-7.0	3.5
82	18.0	-12.0	-3.0	-20.0	-10.0	4.0
83	19.5	-11.5	-6.5	-20.0	-8.5	8.0
84	19.0	-7.5	-6.0	-20.5	-5.0	7.0
85	22.0	-6.5	-3.5	-21.0	-6.5	3.0
86	21.5	-9.5	-1.0	-21.5	-4.0	0.5
87	18.0	-9.0	0.0	-18.0	-5.0	0.0
88	17.5	-11.0	3.5	-18.0	-8.5	-2.5
89	14.5	-13.0	3.0	-15.5	-11.0	-2.5
90	11.5	-12.0	5.0	-12.0	-10.0	-4.0
91	9.0	-13.0	2.0	-9.0	-12.0	-2.5
92	9.5	-16.5	1.0	-11.0	-14.5	0.0
93	6.5	-17.0	-2.0	-8.5	-16.0	2.0
94	4.0	-14.0	-2.5	-5.0	-14.5	2.5
95	0.0	-15.5	-4.0	-1.5	-16.0	3.0
96	-3.5	-14.0	-4.5	1.5	-14.5	3.0
97	-6.0	-16.5	-5.5	5.0	-16.5	3.5
98	-9.0	-14.0	-7.0	7.5	-15.0	6.0
99	-13.0	-15.0	-7.0	9.0	-18.5	7.5
100	-12.0	-16.0	-11.5	10.0	-18.0	11.0
101	-15.5	-16.5	-12.5	14.0	-18.5	12.0
102	-17.0	-17.0	-16.0	15.5	-19.0	15.0
103	-19.0	-13.5	-16.0	17.5	-15.5	15.5
104	-18.0	-14.5	-20.0	18.0	-16.0	19.5
105	-15.5	-13.5	-22.0	14.0	-16.0	21.0
106	-14.0	-13.0	-18.5	11.5	-13.0	19.0
107	-10.5	-11.0	-19.0	10.0	-12.5	18.0
108	-8.5	-10.0	-15.5	7.5	-11.0	15.0
109	-5.0	-8.0	-15.0	5.0	-8.0	15.0
110	-3.0	-8.0	-11.0	3.0	-8.0	11.5
111	-0.5	-6.0	-10.5	0.5	-5.0	10.5
112	2.0	-6.0	-6.5	-1.0	-6.0	7.5
113	4.5	-4.0	-5.5	-4.0	-4.0	5.0
114	7.0	-5.0	-3.0	-7.0	-5.0	3.0
115	9.0	-3.5	-0.5	-9.0	-3.5	0.5
116	12.0	-5.0	2.5	-11.5	-5.0	-2.5
117	9.5	-6.5	5.0	-9.0	-6.0	-5.0
118	7.0	-5.5	3.0	-6.5	-4.5	-3.0
119	4.5	-5.5	0.0	-4.5	-5.5	0.0
120	1.0	-4.5	-1.5	-1.0	-5.0	1.0
121	-1.0	-6.0	-4.5	1.0	-6.0	4.5
122	-4.0	-5.0	-6.5	4.0	-4.5	6.5
123	-6.0	-5.5	-10.0	6.0	-6.0	9.0
124	-8.5	-4.5	-12.5	8.5	-5.5	12.0
125	-11.0	-7.0	-14.0	11.0	-8.0	13.5
126	-13.5	-7.0	-17.0	13.0	-7.5	17.0
127	-16.0	-6.5	-18.0	13.5	-8.0	20.5
128	-18.5	-7.5	-20.0	15.5	-5.5	22.0
129	-18.0	-6.5	-24.0	13.5	-4.5	25.0
130	-14.5	-7.5	-25.5	10.0	-5.5	24.0

STRUCTURE FACTOR TABLES.

The structure factors and protein phase angles for native prealbumin and four derivatives are given to 2.9 Å. Complete listings including standard deviations are available in the Laboratory of Molecular Biophysics, Oxford University. The data is stored in the filestore of the University 1906A computer at Banbury Road, in the format given in Fig 31(11).

The format of the tables is as follows:

H K L F(native) F(PAQ) F(PAUN) F(PAQUN) F(PAQNEW) α_p m*100

where α_p is the protein phase angle in degrees between -180° & $+180^\circ$
 m is the value of the figure of merit.

0 24	12	100	76	49	08	100-170	2	1 5	4	257	478	138	4	463	14	98	1 12	15	397	354	230	341	374	499	60		
0 25	1	201	175	224	103	157-90	49	1 5	5	391	384	222	283	407	50	98	1 12	16	364	306	339	322	329	67	80		
0 25	2	135	152	95	49	157-90	84	1 5	6	81	68	14	74	67	94	20	1 12	17	246	321	265	319	317	446	94		
0 25	3	215	254	92	140	273-90	55	1 5	7	10	50	74	126	94	50	124	20	1 12	18	136	189	197	213	165	19	90	
0 25	4	280	326	269	277	273-90	85	1 5	8	11	155	184	127	130	178	110	77	1 12	19	111	123	85	100	0	71	9	
0 25	5	04	26	07	155	16	40	5	1 5	9	12	254	249	197	244	254-135	97	1 12	20	118	111	75	85	131	83	32	
0 25	6	09	93	123	134	113	91	1 5	10	1	116	151	73	231	117	61	35	1 13	0	277	81	257	100	100	0	100	
0 25	7	200	186	78	45	169-90	100	1 5	11	14	454	301	347	242	330	7	100	1 13	1	159	154	168	197	134-157	39		
0 25	8	133	125	118	05	126-90	80	1 5	12	15	121	156	160	171	140	71	71	1 13	2	230	285	307	307	254-152	95		
0 25	9	131	136	200	187	147-90	95	1 5	13	16	174	117	142	63	147	-87	68	1 13	3	74	113	69	137	97	-97	40	
0 25	10	149	-20	140	0	157-90	74	1 5	14	17	82	114	98	190	80	32	57	1 13	4	209	257	305	244	253	158	93	
0 25	11	144	147	137	0	0	0	1 5	15	18	307	365	236	331	374	-34	97	1 13	5	76	54	76	74	59	-17	42	
0 26	0	176	120	44	311	116-100	100	1 5	16	20	237	363	200	345	365	4	98	1 13	6	256	121	275	203	130-160	100		
0 26	1	135	137	101	108	102	0	97	1 5	17	21	127	113	93	127	120	101	73	1 13	7	154	146	175	135	172	160	92
0 26	2	416	435	410	375	342-100	100	1 6	0	85	117	230	0	102-180	94		1 13	8	259	140	166	19	152-174	100			
0 26	3	290	311	323	292	204	0	98	1 6	1	65	153	310	0	133	7	86	1 13	9	108	87	123	131	117	94	61	
0 26	4	141	123	132	45	113	0	49	1 6	2	167	201	257	0	212	4	74	1 13	10	130	122	87	174	105-118	61		
0 26	5	241	212	196	174	197-100	67	1 6	3	851	725	034	0	744	80	99	1 13	11	136	161	156	256	165	85	66		
0 26	6	242	256	127	222	237	0	100	1 6	4	400	164	76	0	162	112	70	1 13	12	590	511	455	487	545	76	79	
0 26	7	190	166	152	139	160	0	29	1 6	5	465	398	510	0	414	64	95	1 13	13	203	213	177	310	223	59	55	
0 26	8	147	88	188	117	125	0	98	1 6	6	130	111	98	0	95	-84	68	1 13	14	66	82	45	129	113	170	89	
0 26	9	177	150	184	81	153	0	1	1 6	7	138	116	85	0	154	140	51	1 13	15	92	106	59	113	104	-93	3	
0 26	10	45	36	42	0	0	0	0	1 6	8	181	48	109	54	60	-45	93	1 13	16	256	253	221	319	223	-38	62	
0 27	1	53	75	40	05	84	40	89	1 6	9	117	117	101	149	120	24	91	1 13	17	285	292	395	320	311	43	72	
0 27	2	140	146	199	140	124-90	99	1 6	10	94	74	86	101	85	154	32	1 13	18	137	153	179	123	141-112	95			
0 27	3	126	03	193	71	41	-90	100	1 6	11	85	130	158	90	179	99	89	1 14	0	31	22	44	90	44	0	45	
0 27	4	242	263	301	157	243	40	65	1 6	12	180	220	123	204	231	-92	63	1 14	1	263	299	240	237	314	-61	95	
0 27	5	163	257	248	270	150	40	100	1 6	13	379	365	379	450	401	31	67	1 14	2	143	150	147	126	141	-47	10	
0 27	6	212	212	210	150	175	40	36	1 6	14	137	134	94	145	158	-61	70	1 14	3	253	281	293	229	249-142	97		
0 28	0	357	349	104	353	310-100	100	1 6	15	137	82	164	107	100	151	83	1 14	4	128	150	131	172	150-152	53			
0 28	1	141	141	175	156	141-100	99	1 6	16	128	102	75	105	122	34	65	1 14	5	378	305	254	260	308	50	94		
0 28	2	62	73	46	60	112	0	53	1 6	17	333	260	323	297	276	96	94	1 14	6	126	107	103	129	114	4	19	
0 28	3	126	175	207	151	119-140	93	1 6	18	114	81	0	90	92	-98	60	1 14	7	157	139	150	152	133	-78	40		
0 28	4	221	215	228	207	0	0	30	1 6	19	90	131	94	151	124-117	53	1 14	8	352	345	311	399	364	140	33		
0 28	5	100	88	189	0	0	0	80	1 6	20	65	103	17	178	100-151	81	1 14	9	270	241	130	183	237	-78	94		
0 28	6	100	193	211	62	0	0	24	1 6	21	104	63	17	110	81	-97	82	1 14	10	181	159	190	234	169	51	25	
0 28	7	101	65	84	114	0	-90	47	1 7	0	577	535	716	0	584	0	39	1 14	11	185	251	154	235	272	56	97	
1 0	1	34	303	47	0	206	90	95	1 7	1	186	242	150	0	235	145	79	1 14	12	168	166	209	280	158	-61	80	
1 0	2	354	818	808	0	337	40	100	1 7	2	160	115	190	0	133	-57	81	1 14	13	495	416	423	437	485	-97	40	
1 0	3	130	42	62	0	55	-90	-1	1 7	3	190	136	159	0	131	-73	63	1 14	14	199	193	221	236	191	126	51	
1 0	4	500	621	547	0	601	90	100	1 7	4	527	520	515	0	530	-5	69	1 14	15	377	239	383	270	260	116	99	
1 0	5	136	90	55	0	74	90	68	1 7	5	205	229	133	0	230	-60	78	1 14	16	183	203	131	192	180	132	33	
1 0	6	445	354	455	507	330	90	100	1 7	6	151	142	125	37	153-104	77	1 14	17	126	144	96	180	179-133	96			
1 0	7	39	70	18	49	51	90	42	1 7	7	0	135	109	100	183	125	121	8	1 14	18	92	114	122	107	133	127	81
1 0	8	72	252	174	348	225	-90	99	1 7	8	177	157	136	220	164	45	79	1 14	19	151	116	115	124	135	-2	81	
1 0	9	27	32	108	119	51	90	19	1 7	9	11	71	117	42	68	90-130	20	1 15	0	361	289	317	209	268	0	100	
1 0	10	90	224	95	0	245	-90	100	1 7	10	390	371	287	331	398	55	79	1 15	1	249	257	299	255	253	161	40	
1 0	11	50	54	0	71	56	40	19	1 7	11	103	141	131	228	124	72	86	1 15	2	400	411	331	223	375	-62	73	
1 0	12	334	315	424	370	343	-90	100	1 7	12	174	161	201	133	174	138	71	1 15	3	143	145	122	129	139	21	43	
1 0	13	213	196	194	584	195	-90	39	1 7	13	131	128	74	161	145	123	24	1 15	4	179	143	299	240	161	180	89	
1 0	14	175	285	186	572	284	40	101	1 7	14	144	107	201	150	119-110	78	1 15	5	144	114	135	146	118-179	67			
1 0	15	115	115	0	206	124	40	8	1 7	15	76	99	110	94	100-163	64	1 15	6	120	104	0	174	105-116	35			
1 0	16	206	29	200	93	72	-90	100	1 7	16	249	208	336	245	224	-33	78	1 15	7	466	466	342	436	475	-35	50	
1 0	17	05	03	05	268	104	40	93	1 7	17	75	86	73	127	111	50	70	1 15	8	477	332	351	391	527	-48	94	
1 0	18	33	53	07	85	90	-90	88	1 7	18	260	241	197	258	267	-71	82	1 15	9	129	133	113	160	133	35	28	
1 0	19	318	239	175	213	271	90	100	1 7	19	60	103	92	97	88	-56	32	1 15	10	206	221	129	183	214	179	40	
1 0	20	64	61	0	77	71	-90	21	1 8	0	87	99	55	0	85-180	76	1 15	11	109	142	124	167	168	120	74		
1 1	0	253	273	0	0	276	0	100	1 8	1	1080	1107	877	0	1116	61	97	1 15	12	186	240	126	256	224-162	93		
1 1	1	55	05	112	0	49	20	11	1 8	2	223	236	173	0	227	30	43	1 15	13	127	159	126	237	153	100	84	
1 1	2	120	111	420	0	114	-92	90	1 8	3	442	556	362	0	537	128	98	1 15	14	341	344	205	308	300	-76		

1	20	13	238	281	158	256	243	54	41
1	20	14	146	101	119	89	40	-143	0
1	20	15	158	157	113	150	158	-100	31
1	20	16	164	164	143	184	0	-47	21
1	21	0	571	627	627	436	540	0	26
1	21	1	154	125	75	162	43	145	60
1	21	2	100	36	0	145	97	-66	63
1	21	3	272	297	230	229	274	-142	24
1	21	4	313	305	204	278	299	-158	67
1	21	5	410	385	379	327	404	-100	19
1	21	6	201	234	25	154	211	12	95
1	21	7	142	140	118	147	152	-176	67
1	21	8	152	89	110	176	115	164	67
1	21	9	473	493	550	405	650	20	86
1	21	10	192	155	158	234	100	-47	82
1	21	11	180	200	158	180	170	-57	46
1	21	12	96	101	102	149	79	12	69
1	21	13	97	95	34	74	108	98	94
1	21	14	94	104	117	87	110	121	87
1	21	15	110	147	85	135	0	154	21
1	22	0	343	322	320	354	243	0	78
1	22	1	114	69	76	194	92	128	43
1	22	2	56	91	54	78	62	8	78
1	22	3	205	312	222	296	259	-54	32
1	22	4	225	224	237	220	198	-58	62
1	22	5	379	325	301	298	293	116	94
1	22	6	526	506	470	429	401	-61	80
1	22	7	186	142	168	222	177	144	85
1	22	8	273	274	359	222	264	41	65
1	22	9	210	284	218	196	278	-35	97
1	22	10	133	158	124	213	127	176	71
1	22	11	219	237	148	236	237	57	85
1	22	12	72	89	50	95	112	141	87
1	22	13	60	62	28	75	100	42	77
1	22	14	84	96	71	101	125	-58	77
1	23	0	115	117	115	0	80	0	100
1	23	1	153	140	72	99	104	64	7
1	23	2	162	150	184	186	124	3	99
1	23	3	345	371	319	231	314	59	22
1	23	4	200	144	229	206	126	-155	67
1	23	5	228	212	268	197	219	-58	36
1	23	6	174	109	119	240	112	-176	100
1	23	7	317	297	317	182	276	136	63
1	23	8	235	152	243	273	196	-167	94
1	23	9	212	200	202	206	175	-159	67
1	23	10	178	165	151	161	166	-136	34
1	23	11	149	193	171	195	171	112	31
1	23	12	68	67	86	134	104	-94	91
1	23	13	85	108	136	104	109	81	91
1	24	0	369	395	355	326	280	0	19
1	24	1	110	59	81	190	88	155	40
1	24	2	71	99	100	87	80	-56	81
1	24	3	91	84	114	98	90	-82	56
1	24	4	145	154	201	176	150	-156	65
1	24	5	179	176	120	198	180	-28	54
1	24	6	148	109	110	127	119	160	59
1	24	7	175	123	231	151	133	-51	77
1	24	8	154	156	259	159	157	-108	87
1	24	9	146	141	176	109	156	-123	65
1	24	10	171	125	75	132	127	-124	97
1	24	11	191	127	194	187	156	-174	72
1	24	12	155	156	146	145	0	-165	8
1	25	0	149	146	154	166	129	-140	35
1	25	1	114	73	111	179	114	172	46
1	25	2	171	177	230	215	161	-76	73
1	25	3	339	339	390	303	280	-19	71
1	25	4	155	194	221	202	156	1	90
1	25	5	179	158	180	186	149	-174	79
1	25	6	165	165	112	145	157	-2	59
1	25	7	137	124	99	84	147	-40	89
1	25	8	149	71	152	155	106	95	93
1	25	9	147	140	180	157	140	-29	27
1	25	10	158	153	120	0	133	-18	68
1	25	11	79	80	41	0	0	-87	49
1	26	0	114	98	109	174	98	-180	21
1	26	1	209	261	252	251	227	-136	76
1	26	2	322	352	354	275	266	118	93
1	26	3	216	248	216	204	209	65	34
1	26	4	278	260	259	208	223	119	87
1	26	5	200	187	160	217	190	144	67
1	26	6	310	265	369	244	275	-14	85
1	26	7	69	115	131	154	113	-116	33
1	26	8	91	78	168	120	94	79	66
1	26	9	173	132	110	151	114	64	96
1	27	0	256	304	229	261	238	-180	88
1	27	1	155	148	118	168	115	-100	50
1	27	2	187	179	156	145	171	-99	91
1	27	3	247	240	229	176	198	-63	53
1	27	4	211	201	216	176	176	-66	76
1	27	5	123	94	149	116	113	-77	32
1	27	6	162	72	162	148	94	37	96
1	27	7	209	176	306	164	183	53	73
1	27	8	102	36	111	92	0	156	7
1	28	0	460	481	474	540	506	0	60
1	28	1	219	193	193	350	144	-89	85
1	28	2	360	355	358	224	298	-81	93
1	28	3	134	110	118	204	94	91	94
1	28	4	256	256	346	168	220	177	46
1	28	5	177	116	199	187	0	40	23
1	29	0	103	110	197	102	0	0	56
1	29	1	175	154	151	0	0	68	44
2	0	0	1196	884	1125	0	854	-180	100
2	0	1	42	100	35	0	80	-180	77
2	0	2	325	203	100	0	192	-180	100
2	0	3	415	433	245	0	429	0	88
2	0	4	106	120	214	0	132	180	82
2	0	5	151	115	205	0	121	0	23
2	0	6	442	571	615	0	553	-140	100
2	0	7	156	154	81	189	150	-180	76
2	0	8	155	261	89	122	221	180	100
2	0	9	141	143	107	151	137	0	42
2	0	10	46	122	126	143	100	0	4
2	0	11	36	97	67	75	73	0	70
2	0	12	404	562	303	527	554	0	100
2	0	13	381	361	457	402	352	0	83
2	0	14	401	196	406	366	241	-140	100
2	0	15	125	120	51	107	120	0	72
2	0	16	68	66	0	244	83	0	47
2	0	17	61	92	80	116	92	140	48
2	0	18	131	245	165	254	219	-180	100
2	0	19	179	170	120	175	148	-180	57
2	0	20	101	220	129	154	236	-140	100
2	0	21	52	69	65	60	89	0	91
2	0	22	152	62	93	115	0	-140	35
2	1	0	120	110	250	0	103	0	96
2	1	1	521	589	522	0	573	103	97
2	1	2	104	137	138	0	116	-112	18
2	1	3	971	953	644	0	926	-133	94
2	1	4	111	65	115	0	75	-144	42
2	1	5	398	447	554	0	690	59	98
2	1	6	136	175	350	0	173	-23	61
2	1	7	230	150	273	195	178	178	79
2	1	8	168	200	151	110	210	-50	42
2	1	9	426	523	481	681	524	-90	96
2	1	10	55	61	60	52	93	166	46
2	1	11	262	323	255	299	324	-104	97
2	1	12	131	165	129	164	161	114	60
2	1	13	649	732	831	784	801	-63	74
2	1	14	119	40	103	143	41	-116	46
2	1	15	291	241	159	217	266	19	40
2	1	16	155	151	0	206	171	-14	54
2	1	17	390	471	418	457	431	72	94
2	1	18	165	144	222	254	175	-24	36
2	1	19	223	252	112	255	255	-140	90
2	1	20	40	73	47	111	0	-115	44
2	1	21	141	14	105	127	116	46	92
2	1	22	95	42	65	76	0	7	24
2	2	0	359	304	155	0	219	-110	100
2	2	1	168	200	126	0	25	-15	65
2	2	2	224	331	220	0	275	-101	95
2	2	3	174	212	214	0	214	-133	78
2	2	4	254	361	405	0	354	-99	97
2	2	5	249	290	301	0	391	-97	95
2	2	6	292	242	443	0	251	-44	91
2	2	7	131	137	111	157	134	20	54
2	2	8	683	704	524	550	727	59	93
2	2	9	187	224	226	207	215	74	73
2	2	10	207	247	237	257	285	109	94
2	2	11	160	185	114	157	197	125	71
2	2	12	584	587	503	547	625	87	91
2	2	13							

2	16	13	149	178	166	163	146-124	66	2	26	7	302	257	374	243	243	167	93	3	6	9	392	386	382	403	378	16	47
2	16	14	135	93	132	132	98	67	2	26	8	194	124	223	127	125	111	98	3	6	10	271	324	243	207	324	17	70
2	16	13	162	193	137	226	188-157	83	2	26	9	61	33	73	90	3	111	8	3	11	347	0	62	519	698	17	98	
2	16	16	99	87	105	87	128	20	2	27	0	109	125	164	199	94	180	99	3	6	12	140	174	145	192	146	82	63
2	16	17	52	92	113	132	89	123	2	27	1	265	229	162	199	194	170	89	3	6	13	371	402	322	386	416	222	97
2	16	18	104	91	167	133	111	3	2	27	3	237	190	237	172	161	118	98	3	6	14	77	69	70	66	62	140	91
2	17	0	350	386	346	433	309	0	2	27	4	61	93	33	119	93	131	70	3	6	15	338	292	242	271	316	130	97
2	17	1	371	409	403	313	300	34	2	27	5	193	151	207	189	163	106	31	3	6	16	209	200	247	189	216	138	61
2	17	2	448	433	408	401	413	149	2	27	6	319	286	399	231	233	110	78	3	6	18	106	90	142	124	99	59	15
2	17	3	250	306	200	281	294	60	2	27	7	129	92	98	104	0	39	20	3	6	19	102	123	164	199	130	134	93
2	17	4	360	386	436	303	388-111	91	2	28	0	213	108	104	130	113	140	100	3	6	20	42	12	18	126	86	163	71
2	17	5	333	320	317	268	311	163	2	28	1	223	132	136	165	110	176	94	3	6	21	140	119	138	90	133	107	79
2	17	6	76	53	19	86	38	99	2	28	2	233	140	127	102	114	0	83	3	7	0	450	433	483	416	437	180	13
2	17	7	240	242	63	194	233-150	93	2	28	3	233	204	161	132	0	52	63	3	7	1	141	141	110	147	163	154	74
2	17	8	47	112	106	60	121	73	2	28	4	178	176	198	164	0	93	24	3	7	2	133	60	147	30	59	100	100
2	17	9	206	264	0	264	261-149	94	3	0	1	139	141	310	124	144	90	100	3	7	3	147	39	68	91	38	355	97
2	17	10	194	187	167	200	195	22	3	0	2	93	37	237	25	49	90	100	3	7	4	343	281	243	279	284	277	99
2	17	11	240	234	192	173	242	56	3	0	3	228	72	90	0	71	90	100	3	7	5	178	108	369	234	167	127	69
2	17	12	44	80	53	64	82	99	3	0	4	123	87	69	0	81	90	97	3	7	6	180	147	54	82	168	113	72
2	17	13	198	186	176	130	221	24	3	0	5	109	233	54	103	234	90	100	3	7	7	112	41	181	146	40	153	73
2	17	14	291	277	294	309	261	108	3	0	6	55	109	110	199	110	90	80	3	7	8	240	161	193	182	140	49	99
2	17	15	324	318	301	370	314	15	3	0	7	222	240	81	88	230	90	100	3	7	9	180	193	199	184	207	101	84
2	17	16	64	103	138	103	77-172	54	3	0	8	153	185	176	294	189	90	71	3	7	10	332	349	258	323	367	354	95
2	17	17	132	185	133	136	183	134	3	0	9	184	334	133	379	339	90	100	3	7	11	342	0	366	404	428	138	81
2	18	0	334	449	280	330	402	0	3	0	10	193	160	247	228	161	90	97	3	7	12	316	318	442	499	327	114	80
2	18	1	203	197	193	133	185	67	3	0	11	300	449	403	371	467	90	100	3	7	13	316	370	323	368	371	152	83
2	18	2	208	178	172	213	168	149	3	0	12	330	384	301	547	594	90	100	3	7	14	197	180	187	236	179	179	50
2	18	3	303	493	496	467	469	49	3	0	13	313	234	263	241	263	90	93	3	7	15	241	248	228	212	245	94	60
2	18	4	139	171	228	138	178-167	79	3	0	14	369	363	414	429	366	90	87	3	7	16	181	159	238	207	188	275	67
2	18	5	234	239	227	264	257	73	3	0	15	369	363	414	429	366	90	87	3	7	17	181	159	238	207	188	275	67
2	18	6	78	148	27	142	167-170	95	3	0	16	369	363	414	429	366	90	87	3	7	18	181	159	238	207	188	275	67
2	18	7	78	74	67	92	80	62	3	0	17	344	403	270	309	449	90	100	3	7	19	112	98	53	103	107	3	69
2	18	8	192	140	182	136	156	11	3	0	18	137	168	118	148	174	90	92	3	7	20	147	133	59	109	139	100	70
2	18	9	119	141	0	49	132-118	69	3	0	19	231	235	280	381	293	90	100	3	8	1	323	413	359	361	417	143	99
2	18	10	111	133	120	78	130-143	85	3	0	20	65	78	24	106	80	90	94	3	8	2	131	76	181	204	76	162	89
2	18	11	113	117	164	109	133	83	3	0	21	161	222	185	181	202	90	83	3	8	3	311	769	342	302	269	157	90
2	18	12	253	138	210	146	178-159	99	3	1	0	101	42	140	104	41	180	100	3	8	4	103	92	50	166	94	12	63
2	18	13	336	312	303	298	306	72	3	1	1	82	57	136	72	37	34	22	3	8	5	342	484	630	376	491	170	94
2	18	14	225	97	147	121	100-105	99	3	1	2	194	189	164	0	189	152	39	3	8	6	170	289	269	233	292	147	94
2	18	15	204	191	218	233	188	29	3	1	3	279	300	280	0	303	176	93	3	8	7	348	410	204	329	414	134	67
2	18	16	180	147	190	127	186	64	3	1	4	107	87	34	138	87	90	90	3	8	8	201	216	144	172	228	163	69
2	18	17	100	83	39	89	0	99	3	1	5	254	311	254	374	313	147	94	3	8	9	177	128	131	121	120	96	98
2	19	0	273	317	333	362	299-180	99	3	1	6	66	104	197	217	104	89	63	3	8	10	193	189	151	191	209	123	76
2	19	1	270	232	232	163	200	70	3	1	7	121	110	115	36	109	154	93	3	8	11	432	0	308	390	366	109	99
2	19	2	338	324	283	226	311	59	3	1	8	251	131	293	238	177	120	97	3	8	12	144	117	130	153	112	449	82
2	19	3	226	243	263	362	207	37	3	1	9	42	59	125	93	43	146	57	3	8	13	647	519	586	588	573	142	97
2	19	4	628	640	626	504	617	2	3	1	10	893	830	802	837	863	113	63	3	8	14	343	368	345	393	398	42	82
2	19	5	249	193	197	223	183	63	3	1	11	116	119	97	144	127	72	49	3	8	15	136	136	0	121	140	8	87
2	19	6	202	220	220	200	210-123	20	3	1	12	429	399	365	439	431	42	87	3	8	16	10	68	140	94	82	142	43
2	19	7	271	267	411	266	273	88	3	1	13	93	141	249	221	137	107	89	3	8	17	149	103	254	196	83	442	98
2	19	8	214	246	101	109	216	96	3	1	14	303	214	257	233	237	124	96	3	8	18	218	203	284	203	244	120	92
2	19	9	311	322	299	299	328	39	3	1	15	90	104	106	87	137	121	65	3	8	19	142	118	124	106	120	40	63
2	19	10	237	278	243	329	207-123	91	3	1	16	124	187	160	108	164	134	89	3	8	20	98	71	77	81	123	160	81
2	19	11	237	278	243	329	207-123	91	3	1	17	124	187	160	108	164	134	89	3	8	21	98	71	77	81	123	160	81
2	19	12	211	188	221	192	180-105	16	3	1	18	161	156	222	237	173	120	55	3	8	22	98	71	77	81	123	160	81
2	19	13	85	79	60	160	100	81	3	1	19	89	139	70	105	149	120	91	3	8	23	98	71	77	81	123	160	81
2	19	14	171	165	180	120	154	144	3	1	20	72	70	110	143	71	64	23	3	8	24	98	71	77	81	123	160	81
2	19	15	129	14																								

3 13 16 73 94 103 143 0 83 13
3 13 17 289 310 301 279 110 38 73
3 13 13 112 79 138 122 0 111 44
3 13 19 66 98 86 137 0 100 42
3 14 0 39 83 13 130 73-140 3
3 14 1 311 318 353 293 311-145 40
3 14 2 337 341 397 340 331 -8 05
3 14 3 120 113 72 115 137-149 95
3 14 4 263 234 204 270 200 -2 07
3 14 5 219 127 197 127 115 65 99
3 14 6 177 136 147 147 204 93 63
3 14 7 321 346 260 310 350 135 70
3 14 8 767 746 737 731 819 9 66
3 14 9 202 193 214 153 186 -18 60
3 14 10 191 167 131 219 170-133 45
3 14 11 340 310 478 301 367 -49 57
3 14 12 140 106 124 196 115 64 36
3 14 13 273 272 230 297 297 32 33
3 14 14 213 213 290 244 244 -43 62
3 14 15 137 137 131 121 137-179 95
3 14 16 82 100 0 86 115-124 54
3 14 17 42 113 33 132 137-143 95
3 14 18 104 207 229 160 225 70 67
3 15 0 223 174 120 49 104 0 100
3 15 1 213 237 288 236 247 130 79
3 15 2 233 196 191 173 146 -55 98
3 15 3 203 183 181 182 140 -41 32
3 15 4 393 332 324 311 332-112 98
3 15 5 39 62 45 146 63 68 76
3 15 6 334 330 231 267 336 -41 93
3 15 7 602 397 493 341 397 -25 77
3 15 8 314 290 203 230 206 109 40
3 15 9 321 302 329 289 302 -41 49
3 15 10 270 280 223 216 264-136 90
3 15 11 280 327 269 294 304 49 81
3 15 12 523 377 400 334 353-150 88
3 15 13 83 76 37 186 128 154 43
3 15 14 167 152 123 223 155-118 43
3 15 15 76 100 141 147 96 55 49
3 15 16 80 112 97 114 75 161 67
3 15 17 112 90 94 97 93 97 56
3 15 18 74 103 0 121 92 -55 74
3 16 0 738 687 746 377 680-100 84
3 16 1 78 129 60 193 99 42 81
3 16 2 282 291 382 281 303 75 71
3 16 3 119 129 93 103 126 -5 76
3 16 4 194 137 126 139 147 80 90
3 16 5 111 144 43 130 163 143 95
3 16 6 146 133 123 137 141 99 90
3 16 7 208 268 201 249 273 145 97
3 16 8 181 149 134 129 163-125 77
3 16 9 134 98 113 104 106 103 91
3 16 10 138 171 97 184 161 113 90
3 16 11 70 74 0 130 74 -33 70
3 16 12 130 110 147 189 126 133 87
3 16 13 269 297 336 276 283 34 90
3 16 14 174 176 178 207 185-126 83
3 16 15 304 324 310 319 313 30 49
3 16 16 108 133 173 223 133-113 63
3 16 17 80 110 139 80 94 18 92
3 16 18 144 153 153 129 0-101 4
3 17 0 73 179 183 214 230 0 100
3 17 1 131 97 144 99 103 -14 85
3 17 2 41 140 138 110 141 -8 93
3 17 3 118 104 166 132 131 32 76
3 17 4 137 170 97 190 151 -8 73
3 17 5 46 87 37 141 67 79 67
3 17 6 226 189 146 94 173 -10 99
3 17 7 310 289 231 216 270 179 95
3 17 8 113 94 101 193 76 28 93
3 17 9 163 234 192 229 200 -89 95
3 17 10 180 143 194 144 155 69 77
3 17 11 310 301 299 232 336 -87 59
3 17 12 144 104 71 101 112 135 97
3 17 13 230 234 264 240 227 -46 67
3 17 14 203 180 170 229 178-174 65
3 17 15 100 67 92 103 123 93 89
3 17 16 190 158 154 192 187 193 54
3 17 17 100 120 147 73 102-156 94
3 18 0 606 628 557 376 595 0 26
3 18 1 202 227 158 183 211 16 88
3 18 2 388 370 377 333 379 -39 11
3 18 3 163 173 192 237 149 3 56
3 18 4 433 453 373 396 463-120 64
3 18 5 118 113 70 130 136 43 90
3 18 6 304 299 203 237 327-133 91
3 18 7 146 153 111 130 156 89 60
3 18 8 187 183 142 164 188 87 45
3 18 9 237 203 227 191 203-116 96
3 18 10 177 168 64 160 194 -46 75
3 18 11 309 329 261 337 340 -44 96
3 18 12 301 269 336 237 273-171 71
3 18 13 101 68 0 70 94-140 25
3 18 14 217 236 284 226 216 -33 62
3 18 15 127 162 129 100 166 124 92
3 18 16 200 215 240 227 194 114 45
3 19 0 369 317 476 413 330-180 100
3 19 1 404 411 423 313 373-144 47
3 19 2 120 86 104 163 78 137 85
3 19 3 206 208 163 161 190-167 26
3 19 4 211 223 173 262 230 -69 52
3 19 5 88 83 73 107 86-126 69
3 19 6 160 205 110 93 212-143 94
3 19 7 218 208 254 177 215-161 49
3 19 8 200 203 276 164 224 72 90
3 19 9 144 126 136 193 149 -13 13
3 19 10 304 304 313 477 522 75 60
3 19 11 137 134 160 269 167-133 87
3 19 12 403 373 434 396 380 178 61
3 19 13 212 201 217 246 206-136 79
3 19 14 237 196 270 133 213 159 93
3 19 15 191 188 233 184 179 19 49
3 19 16 113 34 120 106 0-141 7
3 20 0 249 307 173 204 264-140 87
3 20 1 190 206 263 263 209 130 85
3 20 2 266 281 330 302 268 144 66
3 20 3 237 207 180 194 159 -8 67
3 20 4 220 212 229 190 212 -48 10
3 20 5 206 153 217 134 100 162 91
3 20 6 233 273 289 231 268 71 14
3 20 7 482 394 427 360 381 179 98
3 20 8 333 347 333 303 335 136 45
3 20 9 324 317 323 293 304 121 48
3 20 10 122 127 110 164 141 -60 76
3 20 11 103 119 0 68 107-100 35
3 20 12 208 191 124 133 162 -46 26
3 20 13 146 149 117 163 140 -17 48
3 20 14 104 41 28 90 102 -42 86
3 20 15 113 149 119 114 126-123 83
3 21 0 36 32 78 74 68 149 45
3 21 1 116 130 144 181 140 -41 91
3 21 2 209 203 190 200 191 -11 59
3 21 3 341 349 330 336 326 73 33
3 21 4 119 71 67 94 70 82 82
3 21 5 389 349 409 306 334 101 97
3 21 6 31 104 102 84 101 -01 47
3 21 7 189 153 233 177 153 2 94
3 21 8 184 166 297 240 167 37 43
3 21 9 64 138 123 101 0 139 51
3 21 10 174 131 190 224 151 -38 47
3 21 11 222 213 229 311 224 -24 90
3 21 12 172 229 167 214 211 -19 83
3 21 13 162 177 127 134 173 -00 73
3 22 0 231 221 60 123 192 0 100
3 22 1 216 203 173 179 194 23 67
3 22 2 274 263 253 267 263 -11 63
3 22 3 232 244 249 212 212 -14 58
3 22 4 457 429 412 422 372-113 95
3 22 5 133 134 106 130 173-124 93
3 22 6 273 260 243 263 263 -77 67

3 22 7 187 164 236 187 172-154 64
3 22 8 69 84 106 134 91 -34 64
3 22 9 380 344 388 314 338 -90 95
3 22 10 84 83 114 161 110 -44 72
3 22 11 195 226 107 171 222 81 89
3 22 12 77 37 -2 100 100-136 84
3 22 13 78 101 87 97 115-127 83
3 23 0 316 333 309 369 317-180 71
3 23 1 240 271 168 219 228 -35 57
3 23 2 130 96 104 111 90 61 95
3 23 3 133 70 67 123 70 128 74
3 23 4 478 508 403 360 434 74 42
3 23 5 163 127 98 198 134 72 51
3 23 6 197 124 162 130 124 137 94
3 23 7 97 92 106 84 111 169 23
3 23 8 185 133 284 140 124 124 92
3 23 9 97 111 171 145 102 21 75
3 23 10 126 115 103 92 123 130 64
3 23 11 76 25 101 84 84 91 91
3 23 12 127 67 60 67 105 5 89
3 24 1 92 87 191 172 94 0 93
3 24 2 171 175 204 140 117 161 91
3 24 3 122 120 206 192 104 123 76
3 24 4 209 153 193 240 162 136 63
3 24 5 75 77 61 130 84 -85 7
3 24 6 250 236 238 237 213 17 74
3 24 7 131 93 90 90 127 165 69
3 24 8 120 70 88 130 124-178 91
3 24 9 228 234 192 174 236 85 5
3 24 10 159 177 161 149 169 -20 91
3 24 11 27 75 62 120 111 -25 75
3 25 0 170 168 114 170 122-180 99
3 25 1 172 142 143 138 120 -48 89
3 25 2 237 217 215 228 194-103 82
3 25 3 185 183 153 191 171 9 73
3 25 4 243 280 280 254 240 15 89
3 25 5 161 127 161 155 156 124 87
3 25 6 97 85 157 93 94-104 82
3 25 7 93 140 150 122 138 4 91
3 25 8 187 181 283 196 180-128 91
3 25 9 158 150 223 131 156 83 69
3 26 0 74 67 120 161 83-180 35
3 26 1 124 142 120 141 124 111 82
3 26 2 218 211 152 166 177 122 95
3 26 3 225 191 239 185 162 -84 94
3 26 4 277 240 246 166 217 70 95
3 26 5 181 144 179 144 145 -54 89
3 26 6 87 84 71 111 82 142 57
3 26 7 230 203 290 193 193 -44 91
3 26 8 103 64 124 78 110-142 26
3 27 0 93 50 44 44 67-186 68
3 27 1 105 103 116 31 73 20 83
3 27 2 157 131 155 110 112 -50 90
3 27 3 177 180 152 176 144-120 80
3 27 4 278 293 296 220 254 90 85
3 27 5 264 240 287 204 227-152 76
3 27 6 227 243 121 224 0 48 59
3 28 0 186 197 107 109 160-180 95
3 28 1 106 77 67 154 67 -10 80
3 28 2 269 247 328 219 0 119 59
3 28 3 0 94 94 114 141 0 179 18
3 28 4 246 227 199 128 0 102 31
4 0 0 676 798 832 862 801 0 100
4 0 1 117 95 59 0 94 0 99
4 0 2 45 77 60 0 78-180 64
4 0 3 103 111 221 0 110-180 68
4 0 4 234 157 353 0 136 0 100
4 0 5 73 63 166 40 60-180 92
4 0 6 223 336 306 469 338-180 100
4 0 7 30 49 235 142 41-180 93
4 0 8 426 466 438 400 470 180 100
4 0 9 103 122 299 310 141 0 100
4 0 10 253 218 267 309 230 160 63
4 0 11 702 727 688 788 741 180 75
4 0 12 304 211 255 220 232-180 98
4 0 13 362 372 304 447 384 180 100
4 0 14 110 91 124 75 97-180 92
4 0 15 370 383 0 460 390-180 95
4 0 16 37 108 128 165 117 0 99
4 0 17 111 30 104 100 87 0 31
4 0 18 106 58 101 30 103 0 4
4 0 19 47 82 88 30 68 0 70
4 0 20 36 97 85 0 90-180 95
4 0 21 143 87 171 107 133 0 88
4 1 0 133 187 136 150 180-180 92
4 1 1 453 427 462 0 429 -65 84
4 1 2 108 160 79 0 161-115 83
4 1 3 103 95 87 0 94 178 85
4 1 4 177 206 274 0 208-141 67
4 1 5 336 243 273 194 243-131 99
4 1 6 156 184 268 302 183 -72 91
4 1 7 249 247 187 224 247 -55 97
4 1 8 90 82 78 168 81 160 55
4 1 9 178 80 194 222 113 92 97
4 1 10 232 190 266 243 203 -63 70
4 1 11 672 624 699 663 663 12 27
4 1 12 66 64 85 51 84 178 72
4 1 13 480 456 433 509 486-174 74
4 1 14 132 108 142 153 116 63 62
4 1 15 423 372 0 431 385 -28 82
4 1 16 116 80 50 150 116 -27 54
4 1 17 135 147 147 192 166 122 70
4 1 18 197 187 184 227 150-145 76
4 1 19 96 55 36 134 138 -58 91
4 1 20 95 95 79 143 107 15 51
4 1 21 128 107 136 75 102 161 51
4 2 0 404 432 541 474 435 0 100
4 2 1 424 259 164 0 257-160 73
4 2 2 66 208 32 0 210-140 86
4 2 3 132 188 270 0 429 -60 98
4 2 4 339 422 257 0 429 -60 98
4 2 5 22 68 81 91 60 -78 34
4 2 6 231 232 222 226 233 166 40
4 2 7 97 131 129 160 130 140 71
4 2 8 287 252 318 370 254-124 93
4 2 9 156 136 333 310 122-168 77
4 2 10 534 536 569 487 570 -96 99
4 2 11 228 220 237 244 225 167 90
4 2 12 95 114 106 144 94 134 33
4 2 13 133 124 231 182 128 4 94
4 2 14 481 396 443 410 427 65 90
4 2 15 157 211 0 221 220 -78 91
4 2 16 173 228 215 237 238 -90 97
4 2 17 130 130 119 152 164-130 38
4 2 18 340 260 268 305 284 110 96
4 2 19 124 140 206 138 134 155 34
4 2 20 219 194 220 258 218-163 63
4 2 21 77 110 81 107 134 -24 94
4 3 0 196 110 154 70 111 0 100
4 3 1 232 356 207 304 358-143 97
4 3 2 115 61 60 115 67 -4 88
4 3 3 155 194 78 169 194 -91 92
4 3 4 64 111 0 58 112 52 97
4 3 5 171 140 130 165 141-135 69
4 3 6 110 130 110 97 140 51 87
4 3 7 336 185 251 168 187 170 100
4 3 8 147 180 235 321 180 -43 80
4 3 9 107 130 181 196 124 130 80
4 3 10 77 72 106 104 70 -23 82
4 3 11 497 514 423 510 551 155 97
4 3 12 94 120 73 30 123-176 84
4 3 13 242 280 190 261 300 144 67
4 3 14 81 37 63 114 97 -38 15
4 3 15 388 336 0 323 369 -33 63
4 3 16 117 117 97 162 144 111 80
4 3 17 124 142 114 157 140 7 81
4 3 18 226 191 228 210 220 148 70
4 3 19 263 260 229 278 281 53 91
4 3 20 35 67 0 114 109 -59 89

4 3 21 91 82 87 98 47-170 72
4 4 0 340 193 290 233 186 0 100
4 4 1 58 148 82 142 110-107 87
4 4 2 112 164 97 178 163-138 89
4 4 3 273 257 421 274 258 -35 36
4 4 4 392 414 299 338 417 -98 74
4 4 5 274 295 320 269 298 166 83
4 4 6 237 256 246 236 253 -87 96
4 4 7 161 94 105 104 95 58 80
4 4 8 16 198 157 279 191 6 94
4 4 9 178 92 230 186 100 -34 79
4 4 10 317 266 242 293 292-139 61
4 4 11 308 324 241 278 361 37 70
4 4 12 167 149 152 156 157-179 91
4 4 13 116 134 123 106 121 -52 89
4 4 14 234 173 162 152 179 4 99
4 4 15 198 200 0 156 219 -1 83
4 4 16 97 122 0 112 172-149 91
4 4 17 206 201 287 212 203 1 57
4 4 18 213 170 208 283 203 123 12
4 4 19 153 171 187 147 163-112 86
4 4 20 225 169 172 166 199-119 93
4 4 21 114 108 71 55 133-162 87
4 5 0 109 115 115 70 113-180 11
4 5 1 146 151 206 190 152-123 84
4 5 2 157 160 169 157 162 105 45
4 5 3 221 120 246 159 120 76 99
4 5 4 172 187 213 243 148 97 77
4 5 5 308 218 342 269 220 98 67
4 5 6 98 102 62 86 104-170 89
4 5 7 363 399 347 412 401 162 80
4 5 8 187 172 227 269 171 159 46
4 5 9 101 75 106 77 50 78 98
4 5 10 320 313 271 331 329 -54 71
4 5 11 127 122 136 122 140 67 91
4 5 12 247 281 218 217 289 -44 87
4 5 13 39 73 97 37 84 -8 76
4 5 14 239 239 0 168 229 41 2
4 5 15 148 205 153 199 180 -91 87
4 5 16 96 155 137 91 138 42 87
4 5 17 213 265 227 253 289 99 93
4 5 18 127 149 165 201 191 57 88
4 5 19 103 0 0 129 96 127 88
4 5 20 162 173 177 136 184 11 57
4 5 21 70 104 110 85 0 -20 74
4 6 0 181 137 253 124 137 0 100
4 6 1 125 127 52 49 129 160 84
4 6 2 151 205 101 192 206 109 93
4 6 3 160 138 327 190 138 -45 93
4 6 4 210 152 200 213 153-110 96
4 6 5 249 232 291 247 233-114 24
4 6 6 262 192 151 86 153-148 97
4 6 7 389 442 373 347 443-133 82
4 6 8 109 65 157 104 65 102 93
4 6 9 276 243 231 138 275 -37 64
4 6 10 402 268 283 283 301 91 99
4 6 11 312 318 287 339 318-110 7
4 6 12 199 185 198 153 174 60 69
4 6 13 219 248 263 256 247 10 65
4 6 14 36 86 0 92 84 131 73
4 6 15 246 334 259 264 317 87 74
4 6 16 161 161 147 81 133 42 -6
4 6 17 113 134 134 117 127 127 98
4 6 18 199 227 157 247 252 85 94
4 6 19 64 0 151 90 50 -65 50
4 6 20 92 77 62 107 69 -17 14
4 7 0 42 191 102 117 192 0 100
4 7 1 317 202 226 142 204-170 100
4 7 2 114 154 85 124 138 -5 74
4 7 3 663 443 443 431 446-122 16

4	11	3	119	22	37	110	103	66	59	4	19	4	286	258	245	214	240	107	31	5	2	4	141	110	276	202	111	168	86
4	11	4	201	257	322	263	314	42	74	4	19	5	230	176	219	217	144	134	62	5	2	5	140	206	86	124	209	50	80
4	11	5	447	431	320	340	400	11	95	4	19	6	391	381	426	358	379	22	10	5	2	6	122	132	56	88	133	108	93
4	11	6	520	514	506	457	501	64	88	4	19	7	368	305	360	287	287	172	91	5	2	7	375	227	316	278	342	174	100
4	11	7	443	508	348	464	539	-3	98	4	19	8	248	285	306	234	298	107	73	5	2	8	38	17	170	158	61	80	55
4	11	8	159	153	160	220	100	22	37	4	19	9	217	222	235	164	220	135	54	5	2	9	220	177	232	249	186	153	82
4	11	9	174	133	246	192	114	-16	93	4	19	10	225	235	240	259	230	74	22	5	2	10	142	151	166	192	161	165	71
4	11	10	422	346	379	307	414	-33	24	4	19	11	207	227	267	288	230	-114	81	5	2	11	558	600	557	586	629	154	97
4	11	11	275	194	271	203	174	-59	94	4	19	12	151	171	204	138	154	-21	61	5	2	12	156	162	164	110	154	132	75
4	11	12	73	137	0	94	110	-103	83	4	19	13	226	166	258	154	197	-21	88	5	2	13	201	239	131	214	233	177	97
4	11	13	352	260	323	349	210	5	99	4	19	14	123	151	128	159	157	-51	77	5	2	14	375	393	438	352	388	-94	65
4	11	14	277	179	206	167	119	52	27	4	19	15	157	133	136	108	138	-41	0	5	2	15	299	251	363	247	259	0	92
4	11	15	343	313	327	394	341	-14	32	4	20	1	151	188	173	194	191	-38	50	5	2	16	102	62	161	78	103	109	23
4	11	16	135	179	219	107	190	-6	80	4	20	2	290	335	337	326	111	-109	95	5	2	17	179	112	142	194	144	178	26
4	11	17	65	92	54	63	110	-1	55	4	20	3	277	245	284	276	241	134	72	5	2	18	370	310	338	283	321	18	41
4	11	18	74	101	24	130	113	-12	89	4	20	4	94	61	60	65	87	-159	70	5	2	19	96	65	112	93	61	161	77
4	11	19	109	126	109	150	123	-1	90	4	20	5	352	314	407	310	308	-68	75	5	2	20	148	110	125	150	127	100	81
4	12	0	179	109	56	146	74	0	100	4	20	6	80	83	137	128	118	-131	67	5	3	0	340	423	245	346	427	180	100
4	12	1	268	246	276	245	245	150	89	4	20	7	119	125	139	171	97	-92	0	5	3	1	10	111	257	81	111	170	74
4	12	2	372	304	402	294	306	76	99	4	20	8	264	272	264	257	279	37	80	5	3	2	103	64	109	82	48	80	100
4	12	3	109	167	64	116	106	-112	94	4	20	9	163	154	177	207	157	44	72	5	3	3	115	154	83	131	157	96	79
4	12	4	364	203	306	214	204	114	99	4	20	10	131	188	145	134	198	125	97	5	3	4	68	173	88	143	176	-117	80
4	12	5	406	512	450	503	526	-140	109	4	20	11	280	274	188	244	273	44	11	5	3	5	133	165	110	109	163	-120	70
4	12	6	257	268	161	191	204	74	94	4	20	12	172	225	158	254	210	47	63	5	3	6	202	155	186	170	157	107	81
4	12	7	492	505	527	512	522	-167	94	4	20	13	79	43	54	114	90	-139	92	5	3	7	100	197	117	225	218	120	83
4	12	8	440	350	374	254	354	-156	99	4	20	14	93	136	97	59	124	-89	74	5	3	8	99	190	110	148	170	74	93
4	12	9	514	537	450	515	554	123	96	4	21	0	88	115	165	0	118	-180	91	5	3	9	197	160	323	278	169	-44	95
4	12	10	271	245	194	212	237	-77	98	4	21	1	172	152	130	159	168	31	52	5	3	10	347	278	294	195	291	-88	100
4	12	11	59	111	103	152	86	-106	90	4	21	2	264	247	212	214	215	-88	88	5	3	11	157	277	328	381	278	33	35
4	12	12	109	129	0	93	97	29	89	4	21	3	295	260	215	237	247	60	93	5	3	12	341	299	343	344	344	-95	39
4	12	13	401	390	404	446	398	-157	60	4	21	4	170	166	127	173	164	-77	58	5	3	13	428	433	284	297	454	47	89
4	12	14	313	317	343	294	323	-120	84	4	21	5	60	85	79	94	98	-94	90	5	3	14	114	109	132	154	108	143	93
4	12	15	290	312	383	352	317	-12	91	4	21	6	240	212	213	181	218	-84	92	5	3	15	179	174	203	201	163	146	28
4	12	16	190	177	186	147	104	-154	39	4	21	7	294	334	336	283	318	21	94	5	3	16	214	235	141	219	248	-64	97
4	12	17	181	178	251	102	101	-154	82	4	21	8	223	235	223	271	230	-13	67	5	3	17	272	262	328	319	297	166	83
4	12	18	73	84	125	97	90	-102	81	4	21	9	455	474	597	429	454	-48	86	5	3	18	165	145	34	183	160	-69	92
4	12	19	93	134	70	145	157	155	55	4	21	10	214	203	155	259	193	-4	47	5	3	19	114	107	85	60	130	52	83
4	13	0	190	234	153	250	214	-140	92	4	21	11	230	229	219	277	229	1	92	5	3	20	86	46	103	117	89	98	74
4	13	1	357	436	398	443	421	-146	96	4	21	12	100	60	98	0	101	154	72	5	4	0	138	214	254	210	213	-180	80
4	13	2	65	52	46	139	49	152	42	4	21	13	74	54	0	84	98	-160	87	5	4	1	142	112	162	93	114	43	87
4	13	3	118	181	101	144	176	-100	94	4	21	14	72	84	67	88	0	87	1	5	4	2	81	67	94	62	66	-23	88
4	13	4	432	425	397	371	437	56	79	4	22	0	37	104	76	164	115	-180	99	5	4	3	129	172	180	139	174	-61	92
4	13	5	300	334	251	313	324	-61	95	4	22	1	254	234	334	278	228	124	38	5	4	4	145	110	214	137	111	-62	91
4	13	6	422	461	394	424	400	16	95	4	22	2	176	139	217	184	142	83	82	5	4	5	116	164	51	149	163	-41	91
4	13	7	465	364	343	350	307	-171	94	4	22	3	370	388	446	387	354	17	8	5	4	6	59	61	108	98	61	-54	74
4	13	8	410	374	400	347	344	117	55	4	22	4	400	443	382	372	394	-15	28	5	4	7	280	264	300	343	260	100	37
4	13	9	293	209	206	260	257	-79	87	4	22	5	280	307	293	240	294	-32	86	5	4	8	292	360	212	223	357	-63	95
4	13	10	216	211	177	179	227	173	89	4	22	6	145	106	149	217	98	122	85	5	4	9	254	244	241	267	264	67	84
4	13	11	192	190	192	240	207	116	91	4	22	7	285	324	299	221	318	-162	81	5	4	10	152	103	62	99	128	91	97
4	13	12	527	500	0	579	583	-124	96	4	22	8	57	103	86	108	123	28	90	5	4	11	694	670	596	681	694	151	83
4	13	13	187	167	126	122	153	-12	91	4	22	9	93	52	83	114	98	94	64	5	4	12	127	122	91	107	117	121	43
4	13	14	112	44	135	119	104	-106	51	4	22	10	260	211	231	146	215	-144	84	5	4	13	503	452	441	432	481	33	73
4	13	15	56	132	104	175	121	-109	57	4	22	11	130	183	155	169	181	78	88	5	4	14	266	249	321	272	258	137	93
4	13	16	79	51	68	73	71	-171	83	4	22	12	207	144	111	118	167	167	83	5	4	15	352	295	308	336	311	60	62
4	13	17	86	103	131	63	118	-51	90	4	22	13	64	72	77	24	0	17	4	5	4	16	250	236	255	225	267	102	87
4	13	18	131	106	71	156	148	45	82	4	23	0	367	371	364	362	302	0											

3 0 11 101 142 0 36 147 113 60
3 0 12 309 310 299 283 307 65 49
3 0 13 143 100 133 114 152-149 45
3 0 14 79 100 0 127 131 6 89
3 0 15 94 167 217 132 134-133 84
3 0 16 40 79 133 115 84 140 75
3 0 17 120 121 137 180 133 7 14
3 0 18 72 99 89 37 133 129 63
3 0 19 130 120 160 131 159 55 79
3 10 1 101 54 123 107 69 57 75
3 10 2 122 131 71 106 116 164 91
3 10 3 213 210 146 211 206 22 73
3 10 4 210 234 218 236 202-142 43
3 10 5 440 470 394 426 403 12 95
3 10 6 103 93 84 126 78 140 95
3 10 7 138 136 103 100 134 44 64
3 10 8 432 394 416 434 379 42 91
3 10 9 432 411 439 444 479 64 68
3 10 10 157 231 107 212 208 113 92
3 10 11 24 99 0 129 90-104 65
3 10 12 218 237 241 272 251 164 70
3 10 13 199 160 249 197 193 65 97
3 10 14 314 358 401 410 351 61 78
3 10 15 139 194 243 230 196-121 91
3 10 16 121 116 140 128 92-107 90
3 10 17 78 74 0 76 106 128 70
3 10 18 57 73 129 101 90 65 54
3 10 19 44 78 0 103 56-159 2
3 11 0 171 132 63 163 151 0 99
3 11 1 203 321 341 333 335 71 47
3 11 2 64 137 0 106 129-115 79
3 11 3 42 30 64 33 52 8 66
3 11 4 147 31 230 121 83 90 100
3 11 5 217 193 270 161 173 46 62
3 11 6 333 434 426 437 508 0 70
3 11 7 601 568 539 571 622 149 73
3 11 8 321 369 417 378 346 43 99
3 11 9 123 134 130 100 104 130 70
3 11 10 193 147 142 136 133 94 100
3 11 11 92 111 0 227 140 65 14
3 11 12 138 101 135 143 152 108 74
3 11 13 377 336 434 439 382 94 82
3 11 14 231 219 219 279 243 71 46
3 11 15 196 204 169 191 207 50 42
3 11 16 93 139 61 114 137 73 94
3 11 17 76 124 102 89 99 69 12
3 11 18 83 119 49 123 125-167 91
3 11 19 73 35 35 126 0 99 22
3 12 0 70 40 71 143 30 0 74
3 12 1 221 323 239 312 325-174 98
3 12 2 434 438 531 436 479 112 94
3 12 3 304 271 224 282 276 100 76
3 12 4 165 173 64 124 170 115 86
3 12 5 624 602 531 564 599 130 76
3 12 6 341 364 287 349 367 98 87
3 12 7 316 341 133 235 300 61 97
3 12 8 344 340 507 521 554 6 65
3 12 9 265 264 224 231 266 111 62
3 12 10 279 314 303 302 317-129 92
3 12 11 161 133 0 200 124 36 95
3 12 12 271 237 293 132 245 75 80
3 12 13 70 113 169 170 119 131 93
3 12 14 101 167 262 164 120-174 87
3 12 15 149 134 201 214 171 149 50
3 12 16 93 147 109 109 131 41 75
3 12 17 180 130 113 103 174 176 67
3 12 18 177 237 139 219 257 23 93
3 13 0 363 639 334 363 404-140 98
3 13 1 71 132 77 237 113 12 90
3 13 2 510 514 523 474 519-147 86
3 13 3 131 180 141 209 180 443 88
3 13 4 441 410 476 407 414 167 67
3 13 5 377 391 611 572 608-116 93
3 13 6 234 216 186 204 224-124 78
3 13 7 724 693 637 701 710-125 38
3 13 8 623 527 507 481 512-141 98
3 13 9 216 173 160 188 195 57 85
3 13 10 304 234 304 286 265-129 83
3 13 11 119 127 0 222 124 0 56
3 13 12 222 270 273 291 272 109 22
3 13 13 260 302 318 317 327 3 96
3 13 14 309 298 348 334 314 177 36
3 13 15 243 214 294 196 237-173 16
3 13 16 165 186 211 190 199-125 89
3 13 17 177 165 151 136 174 23 84
3 13 18 107 103 0 0 93 128 7
3 14 0 228 248 242 193 254-140 71
3 14 1 34 41 34 90 97 54 65
3 14 2 64 125 86 127 135 64 89
3 14 3 271 291 293 313 300 67 84
3 14 4 281 290 271 243 293 26 84
3 14 5 213 203 193 211 216 157 66
3 14 6 123 142 103 180 137 60 37
3 14 7 234 137 120 168 129 42 85
3 14 8 494 446 412 433 404 117 67
3 14 9 225 256 236 231 272 109 86
3 14 10 146 104 142 147 94 121 93
3 14 11 163 202 0 233 236 70 82
3 14 12 249 234 300 240 249 139 62
3 14 13 122 113 104 116 123 67 61
3 14 14 307 255 407 314 296 4 77
3 14 15 394 342 443 343 403 134 47
3 14 16 92 103 149 161 119 156 83
3 14 17 103 107 136 160 141 26 58
3 15 0 61 67 119 161 67 0 85
3 15 1 611 397 606 600 569-120 39
3 15 2 146 243 124 221 236 98 96
3 15 3 234 227 248 230 234 159 11
3 15 4 296 243 233 240 245 95 99
3 15 5 328 314 301 266 317-173 11
3 15 6 432 434 0 419 423 64 9
3 15 7 233 246 199 279 251 156 52
3 15 8 187 230 273 223 237 89 95
3 15 9 301 291 263 298 276 121 35
3 15 10 169 107 178 130 95 92 99
3 15 11 221 243 297 293 234 38 72
3 15 12 88 93 84 136 93 172 81
3 15 13 183 190 176 169 176 172 74
3 15 14 190 212 143 200 212 46 64
3 15 15 208 217 200 134 229 44 64
3 15 16 133 134 162 133 134 13 83
3 15 17 93 66 0 30 76 15 56
3 16 0 339 334 316 312 321-100 66
3 16 1 182 293 246 264 298 27 98
3 16 2 94 113 71 144 122 67 74
3 16 3 128 97 76 100 30-128 67
3 16 4 211 209 143 210 206 137 67
3 16 5 313 376 299 317 377 171 91
3 16 6 387 368 0 333 372 60 60
3 16 7 419 312 302 303 298 42 99
3 16 8 250 273 269 233 276 12 53
3 16 9 377 397 334 332 384-137 63
3 16 10 133 101 162 166 117-159 71
3 16 11 203 233 277 213 210 149 97
3 16 12 316 339 313 294 362 45 90
3 16 13 86 14 107 133 94-140 26
3 16 14 163 191 236 240 192 95 75
3 16 15 69 90 162 136 102 37 67
3 16 16 133 166 60 113 141 160 75
3 17 0 64 30 103 83 72 0 65
3 17 1 333 336 344 316 503 0 71
3 17 2 336 324 270 270 318 76 59
3 17 3 104 118 47 127 52 42 60
3 17 4 132 183 133 170 153 75 91
3 17 5 177 239 174 262 261 31 94
3 17 6 324 333 343 300 323 39 69
3 17 7 178 111 146 92 122 106 97
3 17 8 207 234 114 212 227-117 95
3 17 9 103 131 71 113 100 65 94
3 17 10 216 207 176 128 201 79 68

6 1 111015 964 943 950 964 6 46
6 1 12 378 322 363 394 376 6 66
6 1 13 131 87 52 123 90 6 97
6 1 14 79 75 33 126 68 143 30
6 1 15 146 116 166 74 142 37 54
6 1 16 48 102 31 66 124 60 65
6 1 17 126 128 0 108 121-138 3
6 1 18 108 64 140 191 106-131 51
6 1 19 143 103 140 159 147 48 64
6 1 20 143 132 134 161 144 33 12
6 2 0 221 198 75 197 199-160 100
6 2 1 133 194 151 193 196-139 39
6 2 2 333 219 262 216 221 100 100
6 2 3 91 133 86 149 137-121 49
6 2 4 234 314 230 315 317-143 98
6 2 5 277 291 243 264 290 50 23
6 2 6 114 132 86 129 120 23 43
6 2 7 120 149 101 133 156 15 69
6 2 8 176 77 126 86 86 85 100
6 2 9 81 123 33 87 116 121 54
6 2 10 349 312 423 489 496 58 99
6 2 11 330 367 340 377 382 161 59
6 2 12 296 294 337 385 374 94 73
6 2 13 317 325 313 297 333-166 63
6 2 14 163 175 126 160 186-134 74
6 2 15 463 482 456 358 458-148 71
6 2 16 143 137 123 151 172 65 96
6 2 17 188 196 174 196 196 49 1
6 2 18 270 223 226 237 244 59 92
6 2 19 73 71 104 113 71 7 78
6 3 0 272 170 439 204 171 0 100
6 3 1 33 119 0 71 120 73 95
6 3 2 269 242 143 228 244 28 90
6 3 3 261 224 239 237 225 96 55
6 3 4 225 276 191 195 278 13 92
6 3 5 57 153 125 119 153 20 87
6 3 6 352 309 218 210 322-132 86
6 3 7 141 218 170 209 207 20 96
6 3 8 210 148 149 144 193 43 26
6 3 9 198 216 177 186 201 7 56
6 3 10 312 327 632 621 532-124 96
6 3 11 340 374 346 348 346-170 83
6 3 12 321 476 313 341 469 58 69
6 3 13 233 207 246 218 207 7 99
6 3 14 32 32 46 119 49 30 70
6 3 15 198 168 72 220 193 5 96
6 3 16 246 221 148 173 232 9 87
6 3 17 80 91 91 144 129 30 80
6 3 18 76 98 79 93 91 132 87
6 3 19 113 70 132 119 96 76 47
6 3 20 153 133 133 156 140-136 2
6 4 0 134 98 124 140 101 0 77
6 4 1 59 98 80 79 99 42 20
6 4 2 38 110 40 97 109 99 74
6 4 3 254 248 311 237 251-156 79
6 4 4 273 321 277 376 322 42 89
6 4 5 108 193 142 144 174-151 93
6 4 6 162 94 72 107 97 123 56
6 4 7 172 163 170 165 170 104 91
6 4 8 240 220 268 210 236 36 27
6 4 9 487 421 428 403 401-107 95
6 4 10 303 313 314 334 310 95 82
6 4 11 457 334 390 423 359 48 69
6 4 12 113 162 168 180 149 6 94
6 4 13 202 199 200 191 220-122 58
6 4 14 179 141 165 177 160 4 75
6 4 15 109 112 91 175 140-153 23
6 4 16 130 133 143 185 146-149 39
6 4 17 83 104 142 138 0 73 31
6 5 0 41 89 162 39 87 0 69
6 5 1 132 137 180 130 157 13 77
6 5 2 142 123 237 143 124 64 99
6 5 3 174 116 14 94 118-173 83
6 5 4 69 121 156 173 168 51 90
6 5 5 316 313 237 241 293 45 67
6 5 6 197 151 89 105 172 30 55
6 5 7 381 356 342 312 341 40 96
6 5 8 379 410 363 336 403 39 88
6 5 9 178 114 127 156 83 2 76
6 5 10 173 183 114 168 108 112 99
6 5 11 760 726 699 710 663 90 70
6 5 12 144 201 279 311 144-156 52
6 5 13 311 320 298 316 314 74 47
6 5 14 80 127 0 90 169 166 53
6 5 15 77 101 103 91 103 88 80
6 5 16 269 307 212 234 352 131 93
6 5 17 46 84 105 117 160-111 57
6 5 18 209 229 244 299 249 123 91
6 5 19 163 194 196 179 204-168 94
6 6 0 39 36 44 61 36-180 18
6 6 1 223 282 215 271 287 5 90
6 6 2 344 224 243 223 276 93 97
6 6 3 223 199 244 213 193 61 23
6 6 4 301 356 343 332 342 61 97
6 6 5 107 119 123 143 110 103 75
6 6 6 184 210 179 197 197 44 61
6 6 7 177 219 232 227 209-157 89
6 6 8 173 159 162 170 133 77 96
6 6 9 337 610 345 340 571 162 34
6 6 10 336 333 508 313 521-177 9
6 6 11 279 339 289 310 301 30 51
6 6 12 346 333 337 342 326 88 87
6 6 13 104 97 152 114 85 162 87
6 6 14 108 122 184 154 118-126 83
6 6 15 318 312 303 244 335 33 80
6 6 16 213 176 234 260 267-167 10
6 6 17 118 153 174 163 145 92 19
6 6 18 108 97 117 176 123 2 92
6 6 19 80 104 54 97 132 38 92
6 7 0 124 86 47 91 84-170 100
6 7 1 360 479 446 474 440 2 98
6 7 2 130 101 183 122 164 178 26
6 7 3 328 383 314 348 343 59 70
6 7 4 133 173 234 217 159 179 78
6 7 5 103 81 120 102 79 73 84
6 7 6 243 243 296 253 237 39 53
6 7 7 152 143 141 143 144 730 94
6 7 8 138 134 171 142 126 9 42
6 7 9 491 448 391 434 439 34 93
6 7 10 231 278 240 306 269 64 91
6 7 11 221 202 217 243 147 134 98
6 7 12 136 154 72 157 132-144 96
6 7 13 254 321 274 302 326 12 98
6 7 14 168 190 153 91 107 17 18
6 7 15 129 154 47 127 174 27 93
6 7 16 143 204 233 230 206-112 85
6 7 17 141 147 114 137 147 142 72
6 7 18 133 140 144 229 184 129 83
6 7 19 81 119 121 113 144 128 94
6 8 0 282 229 246 198 218-140 97
6 8 1 108 153 154 141 140-173 93
6 8 2 92 37 47 39 43 0 76
6 8 3 109 110 123 49 141 10 76
6 8 4 233 292 254 272 243 76 96
6 8 5 147 176 222 178 182 73 43
6 8 6 106 144 141 157 174 163 86
6 8 7 330 336 343 311 293-176 98
6 8 8 231 218 243 227 221 34 96
6 8 9 122 123 192 137 130 83 66
6 8 10 377 310 471 303 494 110 96
6 8 11 228 223 193 271 230-131 67
6 8 12 248 223 194 212 218 149 87
6 8 13 46 39 0 14 12 133 87
6 8 14 132 133 124 127 146 65 89
6 8 15 132 89 143 33 0 148 42
6 8 16 133 131 131 191 176 96 87
6 8 17 270 210 234 197 228 74 67
6 8 18 49 41 0 132 138 113 81
6 8 19 141 147 141 141 188 73 64

6	0	1	651	692	539	610	650	0	100	6	17	5	166	136	135	171	151	-70	50	7	2	1	161	164	171	05	155	-113	00
6	0	1	300	347	318	300	316	36	78	6	17	6	66	82	135	136	92	174	84	7	2	2	256	259	260	190	268	-86	60
6	0	2	312	361	328	340	332	-72	96	6	17	7	309	333	313	284	337	-170	95	7	2	3	175	200	216	245	223	161	57
6	0	3	148	208	180	200	191	71	74	6	17	8	284	296	294	254	297	57	69	7	2	4	252	276	225	265	272	53	8
6	0	4	104	82	248	141	84	4	67	6	17	9	104	119	160	136	118	-32	32	7	2	5	366	444	321	407	439	-21	98
6	0	5	241	188	149	195	170	-22	84	6	17	10	138	149	110	125	90	-97	94	7	2	6	201	223	193	231	205	-49	14
6	0	6	375	622	368	579	592	-174	63	6	17	11	521	321	394	292	361	-54	77	7	2	7	142	153	216	132	161	21	50
6	0	7	349	322	309	330	306	-64	32	6	17	12	173	172	91	127	194	12	92	7	2	8	367	351	241	266	349	-57	60
6	0	8	224	216	233	186	217	22	54	6	17	13	81	90	124	90	114	-51	74	7	2	9	175	176	230	217	178	-148	20
6	0	9	338	333	322	296	323	32	27	6	17	14	26	124	134	190	136	-50	83	7	2	10	240	273	268	280	265	-175	48
6	0	10	247	249	393	414	209	-72	87	6	17	15	201	191	198	164	207	120	89	7	2	11	376	370	373	413	357	-21	36
6	0	11	248	213	253	205	208	-7	39	6	18	1	141	97	162	140	54	-180	97	7	2	12	327	334	0	373	367	64	50
6	0	12	149	107	192	190	130	60	27	6	18	2	152	130	167	144	94	-5	97	7	2	13	124	166	120	179	172	143	90
6	0	13	161	149	239	224	177	30	63	6	18	3	196	174	0	163	161	-169	21	7	2	14	126	157	103	107	137	-14	67
6	0	14	140	154	129	72	151	-9	38	6	18	4	342	311	0	294	299	-71	92	7	2	15	167	146	0	213	118	-171	93
6	0	15	92	45	0	85	107	115	42	6	18	5	180	172	213	204	154	79	30	7	2	16	229	230	270	314	257	14	83
6	0	16	186	241	219	270	287	-77	88	6	18	6	395	314	360	254	306	-154	98	7	2	17	158	135	83	158	153	100	71
6	0	17	68	95	0	129	100	112	34	6	18	7	120	105	134	145	127	110	70	7	2	18	236	228	231	204	241	32	53
6	0	18	55	65	0	63	100	-151	85	6	18	8	216	162	207	187	172	171	97	7	3	0	206	174	220	199	171	0	40
6	0	19	55	92	111	79	0	-175	25	6	18	9	286	265	260	245	270	113	41	7	3	1	149	226	156	155	196	-88	55
6	10	0	165	185	232	208	217	-180	46	6	18	10	75	124	103	184	120	-22	90	7	3	2	200	247	220	265	264	-102	90
6	10	1	60	156	72	133	150	178	89	6	18	11	330	293	335	324	304	-43	60	7	3	3	257	266	168	311	254	-121	43
6	10	2	82	128	120	134	123	-73	89	6	18	12	178	180	211	209	181	-87	66	7	3	4	202	113	231	159	117	84	100
6	10	3	509	514	419	407	516	99	69	6	18	13	121	123	160	135	129	-13	45	7	3	5	273	268	248	217	250	-75	74
6	10	4	168	0	111	121	70	97	99	6	18	14	140	132	180	81	136	137	43	7	3	6	331	311	278	261	248	42	55
6	10	5	497	524	508	532	525	43	91	6	19	0	103	79	62	92	162	-180	81	7	3	7	183	193	211	179	203	160	88
6	10	6	371	305	221	275	281	-130	43	6	19	1	124	140	110	103	100	-46	16	7	3	8	371	240	320	304	213	-92	90
6	10	7	398	418	415	409	430	24	92	6	19	2	167	115	0	150	207	-56	64	7	3	9	249	292	196	248	281	111	68
6	10	8	268	237	220	204	221	-72	78	6	19	3	245	272	0	218	237	-32	44	7	3	10	531	450	454	406	-54	-90	62
6	10	9	166	102	107	130	143	114	46	6	19	4	150	206	183	196	180	-102	84	7	3	11	245	279	318	344	290	39	81
6	10	10	167	227	0	219	216	102	95	6	19	5	275	220	282	230	218	157	80	7	3	12	316	299	0	307	300	-75	28
6	10	11	217	264	244	255	255	-113	81	6	19	6	271	263	297	219	270	-59	24	7	3	13	201	208	141	245	207	24	63
6	10	12	244	228	303	246	259	-46	65	6	19	7	431	284	265	246	264	-14	91	7	3	14	253	259	267	294	248	-104	76
6	10	13	336	370	414	340	375	-140	94	6	19	8	126	110	133	194	115	-117	85	7	3	15	276	305	0	260	306	-59	68
6	10	14	89	40	158	122	0	-82	37	6	19	9	243	254	254	215	257	115	46	7	3	16	219	250	131	246	269	-101	97
6	10	15	104	139	73	105	139	101	89	6	19	10	167	166	301	271	157	-173	90	7	3	17	113	139	108	93	157	107	35
6	10	16	104	165	103	155	130	-103	88	6	19	11	172	135	181	102	150	-37	40	7	3	18	275	258	230	291	245	-55	84
6	10	17	161	150	145	149	167	18	6	6	19	12	122	131	167	175	134	81	70	7	3	19	171	168	140	194	144	-135	64
6	10	18	57	114	58	65	114	9	89	6	19	13	50	80	74	103	47	-113	65	7	4	0	53	62	13	50	56	-180	6
6	11	0	399	416	288	302	308	-180	100	6	19	14	132	96	174	74	0	134	24	7	4	1	176	139	175	163	132	-25	79
6	11	1	133	246	107	210	240	-162	97	6	20	0	178	153	170	150	147	0	74	7	4	2	49	126	125	62	123	128	92
6	11	2	325	327	316	341	326	-39	10	6	20	1	135	109	150	120	125	-70	12	7	4	3	311	330	297	317	310	74	0
6	11	3	234	163	81	136	128	28	94	6	20	2	148	91	185	131	82	80	99	7	4	4	98	193	183	199	192	78	94
6	11	4	505	477	581	493	478	-151	73	6	20	3	130	101	89	144	125	-26	40	7	4	5	240	293	247	276	240	9	60
6	11	5	562	548	404	527	544	120	80	6	20	4	461	445	476	384	405	89	95	7	4	6	250	243	248	293	245	125	27
6	11	6	449	448	319	402	457	176	75	6	20	5	212	180	147	153	191	70	36	7	4	7	318	348	314	346	328	-21	63
6	11	7	238	178	129	154	155	109	100	6	20	6	176	175	166	200	195	-37	63	7	4	8	606	566	514	464	533	92	60
6	11	8	404	384	336	318	387	-61	7	6	20	7	274	240	208	286	260	-55	65	7	4	9	274	279	300	225	257	167	45
6	11	9	116	106	155	185	86	-146	63	6	20	8	307	248	211	181	255	-137	57	7	4	10	62	53	170	58	61	77	93
6	11	10	225	235	0	243	227	-135	26	6	20	9	152	126	164	187	121	-36	36	7	4	11	422	399	417	435	392	-97	43
6	11	11	148	146	232	159	142	123	42	6	20	10	257	211	242	224	214	-95	58	7	4	12	139	155	0	143	139	173	68
6	11	12	121	149	287	251	99	-1	79	6	20	11	94	74	144	114	107	-21	57	7	4	13	75	76	135	147	71	19	86
6	11	13	333	379	343	362	420	-175	98	6	20	12	107	71	113	124	86	4	75	7	4	14	89	135	86	53	0	80	65
6	11	14	254	255	282	251	252	-179	37	6	20	13	82	36	54	59	79	12	39	7	4	15	318	311	0	297	307	40	81
6	11	15	104	147	143	169	141	-107	91	6	21	0	124	67	77	239	74	0	99	7	4	16	535	281	278	240	316	-84	84
6	11	16	172	190	173	162	218	-16	60	6	21	1	414	431	430	551	403	-174	81	7	4	17	8						

7 9 14 122 133 96 176 159 138 55
7 9 13 70 89 63 74 129 141 89
7 9 16 81 68 81 89 102 44 24
7 9 17 219 217 239 233 231 34 64
7 9 18 96 46 60 67 117 68 82
7 10 0 373 370 368 368 341-140 78
7 10 1 119 147 74 139 120 138 90
7 10 2 160 226 99 168 216 -69 96
7 10 3 212 233 200 214 249-107 69
7 10 4 49 147 111 130 137 -87 86
7 10 5 292 233 230 227 223-113 96
7 10 6 127 146 160 194 123 37 23
7 10 7 197 133 120 93 93-133 99
7 10 8 390 362 0 337 346 4 92
7 10 9 147 197 213 201 190 67 67
7 10 10 173 123 223 168 118 -73 59
7 10 11 120 114 53 103 91 -27 63
7 10 12 193 202 234 211 190-161 38
7 10 13 147 202 130 171 205-163 93
7 10 14 139 166 163 139 166 -39 83
7 10 15 142 137 133 107 164 -93 79
7 10 16 121 173 182 201 171-143 93
7 10 17 111 86 193 123 110 -43 64
7 10 18 71 124 33 113 0 129 22
7 11 0 244 206 133 193 194-180 15
7 11 1 316 279 242 266 279 34 48
7 11 2 132 143 0 68 111 73 57
7 11 3 234 103 100 190 161 -73 73
7 11 4 172 221 192 212 216-114 47
7 11 5 127 177 168 130 173 62 66
7 11 6 76 99 129 98 103 -61 82
7 11 7 344 371 344 391 337 104 48
7 11 8 140 232 0 217 220 76 93
7 11 9 30 65 119 113 33 -42 28
7 11 10 420 461 394 419 433 40 97
7 11 11 176 191 210 202 163 -17 89
7 11 12 144 171 223 203 179 167 54
7 11 13 93 123 74 79 166 -17 69
7 11 14 140 114 162 99 119 170 53
7 11 15 63 81 101 99 80 162 53
7 11 16 80 140 87 129 148 -97 67
7 11 17 100 130 110 164 168 40 69
7 12 0 96 126 134 139 114-140 9
7 12 1 146 232 120 192 213 174 96
7 12 2 394 371 380 340 351-139 33
7 12 3 244 263 272 263 267 -97 63
7 12 4 134 141 107 113 142-123 77
7 12 5 180 139 192 137 143 164 94
7 12 6 346 321 283 283 296 -65 42
7 12 7 316 279 291 268 234 136 99
7 12 8 113 138 0 230 136-108 61
7 12 9 242 243 273 210 220-120 67
7 12 10 127 88 128 129 51 49 52
7 12 11 119 204 162 189 171 136 92
7 12 12 172 130 172 123 143 126 88
7 12 13 99 82 70 81 93 162 52
7 12 14 103 144 120 90 120 89 15
7 12 15 99 111 146 123 92 -38 31
7 12 16 111 122 69 91 128 -53 93
7 12 17 73 103 96 140 120 -9 81
7 13 0 161 166 221 219 134-180 99
7 13 1 132 129 171 132 121 139 92
7 13 2 218 233 221 213 219 -13 77
7 13 3 301 285 273 233 270 161 78
7 13 4 223 266 241 232 234 -68 72
7 13 5 172 233 193 229 239-173 93
7 13 6 134 209 163 207 202-119 9
7 13 7 117 171 85 187 163 173 95
7 13 8 340 332 0 320 318 -46 59
7 13 9 91 67 38 170 77 -43 29
7 13 10 233 204 307 271 202 -18 80
7 13 11 268 247 244 262 251 -82 59
7 13 12 113 87 33 88 169 26 66
7 13 13 276 236 331 219 234 108 63
7 13 14 126 132 142 133 134 -48 62
7 13 15 133 167 189 202 184 -32 70
7 13 16 78 90 33 0 40 -42 68
7 14 0 202 213 180 191 201-180 13
7 14 1 168 139 131 137 146-164 62
7 14 2 92 64 41 14 62 -51 93
7 14 3 121 108 66 97 112 141 41
7 14 4 297 320 264 249 299 39 83
7 14 5 201 204 223 238 189 -83 25
7 14 6 323 313 290 277 283-117 92
7 14 7 140 120 118 133 106 26 70
7 14 8 207 232 0 188 222 -76 84
7 14 9 193 174 192 231 187 138 26
7 14 10 353 338 392 393 399 -79 97
7 14 11 130 162 134 124 131 46 38
7 14 12 241 230 287 238 260-110 84
7 14 13 97 142 179 124 115 80 64
7 14 14 133 146 237 146 173 67 84
7 14 15 113 127 147 103 123 85 40
7 14 16 171 172 113 108 197 96 94
7 15 0 203 200 238 188 192 0 62
7 15 1 133 170 108 168 140 -13 56
7 15 2 224 262 217 234 239 62 89
7 15 3 163 147 99 162 141 173 23
7 15 4 369 356 330 329 337 -71 91
7 15 5 213 183 138 178 163 32 61
7 15 6 128 136 143 138 146 172 52
7 15 7 303 294 277 198 271 3 63
7 15 8 86 80 0 81 79 66 18
7 15 9 490 514 478 411 309 126 25
7 15 10 179 168 140 138 171 25 33
7 15 11 409 399 344 313 416-179 43
7 15 12 138 177 93 187 161 -3 74
7 15 13 123 143 181 208 133 -31 57
7 15 14 111 123 76 113 116 139 90
7 15 15 168 193 0 147 188 35 28
7 16 0 69 84 80 83 93 0 14
7 16 1 203 209 238 208 211-166 99
7 16 2 163 140 172 133 122 110 16
7 16 3 132 146 133 138 146 18 59
7 16 4 119 113 107 68 113 166 56
7 16 5 188 242 221 209 251 161 96
7 16 6 138 133 143 139 144 16 22
7 16 7 136 177 173 208 187-173 97
7 16 8 113 107 149 143 114 99 5
7 16 9 170 167 209 234 173 147 74
7 16 10 421 418 421 420 415 -26 40
7 16 11 142 164 147 139 163 -12 38
7 16 12 86 53 78 86 68 100 48
7 16 13 172 132 232 178 131-130 67
7 16 14 102 132 126 83 114 67 31
7 16 15 118 76 0 128 119-137 1
7 17 0 204 216 304 220 213-180 100
7 17 1 149 133 146 77 125 -60 64
7 17 2 119 143 89 127 137 113 69
7 17 3 173 172 93 138 161 -83 19
7 17 4 146 112 139 93 133 -61 61
7 17 5 273 238 302 221 269 -71 68
7 17 6 171 131 194 166 143-112 84
7 17 7 97 106 39 146 86-142 51
7 17 8 172 144 184 94 163 132 73
7 17 9 202 266 181 294 292 -32 72
7 17 10 132 122 136 144 115 47 56
7 17 11 138 135 180 216 143 -73 70
7 17 12 220 223 283 223 238-146 60
7 17 13 128 121 178 140 123 98 43
7 17 14 93 120 117 114 123 90 74
7 18 0 171 169 190 134 162-180 24
7 18 1 311 333 369 331 314 106 54
7 18 2 181 104 239 138 123 125 99
7 18 3 197 182 236 180 174 -39 36
7 18 4 243 233 191 180 213 82 99
7 18 5 204 202 122 197 223-164 92
7 18 6 123 110 124 137 0 74 4
7 18 7 176 208 222 207 201 156 92
7 18 8 213 209 312 233 211 44 5
7 18 9 10 37 91 106 64 95 -4 87
7 18 10 161 169 184 201 141 2 39
7 18 11 111 87 32 37 79-170 77
7 18 12 79 94 104 107 101 -12 74
7 19 0 167 137 109 203 133 180 88
7 19 1 194 246 241 221 222 177 91
7 19 2 214 202 173 217 163 102 94
7 19 3 148 134 134 107 137 -34 20
7 19 4 131 91 100 94 113 74 31
7 19 5 237 298 237 244 287 4 94
7 19 6 137 143 147 147 116 101 51
7 19 7 126 144 130 179 133 -21 51
7 19 8 230 226 326 232 229-121 80
7 19 9 307 253 326 204 291 22 24
7 19 10 136 141 141 172 197 104 93
7 19 11 236 220 271 242 229 24 82
7 19 12 128 127 143 128 122-109 96
7 19 13 116 123 107 81 0 103 3
7 20 0 177 183 183 239 192 0 20
7 20 1 204 242 224 232 267 173 97
7 20 2 122 94 140 98 109 17 52
7 20 3 44 38 134 93 69 174 86
7 20 4 234 237 225 269 238 -27 50
7 20 5 248 233 232 283 267 -80 91
7 20 6 335 373 374 313 363 132 43
7 20 7 438 224 198 219 213 90 38
7 20 8 117 103 69 70 104 132 73
7 20 9 166 166 141 192 139 32 57
7 20 10 176 192 168 191 182 77 54
7 20 11 126 133 112 165 0-100 4
7 21 0 216 226 293 288 213 180 100
7 21 1 279 272 241 234 263 62 46
7 21 2 197 149 160 201 148 69 93
7 21 3 364 363 298 267 331 23 79
7 21 4 114 119 112 93 103 177 2
7 21 5 105 94 33 62 109 100 86
7 21 6 108 122 111 139 114 -73 60
7 21 7 246 239 262 200 243 -147 47
7 21 8 126 84 35 143 77 -95 93
7 21 9 217 198 196 141 209 41 60
7 21 10 218 164 181 167 133 -60 96
7 21 11 173 137 144 164 0 79 19
7 22 0 326 269 272 232 272 0 98
7 22 1 340 324 363 274 311 74 43
7 22 2 109 128 133 139 132 128 90
7 22 3 193 178 177 173 172 132 14
7 22 4 238 259 198 187 233 83 85
7 22 5 308 281 319 299 276 37 53
7 22 6 129 134 211 163 143 61 92
7 22 7 88 78 109 144 89 75 78
7 22 8 293 292 281 269 283 45 88
7 22 9 183 167 162 163 160 -19 24
7 22 10 133 107 167 17 0 89 30
7 23 1 193 146 179 188 149 137 98
7 23 2 107 100 132 86 104 -66 30
7 23 3 116 90 170 127 97-141 8
7 23 4 181 190 183 193 164 -54 26
7 23 5 171 181 204 167 168 64 27
7 23 6 149 135 52 143 127 111 83
7 23 7 37 78 83 84 0 81 17
7 24 0 33 33 114 74 37 0 13
7 24 1 195 161 300 164 143-173 98
7 24 2 155 100 163 130 90 -84 67
7 24 3 134 138 221 169 140-174 88
7 24 4 147 111 153 118 110 -72 49
7 24 5 11 53 81 77 33-123 59
7 25 0 127 126 41 20 88 0 92
7 25 1 122 111 134 134 110 74 20
7 25 2 264 283 209 214 234-108 91
7 25 3 263 277 226 219 0-129 17
7 25 4 287 249 171 281 0 122 74
8 0 0 436 301 128 179 287-180 100
8 0 1 427 347 462 311 320-180 100
8 0 2 141 67 71 57 64 0 80
8 0 3 244 261 230 213 267 180 79
8 0 4 214 183 267 184 194-180 94
8 0 5 38 153 139 94 115 0 96
8 0 6 287 257 92 264 204-180 100
8 0 7 302 208 229 147 163-180 100
8 0 8 297 274 274 160 274-180 5
8 0 9 136 190 106 120 171 0 100
8 0 10 293 212 122 217 159 0 100
8 0 11 327 358 468 618 331 180 80
8 0 12 74 40 73 109 63 0 19
8 0 13 36 63 70 119 90 180 85
8 0 14 104 71 86 61 94 0 10
8 0 15 86 168 0 222 153-180 100
8 0 16 224 243 334 253 243-180 100
8 0 17 159 151 169 130 184 0 66
8 0 18 251 219 210 204 222-180 99
8 1 0 35 30 173 123 60 0 93
8 1 1 233 238 190 221 211 37 94
8 1 2 291 181 281 244 177 107 99
8 1 3 250 260 232 250 257 -39 54
8 1 4 173 252 169 261 260 -88 97
8 1 5 223 234 203 183 214-113 71
8 1 6 307 310 363 264 283 -28 67
8 1 7 194 172 184 186 134-164 97
8 1 8 249 187 274 213 192 -34 95
8 1 9 429 383 421 402-146 52
8 1 10 616 539 576 574 551 -80 98
8 1 11 253 244 170 241 233 113 14
8 1 12 249 208 222 178 208-129 68
8 1 13 180 223 257 234 224 179 94
8 1 14 114 84 113 114 92 133 37
8 1 15 143 185 0 207 181 -75 88
8 1 16 131 130 160 137 173 -89 77
8 1 17 93 81 33 110 40 133 64
8 1 18 37 80 40 97 121-121 86
8 2 0 65 30 68 77 95-180 27
8 2 1 239 214 273 218 214 -98 25
8 2 2 288 276 315 334 271 38 59
8 2 3 104 120 76 69 119 15 74
8 2 4 248 328 229 306 322-101 98
8 2 5 203 185 208 146 160-144 73
8 2 6 329 339 277 313 163 28 28
8 2 7 120 112 85 110 103 -19 54
8 2 8 394 331 392 337 348-121 90
8 2 9 660 618 384 390 100 68
8 2 10 446 383 437 436 386 -79 99
8 2 11 171 246 204 196 243 -94 97
8 2 12 226 223 281 168 250 108 89
8 2 13 167 198 237 234 213-107 89
8 2 14 189 221 133 239 230 -49 71
8 2 15 188 190 0 231 194-123 31
8 2 16 77 118 93 160 106 -94 91
8 2 17 110 150 140 91 138 93 44
8 2 18 90 104 114 133 124-173 91
8 3 0 337 366 403 461 327 0 94
8 3 1 223 267 261 264-148 96
8 3 2 441 373 442 387 364-111 96
8 3 3 333 332 323 303 321 7 44
8 3 4 66 70 60 97 67 -27 63
8 3 5 172 97 82 116 70 173 99
8 3 6 230 263 329 232 261 127 83
8 3 7 150 204 233 243 214 12 96
8 3 8 235 234 241 236 230-137 19
8 3 9 173 178 233 203 177 44 41
8 3 10 443 382 387 339 366 77 96
8 3 11 121 136 193 101 126 103 87
8 3 12 363 397 387 331 399-134 76
8 3 13 37 33 93 119 59 172 48
8 3 14 70 48 94 70 81 -56 80
8 3 15 233 206 0 183 217 -34 87
8 3 16 173 160 217 151 183 68 84
8 3 17 67 53 70 74 84 -12 11
8 3 18 238 241 223 243 260 5 51
8 4 0 209 169 209 223 131-180 96
8 4 1 302 397 327 343 360 -14 97
8 4 2 137 136 130 83 133-126 28
8 4 3 197 198 226 213 204 109 48
8 4 4 146 129 213 120 133 144 56
8 4 5 372 311 357 258 284 39 96
8 4 6 232 232 161 211 210 142 63
8 4 7 433 367 418 398 330 0 97
8 4 8 143 92 57 114 91 -78 83
8 4 9 323 301 263 236 273 1 82
8 4 10 268 224 329 274 233 142 74
8 4 11 374 402 342 373 399 0 87
8 4 12 204 194 168 139 211 -81 40
8 4 13 103 127 138 183 148 30 93
8 4 14 114 132 133 66 146 -31 39
8 4 15 244 272 0 176 305 41 91
8 4 16 182 202 123 193 221-123 74
8 4 17 96 81 108 67 92 64 94
8 4 18 97 80 101 101 103 111 36
8 5 0 286 310 233 232 278-180 43
8 5 1 44 100 106 99 80 31 80
8 5 2 438 340 337 332 327 102 97
8 5 3 222 193 183 197 182 -81 82
8 5 4 317 420 338 384 393 94 98
8 5 5 103 101 98 110 97 50 69
8 5 6 178 170 237 177 177 54 43
8 5 7 129 73 90 94 58 23 96
8 5 8 351 299 364 310 273 114 98
8 5 9 157 174 0 132 173-133 69
8 5 10 433 437 443 400 407 126 79
8 5 11 127 130 109 105 114 74 49
8 5 12 252 262 246 198 238-120 47
8 5 13 92 73 162 123 93 -88 46
8 5 14 93 151 97 130 130 80 80
8 5 15 148 109 131 161 137 114 67
8 5 16 64 119 43 108 64 -16 79
8 5 17 63 77 40 100 90-173 83
8 5 18 99 154 142 121 130 33 76
8 6 0 93 110 0 66 118 0 73
8 6 1 379 347 323 337 299 141 97
8 6 2 137 173 104 180 156 33 88
8 6 3 101 114 142 116 113 -94 60
8 6 4 332 339 243 392 337-131 48
8 6 5 309 280 252 280 262 22 94
8 6 6 115 127 153 113 94 -70 89
8 6 7 340 396 342 433 350-170 84
8 6 8 153 134 143 122 132 68 38
8 6 9 233 273 0 313 271 134 72
8 6 10 233 148 96 133 126 103 99
8 6 11 213 256 264 223 233 86 87
8 6 12 90 146 140 160 136 -68 71
8 6 13 278 301 291 234 276 120 23
8 6 14 163 174 169 146 186 116 63
8 6 15 257 249 183 134 264 160 50
8 6 16 49 91 114 90 103 173 86
8 6 17 92 130 94 184 153-123 86
8 6 18 132 170 103 123 0 13 41
8 7 0 360 347 364 388 326 180 99
8 7 1 90 186 194 193 192 -11 93
8 7 2 170 192 217 193 202-137 66
8 7 3 209 170 173 130 163-170 67
8 7 4 226 231 221 202 206 127 18
8 7 5 214 120 114 128 110 1 99
8 7 6 240 253 134 187 241 -43 75
8 7 7 314 304 341 297 311-144 83
8 7 8 276 294 197 232 238-173 74
8 7 9 273 234 0 224 226 39 85
8 7 10 313 332 332 292 313 -36 67
8 7 11 127 192 182 169 198 -4 39
8 7 12 404 370 364 332 409 31 39
8 7 13 73 92 81 104 110 132 96
8 7 14 224 231 328 237 237-136 81
8 7 15 106 136 118 124 1

8 12 7 86 103 103 128 78-148 9
8 12 8 374 436 436 397 436 05 73
8 12 9 231 237 237 193 231-112 39
8 12 10 324 401 422 324 367 47 95
8 12 11 110 104 102 127 113 44 78
8 12 12 230 207 201 243 200 46 52
8 12 13 104 94 138 131 121 19 75
8 12 14 112 123 148 113 137-170 85
8 12 15 73 68 87 80 70 47 58
8 12 16 222 203 211 172 204 68 53
8 13 0 44 35 40 33 46 180 27
8 13 1 144 208 181 170 194 173 13
8 13 2 216 237 211 216 217 04 68
8 13 3 68 110 89 58 70 01 60
8 13 4 374 314 330 294 292 02 44
8 13 5 121 136 0 106 124 113 12
8 13 6 104 140 110 80 126-129 82
8 13 7 126 147 207 192 134 26 16
8 13 8 176 114 136 204 138 05 84
8 13 9 343 336 321 284 346 07 32
8 13 10 258 268 238 204 264-101 45
8 13 11 248 232 262 224 246 108 56
8 13 12 227 293 284 296 266 05 89
8 13 13 132 142 164 214 133 24 77
8 13 14 83 81 96 80 133-176 84
8 13 15 112 120 132 120 137-172 85
8 14 0 144 182 57 174 153 0 100
8 14 1 273 218 300 219 218 107 44
8 14 2 198 197 141 170 183 04 71
8 14 3 98 73 88 117 53 111 90
8 14 4 223 256 286 242 244 43 77
8 14 5 341 382 0 339 373-173 46
8 14 6 151 183 178 173 108 142 78
8 14 7 318 386 302 432 343 174 48
8 14 8 137 176 189 134 170 44 89
8 14 9 153 203 186 184 213 115 93
8 14 10 268 202 187 240 243 104 42
8 14 11 88 134 108 114 113 116 84
8 14 12 158 116 111 133 120 05 83
8 14 13 300 212 353 234 246 107 47
8 14 14 63 112 0 74 82 20 30
8 14 15 87 111 104 119 110 02 71
8 15 0 434 434 429 396 402 0 95
8 15 1 144 127 104 172 134 124 77
8 15 2 278 258 261 234 237 08 42
8 15 3 221 194 209 210 193 08 63
8 15 4 223 193 130 131 203-161 84
8 15 5 197 167 207 119 166 17 84
8 15 6 222 227 190 167 228 7 58
8 15 7 184 160 231 208 137 26 83
8 15 8 263 233 182 217 236 104 86
8 15 9 100 76 66 94 70 14 84
8 15 10 248 213 240 234 216 58 77
8 15 11 153 218 133 134 197 08 66
8 15 12 126 106 0 126 103 38 87
8 15 13 74 91 86 111 88 31 82
8 15 14 177 134 130 117 145 7 77
8 16 0 70 64 36 78 82-180 39
8 16 1 209 176 186 162 190 33 36
8 16 2 94 78 31 103 45 03 67
8 16 3 100 176 303 100 156 115 53
8 16 4 78 05 61 90 69 46 54
8 16 5 178 162 109 110 151 79 75
8 16 6 194 253 274 203 236 136 14
8 16 7 131 110 104 133 113 72 17
8 16 8 105 100 86 128 100 137 85
8 16 9 166 138 134 176 147 10 23
8 16 10 49 143 66 142 117-112 91
8 16 11 98 76 93 67 76 28 89
8 16 12 216 200 217 162 200-131 76
8 16 13 94 102 79 110 114 103 58
8 16 14 172 102 116 189 0 78 41
8 17 0 153 138 209 123 185 0 88
8 17 1 140 107 0 144 83 107 83
8 17 2 238 193 231 222 203 05 80
8 17 3 248 231 187 270 231 97 49
8 17 4 503 494 526 382 494 05 82
8 17 5 67 96 68 174 90-102 84
8 17 6 203 190 97 140 170 81 34
8 17 7 213 186 186 194 198 72 83
8 17 8 160 174 136 121 177 18 87
8 17 9 180 159 134 144 157 09 75
8 17 10 85 52 18 123 83 74 85
8 17 11 184 166 194 188 176 49 23
8 17 12 122 133 43 104 164 110 70
8 17 13 138 124 172 122 130-152 73
8 18 0 103 88 53 172 56 0 100
8 18 1 323 414 498 372 380-178 97
8 18 2 107 94 94 172 102-174 11
8 18 3 347 370 293 332 354 03 85
8 18 4 182 186 240 174 182 04 61
8 18 5 427 414 386 382 426 40 71
8 18 6 236 244 327 283 223-133 84
8 18 7 226 253 247 171 251 05 94
8 18 8 200 216 223 248 216 02 42
8 18 9 146 120 221 106 128 07 82
8 18 10 193 174 103 109 192-126 58
8 18 11 213 234 332 239 200-143 73
8 18 12 94 32 0 62 90 46 95
8 19 0 67 31 70 101 43-140 21
8 19 1 231 206 130 200 226 82 82
8 19 2 163 136 148 188 149 159 64
8 19 3 273 266 230 263 257 147 60
8 19 4 171 94 147 104 97 102 97
8 19 5 202 241 197 198 230 22 84
8 19 6 228 234 208 293 257-164 88
8 19 7 94 110 71 131 112 123 93
8 19 8 320 330 343 244 336 01 26
8 19 9 94 78 73 106 71-170 65
8 19 10 84 57 0 73 75 144 49
8 19 11 130 113 112 121 125 0 43
8 20 0 107 66 33 124 32-100 54
8 20 1 102 110 101 86 103 03 43
8 20 2 152 141 204 183 173 178 92
8 20 3 232 250 256 233 242 35 31
8 20 4 98 81 83 85 05 89 72
8 20 5 62 104 191 134 102 03 87
8 20 6 94 120 55 136 148 26 70
8 20 7 94 96 117 133 141 03 84
8 20 8 84 34 47 94 97 40 77
8 20 9 169 104 92 126 174 15 29
8 20 10 98 124 104 80 104 01 81
8 21 0 111 80 43 112 112 180 34
8 21 1 123 63 76 80 73 21 42
8 21 2 136 178 241 148 172 111 75
8 21 3 240 217 233 208 221 142 73
8 21 4 133 86 131 123 92 25 92
8 21 5 83 97 121 94 112 10 64
8 21 6 90 61 135 143 93-174 44
8 21 7 132 107 133 134 123 113 91
8 21 8 83 143 113 64 139-170 46
8 21 9 137 172 180 181 0 03 14
8 22 0 261 230 346 230 230-100 100
8 22 1 143 110 233 133 120 134 94
8 22 2 124 120 48 133 124-135 31
8 22 3 243 224 201 212 244 23 42
8 22 4 133 132 172 170 101-102 37
8 22 5 132 94 114 100 123 00 27
8 22 6 113 144 170 144 133-163 46
8 22 7 173 143 253 178 191-171 92
8 22 8 207 211 196 202 0-137 1
8 23 0 197 111 131 136 123 0 100
8 23 1 204 131 204 144 148-122 40
8 23 2 184 107 93 121 124-103 84
8 23 3 144 68 111 81 85 02 49
8 23 4 132 114 173 133 146 08 54
8 23 5 146 123 146 131 133 17 16
8 23 6 142 133 53 104 0 6 16
8 24 0 100 106 94 121 0-100 9
8 24 1 127 136 118 114 0 132 3

8 24 2 114 20 35 57 0 54 75
8 24 3 151 143 183 194 0 153 31
8 24 4 230 283 280 380 274 00 89
8 24 5 153 84 213 104 87 90 100
8 24 6 112 216 190 154 221 00 100
8 24 7 151 240 194 283 244 00 100
8 24 8 93 94 232 74 117 90 55
8 24 9 67 64 81 53 54 90 31
8 24 10 7 95 87 111 113 40 00 84
8 24 11 384 306 322 315 201 00 100
8 24 12 80 66 92 124 34 00 90
8 24 13 60 132 84 103 120 90 99
8 24 14 62 80 83 170 84 0 66
8 24 15 180 370 353 381 362 00 30
8 24 16 92 60 49 134 84 00 40
8 24 17 32 65 51 72 110 00 90
8 24 18 104 104 0 74 120 00 87
8 24 19 36 126 51 93 00 63
8 24 20 90 137 104 121 122 00 51
8 24 21 126 60 54 65 57 00 98
8 24 22 111 169 147 124 144 137 93
8 24 23 101 120 162 121 143-172 93
8 24 24 370 310 304 294 284 70 95
8 24 25 217 231 171 131 211 00 75
8 24 26 216 205 226 197 194-160 86
8 24 27 165 187 143 162 170 2 54
8 24 28 103 70 38 80 44 170 90
8 24 29 175 136 199 178 56 71
8 24 30 171 214 248 221 225-104 81
8 24 31 102 106 163 134 119 136 88
8 24 32 111 396 424 409 342 437-111 91
8 24 33 130 119 204 181 141 178 84
8 24 34 352 334 337 343 342 134 4
8 24 35 168 123 133 131 137 57 54
8 24 36 73 93 0 85 80-141 60
8 24 37 143 134 163 122 134 32 73
8 24 38 121 184 179 173 175 51 65
8 24 39 234 153 137 184 162-180 100
8 24 40 92 93 97 74 111-168 91
8 24 41 119 161 93 131 134 21 73
8 24 42 220 204 274 181 241 124 53
8 24 43 286 310 254 261 304-131 44
8 24 44 144 230 101 116-166 81
8 24 45 119 146 37 80 90 50 46
8 24 46 153 131 130 192 150-107 78
8 24 47 267 346 303 333 333 152 95
8 24 48 244 234 286 282 250 100 60
8 24 49 190 211 258 253 214 20 8
8 24 50 211 219 185 209 172 184 52
8 24 51 291 255 315 283 275 147 39
8 24 52 123 79 88 184 42 03 85
8 24 53 210 158 238 180 177 171 90
8 24 54 62 81 0 73 81-167 60
8 24 55 261 233 200 162 257-144 55
8 24 56 140 154 188 161 133 72 31
8 24 57 295 344 422 363 363 0 100
8 24 58 110 131 97 161 124 30 70
8 24 59 264 243 321 282 273 155 64
8 24 60 226 252 170 196 227-178 35
8 24 61 227 234 194 239 212 152 69
8 24 62 94 118 88 104 94 144 88
8 24 63 78 90 113 103 88-110 40
8 24 64 207 227 209 201 233-177 88
8 24 65 140 124 157 148 147 105 9
8 24 66 107 08 0 83 54 40 85
8 24 67 164 183 183 189 178 122 76
8 24 68 80 98 93 110 83 23 72
8 24 69 130 133 142 141 154 155 61
8 24 70 360 557 527 563 533 27
8 24 71 98 74 80 80 75-133 37
8 24 72 162 145 0 100 169 127 74
8 24 73 148 100 84 113 115 132 92
8 24 74 140 144 152 153 172 147 86
8 24 75 414 433 396 414 393-180 11
8 24 76 230 260 375 337 304 11 59
8 24 77 176 111 166 161 111 94
8 24 78 180 182 256 201 201 49 32
8 24 79 272 363 282 321 353 77 98
8 24 80 114 114 174 107 82-154 78
8 24 81 171 186 165 150 183-177 52
8 24 82 147 162 160 101 134-171 67
8 24 83 36 70 84 117 74 25 73
8 24 84 216 268 278 294 264-105 83
8 24 85 281 274 268 292 298-124 95
8 24 86 205 266 198 113 256 75 58
8 24 87 287 294 267 283 299-150 75
8 24 88 90 73 50 124 74 25 16
8 24 89 115 124 117 69 150 125 82
8 24 90 67 83 0 70 156 67
8 24 91 224 170 192 150 196 04 91
8 24 92 112 97 117 147 0-107 3
8 24 93 301 301 323 293 338 0 99
8 24 94 420 483 436 513 483 17 98
8 24 95 84 91 83 88 72 26
8 24 96 384 428 395 404 421 05 84
8 24 97 161 166 148 154 190 179 79
8 24 98 340 322 453 484 504 27 92
8 24 99 132 140 100 134 122 03 71
8 24 100 205 138 0 114 108 0 100
8 24 101 200 211 227 232 207 178 73
8 24 102 275 291 287 302 271 67 48
8 24 103 290 313 317 304 305-162 46
8 24 104 198 114 152 90 110-142 98
8 24 105 113 94 81 82 101 73 53
8 24 106 187 232 110 191 242 21 97
8 24 107 142 105 0 111 111-117 61
8 24 108 216 246 167 210 255 11 83
8 24 109 268 293 267 252 276-161 76
8 24 110 190 250 221 197 0 82 27
8 24 111 212 181 201 180 187 0 80
8 24 112 420 358 351 309 357-151 98
8 24 113 325 267 312 286 234 48 98
8 24 114 170 137 34 140 100-101 96
8 24 115 80 92 47 120 108 26 84
8 24 116 133 224 223 235 248 157 94
8 24 117 240 176 194 215 33 91
8 24 118 375 410 0 412 407-164 94
8 24 119 320 414 390 370 398 31 96
8 24 120 130 141 95 118 122 03 60
8 24 121 170 247 218 244 234 75 94
8 24 122 63 98 90 20 102 14 73
8 24 123 350 359 330 301 354-147 14
8 24 124 220 265 233 231 220 47 94
8 24 125 138 77 142 84 94 00 90
8 24 126 191 204 218 201 204 57 43
8 24 127 134 124 124 124 0 72 4
8 24 128 87 74 112 64 0 105 2
8 24 129 313 501 470 517 478-180 48
8 24 130 160 181 213 221 172 37 37
8 24 131 236 236 224 260 210-124 71
8 24 132 354 535 535 511 502 162 62
8 24 133 163 155 207 182 160 98 55
8 24 134 477 468 338 394 390 113 95
8 24 135 293 266 325 251 288 57 64
8 24 136 226 247 0 282 232 2 4
8 24 137 50 67 84 74 76-174 81
8 24 138 67 120 70 67 113 84 85
8 24 139 111 178 215 200 168-114 94
8 24 140 150 107 0 158 194 137 88
8 24 141 140 180 182 167 168 77 80
8 24 142 273 259 244 213 265-157 54
8 24 143 140 167 174 197 172 33 35
8 24 144 114 125 114 167 144-121 20
8 24 145 161 161 0 134 145 123 68
8 24 146 176 206 155 134 162 0 70
8 24 147 251 251 265 271 210-113 71
8 24 148 193 234 217 201 214 74 61
8 24 149 114 178 117 132 05 68
8 24 150 232 172 245 207-124 81
8 24 151 217 154 240 233-147 76

8 24 152 240 299 299 308 286 059 14
8 24 153 49 94 0 123 76 132 50
8 24 154 83 109 133 73 95 09 82
8 24 155 31 74 104 73 90 115 71
8 24 156 366 373 307 363 359 146 48
8 24 157 271 235 289 188 248 127 79
8 24 158 74 137 118 117 121 82 36
8 24 159 192 120 114 146 163-143 78
8 24 160 228 181 182 227 230 23 31
8 24 161 182 181 206 163 212 77 75
8 24 162 153 176 191 167 192 05 88
8 24 163 370 338 342 333 301 0 100
8 24 164 216 277 177 224 222 148 87
8 24 165 213 194 178 202 198 4 24
8 24 166 246 222 254 224 214 17 84
8 24 167 173 107 106 54 84 083 53
8 24 168 75 89 162 110 90 1 91
8 24 169 46 90 71 84 47 68 80
8 24 170 378 400 0 307 372 21 44
8 24 171 123 108 178 141 113 137 91
8 24 172 229 234 226 209 227 075 30
8 24 173 125 101 115 110 107-113 70
8 24 174 176 210 240 203 182 119 78
8 24 175 225 234 243 192 228 172 54
8 24 176 235 193 202 129 197 44 91
8 24 177 14 70 71 66 49 102 89 92
8 24 178 33 86 95 74 108-121 87
8 24 179 144 164 194 107 0 143 42
8 24 180 41 74 77 74 28-180 99
8 24 181 106 142 87 123 128-134 87
8 24 182 208 137 154 108 121 109 99
8 24 183 242 281 291 239 284-122 80
8 24 184 346 278 327 329 248 91 99
8 24 185 350 331 184 269 290 127 98
8 24 186 206 182 156 165 143 55 83
8 24 187 7 89 46 110 108 81 60 39
8 24 188 273 277 220 262 266 47 12
8 24 189 176 250 257 218 235 118 93
8 24 190 143 193 129 149 191 103 95
8 24 191 239 208 173 212 244 109 47
8 24 192 183 183 175 169 199 46 63
8 24 193 156 141 101 122 148 82 81
8 24 194 217 203 178 232 217 14 51
8 24 195 83 77 119 103 102 47 89
8 24 196 235 225 282 227 234-180 29
8 24 197 274 278 253 246 265 7 28
8 24 198 248 286 238 265 286-106 90
8 24 199 149 131 143 148 123-107 76
8 24 200 133 174 177 193 171 02 93
8 24 201 247 270 336 285 260 33 74
8 24 202 6 69 96 90 102 93-112 83
8 24 203 269 294 268 256 281 33 50
8 24 204 85 94 131 141 105 22 75
8 24 205 146 115 156 149-114 90
8 24 206 267 237 242 282 235 167 32
8 24 207 94 160 190 177 151 06 93
8 24 208 181 201 133 177 190 98 86
8 24 209 283 243 293 270 266 101 71
8 24 210 182 213 213 235 219-147 75
8 24 211 94 129 115

9 1A 6 232 213 247 141 210 3 0 0
9 1A 7 121 42 197 40 73-140 58
9 1A 8 66 114 07 41 123 103 86
9 1A 9 303 284 326 237 273 48 64
9 1A 10 203 241 213 240 222 111 87
9 1A 11 102 113 134 152 0-100 21
9 1A 0 43 80 67 62 47 0 40
9 1A 1 234 246 224 290 205-156 43
9 1A 2 176 188 203 181 187-133 63
9 1A 3 54 67 130 43 44 78 88
9 1A 4 297 292 320 310 275 -48 74
9 1A 5 123 154 19 42 100 31 94
9 1A 6 208 180 178 200 191 101 51
9 1A 7 128 74 51 39 77-159 46
9 1A 8 104 76 160 73 98 151 86
9 1A 9 111 90 133 102 131-113 76
9 2A 0 38 74 103 34 53 0 48
9 2A 1 70 62 72 96 84 356 88
9 2A 2 134 131 100 112 132-131 31
9 2A 3 247 248 142 284 247 -21 51
9 2A 4 223 223 170 233 224 171 71
9 2A 5 223 147 146 203 193 113 79
9 2A 6 103 34 0 101 106 75 93
9 2A 7 140 207 143 167 140 12 52
9 2A 8 126 134 109 164 1 42 20
9 2A 9 100 64 123 33 71 0 16
9 2A 1 184 147 214 190 116 159 74
9 2A 3 173 124 42 84 152 -45 63
9 2A 4 139 44 108 43 108 174 24
9 2A 5 107 47 137 84 103-163 64
9 2A 6 117 110 0 147 140 -36 41
9 2A 7 272 264 293 283 0-123 33
9 2A 8 104 44 48 121 114-140 23
9 2A 9 33 67 61 70 46 -42 76
9 2A 1 159 174 143 113 101 143 50
9 2A 2 3 89 110 126 132 75 -73 68
9 2A 3 162 198 152 193 0 44 3
9 2A 4 171 143 221 164 0-110 37
9 2A 5 74 96 237 84 0 0 59
9 2A 6 102 46 103 124 0-106 3
10 0 0 249 158 213 243 119 0 100
10 0 1 711 627 633 607 579 0 100
10 0 2 464 547 441 622 523-140 51
10 0 3 690 667 636 671 612 0 24
10 0 4 77 43 78 43 93 0 43
10 0 5 60 44 0 42 50-160 83
10 0 6 248 276 233 173 224-140 66
10 0 7 66 40 43 62 66-100 73
10 0 8 330 333 242 430 330 0 46
10 0 9 34 73 112 73 113 0 44
10 0 10 62 43 38 107 87-140 27
10 0 11 73 74 61 74 0 23
10 0 12 67 77 43 176 43 140 86
10 0 13 304 230 227 204 276 0 49
10 0 14 83 73 141 71 84 0 82
10 0 15 83 43 0 0 71 0 38
10 0 16 42 33 63 77 0 9
10 1 0 137 31 126 61 78-100 100
10 1 1 370 410 382 388 393-111 68
10 1 2 206 192 167 220 205 -43 63
10 1 3 304 298 412 330 292 152 73
10 1 4 496 321 443 309 456 62 42
10 1 5 303 311 0 289 292-158 27
10 1 6 173 241 192 221 243-106 80
10 1 7 132 107 100 128 114 -16 28
10 1 8 270 336 334 347 361 131 97
10 1 9 248 223 213 166 210 -74 68
10 1 10 39 76 44 82 66 73 67
10 1 11 117 120 127 100 134 31 64
10 1 12 341 333 320 392 302 37 76
10 1 13 131 194 78 206 171-178 64
10 1 14 214 172 207 196 177 136 40
10 1 16 224 194 224 198 0-100 6
10 2 0 223 223 280 134 241 0 49
10 2 1 304 337 284 331 302 -44 40
10 2 2 287 237 236 188 197 49 96
10 2 3 233 233 138 246 220 -99 92
10 2 4 243 263 220 214 233 -48 13
10 2 5 201 203 0 283 221-104 98
10 2 6 332 364 418 429 332-158 82
10 2 7 194 207 170 163 167 158 31
10 2 8 76 74 92 133 78 41 19
10 2 9 263 213 277 196 223 43 46
10 2 10 182 192 213 72 187 49 54
10 2 11 287 227 271 206 243 48 46
10 2 12 140 162 163 248 150 179 44
10 2 13 449 421 413 437 460 162 66
10 2 14 41 48 100 43 81 67 12
10 2 15 111 83 0 113 110 48 14
10 2 16 112 42 121 64 0 -44 7
10 3 0 212 246 221 227 303 0 100
10 3 1 378 338 384 376 340 -76 60
10 3 2 244 283 237 293 242 -48 50
10 3 3 310 316 376 368 490 10 48
10 3 4 99 113 94 140 0 34 22
10 3 5 376 293 0 233 283 70 40
10 3 6 70 160 147 143 149-158 93
10 3 7 204 192 288 220 214-147 80
10 3 8 446 502 412 493 501-140 46
10 3 9 124 65 0 83 68 145 67
10 3 10 100 104 150 133 176 -97 46
10 3 11 98 63 64 93 88 -25 66
10 3 12 248 268 233 228 233 10 30
10 3 13 16 33 97 122 76 67 86
10 3 14 91 92 130 71 102 -44 30
10 3 15 70 62 0 81 62 152 16
10 3 16 109 119 122 67 0 -23 14
10 4 0 262 254 164 293 180-140 97
10 4 1 124 83 144 83 100-133 71
10 4 2 337 331 320 393 328-124 63
10 4 3 148 191 104 190 143 -70 76
10 4 4 367 396 312 343 347 6 52
10 4 5 262 199 0 190 181 -25 48
10 4 6 141 176 168 162 185-172 41
10 4 7 88 124 64 138 99 166 67
10 4 8 132 114 120 180 125 70 66
10 4 9 104 143 145 122 156 127 26
10 4 10 167 224 243 197 204 -19 74
10 4 11 197 251 133 192 233 -18 93
10 4 12 140 126 67 177 153 -23 91
10 4 13 111 124 0 72 114 -10 77
10 4 14 160 157 130 139 140 113 68
10 4 15 67 102 0 108 107 116 69
10 4 16 142 104 218 184 0-117 29
10 5 0 171 214 0 183 213-140 46
10 5 1 147 174 0 190 172 -3 75
10 5 2 141 100 0 143 167 43 44
10 5 3 243 312 348 314 291 106 43
10 5 4 131 231 154 239 144 73 96
10 5 5 137 203 243 162 141 133 44
10 5 6 90 63 0 81 66 110 57
10 5 7 180 220 141 216 231-134 49
10 5 8 237 336 203 300 333 -75 47
10 5 9 168 167 103 134 111 -13 59
10 5 10 316 343 444 438 323 -73 64
10 5 11 94 126 137 143 113 -53 54
10 5 12 63 46 42 42 42 42 42
10 5 13 249 213 127 213 229-124 40
10 5 14 177 138 223 196 106 172 73
10 5 15 103 72 134 37 62 -65 42
10 6 0 341 374 0 343 333-140 1
10 6 1 201 233 0 277 271 -56 66
10 6 2 183 170 0 132 171 -25 67
10 6 3 302 340 214 324 206 49 63
10 6 4 64 73 100 48 103 29 44
10 6 5 323 300 292 341 323-127 53
10 6 6 203 194 183 200 194-130 67
10 6 7 147 174 143 163 197-150 84
10 6 8 238 312 301 324 313-155 73
10 6 9 147 233 212 236 231 134 43

11	3	77	94	74	91	94	14	25	12	2	6	122	122	102	218	124	131	25	13	0	1	253	244	212	284	247	990	32
11	3	6	237	166	173	181	191	160	12	2	7	131	183	181	131	184	41	72	13	0	2	199	193	344	226	188	90	100
11	3	7	217	232	221	212	243	129	12	2	8	136	114	123	173	144	456	6	13	0	3	265	262	267	220	233	90	79
11	3	8	197	164	193	193	164	109	12	2	9	221	321	220	278	294	107	9	13	0	4	134	174	164	136	127	90	24
11	3	9	120	164	103	167	137	123	12	2	10	296	207	237	317	276	446	48	13	0	5	373	303	299	293	352	990	72
11	3	10	113	139	124	152	152	130	12	2	11	103	133	87	123	127	111	92	13	0	6	62	71	30	86	14	90	15
11	3	11	112	104	143	142	135	23	12	2	12	77	33	70	47	61	160	45	13	0	7	166	181	186	257	173	990	28
11	3	12	215	162	162	154	177	0	12	3	0	108	74	158	63	62	180	100	13	0	8	66	67	124	81	77	90	91
11	3	13	233	194	163	221	0	90	12	3	1	454	452	404	447	444	169	25	13	0	9	82	101	139	81	118	90	90
11	3	0	137	201	227	224	167	140	12	3	2	188	236	211	228	237	41	93	13	1	0	280	280	296	286	271	0	8
11	3	1	207	232	102	216	243	170	12	3	3	246	273	266	264	263	93	52	13	1	1	161	226	214	257	210	128	90
11	3	2	141	166	102	119	166	138	12	3	4	125	103	67	97	123	38	58	13	1	2	141	173	169	186	167	121	50
11	3	3	111	80	96	103	70	74	12	3	5	170	212	210	183	194	172	10	13	1	3	234	260	270	266	253	94	93
11	3	4	64	33	64	81	76	32	12	3	6	121	50	109	84	77	0	99	13	1	4	273	248	203	209	246	90	64
11	3	5	71	73	76	101	90	81	12	3	7	94	80	20	72	114	138	41	13	1	5	127	133	187	167	131	143	71
11	3	6	130	183	72	163	176	61	12	3	8	117	114	80	122	131	149	87	13	1	6	181	196	133	188	183	554	3
11	3	7	106	131	98	137	172	152	12	3	9	210	217	222	204	223	93	56	13	1	7	136	73	0	86	107	33	31
11	3	8	93	136	84	127	126	13	12	3	10	106	61	98	117	113	76	11	13	1	8	171	166	250	212	169	180	77
11	3	9	57	73	63	113	77	45	12	3	11	77	81	108	81	32	144	35	13	1	9	31	87	61	87	89	339	41
11	3	10	113	62	0	67	68	7	12	3	12	43	91	119	87	81	33	76	13	1	10	120	114	92	83	0	89	29
11	3	11	81	64	113	37	94	114	12	4	0	83	50	41	33	81	0	100	13	2	0	73	172	129	148	158	0	100
11	3	12	167	127	142	107	130	84	12	4	1	160	139	132	113	167	131	61	13	2	1	149	123	120	147	153	123	19
11	3	13	111	100	43	122	0	138	12	4	2	207	200	190	209	217	114	62	13	2	2	237	269	271	263	249	0	85
11	10	0	493	400	374	373	480	0	12	4	3	233	246	144	219	241	163	75	13	2	3	223	200	100	214	213	85	80
11	10	1	199	219	240	280	201	154	12	4	4	353	326	338	307	343	61	34	13	2	4	346	299	344	311	303	2	81
11	10	2	303	304	233	293	316	8	12	4	5	223	208	166	203	222	47	83	13	2	5	196	197	188	181	187	126	33
11	10	3	113	103	44	103	97	112	12	4	6	107	116	78	73	111	79	85	13	2	6	109	138	169	138	133	139	94
11	10	4	118	193	192	139	177	94	12	4	7	243	271	214	239	267	94	83	13	2	7	230	251	234	267	222	108	37
11	10	5	157	180	161	173	173	143	12	4	8	94	127	130	133	127	169	60	13	2	8	150	91	93	106	123	6	79
11	10	6	120	110	113	86	86	16	12	4	9	203	234	238	234	241	158	76	13	2	9	198	182	152	148	168	136	13
11	10	7	200	232	219	174	232	14	12	4	11	111	133	121	124	113	101	40	13	2	10	262	257	253	249	0	113	3
11	10	8	78	110	179	142	102	157	12	4	12	76	93	131	87	0	6	16	13	3	0	93	70	137	73	72	180	23
11	10	9	132	183	103	211	171	60	12	5	0	178	190	100	177	199	180	76	13	3	1	96	130	0	98	146	164	91
11	10	10	58	96	100	61	91	137	12	5	1	173	103	149	101	138	162	80	13	3	2	218	192	103	151	213	91	83
11	10	11	53	113	110	110	104	21	12	5	2	301	266	224	264	281	128	58	13	3	3	338	419	343	416	391	96	98
11	10	12	103	93	0	91	109	63	12	5	3	181	157	163	133	163	170	62	13	3	4	109	93	72	110	108	98	70
11	11	0	613	671	592	648	641	180	12	5	4	191	173	143	152	181	170	66	13	3	5	344	240	231	269	283	50	97
11	11	1	173	136	133	239	153	82	12	5	5	336	333	338	302	323	44	71	13	3	6	177	151	121	150	167	106	18
11	11	2	323	333	289	266	330	68	12	5	6	69	63	78	98	94	101	87	13	3	7	160	171	132	139	164	98	27
11	11	3	217	119	137	130	140	83	12	5	7	118	67	124	108	104	430	30	13	3	8	166	176	224	187	197	141	78
11	11	4	143	134	83	162	143	157	12	5	8	229	215	159	242	243	114	63	13	3	9	73	69	123	101	83	141	79
11	11	5	341	330	310	348	336	121	12	5	9	146	120	23	136	163	153	89	13	3	10	173	191	230	176	0	92	53
11	11	6	417	421	394	411	422	10	12	5	10	186	151	139	127	179	4	65	13	4	0	192	148	107	168	164	180	33
11	11	7	108	103	0	96	82	100	12	5	11	66	63	0	78	102	4	87	13	4	1	109	114	63	103	102	104	48
11	11	8	143	134	202	174	166	78	12	5	12	104	108	88	115	0	4	6	13	4	2	148	142	172	123	130	103	12
11	11	9	218	133	136	132	182	106	12	6	0	98	93	83	137	60	180	100	13	4	3	93	89	131	92	12	138	57
11	11	10	49	92	0	40	86	173	12	6	1	462	473	466	488	480	31	80	13	4	4	373	364	304	343	352	73	34
11	11	11	231	236	191	267	263	136	12	6	2	240	254	202	263	233	20	37	13	4	5	107	108	92	40	91	8	64
11	12	0	140	290	304	254	226	0	12	6	3	74	69	139	67	63	123	74	13	4	6	128	142	99	156	153	137	74
11	12	1	332	344	432	421	338	75	12	6	4	273	291	318	289	278	147	64	13	4	7	104	102	117	90	112	64	64
11	12	2	134	208	210	203	193	21	12	6	5	117	171	72	164	158	124	94	13	4	8	240	233	168	216	237	80	76
11	12	3	201	200	193	128	194	46	12	6	6	143	166	160	137	134	88	61	13	4	9	181	159	167	141	176	173	80
11	12	4	53	68	29	101	81	96	12	6	7	210	231	200	207	216	71	63	13	4	10	72	73	12	91	0	67	8
11	12	5	247	239	334	246	243	93	12	6	8	204	227	238	213	216	142	17	13	5	0	208	220	215	250	223	0	57
11	12	6	214	193	210	183	174	2	12	6	9	114	181	113	163	171	93	93	13	5	1	147	137	134	164	141	157	70
11	12	7	100	84	64	86	98	13	12	6	10	156	190	116	208	168	80	60	13	5	2	116	112	117	98	110	64	63
11	12	8	76	86	121	107	77	174	12	6	11	107	134	110	144	149	8	93	13	5	3	164	150	44	136	162	133	73
11	12	9	94	79	100	31	80	130	12	6	12	84	157	143	143	0												

14	3	5	64	98	112	111	82	-68	26
14	3	6	135	96	109	108	0	-170	59
14	4	0	59	70	0	90	62	0	40
14	4	1	91	76	71	93	100	82	48
14	4	2	135	145	156	148	153	45	12
14	4	3	122	130	104	110	121	96	16
14	4	4	60	72	128	72	83	156	50
14	4	5	146	96	53	74	0	-138	35
14	4	6	224	227	194	202	0	-168	32
14	5	0	201	170	150	148	179	0	74
14	5	1	73	56	29	93	94	2	17
14	5	2	83	77	116	61	89	103	59
14	5	3	91	63	81	60	80	170	65
14	5	4	157	156	167	170	156	-32	13
14	6	0	113	106	115	90	110	-180	16
14	6	1	94	121	148	116	102	-59	72
14	6	2	113	147	107	140	132	-34	40
14	6	3	84	82	57	95	96	-55	83
14	6	4	134	156	151	146	0	38	12
14	7	0	99	73	63	105	86	0	97
14	7	1	102	79	104	110	87	103	9
14	7	2	153	64	99	70	95	33	95
14	7	3	194	161	161	155	0	-26	21
14	8	0	97	81	151	111	0	0	83
14	8	1	91	104	101	111	0	52	8
14	8	2	133	146	134	152	0	94	5
14	8	3	90	117	98	74	0	-160	7
14	9	0	73	109	36	84	0	0	89
14	9	1	108	97	109	106	0	70	3
11	10	6	120	110	115	66	86	16	87

Bibliography

- Adams, M.J. (1968) Oxford D.Phil. thesis.
- Adams, M.J., Blundel, T.L., Dodson, E.J., Dodson, G.G., Vijayan, M.,
Baker, E.N., Harding, M.M., Hodgkin, D.C., Rimmer, B., and Sheat, S.
(1969) *Nature* 224, 491-495
- Alvsaker, J.O., Haugli, F.B. and Laland, S.G. (1967) *Biochem.J.* 102, 362-366
- Arnone, A., Bier, C.J., Cotton, F.A., Day, V.W., Hazen, E.E.,
Richardson, D.C., Richardson, J.S. and Yonath, A. (1971)
J. Biol. Chem. 246, 2302
- Banyard, S.H. (1973) Oxford D.Phil. thesis.
- Bijvoet, J.M. (1954) *Nature* 173, 888
- Blake, C.C.F. (1968) *Adv. Protein Chem.* 23, 59
- Blake, C.C.F. (1972) *Progress in Biophysics and Molecular biology*
25, 83-130
- Blake, C.C.F., Fenn, R.H., North, A.C.T., Phillips, D.C. and Poljak, R.J.
(1962) *Nature* 196, 1173
- Blake, C.C.F., Swan, I.D.A., Rérat, C., Berthou, J., Laurent, A.
and Rérat, B. (1971) *J. Mol. Biol.* 61, 217-224
- Bloomer, A.C. (1972) Oxford D.Phil. thesis.
- Blow, D.M. and Crick, F.H. (1959) *Acta Cryst.* 12, 794
- Blow, D.M. and Matthews, B.W. (1973) *Acta Cryst.* 29, 56
- Branch, W.T. and Edelhoch, H. (1972) *Arch. Biochem. Biophys.* 152, 144-151
- Branch, W.T., Robbins, J. and Edelhoch, H. (1971) *J. Biol. Chem.* 246
6011-6018
- Buerger, M.J. (1962) "X-ray Crystallography" John Wiley and Sons, Inc. N.Y.
- Camerman, N. and Camerman, A. (1971) *Science* 175, 764-766
- Christensen, L.K. (1959) *Nature* 183, 1189
- Chothia, C. (1973) *J. Mol. Biol.* 75, 295-302
- Crick, F.H.C. and Magdoff, B.S. (1956) *Acta Cryst.* 2, 901
- Cruickshank, D.W.J. (1956) *Acta Cryst.* 2, 747

- Cullis, A.F., Muirhead, H., Perutz, M.F. and Rossman, M.G. (1961)
 Proc. Roy. Soc. 265, 15-33
- Davis, P.J., Handwerger, B.S. and Gregerman, R.I. (1972) J. Clin. Invest.
51, 515-521
- Dickerson, R.E., Kendrew, J.C. and Strandberg, B.E. (1961) Acta Cryst.
14, 1188
- Dickerson, R.E., Kopka, M.L., Varnum, J.C. and Weinzierl, J.E. (1967)
 Acta Cryst. 23, 511
- Fridborg, K., Kannan, K.K., Liljas, A., Lundin, J., Strandberg, B.,
 Strandberg, R., Tilander, B. and Wirén, G. (1967) J. Mol. Biol.
25, 505
- Gonzalez, G. (1972) Oxford D.Phil. thesis.
- Gonzalez, G. and Offord, R.E. (1971) Biochem. J. 125, 309-317
- Goodman, D.S. and Leslie, R.B. (1971) J. Biol. Chem. 246, 7544
- Gotto, A.M., Lux, S.E. and Goodman, D.S. (1972) Biochim. Biophys. Acta
241, 429-435
- Green, A.M., Marshall, J.S., Pensky, J. and Stanbury, J.B. (1972)
 Science 175, 1378- 1380
- Hamilton, W.C., Rollett, J.S. and Sparks, R.A. (1965) Acta Cryst. 18, 129-130
- Hardman, K.D. and Ainsworth, C.F. (1972) Biochem. 11, 4910-4919
- Harker, D. (1953) Acta Cryst. 6, 731
- Hart, R.G. (appendix to) (1961) Acta Cryst. 14, 1195
- Haupt, H. and Heide, K. (1966) Experientia 22, 449
- Hedges, T. (1972) Oxford Part II B.Sc. thesis
- Hershman, J.M. (1963) Endocrinology 72, 799
- Ingbar, S.M. (1963) J. Clin. Invest. 42, 143
- International Tables for X-ray crystallography, vols. 1, 2, 3.
 Birmingham: The Kynoch Press.
- Jaarsveld, P.P.Van., Branch, W.T., Edelhoch, H. and Robbins, J. (1973b)
 J. Biol. Chem. 248, 4706-4712

- Jaarsveld, P.P.Van., Edelhoch, H., Goodman, D.S. and Robbins, J. (1973a)
J. Biol. Chem. 248, 4693-4705
- Kanai, M., Raz, A. and Goodman, D.S. (1968) J. Clin. Invest. 47, 2025-2044
- Kartha, G. (1965) Acta Cryst. 19, 883
- Kirby, W., White, G. and Glover, J. (1971) Biochem. J. 123, 31
- Kraut, J. (1968) J. Mol. Biol. 35, 511-513
- Kraut, J., Sieker, L.C., High, D.F. and Freer, S.T. (1962) Proc. Nat.
Acad. Sci. U.S.A. 48, 1417
- Larsen, P.R. (1972) J. Clin. Invest. 51, 1125-1134
- Liljas, A., Kannan, K.K., Bergsten, P.C., Waara, I., Fridborg, K.,
Strandberg, B., Carlbom, U., Järup, L., Lövgren, S. and Petef, M.
(1972) Nature New Biol. 235, 131-136
- Lundberg, B. (1965) Acta Cryst. 18, 576
- Luzzati, V. (1953) Acta Cryst. 6, 142
- Luzzati, V. (1955) Acta Cryst. 8, 795
- Matthews, B.W. (1966a) Acta Cryst. 20, 82-86
- Matthews, B.W. (1966b) Acta Cryst. 20, 230-239
- Morgan, F.J., Canfield, R.E. and Goodman, D.S. (1971) Biochim. Biophys. Acta
236, 793-801
- Nilsson, S.F. and Peterson, P.A. (1971) J. Biol. Chem. 246, 6093-6105
- North, A.C.T. (1965) Acta Cryst. 18, 212
- North, A.C.T. and Phillips, D.C. (1960) Progress in Biophysics and
Molecular biology 19, 1
- North, A.C.T., Phillips, D.C. and Matthews, F.C. (1968) Acta Cryst.
24, 351-359
- Osorio, C. (1962) J. Physiol. (London) 163, 151
- Pages, R.A., Robbins, J. and Edelhoch, H. (1973) Biochem. 12, 2773-2779
- Patterson, A.L. (1934) Phys. Rev. 46, 372
- Perutz, M.F. (1970) Nature 228, 734

- Peterson, P.A. (1971) J. Biol. Chem. 246, 44
- Peterson, P.A. (1972) Eur. J. Clin. Invest. in press
- Petsko, G.A.P. (1973) Oxford D.Phil. thesis
- Phillips, D.C. (1964) J. Sci. Instr. 41, 123-129
- Phillips, D.C. (1966) Adv. in Research by Diffraction Methods 2, 75
- Rask, L., Peterson, P.A. and Björk, I. (1972) Biochem. 11, 264
- Rask, L., Peterson, P.A. and Nilsson, S.F. (1971) J. Biol. Chem. 246, 6087-6097
- Rask, L., Vahlquist, A. and Peterson, P.A. (1971) J. Biol. Chem. 246, 6638
- Raz, A. and Goodman, D.S. (1969) J. Biol. Chem. 244, 3230-3237
- Raz, A., Shiratori, T. and Goodman, D.S. (1970) J. Biol. Chem. 245, 1903
- Rérat, C. and Schwick, H.G. (1967) Acta Cryst. 22, 441
- Ross, J.E. and Tapley, D.F. (1966) Endocrinology 79, 493-504
- Rossman, M.G. (1960) Acta Cryst. 13, 221
- Schiffer, H., Girling, R.L. Ely, K.R. and Edmundson, A.B. (1973)
Reprint ANL/BIM-73-1. Argonne National Laboratory, U.S.A.
- Schultze, H.E., Schönenburger, M. and Schwick, G. (1956) Biochem. Z. 323, 267
- Sigler, P.B., Blow, D.M., Matthews, B.W. and Henderson, R. (1968)
J. Mol. Biol. 35, 143
- Swan, I.D.A. (1970) Oxford D.Phil. thesis
- Tsernoglou, D., Hill, E. and Banaszak, L.J. (1972) J. Mol. Biol. 69, 75-87
- Vand, V. (1957) Acta Cryst. 10, 303
- Watson, H.C., Shotten, D.M., Cox, J.M. and Muirhead, H. (1970)
Nature 225, 806-815
- Williams, R.J.P. and Petsko, G.A.P. to be published.
- Wilson, A.J.C. (1942) Nature 150, 152
- Woebbar, K.A. and Ingbar, S.M. (1964) J. Clin. Invest. 43, 931-942