Structural studies on two lipocalins

João Henrique Resende de Oliveira de Morais Cabral

Submitted in satisfaction of the requirements
for the degree of PhD in the
University of Edinburgh
Year of presentation 1993

I thank first and foremost all of my friends who managed to endure my moods, dissuaded me from acts of irreparable violence and encouraged me through these hard fives years, Mary, Elspeth, LLuis, Jean-Michel, Myrtle, Alan, Nia, Carol, Thales, Paul T., Paul A., Fred, Dima, Peter and Stella.

I thank Dr. Lindsay Sawyer, my supervisor, for his support and encouragement.

I thank too, the essential scientific support and teachings from Dr. Jean-Michel Sallenave, Dr. Steve Johnstone, Dr. Mary Turner and Dr. Dimitri Alexeev.

I thank Dr. Margarida Damas for her friendship.

I acknowledge the financial support provided by the EEC and SERC.

Last but not least, I thank my parents, my sister, my grandmother and my extended family (the thousands of cousins, uncles, aunts and all my donkeys) for their love.

Declaration

I hereby declare that this thesis was composed by me, that the work of which this is a record was done by me, except where stated in the thesis. This work has not been accepted elsewhere in any previous application for a degree. All of the sources have been acknowlegde.

Abstract

The number of proteins being included in the lipocalin structural family has been steadily increasing since the solution of the structures of beta-lactoglobulin and retinol-binding protein. The family now comprises some 23 proteins and these come from many different sources and display a variety of characteristics. The work described here focused on two of those proteins, beta-lactoglobulin (Blg) and apolipoprotein D (apo D).

The structure of bovine beta-lactoglobulin (Monaco, *et al.*, 1987)from the trigonal crystals grown at pH 7.8, space group $P3_221$, presents several aspects that are in disagreement with previous work, in particular the positioning of the bound retinol in the surface of the protein and the observation that the protein was in its monomeric form. We have checked these findings independently by extending the resolution of a 6 Å x-ray structure (Green, *et al.*, 1979). A 3.0 Å model is now available though still in the refinement stages. Not surprisingly it was found that the protein is indeed in the dimeric form and that the arrangement of the dimer is very similar to the one found in BlgY (orthorhombic crystal form)(Papiz, *et al.*, 1986), the other independently determined structure. Changes were observed in the threading of the sequence, in particular between residues 75 and 32 where movements of as much as five residues are found, and in the C-terminus between 141 and 150 where a shift of two residues is observed along a β-strand. These changes result in an overall increase of the hydrophobicity of the pocket.
In parallel, cocrystallization of Blg with a variety of ligands has been successful in at least one case where density is evident in the binding cavity.

Human apolipoprotein D is present in the fluid of cysts formed during gross-cystic-disease, which is the most common breast disease in premenopausal women, and is a potential biological marker for breast cancer. The protein is found as well, associated with the lipoproteic system in the blood and is produced in high amounts in regenerating rat peripherial nerves. Several structural aspects of the protein from breast cysts were investigated; four of the cysteines were shown to be forming disulphide bridges and the fifth is not present as a free-cysteine. The influence of the pH on the conformational transitions (investigated by CD) and monomer-multimer balance (investigated by gel filtration) shows occurring parallelism with changes on Blg and in particular the transition seen between pH 6.5 and 5.5 could be due to the onset of the dissociation of the tetramer that populates the higher pHs. From the CD studies a 15% α-helix content was determined, more than the 7-10% found for Blg or the 7% for Rbp but very close to the content on insectocyanin, with which shares 30-40% residue identity. We examined too, the association of this globular protein with lipid vesicles, and electron-micrographs showed the formation of rouleaux structures, an effect commonly observed with other apolipoproteins; CD showed a small change in the amount of α-helix of the bound-protein .
The binding of several hydrophobic molecules was monitored by fluorescence and it was shown that arachidonic acid, a fatty acid with intense biological activity, binds with a association constant of $4.1 \times 10^7 M^{-1}$. This ligand provides a possible link between the diverse biological situations where apo D is present. A synthetic antagonist to the thromboxane A_2 receptor was shown to bind too but with slightly less affinity.

List of abbreviations

apo AI	apolipoprotein AI
apo D	apolipoprotein D
Blg	beta-lactoglobulin
BlgY	Blg crystal form Y (spacegroup $B22_12$)
BlgZ	Blg crystal form Z (spacegroup $P3_221$)
EP092	(see page 38 for full name)
GCDF	gross-cystic disease fluid
GnCl	guanidinium chloride
12-Hete	(see page 38 for full name)
5,15-diHete	(see page 38 for full name)
HDL	high-density lipoprotein
HgI	any of the several possible complexes formed between mercury and iodine
Ins	insecticyanin
LCAT	lecithin-cholesterol acyltransferase
MMA	momercuric acetic acid
Mup	major urinary protein
Rbp	retinol binding protein
SDS-PAGE	denaturing polyacrylamide gel electrophoresis

Table of contents

Chapter 1 Introduction

1. Introduction

This chapter describes the characteristics of the lipocalin family of which beta-lactoglobulin and apolipoprotein D are members. Blg will be discussed only briefly in the next sections because several extensive reviews are available (Tilley, 1960; Townend, *et al.*, 1969; McKenzie, 1971; Hambling, *et al.*, 1992). The introduction to apolipoprotein D is more extensive since no publication collecting all of the work done so far is yet available.

1.1 Lipocalin family

The continuous emergence of three-dimensional protein structures from nuclear magnetic resonance or X-ray crystallographic studies has increased the knowledge of the structural organization of these biological molecules and has brought insight to the relationship between function and structure. In the particular case of the lipocalins, it was with the solution of the structures of human retinol-binding protein (Rbp) (Newcomer, *et al.*, 1984) and of bovine beta-lactoglobulin (Blg) (Sawyer, *et al.*, 1985) that the then disparate and apparently unrelated proteins were recognized as sharing not only the same fold but were probably also functionally and genetically related.

A trait of the lipocalin family is the polypeptide fold. The structures share the same calyx shape composed of two orthogonal β-sheets of four antiparallel strands each and an α-helix, forming a "sheltered" pocket, shown in FIG 1.1. There are six such structures described in the literature: the above mentioned Rbp and Blg, insecticyanin (Ins) (Holden, *et al.*, 1987), bilin-binding protein (Bbp) (Huber, *et al.*, 1987a,b), major urinary protein (Mup) and urinary α_2-globulin (A2u) (Böcskei, *et al.*, 1992). These structures are highly superimposable, FIG. 1.2, and for example the r.m.s. deviation of 97 core Cα between Bbp and Rbp is 1.65 Å (Cowan, *et al.*, 1990).

A simple gathering of the known and proposed functional roles of these proteins raises another common characteristic in that all of the members bind or are thought to bind small hydrophobic molecules. The crystal structures of Rbp, Ins, Bbp, Mup and A2u have in fact revealed the presence of a protein specific ligand in the pocket, establishing a strong relationship between the fold and the functional role of the proteins.

Another distinct structural family of hydrophobic molecule-binding proteins exists, to which the fatty acid binding proteins (Veerkamp, *et al.*, 1991) belong, that has a similar arrangement of pleated sheet surrounding a pocket but differing from the lipocalin family in that they are intracellular with an average molecular weight of 12-15 kDa and are formed by 10 strands of β-sheet whereas the lipocalins are extracellular, have an average weight of 18-20 kDa and contain 8 strands of sheet.

Another common trait among the lipocalins is the primary structure. Although

FIG. 1.1 Rbp - the structural characteristics of lipocalins

Top: view of the two crossed β-sheets and the α-helix.
Bottom: view into the pocket.
Green arrows are the β-strands, red cylinder is the α-helix, yellow tube are non-defined secondary structure stretches and space-filled molecule is the bound retinol.

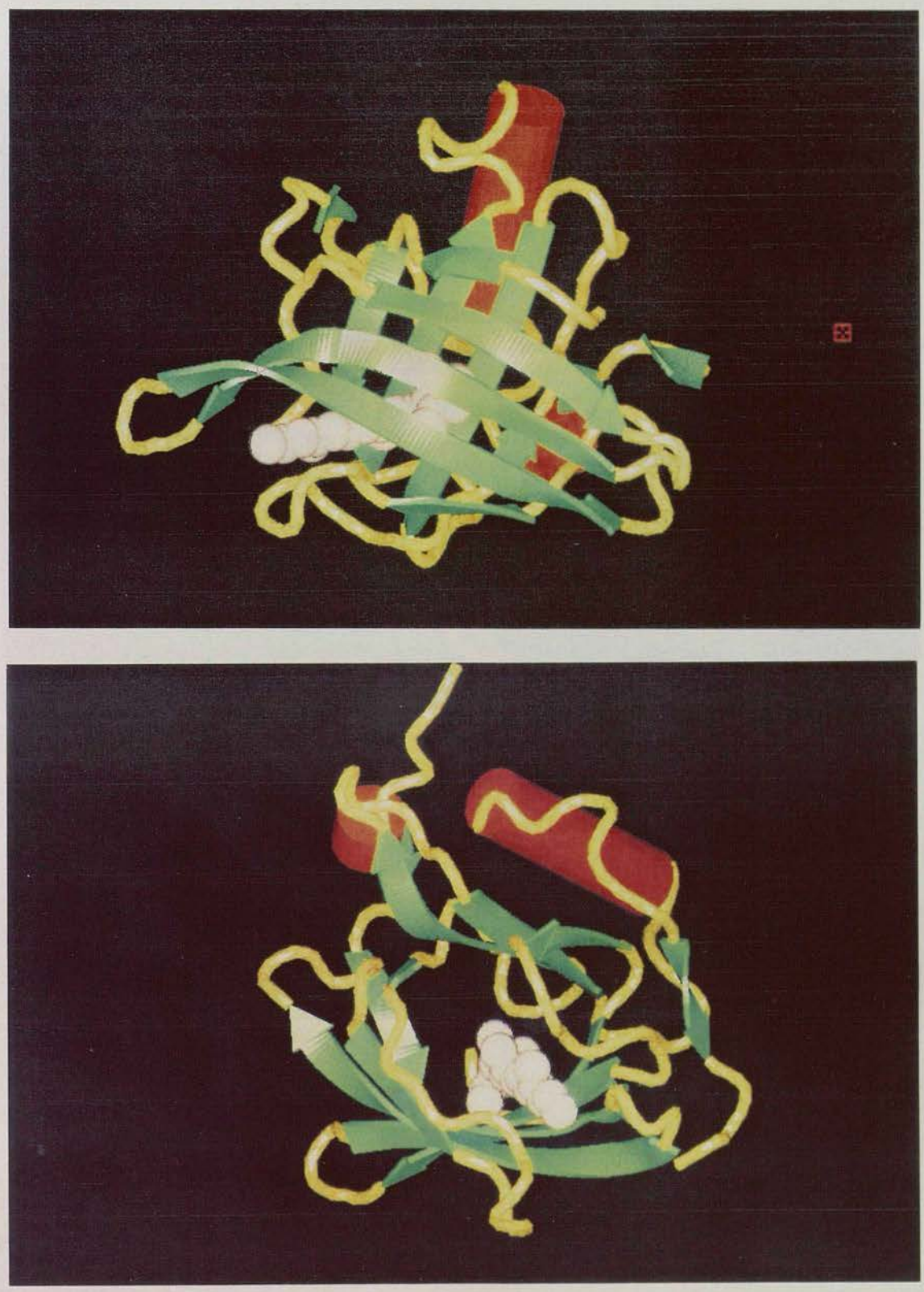
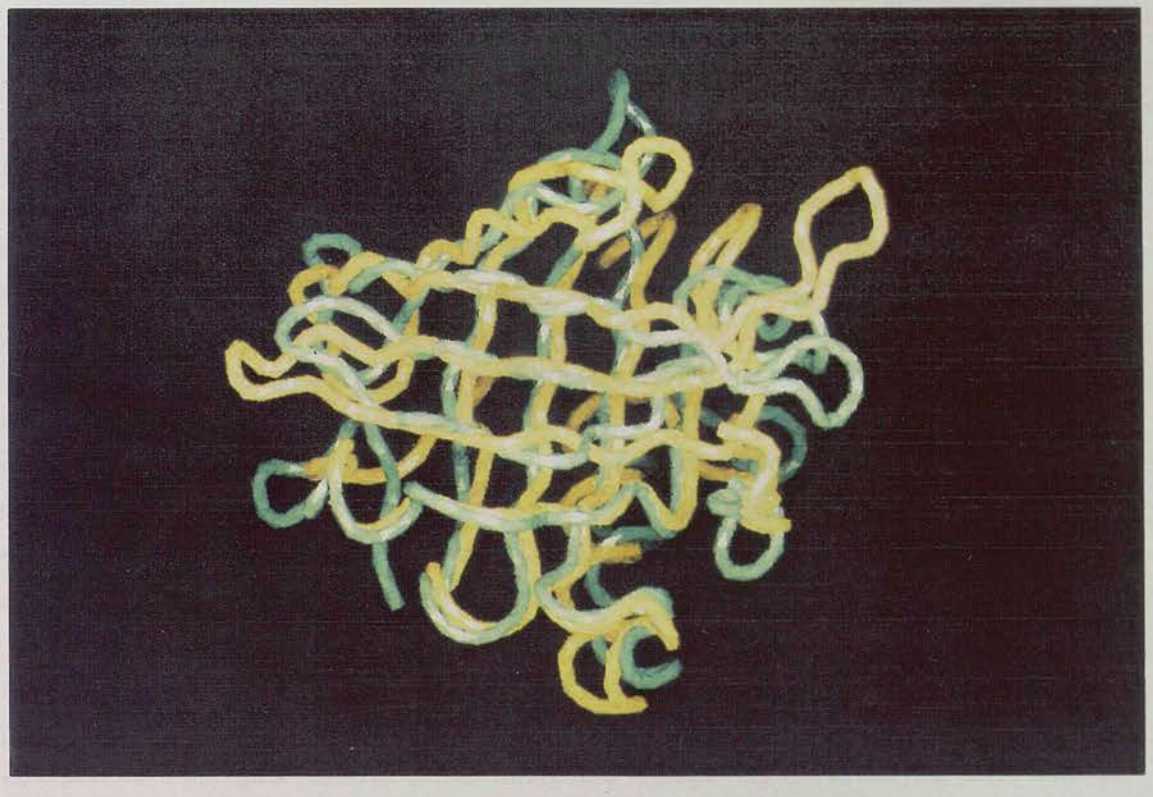


FIG. 1.2 Rbp and Mup superimposed.

Rbp and Mup main-chain are represented by a yellow and a green tube, respectively . The two structures were superimposed with the LSQ facility in O.



sequence alignment of the proteins shows low homology with an the average identity of 25-30% (Godovac-Zimermann, 1988) the following sequence motif is detected (Sawyer, 1987).

--u-x-x-G-x-W-y--//--C--//--T-D-Y-y-x-x-y--

u - basic residue
y - aromatic residue
x - either

Interestingly these conserved residues are found to be clustered at the base of the calyx (North, 1989) raising the hypothesis that this arrangement forms the recognition zone for the protein receptor, while demonstrating enough changes to permit specificity.

The motif, when applied to a search of protein sequence data bases, has allowed the extention of the lipocalin family to include other proteins of unknown three-dimensional structure. The family currently comprises some 23 different proteins listed in TABLE 1.1, of diverse provenance and role.

A further common aspect of the lipocalin family members is the primary organization of the genes. Work done by Ali and Clark, (1988) on the intron/exon organization of ovine Blg and comparison with four other proteins of the family demonstrated the parallelism of the exon size and on the relation between the exons and the particular tertiary elements encoded. It was suggested that the proteins compose a divergent but evolutionarily related gene family.

Only common features have been mentioned, however the group is perhaps more interesting in its diversity some details of which will now be described. All of the proteins are soluble in aqueous media but apolipoprotein D is capable of forming interactions with lipid vesicles (McConathy and Alaupovic, 1976; Steyrer and Kostner, 1988) and therefore possibly exists in equilibrium between lipidic and aqueous media (Holmquist, 1990). Also the very diverse sequences allow for different post-translational modifications, for example, α1-acid glycoprotein (Bennett and Schmid, 1980) and apolipoprotein D (McConathy and Alaupovic, 1976) are highly glycosylated. Some ligands are specific for a particular protein but that protein may present a broad affinity (Blg has numerous known ligands as outlined in section 1.2). Finally prostaglandin D synthase (Nagata, *et al.*, 1991) is known to have enzymatic activity - a property not observed for any of the other members.

Thus, a particular fold seems to have evolved and been applied with a "free hand" to the role of binding small hydrophobic molecules.

TABLE 1.1 Members of the Lipocalin family

Protein	Source	Putative function	Ref.
Bovine Blg	Milk	Transport/transfer in gut of young?	(1)
α-PEG	Human amniotic fluid	?	(2)
Human Rbp	Serum	Retinol transport in complex with transthyretin	(3)
Purpurin	Neural retina	Transport retinol across interphotoreceptor matrix	(4)
α1-Microglobulin	Human serum,urine	Regulation of immune response	(5)
Frog Bowman's gland protein	Nasal mucosa	Presentation of odorants to receptors	(6)
Cow nasal protein	Nasal mucosa	Presentation of odorants to receptors	(7)
Rat odorant binding protein	Nasal mucosa	Presentation of odorants to receptors	(8)
Aphrodisin	Hamster vaginal discharge	?	(9)
Apolipoprotein D	Human blood and breast cyst	Growth modulation?	
α1-Acid glycoprotein	Human serum, urine	Acute phase protein	(10)
Rat epididymal protein	Epididymal luminal	Binds retinoic for sperm maturation	(11)
A2u	Rat urine, serum	Binds pheromones?	(12)
Mup	Mouse urine, serum	Binds pheromones?	(13)
Complement--C8γ	Serum	Involved in complement complex	(14)
Crustacyanin-A and -C	Crustaceous carapace	Colouration by binding astaxanthin	(15)
Insectocyanin	Insect hemolynph	Binds biliverdin IX	(16)
Bbp	Insect hemolynph	Binds biliverdin IX	(17)
Prostaglandin--D sinthase	Rat and human brains	Converts prostaglandin-H2 to prostaglandin-D2	(18)
Tear pre--prealbumin	Human tear	Involved in the formation of the air/tear thin lipid film?	(19)
Condrocyte 21 protein	Skeletal tissue	Stabilization of mature cells	(20)
Probasin	Prostate epithelial cells	Linked to the state of cell differentiation	(21)

(1) Godovac-Zimmermann and Braunitzer, (1987);(2) Julkunen, *etal.*, (1988);(3) Rask, *et al.*, (1979);(4) Berman, *et al.*, (1987);(5) Åkerström and Lögdberg, (1990);(6) Pervaiz and Brew, (1985);(7) Lee, *et al.*, (1987);(8) Pevsner, *et al.*, (1988);(9) Henzel, *et al.*, (1988);(10) Cooper, *et al.*, (1987);(11) Brooks, *et al.*, (1986);(12) Unterman, *et al.*, (1981);(13) Clark, *et al.*, (1987);(14) Haefliger, *et al.*, (1987);(15) Keen, *et al.*, (1990);(16) Riley, *et al.*, (1984);(17) Sutor, *et al.*, (1988);(18) Nagata, *et al.* (1991);(19) Redl, *et al.*, (1992);(20) Cancedda, *et al.*, (1990);(21) Spence, *et al.*, (1989)

1.2 Beta-lactoglobulin

This protein is secreted as a major component of the milk of several ruminants like cow and sheep, as well as other mammals like pig and dolphin (Hambling, *et al.*, 1992).

Bovine Blg is a protein of 162 residues (Braunitzer, *et al.*,1972), 18kDa. It does not present any chemical post-translational modification in the most common variants. The sequence variability induces differences in its behaviour in solution in that some variants form dimers under physiological conditions while others are present as monomers. In fact, in the same animal, genetic polymorphism can produce molecules that have different properties as is the case with bovine variant B which is five times more soluble than variant A.

Despite the availability of large amounts of pure protein since the first isolation of Blg in the 1930's and subsequent numerous studies to which it was subjected, the function of the protein is still unknown although there is a whole array of evidence that suggests a role in the transport of some essential metabolite for the new-born. Of particular interest are its resistance to acidic conditions and gastric digestion (McAlpine, 1991), the discovery that Blg enhances the retinol uptake in the jejunum and ileum of suckling rats (Said,*et al.*, 1989), the detection of a specific receptor in the neonate calf (Papiz, *et al.*, 1986) and its capacity to bind small hydrophobic ligands.

The protein is known to bind a wide variety of compounds. All are small molecules with low aqueous solubility, some with high biological relevance, see TABLE 1.2. Binding studies have been undertaken using a variety of techniques and attempts have been made to characterize the site, in terms of residues and environment involved in the interaction with the ligand. Fluorescence of the protein or of the ligands (Dufour, *et al.*, 1990; Fugate and Song, 1980), site-direct mutagenesis (Cho, *et al.*, 1993) and X-ray crystallography (Papiz, *et al.*, 1986. Monaco, *et al.*, 1987) were the most promising approaches but disappointingly have not been able to pinpoint the site or sites of binding and so the relevance of the pocket, in contrast to other members of the lipocalin family, is still a matter of discussion. The protein shows conformational and aggregation changes that are pH dependent (Hambling, *et al.*, 1992). They are of interest mainly because the protein can resist acidic conditions without denaturation, allowing it to pass through the stomach upon ingestion. The bovine protein is a monomer below pH 3.5, then in the range 3.7 to 6.5 it exists in equilibrium between monomer, dimer and octamer. Above pH 6.5, the predominant forms are the dimer and monomer. These changes are coincident with conformational changes detected by circular dichroism and optical rotatory dispersion in particular the transitions occurring between pH 4 and 6, and pH 6.5 and 7.8 - often called the Tanford transition. Above pH 8.5 irreversible denaturation is initiated.

TABLE 1.2 Blg ligands

Ligand	Numb./monomer	Association[a] constant (M^{-1})	Ref.
Retinol	1	$5x10^7$	(1)
Stereate	1	$1.7x10^5$	(2)
Palmitate	1	$6.8x10^5$	(2)
Laurate	1	$0.5x10^5$	(2)
Oleate	1	$0.4x10^5$	(2)
Heptane	1	$0.5x10^5$	(3)
Butane	1	$1.7x10^3$ ($5.8x10^2$)	(4)
Pentane	1	$7.1x10^3$ ($6.2x10^3$)	(4)
Iodobutane	1	$2.8x10^3$	(4)
SDS	1	$3.1x10^5$	(5)
2,6-MANS[b]	1	$3.4x10^5$	(6)
Methyl orange	1	$0.2x10^4$	(5)
n-Octylbenzene-*p*-sulphonate	1.5	$6.3x10^4$	(5)
p-Nitrophenol	1	$1.9x10^4$	(7)
p-Nitrophenylacetate	1	$3.1x10^4$	(7)
p-Nitrophenyl-β-glucuronide	1	$1.6x10^4$	(7)
p-nitrophenyl sulphate	1	$2.0x10^3$	(7)
p-Nitrophenyl pyridoxal phosphate	1	$3.1x10^3$	(7)
2-Heptanone	1	$0.2x10^3$	(5)
2-Octanone	1	$0.5x10^3$	(5)
2-Nonanone	1	$2.4x10^3$	(5)
Toluene		$4.5x10^3$ ($5.9x10^1$)	(8)
Trifluorotoluene		$4.2x10^2$ ($3.1x10^1$)	(8)
Hexafluorobenzene		$1.6x10^3$	(8)
Protoporphyrin IX	1	$4x10^7$	(9)
Hemin	1	$0.4x10^7$	(9)
Ellipticine	0.5	$7x10^5$	(10)

a- Figures in parentheses represent the association constant for the binding of a second ligand at a second site. b- N-methyl-2-anilino-6-naphthalene sulphonate. (1) Fugate and Song (1980);(2) Spector and Fletcher, (1970);(3) Mohammadzadah-K., *et al.*, (1969);(4) Wishnia and Pinder, (1966);(5) O'Neill and Kinsella, (1987);(6) Lovrien and Anderson, (1969);(7) Farrel, *et al.*, (1987);(8) Robillard and Wishnia, (1972a,b);(9) Dufour, *et al.*, (1990);(10) Dodin, *et al.*, (1990)

The crystallographic studies of Blg have been on going for more than thirty years and much of the initial work was performed by D. Green and coworkers. It culminated in the publication of the low resolution structures of four different crystal forms reported in Green, *et al.*, (1979). The work revealed the low content of α-helix and the presence of a molecular dyad, both facts in agreement with other studies. At 6 Å however it did not allow much insight into the detailed structure of the molecule.

The extension of the resolution to 2.8 Å in the so-called crystal form Y (space group $B22_12$) was achieved a few years ago and is described in Papiz, *et al.*, (1986). It was shown then that the overall folding consisted of the previously described (in section 1.1) calyx shape and the close resemblance to the Rbp structure was noted. The point is clearly made that the structure suffers from a few problems like the lack of definition of the density of one external loop, as well as poor density for the N- and C-termini and one of the disulphide bridge (Cys66-160). The existence of a molecular dyad is reaffirmed and the structure shows that the contact surface is established by one of the β-sheet strands. No density identifying a ligand molecule was found in the pocket but a model for the binding of retinol based on the Rbp binding site was discussed. This model places the retinol in a deeper position inside the calyx than is found in Rbp.

Some of these findings were contested by work on the Z crystal form (space group $P3_221$) (Monaco, *et al.*, 1987) in particular in the assignment of density for a retinol molecule not in the putative binding pocket but on the surface of the protein. The molecule was found to interact with hydrophobic residues in the proximity of Lys141 with which, and in agreement with biochemical studies (Horwitz and Heller, 1974), it is postulated it will form a Schiff-base interaction. The crystal packing revealed no close interaction between single protein molecules, as found in the Y form, and it was considered that the monomer had been "selected" during crystallization. This is corroborated by the existence of a monomer-dimer equilibrium under the conditions of crystallization but it contradicts the findings described in the low resolution studies (Green, *et al.*, 1979).

1.3 Apolipoprotein D

The following sections present a comprehensive literature review of apolipoprotein D. The fields of apo D and GCDFP-24 have evolved separately but these two molecules have since been determined as one protein. This review will attempt to integrate the two fields.

1.3.1 Biochemical characteristics

Apolipoprotein D (apo D) from human plasma has been biochemically

characterized as a glycoprotein, with 18% (in dry weight) content of carbohydrate (McConathy and Alaupovic, P., 1976) comprising the following sugars: 1.89% D-glucose, 4.42% D-galactose, 2.96% D-mannose, 4.49% D-glucosamine and 4.82% neuraminic acid. Two possible glycosylation sites at Asn45 and Asn75 have been interpreted from the sequence (Drayna, *et al.*, 1986) and later confirmed from peptide sequencing (Balbín, *et al.*, 1990).

The sequence is 169 amino acids long and a molecular weight of 19302 kDa is expected but the associated sugars increase this to ~22 kDa (McConathy and Alaupovic, 1976) and present possible migration problems in many of the normally used systems for molecular weight determination (See and Jackwoski, 1990; Balbín, *et al.*, 1990). It presents an apparent molecular weight by SDS-PAGE of ~30 kDa in the blood form (purified from High Density Lipoproteins - HDL) and ~24 kDa in the cyst form (purified from Gross-Cystic Disease Fluid - GCDF - in human breast).

Apo D purified from GCDF migrates in a gel filtration column as a 100 kDa protein (Balbín, *et al.*, 1990), leading to the assumption that it is possibly present in the GCDF as a tetramer. In the lipoprotein system this protein is found associated to apo AI and lecithin-cholesterol acyltransferase (LCAT) as demonstrated by the simultaneous removal of 96% of apo D, 64% of LCAT and 11% of apo AI from high-density lipoproteins particles, with an anti-apo D immunoadsorbent (Albers, *et al.*, 1981). However, Hölmquist, (1989) using gel filtration to purify LCAT, found that only 33% of this protein is associated with apo D. Also Fielding and Fielding, (1980) suggest the existence of two complexes with the following compositions, apo AI : LCAT : apo D with molar ratio 1.0 : 0.9 : 1.9 and apo D: apo AI : apo AII (1.0 : 3.8 : 2.2). Blanco-Vaca, *et al.*, (1990; 1992) have found evidence for disulphide interchange with other apolipoproteins and suggest that the covalently associated heterodimer apo D : apo AII accounts for 67% of the apo D in the plasma, while other minor heterodimers are formed with apo B100 in VLDL (very-low density lipoproteins) and LDL (low-density lipoproteins). The reasons for this dimerization are not clear. There is evidence as well, for the existence of free apo D in the plasma as it was found in human urine at a concentration of 1.4 mg/l (Hölmquist, 1990). This is a higher value than any described for other apolipoproteins which are tightly associated to lipidic particles.

The protein is capable, as well, of establishing interactions with lipids. In particular in the plasma, by disruption of the HDL a lipoprotein sub-particle, the Lp-D, is formed that consists of 65-75% protein and the rest lipids (McConathy and Alaupovic, 1973; 1976; 1986).

1.3.2 Genetic characteristics

The screening of a human cDNA liver library has allowed the determination of the whole gene sequence, translated to 169 amino-acids for the mature apo D plus a 20-

residue amino terminal secretion peptide signal (Drayna, *et al.*, 1986). This work confirmed that the N - terminus is chemically blocked (McConathy and Alaupovic, 1976) by cyclization of the glutamic acid.

The sequencing of several clones indicated a possible polymorphism at position 110, Leu for Phe (Drayna, *et al.*, 1986). Polymorphism was also detected in the blood apo D by isoelectric focusing (IEF) gel where several bands were revealed between pH 4.2-4.9 (Kamboh, *et al.*, 1989). The number of bands is reduced by digestion of the protein with neuraminidase proving that some of the polymorphism is due to differing carbohydrate content and in particular of neuraminic acid. It was also demonstrated (Kamboh, *et al.*, 1989) that the IEF pattern was different for some individuals of African ancestry and that this different phenotype was due to the existence of two alleles for the protein, APOD*1 and APOD*2. The frequency of APOD*2 in Nigerian blacks is 2.2% and in U.S. blacks 1.3%. This allele was not detected in U.S. whites, Dogrib Indians, Mayan Indians, Aleuts, Kodiak Island Eskimos and St. Lawrence Island Eskimos.

The apo D human gene is divided into at least five exons; three introns are found in the protein-coding region and at least one in the 5' untranslated region (Drayna, *et al.*, 1987). The intron positions are similar to those of the Rbp (retinol-binding protein) gene. The unique gene was localized in human chromosome 3, with 41% of the signal clustered over the distal long arm.

The rat amino acid sequence deduced from a cDNA library of sciatic nerve showed 73.4% identity to the human gene (Spreyer, *et al.*, 1990), while the sequence from rabbit testis cDNA library is 80% identical to the human (Provost, *et al.*, 1990).

1.3.3 General distribution

In the work described by Drayna, *et al.*, (1986), the human tissue distribution of apo D mRNA was examined. The mRNA was detected in the pancreas, adrenal gland, kidney, small intestine, placenta, spleen and fetal brain, all of which presented a higher signal of expression than the liver. It was not detected in white blood cells or in monocytes (cell lines U937 and HL60).

The protein has been immuno-localized in the supranuclear area of enterocytes (human intestine cells) and in the perinuclear area of hepatocytes (human liver) (Bouma, *et al.*,1988). The protein was not detected in the colon (Mazoujian and Haagensen, Jr., 1990).

Apo D has been detected by immunoreaction in HDL (Provost, *et al.*, 1990) of rabbit, pig, dog, cow, goat, sheep, Cynomolgus and Rhesus monkey; it was not detected in guinea pig, cat or rat HDL. In rabbit the apo D mRNA was present in, from highest to lowest level of expression, spleen, adrenal gland, lung, brain, testis, kidney, heart, small intestine, bone marrow, thymus, pancreas, skeletal muscle, liver and

lymph node. The level of expression in the spleen is 59-fold higher than in the liver.

In the Rhesus monkey (Smith, *et al.*, 1990), apo D mRNA was detected in cells of mesenchymal origin, which were identified as fibroblasts and interstitial cells, and in most of the peripherial tissues examined like spleen, testes, liver, pancreas, skeletal muscle, kidney, jejunum, pituitary, peripherial nerve and brain.

1.3.3.1 Distribution in the nervous system

Apo D was detected (Boyles, *et al.*, 1990) in rat sciatic nerve extracts as well as in rat spinal cord and rat dorsal root ganglia but the detection system used could not detect apo D in rat plasma due to very low levels. In particular, apo D was detected by immuno-cytochemistry in spinal cord fibrous astrocytes, oligodendrocytes and protoplasmic astrocytes, though in smaller quantities in the last two cell types. A signal was also detected in giant motor neurons of the ventral horn and in the neuropil of the dorsal horn. In the dorsal root ganglia and sciatic nerve, apo D was detected in cells thought to be neurolemmal or fibroblasts.

Electron microscopy was used to observe the apo D-containing cells in the rat neural tissue (Boyles, *et al.*, 1990). It was found in the endoplasmic reticulum (part of the secretory apparatus) of both spinal cord astrocytes and oligodendrocytes and of the neurolemmal or fibroblasts of the dorsal root ganglia and sciatic nerve. The protein was detected too on the cell surface of neurolemmal cells and associated with the matrix of the nerve. In neurons of both spinal cord and dorsal root ganglia it was observed in small, membrane-bounded compartments and never in the secretory apparatus. The macrophages of the sciatic nerve contained apo D only in lysosomal structures and infoldings or projections of the cell surface. All of this suggests that apo D is a secretory product of astrocytes, oligodendrocytes and neurolemmal or fibroblastic cells while being taken up by macrophages and specific neurons.

Spreyer, *et al.*, (1990), working with regenerating rat sciatic nerve, identified the endoneural fibroblast as the major type cell expressing apo D mRNA; no mRNA was found in macrophages or Schwann cells.

In Rhesus monkey (Smith, *et al.*,1990) synthesis was detected in cells of the peripherial nerve and brain, like neuroglial cells, cells in the subarachnoid space of the surface of the brain as well as perivascular cells and scattered neurons in the brain .

1.3.3.2 Localization in tumour cells

Apo D is in general found at higher concentrations in the cytosol of well differentiated breast tumour than in aggressive poorly differentiated ones (Silva, *et al.*, 1982).

Using polyclonal antibodies and the immunoperoxidase technique (Mazoujian

and Haagensen, Jr., 1990) apo D was localized in sweat gland tumours of apocrine differentiation. It was found as well in metaplastic epocrine epithelium of the breast and in the fluid contained within the cysts though normal breast epithelium did not stain except for a few sporadic cells. Normal adrenal cortex and corpus luteum, peripherial nerves, pituitary and renal tubules demonstrated immunoreactivity. Colon, oesophagus, lung, pancreas, parathyroid glands, stomach and thyroid revealed no staining.

The same study showed that 13 out of 25 cases of breast carcinoma, as well as endometrial (6 out of 8), ovarian (4 out 10), prostatic (5 out of 8) adenocarcinomas and leiomyosarcoma (1 out 1) stained positively for apo D. Some staining was found for carcinomas from the colon, kidney, liver, lung, pancreas and stomach. It was concluded that apo D is localized in steroid-responsive tissues.

In the particular case of sweat glands (Mazoujian, 1990), immunocytochemistry experiments revealed the presence of apo D in normal apocrine glands and ducts but not in eccrine ones. In the sweat gland tumours studied, many stained for apo D. No detection was demonstrated in tumours of eccrine differentiation. It was concluded that apo D is useful for the identification of apocrine differentiation and demonstration of functional activity on this subset of sweat gland tumours.

Several cancer lines excrete apo D. The LNCaP prostate cell line (Simard,*et al.*, 1991) and the human breast ZR-75-1, MCF-7 (Simard, *et al.*, 1990) and T47D (Haagensen, *et al.*, 1992) are ideal models for the study of the physiological function of apo D and of the regulation of its expression.

Though not constituted of tumour cells, cysts are considered to be the consequence of aberration in the environment of the particular tissue leading to hypersecretion (Dogliotti, *et al.*, 1990). Of the proteins found in the fluid of cysts from the human breast gross-cystic disease apo D is the major component (Balbín, *et al.*, 1990)

1.3.3.3 Apo D in the lipoprotein system

The protein was first detected in plasma as a distinct apolipoprotein by McConathy and Alaupovic, (1973). That study concluded that in terms of amino acid composition and immunological criteria, the protein is different from other apolipoproteins.

Apo D (McConathy and Alaupovic, 1986) is present as a minor component in the plasma (5% of the proteins) of normolipidemic subjects, at concentrations ranging from 60-100 mg/l of plasma. It is mainly (60 - 65%) localized in the HDL (high density lipoproteins) with the following relative distributions; 43% in the HDL_3 subparticle and 21% in the HDL_2 subparticle (Curry, *et al.*, 1977) with trace amounts in the VLDL (very low density lipoproteins) and LDL (low density lipoproteins), and the remainder (36%) in the VHDL (very high density lipoproteins) - which can be considered to be a

lipoprotein-free fraction. It was already mentioned above, that by the break-up of the HDL_3 a sub-particle is formed designated as Lp-D. It consists of 65-75% protein, while the remainder is composed by 8.2% triglycerides, 18.1% cholesterol, 27.7% cholesterol ester and 46.5% phospholipid of which 26.4% is lysolecithin, 33.3% sphingomyelin and 40.1% lecithin. These are unusually high relative values for the first two phospholipids and low for the last when compared with other lipoproteic particles (McConathy and Alaupovic, 1986 and 1976).

Baboon plasma contains apo D which has been partially characterized with respect to the amino acid and carbohydrate (Bojanovski, *et al.*, 1980). Specifically it differs from the human form in that it posesses relatively more basic amino acids and more mannose in its carbohydrate composition.

1.3.4 Functional role

The function of apo D is far from clear. The discovery of circumstances *in vivo* where a particular protein is present, especially conditions where either specific or high expression occur, or where a defined pathology is concurrent with biochemical alterations in the behaviour of a protein (genetic alterations or unbalanced regulation), are of immense use in the determination of the functional role. At the moment, four main manifestations of protein occurrence can be delineated and studied: the association of apo D with other apolipoproteins and lipids in the plasma, the overproduction of protein in injured peripheral nerves of rat, the excretion of the protein by some tumour cells and finally the production of apo D in cysts. These four situations will be outlined separately and the common points among them summarized.

1.3.4.1 Plasma

The plasma lipoprotein system is a very dynamic one that fulfils the task of transporting triacylglycerols and cholesterol through the organism (Erkelens, 1989). It is composed of macromolecular structures designated as lipoprotein particles and these have been categorized according to their separation by ultracentrifugation in increasing density as chylomicrons, very-low-density-lipoproteins (VLDL), intermediate-density-lipoproteins (IDL), low-density-lipoproteins (LDL) and high-density-lipoproteins (HDL) (Chapman, 1986). Each displays distinct physical properties, see TABLE 1.3, and is composed of protein and lipids, cholesteryl-esters, triglycerides, phospholipids and cholesterol in the proportions described in TABLE 1.4.

The lipoprotein particle is basically a bilayer system: a "sack" composed of phospholipids, cholesterol and protein, ready to receive neutral lipids and consequently to swell. The apolipoproteins stabilize the structure and size of particular lipoproteins (Erkelens, 1989), see TABLE 1.5, but some have further functions. For example, apo

B48 enables chylomicrons to move through the cellular membrane and apo B100, does the same for the VLDL. Some act as ligands to receptors and therefore "guide" the lipoproteins to catabolic sites, like apo B100 that interacts with apo B/E receptor, and

TABLE 1.3 Physical characteristics of lipoproteins (Gotto,Jr., *et al.*, 1986)

Lipoprotein	Particle size (nm)	Density (g/ml)	Molecular weight (Da)
Chylomicrons	75-1200	0.93	~$4x10^8$
VLDL	30-80	0.93-1.006	1-$8x10^7$
IDL	25-35	1.006-1.019	5-$1x10^7$
LDL	18-25	1.019-1.063	$2.3x10^6$
HDL2	9-12	1.063-1.125	$3.6x10^5$
HDL3	5-9	1.125-1.210	$1.8x10^5$

TABLE 1.4 Composition of lipoproteins (Gotto,Jr., *et al.*, 1986)

	Surface components (mol%)			Core lipids (mol%)	
	Cholesterol	Phospho-lipids	Apolipo-proteins	Triglycerides	Cholesteryl esters
Chylomicrons	35	63	2	95	5
VLDL	43	55	2	76	24
IDL	38	60	2	78	22
LDL	42	58	0.2	19	81
HDL2	22	75	2	18	82
HDL3	23	72	5	16	84

apo E with the apo E and apo B/E receptors. Others are cofactors for the enzymes involved in lipid transport such as apo AI for lecithin-cholesterol acyltransferase (LCAT) or apo CII for lipoprotein lipase, or even as inhibitors like apo CIII in the catabolic pathway of the VLDL and chylomicrons. A number of enzymes (Erkelens, 1989) are associated with, or act upon, the system, such as LCAT (which esterifies cholesterol), lipoprotein lipase (which hydrolyses triacylglycerols from chylomicrons and VLDL and is attached to the luminal side of endothelial cells), the hepatic lipases

(which probably hydrolyse triacylglycerols from IDL and interconvert HDL subparticles - HDL_2 and HDL_3) and there are as well some transfer proteins like cholesteryl-ester transfer protein which are involved in the exchange or transfer of neutral lipids. Essential to the whole process are the already-mentioned receptors that exist on the surface of cells of tissues like the liver, involved in the metabolism of lipids.

TABLE 1.5 Protein composition of the lipoproteins (Gotto,Jr., *et al.*, 1986)

	HDL	LDL	IDL	VLDL	Molecular weight
		(mol%)			(Da)
Apo AI	100				28016
Apo AII	100				17414
Apo AIV					44465
Apo B48					264000
and		90	8	2	
Apo B100					550000
Apo CI	97		1	2	6630
Apo CII	60		10	30	8900
Apo CIII	60	10	10	20	8800
Apo D	100				22000
Apo E	50	10	20	20	34145

The close association of apo D, in the HDL_3, with apolipoprotein AI (apo AI) and especially with LCAT (Albers, *et al.*, 1981; Fielding and Fielding, 1980) which catalyzes the formation of cholesteryl-esters and lysophosphatidyl choline from cholesterol and phospholipids, led to the proposition of the involvement of apo D in the esterification of cholesterol either as an LCAT activator (Kostner, 1974) or as a transfer protein of cholesteryl-ester (Chajek and Fielding, 1978). However, these two possible roles were proved not to be real. The involvement of apo D was verified not to be of a regulatory kind, as it is for apo AI, because very little change in the activity of LCAT was noted in the presence of apo D (Steyrer and Kostner, 1988), and the authors postulated instead, a stabilization effect of apo D on LCAT. Furthermore, the separation of apo D from a protein with cholesteryl-ester transferring properties (Morton and Zilversmit, 1981) and the verification that the removal of most of the apo D from plasma had no effect on the transferring activity (Albers, *et al.*, 1981) eliminated the idea of a role in the direct transport of cholesterol, in the lipoprotein system. The fact

remains however, that a physiological complex exists suggesting that the roles of these three proteins, apo AI, apo D and LCAT are connected.

Several studies have been undertaken to find factors and circumstances that can be correlated to abnormal concentrations of apo D in the plasma. Thus, Albers, *et al.*, (1981), have found that the levels of apo D are slightly higher in men than in women both for normolipidemic and hyperlipidemic adults. This was confirmed by Haffner, *et al.*, (1985) for a random population, though no such correlation had been found by Curry, *et al.*, (1977). Hypertriglyceridemics showed lower levels of apo D (Albers, *et al.*, 1981) while hypercholesterolemics and dysbetalipoproteinemics presented normal concentrations (all these clinical conditions are associated to genetic disorders or disorders like daibetes, hypothyroidism, anorexia, etc. See TABLE 1 in Gotto,Jr., *et al.*, 1986). The same study presented evidence for half the normal apo D levels in subjects with familial LCAT deficiency (having little or no LCAT in the plasma). Curry, *et al.*, (1977) also found that the phenotype designated as hyperchylomicronemia showed half the normal levels of apo D.

Two alleles have been detected (Kamboh, *et al.*, 1989) in the human population (see section 1.3.2) and an attempt was made to find changes associated with the the presence of the very rare allele APOD*2. Though noting that due to the small population examined the results are not statistically relevant, it was concluded that no significant change in the lipid or lipoprotein levels could be found to be associated with the presence of APOD*2.

A thorough collection of data reported by Haffner, *et al.*, (1985) examined the correlation of factors like smoking, intake of alcohol, age, gender and adiposity with apo D, apo AI, apo AII levels and other components of the lipoproteins. It was observed that apo D was positively correlated to the cholesterol content in the subparticle HDL_3 (HDL_3C), apo AI and apo AII, while LCAT had a moderate positive correlation with HDL_3C and a negative correlation with HDL_2C. It was found as well that age was positively correlated with apo D in women while no other factor was correlated to age. The triglycerides in the VLDL were correlated to HDLC, especially HDL_2C, and with LCAT; the same pattern was seen for the body mass index (BMI). Alcohol was positively correlated to HDLC as well as to apo AI and apo AII, both in men and women, while correlated to apo D only in men. Negative correlations were observed for males between smoking and apo AI, D, AII, and in both genders for smoking and HDL_3C and LCAT.

1.3.4.2 Neural tissue

The regeneration of peripheral nerves (Boyles, *et al.*, 1990; Spreyer, *et al.*, 1990) demands an influx of "raw materials" for the re-establishment of the tissue. Schwann cells distal to the injury site, and macrophages, are involved in cleaning the

destroyed cells and in the recuperation of lipids, such as cholesterol, from the cell membrane. In addition, locally produced lipids and lipids entering from the plasma are used for the regeneration.

Lipoproteins seem to have a role in the transport of the lipids during denervation and regeneration (Boyles, 1989) as large amounts of cholesterol-rich lipoproteins with apo E and apo AI were identified in regenerating rat sciatic nerve. In the same work the receptor for apo E was found to be present in regenerating axons. While apo E was shown to be produced by both locally resident macrophages and endothelial cells, apo AI entered the nerve from the plasma as a component of plasma lipoproteins when the blood/nerve barrier is destroyed during injury.

Later, apo D and apo AIV were identified also; they accumulate in the regenerating peripheral nerve (Boyles, *et al.*, 1990), but while apo AIV seemed not to be produced locally as concluded by absence of signal in metabolic labelling experiments with [^{35}S]methionine, apo D did incorporate the label. A fraction of apo D, 13-15%, was found in the lipoprotein-free fraction indicating that some of the protein is in the free form.

These proteins (apo's AI, AIV, D, E) and albumin are seen (Boyles, *et al.*, 1990) to increase from the first day after injury over several weeks while the regeneration occurs, FIG. 1.3. Apo AI and AIV increased for 2-3 days before decreasing briefly and then peaking at 3 weeks at 26-fold and 14-fold respectively of the basal concentration. The concentration decreases to around 3-fold above normal by 6 weeks after injury. Apo D and apo E showed a fast increase and their concentrations peaked at 500-fold and 250-fold respectively at 3-4 weeks after injury, then the concentrations decrease and at 6 weeks the levels are still 200-fold and 150-fold respectively of the normal. The increase coincides with the period of axon growth (1-2 weeks after injury) and active myelination (2-8 weeks after injury).

Apo E was demonstrated (Handelmann, *et al.*, 1992) to have a double direct effect *in vitro*, on the growth of the rabbit dorsal root ganglion neurons, when complexed as a lipoprotein it stimulates neurite growth by increasing the uptake of cholesterol (an essential component of membranes) in the form of β-VLDL. When in its free-form it decreases branching while promoting extension of the cells.

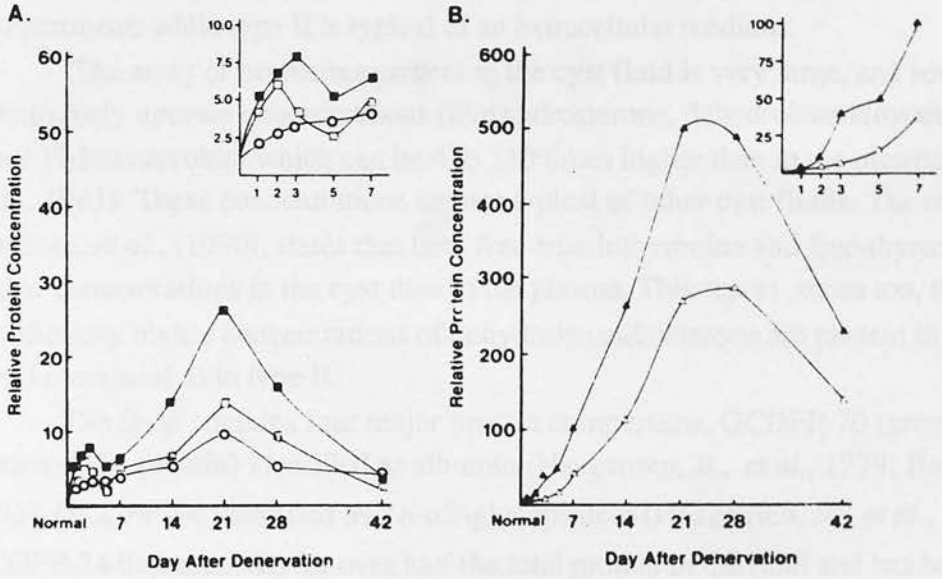
Another study (Spreyer, *et al.*, 1990) found an increase in apo D mRNA in endoneural fibroblasts present in regenerating rat sciatic nerve. The increase was noted from the second day after injury, peaking at to 40-fold by the sixth day and then slowly decreasing to 5-fold after 12 weeks. It is noted that the increase cannot only be accounted for by the 4- to 8-fold increase in the number of fibroblasts. The protein was found in the lipoprotein fraction of the medium used to culture the crushed sciatic nerves.

Despite this involvement in the neural tissue, patients with apolipoprotein

deficiencies, like familial hypercholesterolemia where there is a diminished expression of normal functional low density lipoprotein receptors (in particular receptors for apo E) or with lower levels of apo E, suffer no neurological problems (Boyles, *et al.*, 1990).

FIG. 1.3 Variation of the concentration of albumin, apo's AI, AIV, D and E in regenerating rat sciatic nerve. (Boyles, *et al.*, 1990)

The points were generated by densitometry of immunoblots of reduced SDS-PAGE. The concentrations are relative to the level in normal nerve. Graph A shows the changes for albumin (circles), apo AIV (open squares) and AI (filled squares). Graph B shows variations of apo D (filled triangles) and apo E (open triangles). Insets represent detailed variation in the first seven days.



1.3.4.3 Breast cysts

Gross-cystic-disease (GCD) is the most common mammary pathology of the human breast tissue occurring in 10% of the population of premenopausal women. The peak age for full development of the disease is around 40-44 years (Haagensen, 1971). The disease is characterized by the spontaneous formation of cysts in the mammary gland (Molina, *et al.*, 1990). These structures are filled with abundant fluid and lined

by a single layer of epithelium or connective tissue (Haagensen, Jr., *et al.*, 1979).

Though the cysts are not considered to be premalignant lesions (Dogliotti, *et al.*, 1990), several studies have examined the correlation of GCD to carcinoma (Haagensen, 1971; see references in Balbín, *et al.*, 1990) and have found that there is an increase in the risk of developing breast cancer in women exhibiting GCD.

The fluid has been analysed extensively and it has been possible to categorize the cyst into two main groups. The components reported to be present in the cysts were hormones (Dogliotti, *et al.*, 1990; Bradlow, *et al.*,1981), electrolytes (Dogliotti, *et al.*, 1990) and proteins (Dogliotti, *et al.*, 1990; Balbín, *et al.*, 1990; Haagensen, Jr., *et al.*, 1979). Another approach to categorize cysts has been to examine the morphology of the lining epithelium cells (Dogliotti, *et al.*, 1990).

The determination of the concentration of monovalent cations, Na^+ and K^+, and their charge counterpart Cl^-, allows an immediate separation of cysts in two groups. Type I presents high K^+/Na^+ ratios, low concentration of Cl^- and low pH, while for type II these factors are present as low K^+/Na^+, and higher Cl^- concentration and pH. It is noticeable that the type I presents electrolyte characteristics of an intracellular compartment, while type II is typical of an extracellular medium.

The array of hormones present in the cyst fluid is very large, and some present in extremely unusual concentrations (like androsterone, dehydroisoandrosterone and other 17-ketosteroids) which can be 4 to 110 times higher than in the plasma (Bradlow, *et al.*, 1981). These concentrations are not typical of other cyst fluids. The report by Dogliotti, *et al.*, (1990), states that both free-triiodothyronine and free-thyroxine had higher concentrations in the cyst than in the plasma. This report states too, that significantly higher concentrations of dehydroisoandrosterone are present in cysts of type I compared to in type II.

The fluid contains four major protein components, GCDFP-70 (gross-cystic-disease-fluid-protein) identified as albumin (Haagensen, Jr., *et al.*, 1979; Balbín, *et al.*, 1991), GCDFP-44 identified as Zn-α2-glycoprotein (Haagensen, Jr., *et al.*, 1979), GCDFP-24 that accounts for over half the total protein in the fluid and has been identified as apo D (Balbín, *et al.*, 1990) and GCDFP-15. Some other proteins are present, though in much lower concentrations, like α-amylase, epidermal growth factor (Dogliotti, *et al.*, 1990), pepsin C (Sánchez, *et al.*, 1992) and many others (Dogliotti, *et al.*, 1990). Balbín, *et al.*, (1991) have demonstrated that the distinction into the two cyst types can be done by determination of albumin content, as this protein is more concentrated in type II cysts than in type I. Interestingly, two reports have identified enzymatic activity associated with apo D from cyst fluid; Kesner, *et al.*, (1990) discusses that there is associated proteolytic activity and Banerjee, *et al.*, (1990) show evidence for associated esterase activity.

Dogliotti, *et al.*, (1990) discuss the correlation between the cyst type and the

morphology of the epithelium lining the cysts and find that epithelia of type I cysts have higher probability of being constituted by apocrine cells.

The knowledge about breast cysts is still not advanced enough to understand the mechanism of formation and the relevance of these different characteristics. However, some hypotheses have been presented, such as Balbín, *et al.*, (1991) who have proposed that type I cysts are the initial stage of development characterized by activity of the apocrine epithelium explaining the intracellular charateristics of the fluid. It follows that type II would be a later stage of development when the epithelium has been flattened and become more permeable.

Of utmost interest was the discovery of a progesterone binding protein existing in the breast cyst fluid with different properties from other known steroid binding proteins (Pearlman, *et al.*, 1973). This protein was later designated as GCDFP-24 (Haagensen, Jr., *et al.*, 1979) and then shown to be apo D (Balbín, *et al.*, 1990). Much work has been done on the binding properties of apo D (as GCDFP-24), in particular the work by Lea, (1988) and Dilley, *et al.*, (1990) who have measured the affinities for several ligands, see TABLE 1.6. These studies have concentrated upon the metabolites of cholesterol as these are so common in breast cyst fluid.

1.3.4.4 Tumour cells

The finding of high concentrations of apo D in the cyst fluid and later the discovery, as discussed in section 1.3.3.2, that the protein is present in several cases of breast carcinoma has raised the possibility of its use as a biochemical marker for breast carcinoma.

Søreide, *et al.*, (1987; 1991a; 1991b) have looked at the validity of using apo D as a biological marker for providing diagnostic information. It was concluded that there is significant negative correlation between the presence of apo D and the advance of the tumour towards more aggressive stages, and between the concentration of apo D and the presence of malignant tissue. However it was pointed out that the apo D as a marker of tumour differentiation, is only useful as part of a group of tumor parameters.

In section 1.3.3.2 it was pointed out that apo D was secreted by several cancer cell lines. Work done with breast cancer cell lines ZR-15, MCF-7 (Simard, *et al.*, 1990; Labrie, *et al.*,1990), and T47D (Haagensen, *et al.*, 1992) and the prostate cancer cell line LNCaP (Simard, *et al.*, 1991) have demonstrated that there is an inverse relation between apo D secretion and cell growth, modulated by steroid hormones. As an example, dihydrotestosterone when added in increasing concentrations to the LNCaP cell line only induces growth between 10^{-11} and 10^{-9} M, while the amount of apo D expressed in that range is diminished. Above concentrations of 10^{-9} M, the level of apo D is raised sharply by 3.2-fold. Very similar effects are seen with the other proteic

TABLE 1.6 Apo D ligands

Ligand[a]	Association constant (M^{-1})	Ref.
Cnolesterol	0	(1)
Progesterone	$1.3x10^6$	(1)
	$1.0x10^6$	(2)
R-5020	$5.5x10^5$	(1)
Org 2058	$1.3x10^5$	(1)
R 1881	$8x10^4$	(1)
Dehydrotestosterone	$1.8x10^5$	(1)
Androstenedione	0	(1)
Dehydroisoandrosterone	$1x10^5$	(1)
Tamoxifen	$4.5x10^5$	(1)
Estriol	0	(1)
Estradiol	$3x10^4$	(1)
	$<2.0x10^5$	(2)
Dehydroisoandrosterone-sulphate	$5x10^4$	(1)
Estrone sulphate	$1x10^4$	(1)
Pregnenolone	$1.3x10^6$	(2)
5β-pregnan-3-ol-20-one	$9.1x10^5$	(2)
5α-pregnan-3,20-dione	$8.3x10^5$	(2)
5β-pregnan-3,20-dione	$8.0x10^5$	(2)
Promegestone	$4.2x10^5$	(2)
Megestrol acetate	$2.2x10^5$	(2)
Medroxyprogesterone acetate	$2.2x10^5$	(2)
	$1.3x10^5$	(1)
17α-hydroxy-4-pregnene-3,20-dione	$<2.0x10^5$	(2)
17α-hydroxy-5-pregnene-3-one-20-ol	$<2.0x10^5$	(2)
20α-hydroxy-4-pregnene-3-one	$<2.0x10^5$	(2)
Deoxycorticosterone	$<2.0x10^5$	(2)
	$5x10^4$	(1)
Estrone	$<2.0x10^5$	(2)
	$7x10^4$	(1)
Cortisol	$<2.0x10^5$	(2)
Testosterone	$<2.0x10^5$	(2)
	$1x10^5$	(1)
4-androstene-3,17-dione	$<2.0x10^5$	(2)

a- for all the ligands there are four binding sites per tetramer of protein. (1) Lea, (1988);(2) Dilley, *et al.*, (1990)

components of the gross-cystic-disease, GCDFP-44 and GCDFP-15 (Haagensen, *et al.*, 1992) in T47D breast cancer cell line. The expression of these proteins and in particular of apo D is closely controled by steroid hormones, and it seems to be connected to the growth of the cells that excrete them.

It is also noted that a report (Albers, *et al.*, 1984) on the effects of the anabolic steroid stanozolol on the proteins of the lipoprotein system LCAT, apo B, Lp(a) (Lipoprotein (a)) and in particular apo D, finds that there is a decrease in the *in vivo* concentrations of all the proteins except for apo B.

As a note, expression of apo D can be detected in normal human diploid fibroblast cultures as well (Provost, *et al.*, 1991).The expression was dependent on the growth state of the culture and it was concluded that it is probably a growth arrest consequence.

The importance of apo D for the modulation of growth is a relevant factor raised by these experiments but that has not yet, to our knowledge, been explained. Nevertheless the relevance of apo D as a biochemical marker of cell development (growth/expansion) in general and tumour cells in particular is, once again, emphasized by these findings.

1.3.4.5 Lipocalin member

The determination of the sequence (Drayna, *et al.*, 1986) and the recognition of apo D as a member of the lipocalin family has revealed the probable role of the protein in binding small hydrophobic molecules.

A three-dimensional model of apo D based on the structure of insecticyanin and bilin-binding protein (Bbp) (Peitsch and Boguski, 1990) has led to the prediction that apo D binds the same sort of ligands as insecticyanin and Bbp. In fact the ligand-binding experiments described in this article indicate that apo D does bind bilirubin. Most interesting was the fact that the affinity for cholesterol was shown to be extremely low. This is supported by analysing the modelled "binding pocket" where there seems to be enough space for two molecules of cholesterol and insufficient for cholesteryl esters due to their elongated form. It is proposed as well that the binding of the protein to the HDL is favoured by the hydrophobic loop at the entrance to the binding cavity.

1.3.4.6 Summary

In the previous sections the properties and characteristics of apo D have been described. They can be summarized in the following way:

a) in the blood - involvement of apo D with the lipoprotein system, which has the function of transporting lipids

b) in the neural system - involved in the regeneration process

	parallelism with apo E
c) in breast cysts -	binding of some steroids in wich the cysts are particular rich
	no binding of cholesterol
d) in tumor cells -	probable biological marker
	steroids regulate its expression
	involvement in growth modulation ??
e) as a lipocalin member -	a small hydrophobic molecule carrier

So it is possible to conclude that apo D is probably involved in the control of or is at least part of growth/regeneration processes. It probably fulfils this role by either removing unecessary or inhibiting metabolites or by bringing metabolites essential for these processes.

1.4 Thesis rationale

The work described in this thesis involves the examination of different structural aspects of apo D and Blg.

The structure of Blg presents many problems, as mentioned by Papiz, *et al.*, (1986), and in particular there were several aspects of the trigonal crystal form, BlgZ, that did not seem to agree with previous work (see section 1.2). It was our purpose to obtain a better X-ray structure and at the same time to try to obtain proof of the importance of the "pocket" for the binding of the ligands.

For apo D the questions were of a more general nature with the aim of finding promising leads based on the intention of bringing the different known aspects of the protein, as GCDFP-24 and as apo D, to some sort of unity. The questions asked were of the sort: how is it that a globular protein is involved with and binds to lipoproteins? what ligands does apo D bind when associated to HDL and are they relevant in any of the different situations where apo D is present? what is the three-dimensional structure of apo D and what are the structural characteristics common to other lipocalins?

Chapter 2 Materials and methods

2. Materials and Methods

This chapter will describe first, the methods applied to apo D: purification, antibody techniques and general biochemical techniques, deglycosylation procedures, CD studies, aggregation experiments, studies of the cysteine residues, lipid interaction experiments, ligand binding studies and crystallization attempts. Followed by the methods used with Blg: crystallization conditions, heavy atom soaking, crystallographic data collection and programmes used.

2.1 HDL preparation

The rationale of the HDL preparation is based on the different densities of the various plasma components and on the use of ultra-centrifugation flotation for their separation. The procedure described (Dr. H.S. Simpson - personal communication) is similar in principle to others reported in the literature (McConathy and Alaupovic, 1973; Kostner and Alaupovic, 1972).

Whole, out-dated blood (around 30 days after collection), obtained from the Scottish National Blood Transfusion Service, Royal Infirmary, Edinburgh was centrifuged at 3000 rpm (Beckman rotor JA-14), 10°C for 1h with the following additives; 0.01% EDTA , 0.5 g/l ascorbic acid, 1 mM PMSF and 0.01% thimerosal. The supernatant (plasma), which usually presented a red-brown colour from freeze-thawing haemolysis was removed and the pellet discarded. The density of the plasma, after adding PMSF to 100 µM and ascorbic acid to 0.5 g/l, was assumed to be 1.006 g/ml and adjusted to 1.063 g/ml by adding 0.0834 g of solid KBr per millilitre.

In centrifugation tubes of 60 ml volume, 50 ml of the prepared plasma were overlayed with 1.003 g/ml of KBr solution to a total volume of 60 ml. The tubes were spun for 22 h at 45000 rpm (Beckman rotor 45Ti) at 4°C. This resulted in a separation into three fractions; the yellow top layer contained all the low density plasma components (<=1.003 g/ml) except for the VLDL which formed a white skin on the wall of the tube. The bottom layer presented a gradient of colour from yellow to red, containing all the components denser than 1.003 g/ml, of which HDL are part.

The top layer was removed with care, using a thin Pasteur pipette, positioned against the wall of the tube at the liquid-air interface, drawing up the liquid without disturbing the layers. The white skin was subsequently removed from the wall with white tissue paper.The bottom layer was homogenized by gentle mixing with a wide-bore plastic pipette without disturbing the pellet of cell membrane fragments usually found at the bottom of the tube.

The additives PMSF and ascorbic acid were added to the bottom solution in concentrations of 100 µM and 0.5 g/l respectively and the pH was adjusted to >7 with

NaOH, to avoid precipitation.

The density was then adjusted by the following procedure:

a) The solution was weighed and the density determined.

b) The following formula was used to determine the amount of solid KBr added to attain a final density of 1.23 g/ml.

$$m = V_o (d_o - d) / 1 - d v$$

m - amount of KBr to be added (g)

V_o - total initial volume of the solution (ml)

d_o - initial density (g/ml)

d - final desired density (g/ml)

v - specific volume of the salt (for KBr =0.312 ml/g)

c) It was usually necessary to rectify the density further. This was done by adding small amounts of solid KBr until it was possible to have this solution overlayed by a 1.225 g/ml KBr solution.

The above setup was centrifuged, using the same conditions as before, inducing the HDL to float to the top of the tube. This layer was collected with a thin Pasteur pipette as described above for the low density components.

Ascorbic acid was added to 0.5 g/l and the pH was adjusted to >7 with NaOH. The solution was diluted to a volume of ~35 ml with a 1.225 g/ml KBr solution and the density adjusted to 1.23 g/ml as described before. After overlaying with a 1.225 g/ml KBr solution, the sample was spun for 16h at 30000 rpm (Beckman rotor 45Ti), 4°C. This last step has the purpose of further cleaning the HDL.

2.2.1 Apo D purification by hydroxylapatite chromatography from plasma

The procedures described are derived from the method presented by McConathy and Alaupovic, (1973). The method exploits the capacity of hydroxylapatite to retain any proteic component of the HDL other than apolipoprotein D (apo D). The modifications are basically a scale-up of the procedure and the removal of any steps that could provoke the denaturation of the protein. This was not a major concern of the original method as it had been conceived to characterize the protein biochemically - amino acid analysis, carbohydrate analysis, immune cross-reactions (McConathy and Alaupovic, 1973 and 1976) - while the structural characterization by X-ray crystallography requires larger quantities of native-state protein.

The HDL, isolated as described above, was dialysed extensively against 1 mM potassium hydrogen phosphate (K_2HPO_4), pH 8, 0.5 g/l ascorbic acid, 100 µM

PMSF, 0.01% thimerosal. The dialysate was concentrated approximately 5 fold on an Amicon concentrator using a YM10 ultrafiltration membrane (Amicon) and loaded into a 3.0x60 cm column containing hydroxylapatite (BioRad HTP) which had been extensively equilibrated with 1 mM K_2HPO_4, pH 8. An amount of 100-200 mg of total concentrated protein solution, evaluated by the Lowry method modified for the presence of lipids (section 2.6), were routinely loaded into the column at a flow rate of 40 ml/h. This procedure was performed at 4°C.

A non-bound fraction containing apo D was eluted with the equilibration buffer. The elution buffer was then substituted by 1 M K_2HPO_4, pH 8 and a second peak emerged which contained all the other components of HDL.

2.2.2 Apo D purification by hydroxylapatite chromatography from GCDF

The procedure for isolation of apo D from gross-cystic disease fluid is similar to that for the plasma (section 2.2.1), except that the raw fluid after dialysis against 1 mM K_2HPO_4, pH 8 is spun at 45000 rpm (Beckman rotor 50Ti), 4°C for 2 h before being loaded into 2.6x70 cm hydroxylapatite column and eluted at a flow rate of 40 ml/h.

2.3 Apo D purification from GCDF by gel filtration

The procedure followed was as described by Balbín, *et al.*, (1990).

Several batches of GCDF were spun for 1h at 30000 rpm (Beckman rotor 50Ti), 4°C to pellet cell debris. The supernatant was subsequently concentrated 5-fold with a Centricon-10 system (Amicon), at 4°C. Between 0.5 and 0.8 ml of the concentrate were loaded into a preparative gel filtration column (Beckman TSK300SWG 21.5x300 mm, precolumn Beckman Spherogel 7.5x75 mm), attached to a HPLC system (Waters), equilibrated and eluted at 0.2 ml/min with 0.1 M ammonium acetate, pH 6. The runs were performed at room temperature and the 1 ml fractions collected were analysed by SDS-PAGE (section 2.4).

2.4 SDS-PAGE

The gel electrophoresis was performed according to the method of Laemmli (Laemmli, 1970), usually at a concentration of 12.5% acrylamide. The mini-gel system from BioRad was routinely used.

The gels were routinely stained with coomassie blue.

2.5 Methods for evaluating the protein concentration

For evaluating the concentration of apo D in solutions of pure protein, the absorption at 280 nm was determined. Both the theoretical molar absorption coefficient of 32320 $M^{-1}cm^{-1}$, calculated according to (Gill & von Hippel, 1989), and the molar mass of 19400 g (the carbohydrate components ignored) were used.

For Blg the values used were a molar absortion coefficient at 280 nm of 17600 $M^{-1}cm^{-1}$ (Dufour and Haertlé, 1990), and a molar mass of 18000 g.

For the determination of protein content in the HDL or protein concentration in the presence of lipid, a modification of the Lowry method was used. The modifications from the standard method (Peterson, 1979) require the presence of 1% SDS in the solution A, the incubation time after adding the Folin-Ciocalteau reagent is of 45 min not of 30 min, and the measurement of the absorption is done at 660 nm not at 750 nm. A standard curve was prepared with bovine serum albumin or with apo D, when required.

2.6 Delipidation of the HDL

The delipidation of the HDL (Osborne,Jr., 1986) was performed by dropwise addition of 1 ml of methanol followed by 2 ml of diethylether to 250 µl of lipoprotein solution while vortexing. The mixture was allowed to stand in ice for 10 min. This was followed by centrifugation for 4 min at 1000 rpm in a bench-top centrifuge to remove the supernatant. Diethylether (3 ml) was added while vortexing the slurry, followed by a new centrifugation period and discarding the supernatant. The procedure was repeated twice after which the remaining organic solvent was removed under a stream of nitrogen.

2.7 Polyclonal antibody preparation

The protein was prepared (McConathy and Alaupovic, 1986) by extensive delipidation of HDL, as described in section 2.6, followed by resolubilization with 8 M urea in 1 mM K_2HPO_4, pH 8. The solution was then diluted to 2 M urea in the same buffer and loaded directly onto the hydroxylapatite column using the conditions described in 2.2.1.

A 5 mg sample of apo D obtained in this way was further purified by several electrophoresis runs on 12.5% SDS-PAGE. The gels (17x14 cm) were stained with regular Coomassie stainer for 20 min and destained for 20 min so as to detect the transversal band containing apo D. This band was cut and reduced to smaller pieces before being frozen at -20°C.

After several gels had been run and the apo D bands cut and pooled, the protein was electroeluted from the gels. The electroelution was achieved by placing the gel pieces in small wells with one open end and the other (in the direction of the elution) closed with dialysis tubing, filling the well and the two tanks with electrophoresis buffer and applying 125V for 5 hours.

The prepared apo D was dialysed against water and 100 µg (in 300 µl) were mixed with 200 µl of of sodium phosphate buffer (10 mM sodium phosphate, 9 g/l NaCl, pH 7.4) and 500 µl of Freund's adjuvant. This mixture was injected intracutaneously in the back quarters of a rabbit (not used previously for antibody raising). Before injection some blood was removed and used to produce pre-immune serum. A boost of 50 µg was administered 10 days later and then every 15 days for two months or until the amount of anti-apo D serum collected was considered to be enough. The animal was not sacrified.

2.8 Western blots

The electrophoresis gels were blotted in a semi-dry system (LKB semi dry blotting system) into nitrocellulose. Six pieces of 3MM paper and one piece of nitrocellulose were cut to the same size as the gel to be blotted. Two of the pieces of paper were soaked in anode buffer 1 (0.3 M Tris.HCl (pH 10.4), 20% (v/v) methanol, 0.1% SDS) and placed, one on top of the other, on the anode plate. A piece of paper soaked in anode buffer 2 (25 mM Tris.HCl (pH 10.4), 20% (v/v) methanol, 0.1% (w/v) SDS) was placed on top of these, followed by the nitrocellulose which had been soaked in water and then the gel which had been soaked in cathode buffer (25 mM Tris.HCl (pH 9.4), 20% (v/v) methanol, 0.1% (w/v) SDS, 40mM 6-amino-n-hexanoic acid). The remaining 3 pieces of paper were soaked in cathode buffer and layered on the top of the gel. The cathode plate was placed on top of the stack and transfer was achieved by applying a current of 0.8 mA per cm^2 gel area.

Depending on the procedure to be used for the development of the Western blot, two different methods, described in the following sections, were applied for the reaction with the antibodies.

TBS buffer (1.211 g/l Tris, 9 g/l NaCl, pH 7.4) was routinely used for the western blots.

2.8.1 Development with 4-chloro-1-naphthol

The membrane was blocked with TBS plus 0.5% Tween-20 for 2 h at room temperature or overnight at 4°C. Incubation with the anti-serum was done at a dilution of 1/100 in TBS for 2 h of gentle agitation at room temperature or overnight at 4°C. The

membrane was washed for three periods of 15 min with TBS at room temperature.

Incubation with goat anti-rabbit second antibody coupled to horseradish peroxidase (BioRad) was performed at room temperature with gentle agitation for 1 h. This was followed by three washings of 20 min each. The development was done by incubating the membrane with 5 ml of 3 mg/ml 4-chloro-1-naphthol in methanol, 25 ml of TBS and 8 µl of 30% hydrogen peroxide for a few minutes, the reaction was stopped with water when it was judged appropriate.

2.8.2 Development with the ECL system (Amersham)

The ECL development system (Amersham) was used when permanent records of the western blots were required.

The blocking of the membrane was done by incubation with TBS plus 0.5% Tween-20 and 5% dried skimmed milk (Marvel) for 2 h at room temperature. The incubation with the anti-serum was done at a 1/1000 dilution in TBS, 0.5% Tween-20, 5% Marvel at room temperature for 1 h. This was followed by washing once with TBS, 0.5% Tween-20 for 30 min plus three periods of 5 min. The membrane was then incubated for 20 min with a dilution of 1/5000 of the goat anti-rabbit second antibody, coupled to horse-radish peroxidase (BioRad) in TBS, 0.5% Tween-20, 5% Marvel. The final washes were done with one period of 20 min and 3 periods of 5 min with TBS, 0.5% Tween. The development was performed according to the manufacturers instructions.

2.9 Deglycosylation of apo D

The attempts to remove sugars from apo D were split in two branches, the removal of neuraminic acid, which is easily available to enzymes, and the total removal of the sugar components, which is not allways attained without denaturation ofthe protein.

2.9.1 Digestion of sialic acid

The digestion of the sialic acid present on apo D was done with *Vibrio cholerae* neuraminidase, acylneuraminyl hydrolase (EC 3.2.1.18), (Boehringer) as described in (Walsh, *et al.*, 1990).

An amount of 0.5 mg of apo D in 20 µl of water was diluted to 40 µl with buffer containing 100 mM sodium acetate, 14 mM calcium chloride, 0.04% sodium azide, 0.2 µM TPCK (tosyl-L-phenylalanine chloromethylketone), pH 5.5. Approximately 0.8 mU of enzyme were added to the solution and incubated at 37°C,

for 24 h. After 18 h a fresh aliquot of enzyme was added.

2.9.1.1 Evaluating the sialic acid digestion - Warren's assay

Warren's assay was used to monitor the extent of sialic acid digestion. This method, as described by Chaplin, (1986) detects free sialic acid with a sensitivity of 3-300 μM.

The protein was dialysed to eliminate any existing free sialic acid. The protein concentration was determined by aborption at 280 nm and the required amount for the assay was chemically desialated by prior incubation in 0.05 M sulphuric acid at 100°C for 1 h, freeze-dried and then resuspended in 80 μl of water.

The composition of the necessary reagents is as follows;

Cyclohexanone (BDH)

Reagent A - 4.278 g of sodium metaperiodate (BDH) were dissolved in 4 ml of water. Orthophosphoric acid (BDH) (58 ml) was added and the solution made up to 100 ml with water.

Reagent B - 10 g of sodium arsenite (Sigma) and 7.1 g of sodium sulphate (BDH) were dissolved in 0.1 M sulphuric acid (BDH), to a final volume of 100 ml.

Reagent C - Thiobarbituric acid (Aldrich) (1.2 g) and sodium sulphate (BDH) (14.2 g) were dissolved in water to a final volume of 200 ml. This solution is stable for a few weeks only.

To 80 μl of protein solution, 40 μl of reagent A were added and mixed well. The mixture was left at room temperature for 20 min followed by the addition of 400 μl of reagent B while vortexing. After incubating another 5 min at room temperature, 1.2 ml of reagent C were added and vortexed. The vial was capped and put into a boiling bath for 15 min. The solution was rapidly cooled to room temperature and the chromophore was extracted with 1ml of cyclohexanone with vigourous shaking. The layers were separated by centrifugation and the absorbance of the top layer was measured at 549 nm. The extinction coefficient of sialic acid (in the assay conditions) is 57000 $M^{-1}cm^{-1}$. In the blank, water replaces the protein solution.

2.9.1.2 Extraction of neuraminidase

Neuraminidase was eliminated from the protein solution by the use of affinity chromatography according to the following protocol (Cuatrecasas and Illiano, 1971).

N-(p-aminophenyl) oxamic acid matrix (Sigma) (2 ml) was equilibrated with 0.05 M sodium acetate, 0.2 mM EDTA, 2 mM calcium chloride, pH 5.5. The protein mixture was applied to the column in the digestion buffer (section 2.9.1.1).

Neuraminidase was retained in the column and the rest of the protein components were eluted with the equilibration buffer. Neuraminidase was removed from the column with 0.1 M sodium hydrogen carbonate, pH 9.1.

2.9.2 Deglycosylation with endo F / GPase F

The enzyme cocktail contains endo-ß-N-acetylglucosaminidase F (EC 3.2.1.96) and peptide N-glycohydrolase F (EC 3.2.2.18) from *Flavobacterium meningosepticum* (Sigma). The different procedures attempted were based on the method described in Biochemica Information, Boehringer Mannheim.

Apo D (25 µg) in 50 µl of 0.25 M sodium acetate, 20 mM EDTA, pH 6.5 was incubated with 0.10 U of enzyme for 24 h at 37°C. Various additives such as 5 mg/ml n-octylglucoside, 10-20 mM mercaptoethanol, 1-10 mg/ml SDS, 5-6 M guanidinium chloride, 5 mg/ml CHAPS were included in the buffer, either before or during the reaction in order to optimize deglycosylation conditions.
The extent of deglycosylation was evaluated by SDS-PAGE.

2.9.3 Chemical deglycosylation

The method used is that described in Beeley, (1985); Edge, *et al.*, (1981) and is recommended for removal of Ser/Thr-linked sugar units and all but the sugar linked directly to the protein in Asn-liked chains.

Warning: The reagent trifluoromethane sulphonic acid (TFMS) is an extremely corrosive volatile acid and should be handled with care and kept in ice all times during handling (storage at -20°C in a vial with teflon-lined cap).

Anisole (Sigma) (1ml) was mixed with 2 ml of TFMS (Sigma) in a 5 ml Reactivial (Pierce) and cooled to 0°C. Thoroughly dried protein (5-10 mg), was dissolved in 1 ml of the above mixture in the reactivial and nitrogen was bubbled through for 30 seconds at 0°C. The mixture was then incubated for at least 1 h at different temperatures. The protein was recovered by the dropwise addition of 2 volumes of diethylether (precooled to -40°C), followed by 3 ml of ice-cold 50% (vol/vol) water / pyridine. The precipitate formed was redissolved by vortexing and the ether phase was later discarded.The extraction with ether was repeated 3-4 times. The aqueous phase was dialysed against 2 mM pyridine / acetic acid, pH ~5.5 overnight.

The protein precipitated during dialysis probably due to the fact that the pH used is close to the pI of apo D. To redissolve the protein, a subsequent dialysis against 10 mM potassium phosphate, pH 7.3 followed. The samples were analysed by SDS-PAGE.

2.10 Circular dichroism experiments

The circular dichroism experiments were kindly done at the Scottish CD facility, Stirling University, by S. Kelly in the laboratory of Dr. N. Price. A Jasco J-6000 spectropolarimeter was used and all the experiments were done at 21°C. The buffers used (all at a 10 mM concentration) were the following: cacodylate, pH 5.5; sodium acetate, pH 4.5; sodium phosphate, pH 6.5 and pH 7.5; potassium hydrogen phosphate, pH 7.0 and Tris, pH 8.5. All experiments were corrected for protein concentration (evaluated by absorbance or modified Lowry (section 2.5) at an optimum value of 0.3 mg/ml for apo D) and path length. The CONTIN procedure (Provencher and Glöckner, 1981) was applied to the far UV spectra to evaluate the secondary structure content.

2.11 Apo D - apo AI aggregation experiments

Apo AI (Sigma) was resuspended to 0.5 mg/ml in 0.2 M Tris, pH 8.4, 1 mM PMSF, 0.01mM TPCK, 1% ethanol and stored at 4°C.

The required aliquot of apo AI was diluted to 0.2 mg/ml with each of the following three buffers; 0.2 M Tris, pH 7.4, and 0.1 M citric acid / sodium hydrogen phosphate at pH 7.5 and 5.5. The same procedure was used to prepare samples of apo D. The solution of apo D and apo AI were combined at a ratio of 1:1 (v:v). The pH 7.4 sample was incubated for a couple of hours at room temperature while pH 7.5 and pH 5.5 solutions were incubated at 4°C. The latter mixtures were incubated in the presence of 3 M guanidinium chloride (GnCl) for 4 h and then dialysed against decreasing concentrations of GnCl to 0 M in 0.5 M decrements for 4 h each. The samples were analysed in the final incubation buffer by gel filtration as described in 2.12.

2.12 Gel filtration experiments

Analytical gel filtration was used to analyse the mobility of apo D under different conditions. For most cases a G3000 SWXL (7.8x300 mm) with SWXL guard column was used, attached to a Gilson FPLC system. For the apo D / lipid vesicle association experiments, described in section 2.11, a G4000 SWXL columnn (7.8x300 mm) was used. The absorption at 280 nm was registered.

2.13.1 Ellman's assay for thiols

This assay measures the release of a chromophore nitrothiobenzoate (NTB) upon the reaction of a thiol with 5,5-dithiobis-2-nitrobenzoic acid (DTNB) (Sigma).

The protein was dissolved in 0.1 M potassium phosphate, 1 mM EDTA, pH 7.3 with or without 6 M guanidinium chloride (GnCl) at a minimum free thiol concentration of 2 μM (Creighton, 1989). After zeroing the system, 50 μl of 3 mM DTNB (0.1 M potassium phosphate, pH 7.3) were added to 1 ml of sample and blank. Absorption measurements were made at 412 nm. The molar extinction coefficient of NTB is 13700 $M^{-1}cm^{-1}$ in 6 M GnCl and 14150 $M^{-1}cm^{-1}$ in its absence.

2.13.2 Disulphide analysis

The photometric determination of the number of free thiols and disulphide bonds present in apo D was achieved by the method described by Thannhauser, *et al.*, (1984; 1987).

The reagent 2-nitro-5-thiosulphobenzoate (NTSB), was prepared by dissolving 100 mg of Ellman's reagent, 5,5-dithiobis-2-nitrobenzoic acid (Sigma), in 10 ml of 1 M sodium sulphite. The pH was adjusted to 7.5. This bright red solution was warmed to 38°C and oxygen bubbled through, till the colour changed to pale yellow. The solution is stable for 1 year at -20°C.

The assay was performed by adding at least 8 nmol of protein to 3 ml of 1:100 dilution of NTSB stock solution in 50 mM glycine, 100 mM sodium sulphite, 3 mM EDTA, with or without 2 M guanidinium thiocyanate, and the pH was adjusted to 9.5. The reaction mixture was incubated for 25 min in the dark. After which the absorbance at 412 nm was recorded. The molar extinction coefficient of the chromophore in the presence of guanidinium thiocyanate is 13900 $M^{-1}cm^{-1}$ and in its absence it was considered to be 14150 $M^{-1}cm^{-1}$. The stoichiometry of the reaction is such that one mole of chromophore is produced per mole of disulphide bond or mole of free thiol.

2.14 Protein / Lipid interaction studies

The lipid vesicles (with and without protein) were prepared by a slight modification of the procedure described in Sizer, *et al.*, 1987.

Typically, 3.8 mg (5.2×10^{-6} mol) of dipalmitoyl phosphatidylcholine (DPPC) (Sigma) were dissolved in 380 μl of chloroform. A stock solution of cholesterol (Sigma) was prepared at 10 mg/ml in chloroform. The protein solution was obtained by dissolving 2.6×10^{-8} mol of apo D in 400 μl of 17 mM n-octylglucoside (Sigma), 10 mM sodium hydrogen phosphate, 0.1% sodium azide, 1 mM ascorbic acid, pH 7.5.

The lipid vesicles were prepared by mixing the DPPC and cholesterol solution

to a final molar ratio of 150:7.5 before evaporating the chloroform in a glass vial under a nitrogen stream. This was followed by freeze-drying for 1 h. The film formed was resuspended in 34 mM n-octylglucoside, 10 mM sodium hydrogen phosphate, 0.1% sodium azide, 1 mM ascorbic acid, pH 7.5. Usually some aggregates remained. This preparation was sonicated for 10 short bursts in an ice-bath while the protein-detergent solution was sonicated for five short bursts under the same conditions (ice bath and glass vial for rapid cooling). It was necessary to centrifuge the solutions after sonication to eliminate the froth which sonication typically produced. In experiments where protein was going to be incorporated, the sonicated suspensions were mixed together at a molar ratio of 1:200 (protein:DPPC) and 5 bursts of sonication in an ice bath were applied.

The samples, lipid vesicles with or without protein, were dialysed against 10 mM sodium hydrogen phosphate, 0.1% sodium azide, 1 mM ascorbic acid for 48 h (two to three changes of 5 l volumes of buffer). The vesicles were isolated by FPLC gel filtration as described in section 2.12.

2.15 Apo D/Triton X114 temperature-induced phase partitioning

Apo D and Blg, both at the concentrations of 100 µg/ml and 10 µg/ml concentration were incubated separately at 4°C for 15 min in 50 mM KH_2PO_4, pH 7.0, 1% v/v Triton X114 (Sánchez-Ferrer, *et al.*, 1990). 5 and 15 µl aliquots from the 100 and 10 µg/ml protein solutions respectively, were removed at the end of the incubation and kept at 4°C. The temperature was then raised up to 30°C for 30 min at the end of which the samples were centrifuged for 10 min at 11000 g (bench-top centrifuge). Again aliquots (5 and 15 µl) were removed from the supernatant. The different aliquots and a small amount of the pellet were run in a 12.5% SDS-PAGE.

2.16 Electron microscopy

The buffer of the lipid vesicles (with or without protein) was changed to the following volatile buffer: 0.125 M ammonium acetate, 2.6 mM ammonium carbonate, 0.26 mM EDTA, pH 7.4. A 5 µl drop of 2-10x diluted vesicle preparation was laid on the microscope grid, previously coated with collodion and carbon. The adsorption was allowed to occur for 1 min and then the liquid was removed with filter paper. Staining was done by incubating a 10 µl drop of 2% PTA (phosphotungstic acid), pH 7.4 on the grid for 30 seconds. The grids were observed with a Phillips CM12 transmission electron microscope with the kind assistance of Mrs.S. Bury.

2.17 Fluorescence ligand-binding studies.

The ligand-binding experiments were performed in a Perkin-Elmer Luminescence Spectrometer L-50 at 25±0.1°C, controlled by a temperature controler Grant LTD6 at the laboratory of Dr.D. Dryden, University of Edinburgh. All the experiments were done in 50 mM potassium dihydrogen phosphate, pH 7.0 and the ethanol used for dissolving the ligands was of chromatographic grade (Rathburn). The slits were set at 5nm bandwidth, the sample was excited at 280 nm and the emission was recorded at 340 nm (the maximum wavelength of emission for the protein under these conditions).

To correct for the inner filter effect (Birdsall, *et al.*, 1983), the titration of N-acetyl-tryptophanamide (Sigma) with ligands and ethanol was measured under the same conditions as the protein. Specifically, it is important that the absorbance at the excitation wavelength will be the same as that of the protein.

The protein (in a weakly buffered stock solution) was diluted to a total volume of 1 ml with the buffer described above to attain a concentration of 2.0×10^{-6} M. This was centrifuged for 10 minutes in a Eppendorf centrifuge an absorbance spectrum was recorded between 240 nm and 350 nm on 950 µl of the solution. The protein concentration was determined by subtracting the baseline absorbance measured between 300 and 350 nm (where no absorbance due to the protein is detected) from the maximum absorbance at 280 nm and using the molar extinction coefficient indicated in section 2.6. An aliquot of 900 µl of this solution was then put into the quartz fluorimeter cuvette and 100 µl of buffer were added. The solution was allowed to equilibrate at the experimental temperature.

Ligands were added to 1ml of protein solution, at a concentration of 100 µM in ethanol, using a 10 µl Hamilton syringe. The ligand was equilibrated with the protein in 2 µl additions up to 12 µl and then in 4 µl additions up to a final added volume of 40 µl. The solution was mixed thoroughly after each addition with a 200 µl automatic pipette. The readings were normally stable after the mixing and at least 3 fluorescence values were registered; more if the reading was considered noisy. At least two assays per ligand were carried out and when the ligand appeared to bind, three were registered.

A base line for the effect of the ethanol on the fluorescence of apo D was produced by titrating the protein with ethanol, and the emission spectra were recorded at different stages of the titration.

In each assay the values were averaged and rounded to the nearest unit. The fluorescence difference between the initial and each point was calculated and corrected for concentration, so that all the measurements were extrapolated to a protein concentration of 2.0×10^{-6} M. The corrected values were averaged between assays for the same ligand and finally subtracted from the ethanol baseline (corrected and averaged

as per the ligand titrations). These values were then corrected if the N-acetyl-tryptophanamide titrations showed significant inner filter effect.

The ligands used were the following: progesterone (4-pregene-3,20-dione), cholesterol (5-cholesten-3β-ol) , prostaglandin E1 ((11α,13E,15S)-11,15-dihydroxy-9-oxoprosta-13-enoic acid), prostaglandin F2α ((5Z,9α,11α,13E,15S)-9,11,15-trihydroxyprosta-5,15-dienoic acid), prostaglandin D2 ((5Z,9α,13E,15S)-9,15-dihydroxy-11-oxoprosta-5,13-dien-1-oic acid), arachidonic acid (eicosa-5Z,8Z,11Z,14Z-tetraenoic acid), palmitic acid (hexadecanoic acid), DPPC (L-α-dipalmitoyl phosphatidyl choline) all from Sigma.

12-Hete (12-hydroxy-[S-(E,Z,Z,Z)]-5,8,10,14-eicosatetraenoic acid), 5,15-diHete (5,15-dihydroxy-[S-[R*,R*-(E,E,Z,Z)]]-6,8,11,13-eicosatetraenoic acid, leukotriene D4 ((5S,6R,7E,9E,11Z,14Z)-5-hydroxy-6-(S-cystainylglycinyl) eicosatetraenoic acid), linoleic acid(octadeca-9Z,8Z,11Z,14Z-tetraenoic acid), oleic acid (octadeca-9Z-enoic acid) from Cascade Biochem, Reading.

EP092 ((+5)-endo-(6'-carboxyhex-2'Z-enyl)-{1-[(N-phenylthiocarbamoyl) hydrazono]-ethyl} bicyclo [2.2.1] heptane was provided by Drs Wilson, Dawson and Jones from the Department of Pharmacology, University of Edinburgh.
Other ligands, in particular cholesteryl esters, which we wished to test could not be used because of their very low solubility in ethanol, DMSO or DMF.

The analysis of the binding parameters was kindly done by Dr.G. Atkins, University of Edinburgh.

2.18 Crystallization conditions for apo D

It was known from the beginning that the crystallization of apo D would not be without problems since its heterogeneity was well described in the literature. In particular, the high level of glycosylation (see section 2.9) induces observable heterogeneous behaviour in SDS-PAGE (diffuse band) and isoelectrofocusing (high number of bands) (Kamboh, *et al.*, 1989). It was shown as well, that the removal of the neuraminic acid would reduce that heterogeneity (Kamboh, *et al.*, 1989), though not totally. Despite all this, it was decided to try to crystallize the native protein isolated from gross-cystic disease fluid and later the neuraminic acid-free protein.

Two approaches were used, the "sparse matrix" described by Jancarik and Kim, (1991) and the "incomplete factorial experiment" described by Carter,Jr. and Carter, (1979); Carter,Jr., *et al.*, (1988); Betts, *et al.*, (1989).

The "sparse matrix " precipitant solutions were prepared as described in the article but two more solutions were added. These had numbers 51- and 52- in sequence with the reported ones.

51- 0.1 M cacodylate, 50 mM cadmium acetate, pH 6.5

52- 1 mM cadmium acetate, 0.1 M cacodylate, 30% PEG 6000, pH 6.5

These two solutions were prepared because some of the members of the superfamily have been crystallized (high quality crystals) in the presence of cadmium or even with cadmium as precipitant (Newcomer, *et al.*, 1984; Böcskei, *et al.*, 1991).

The "incomplete factorial experiment" procedure is well described in the literature referenced above.

2.19 Crystallization conditions for BlgZ

The conditions used were based on the ones described by Aschaffenburg, *et al.*, (1965); Bolognesi, *et al.*, (1979); Monaco,*et al.*, (1987).

Dialysis crystallization - pH 7.8 - the protein (Sigma) was dissolved in 0.2 M phosphate buffer (sodium phosphate/potassium phosphate 1:1 (mol:mol)), pH 7.8 at 30 mg/ml. It was put into a 50 µl dialysis button (Cambridge Repetition Engineers) and dialysed at 20°C against 20 ml of 0.2 M phosphate buffer, 1.5 M ammonium sulphate, pH 7.4. The buttons were checked regularly and every 3-4 days the concentration of ammonium sulphate was increased by 0.2 M and the pH readjusted. This continued up to the concentration of 1.9 M ammonium sulphate when the increases were reduced to 0.1 M at a time, up to around 2.1 M when crystals would start to appear.

Dialysis crystallization - pH 8.5 - the protein (Sigma) was dissolved in 0.1 M Tris, pH 8.5 at 30 mg/ml. The button (50 µl) was dialysed against 0.1 M Tris, 1 M sodium citrate, 0.20 M, pH 8.5 at 20°C. The concentration of ammonium sulphate was increased by 0.2 M and pH was controlled. At around a concentration of 0.27 M crystals would start to appear.

Sitting drop - pH 7.8 - To 50 µl of a solution of Blg (30 mg/ml) in 0.2 M phosphate buffer (sodium phosphate/potassium phosphate 1:1 (mol:mol)) pH 7.8, 10 µl of ammonium sulphate 2.17 M, 0.2 M phosphate buffer, pH 7.8 and 15 µl of 0.2 M phosphate buffer pH 7.8, were added. This mixture was equilibrated against 2.17 M ammonium sulphate, 0.2 M phosphate, pH 7.8. The crystals would take typically around a week to show up and 2-2.5 weeks to attain maximum size (0.3-1 mm on the long diagonal of the faces). The crystals would not last more than a month if left in these conditions, but they could be stabilized for two months if changed to 0.2 M phosphate buffer, 3 M ammonium sulphate , pH 7.8.

Note: the sitting drop technique was routinely used for the growth of BlgZ crystals

2.20 BlgZ co-crystallization experiments

The co-crystallization experiments were done both with the sitting-drop (Charles Supper Company) and dialysis button (Cambridge Repetition Engineers) conditions described above. Typically 3 mg of Blg were incubated with the ligand (1-10x the molar quantity of the protein - if it was an aqueous immiscible liquid, up to 1 ml of the ligand would be added to the incubation mixture while if it was an insoluble solid then an amount as much as 1000x the molar quantity of protein would be equilibrated with the incubation mixture) in 1ml of 0.2 M phosphate buffer (pH 7.8) at 4°C for 24 h in a slow rotating wheel. The mixture would then be concentrated to 30 mg/ml with a Centricon-10 microconcentrator (Amicon), after the immiscible or insoluble ligands, if present, had been separated.

In the initial sitting drop trials, the concentration of ammonium sulphate in the well solution would range from 12% to 60% saturation. The drop would be constituted from 1 volume of protein plus ligand and 1 volume of well solution.

The number of known Blg ligands is quite large, as described in section 1.2. For the co-crystallization experiments only a few were used, with the only criteria of being reasonably stable. Attempts were made with other molecules that are not known ligands but similar to known ones. The list is the following: heptane, 2-nonanone, iodobutane, palmitate, hemin, iodonitrobenzene, iodophenol, p-nitrophenol, ascorbic acid, n-octylglucoside.

2.21 Heavy atom soaking with Blg

All the heavy atom soaking experiments were performed with the heavy atom dissolved in 3 M ammonium sulphate, 0.2 M phosphate buffer (see section 2.19), pH 7.8 in Linbro plates (normally used for Elisas). There was a wealth of experiments done by Dr.D.W. Green and collaborators and it was possible to repeat or explore further the conditions used as we had access to their data.

The quality of the soak was evaluated by the traditional techniques (Blundell and Johnson, 1976) of observing the aspect of the crystal and in most cases by precession photograph.

Only the conditions for heavy atom soaks where a three-dimensional reflection set was collected and used for the determination of the structure phases are described:

MMA (monomercuric acetic acid) - 2 mM for 2 weeks (longer than that leads to an obvious deterioration of the crystals)

HgI_2.KI 1:10 (mol:mol) previously dissolved in water and then mixed together in the final buffer - 50 μM overnight

$K_2Pt(NO_2)_4$ - 3 mM for 4 days

2.22 Data collection and processing

The three-dimensional data sets were collected at the University of Glasgow (Chemistry Department) and Oxford (Laboratory of Molecular Biophysics) with the kind assistance of Dr.A. Freer, Prof.N. Isaacs and Dr.D. Stuart. In both places, a Xentronics area detector mounted on a rotating anode was used.

The BlgZ crystals present rhombohedral morphology with the c*-axis along the long diagonal, so they were mounted with this diagonal down the Lindemann tube (though due to their shape, the diagonal is never colinear with the Lindemann tube axis) which allows for a minimum of 60° rotation needed to obtain a complete data set. When the collection of Bijvoet pairs was a specific requirement of the data collection, then the crystals were mounted with the long diagonal perpendicular to the tube axis. This permits the collection of the Bijvoet pairs as close in time as possible or, best of all in the same frame; in this orientation a rotation of 180° is needed for the collection of a complete set of reflections with reasonable redundancy. The data collection was always done at room temperature.

The data were processed with the packages Xengen or XDS (Kabsch, 1988a; 1988b)

Some data were collected at the Chemistry Department, Edinburgh University using a Siemens-Stoe AED2 four-circle diffractometer installed on a sealed tube generator, fitted with a graphite monochromator. This system was used in the later stages of the project to evaluate the quality of the three-dimensional reflection sets from heavy atom soaking experiments.

2.23 Precession photography

To evaluate the quality of the crystals after heavy atom soaking or obtained from co-crystallization experiments, precession photography was performed with a Huber precession camera mounted either on a sealed tube or rotating anode generator, as available. The photographs were taken at room temperature, routinely for 24-36h with a 10° precession angle.

2.24 Programs used

Most of the crystallographic handling of the data was done with the CCP4 (LCF version) package (SERC, Daresbury Laboratory) installed in a microVax 3100. The molecular replacement, part of the crystallographic refinement and analysis of the models was performed with the the X-PLOR package (Brünger, A.T., 1992) installed in a ESV10. The display of the molecular models and electron-density maps was done

with O (Jones, *et al.*, 1991) also on an ESV10 workstation. The final model was also analysed with DSSP (Kabsch and Sander, 1983) and PROFILE (Lüthy, *et al.*, 1992).

Chapter 3 Crystallographic work with Blg

3. Crystallographic work with Blg

This chapter will deal with the crystallographic work in three general parts. The first (including sections 3.1 to 3.2) reports the data collection and analysis, the second (sections 3.3 to 3.7) describes the heavy atom phasing and model building. The final part (section 3.8) reports the molecular replacement work.

3.1 Data quality

In this section several aspects of data processing will be discussed in light of the procedures used and results achieved at different stages of the work.

3.1.1 Indexing problems

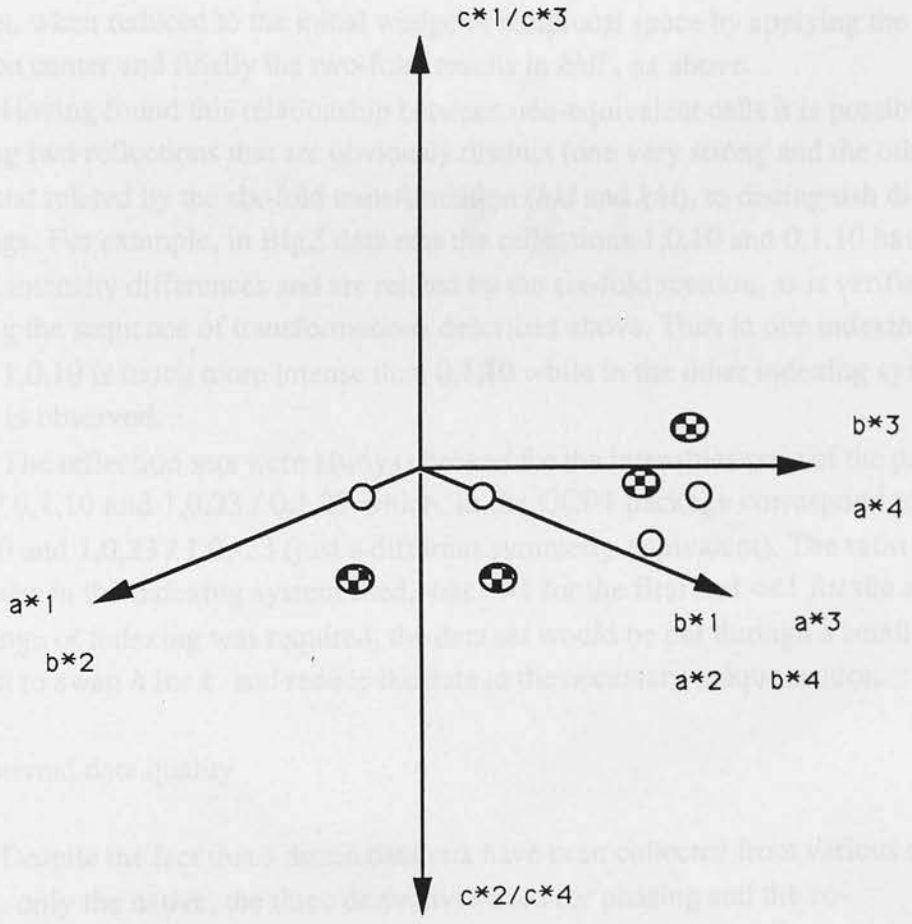
The processing of the data was done in space group $P3_221$, and there are aspects of this space group that deserve detailed attention, in particular, the possibility of processing reflections in two indexing systems. Both these systems are acceptable within the symmetry restrictions of the space group but the same one has to be used at all times.

The indexing problem arises from the symmetry of reciprocal space, which is described by Laue group $P\bar{3}m1$. This Laue group generates an hexagonal system where the hexagonal relationship arises from an improper three-fold rotation axis. This means that if a right handed axis system is to be defined in two sequential 60° of data or hexagonal cells, there are four ways to do so (1 to 4, as shown in FIG. 3.1), and these four cells are not equivalent. For example, if two batches corresponding to sequential portions of data, were being compared and one had been indexed with cell 1 and the other with cell 4, then equivalent reflections (related by a two-fold in this example) would be indexed with same *hkil* indices and they would be directly comparable. However if the indexings used were 1 and 3, then reflections with the same indices would not be equivalent and they should not be merged.

This "misindexing" is easily noticeable for batches of the same data set as the merging scale factors will be quite different from the internal scaling factors and the merging residual will be quite poor: for example the BlgZ native data was originally collected in two batches, each with internal residuals of around 6%. In spite of the fact that no changes in the diffraction pattern were observed from one batch to the other, the merging residual was 14%. This problem was solved by choosing a different orientation matrix for the second batch; the final indexing, related to the first by 180°, gave a merging residual of 6.3%. Further, when comparing different data sets, the isomorphous residual will be extremely high for the different indexing systems. For example the MMA derivative data set processed in an inconsistent indexing system to

FIG. 3.1 Four possible axial systems for Laue group $P\bar{3}m1$

Each axis sytem is defined by a number, 1 to 4. Represented with different filling patterns are two different reflections and their symmetry equivalents. In the same indexing system, the two different reflections have the same h and k but opposite sign for l.



the native data, had an isomorphous residual of 40% when compared to the latter, while in the same indexing system this value dropped to 25%.

In the cases described, the data were reprocessed in order to change the indexing. However, this is not necessary since the relationship between two equivalent reflections in different indexing systems can be described in the following way; the reflection that in cell 3, Fig. 3.2, is indexed as *hkil* corresponds in cell 1 to *-k-i-hl* or by the application of the inversion centre, the three-fold and two-fold symmetry operations, to *khil*. Thus equivalent reflections are indexed as *hki* in one system and *khi* in the other, which means that a simple change of *h* for *k* will convert one cell into the other. The same result is achieved by reasoning that, as the problem with the indexing is the application of an hexagonal system where there is no six-fold, it should be possible to convert from one cell to another non-equivalent, by the application of that same six-fold. This means that *hkil* becomes *-h-k-il* with the six-fold proper rotation and then, when reduced to the initial wedge of reciprocal space by applying the inversion center and finally the two-fold, results in *khil* , as above.

Having found this relationship between non-equivalent cells it is possible by choosing two reflections that are obviously distinct (one very strong and the other very weak) and related by the six-fold transformation (*hkl* and *khl*), to distinguish different indexings. For example, in BlgZ data sets the reflections 1,0,10 and 0,1,10 have very marked intensity differences and are related by the six-fold rotation, as is verifiable by applying the sequence of transformations described above. Thus in one indexing system 1,0,10 is much more intense than 0,1,10 while in the other indexing system the reverse is observed.

The reflection sets were always checked for the intensities ratio of the pairs 1,0,10 / 0,1,10 and 1,0,23 / 0,1,23 which, in the CCP4 package correspond to 1,0,10 / 1,0,-10 and 1,0,23 / 1,0,-23 (just a different symmetry equivalent). The ratio for these pairs in the indexing system used, was >>1 for the first and <<1 for the second. If a change of indexing was required, the data set would be put through a small program to swap *h* for *k* and reduce the data to the necessary unique section.

3.1.2 Internal data quality

Despite the fact that a dozen data sets have been collected from various different crystals, only the native, the three derivatives used for phasing and the co-crystallization with ascorbic acid will be analysed.

A series of internal statistics considered to be indicators of the precision of data, are presented in TABLES 3.1 to 3.4. Specifically, completeness (observed percentage of total number of reflections), average redundancy (average number of observations per unique reflection), percentage of unique reflections with intensity over sigma of intensity higher than 2 or 3 (both for intensities and for anomalous differences)

FIG. 3.2 Schematic representation of reciprocal space for Laue group $P\bar{3}m1$.

Viewed down the three-fold axis. Two possible indexing systems are represented, 1 and 3 (see FIG. 3.1 and text). Circles represent one reflection and its symmetry equivalents. The reflections have been labeled with the indexes corresponding to system 1.

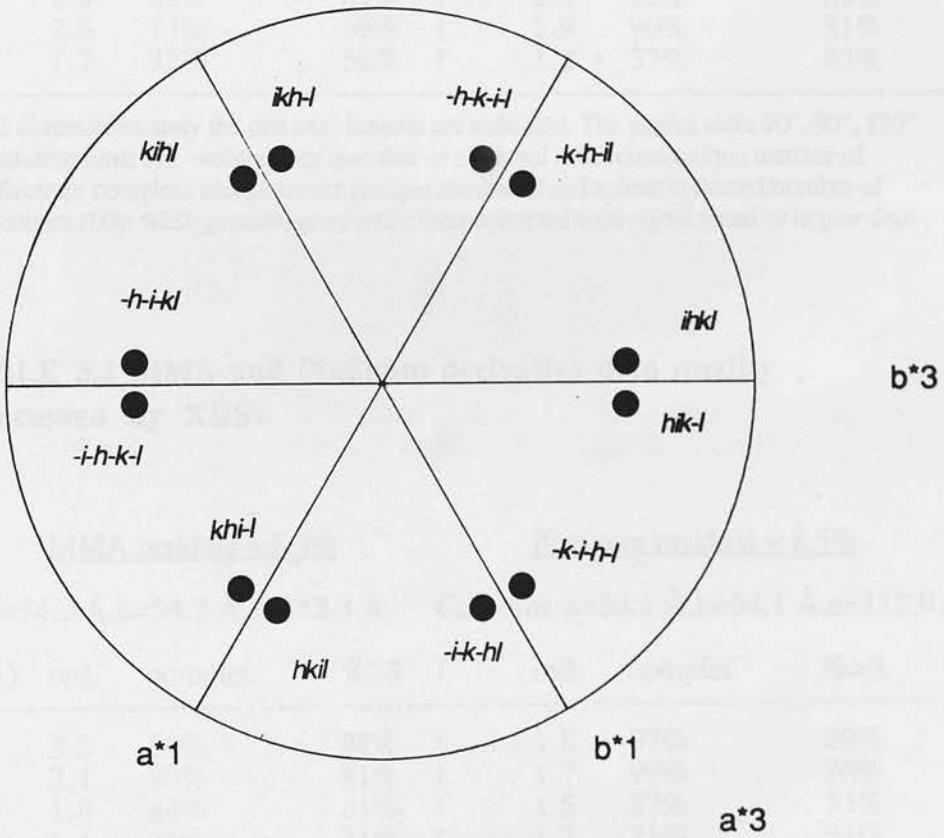


TABLE 3.1 Native and "ascorbic" data quality (processed by XENGEN)

	Native residual = 6.3%			"Ascorbic"residual = 4.3%		
Cell dim.	a=54.0 Å,b=54.0 Å,c=112.7 Å			Cell dim. a=54.3 Å,b=54.3 Å,c=113.3 Å		
Res.bins (Å)	red.	complet.	%≥2	red.	complet.	%≥2
∞ -5.1	2.8	96%	97%	4.4	99%	97%
5.1-4.0	2.7	91%	96%	1.9	99%	97%
4.0-3.5	2.9	83%	88%	1.9	97%	95%
3.5-3.2	2.8	81%	81%	2.0	95%	88%
3.2-3.0	2.6	73%	69%	1.9	90%	81%
3.0-2.8	1.7	32%	56%	1.2	37%	63%

Cell dim.- cell dimensions, only the cell axes lengths are indicated. The angles were 90°, 90°, 120°. Res.bins- Resolution bins; red.- redundancy (number of observed reflections/unique number of reflections collected); complet.- completeness (unique number of reflections collected/number of possible reflectionsx100); %≥2- percentage of reflections collected with signal equal or higher than 2xsigma

TABLE 3.2 MMA and Platinum derivative data quality (processed by XDS)

	MMA residual = 6.3%			Platinum residual = 8.5%		
Cell dim.	a=54.3 Å,b=54.3 Å,c=113.3 Å			Cell dim. a=54.1 Å,b=54.1 Å,c=112.9 Å		
Res.bins (Å)	red.	complet.	%>3	red.	complet.	%>3
∞ -6.3	3.2	84%	92%	1.8	97%	89%
6.3-4.5	2.1	90%	81%	1.7	99%	76%
4.5-3.7	1.8	84%	61%	1.5	87%	53%
3.7-3.2	1.1	25%	31%	1.2	71%	34%

Cell dim.- cell dimensions, only the cell axes lengths are indicated. The angles were 90°, 90°, 120°. Res.bins- Resolution bins; red.- redundancy (number of observed reflections/unique number of reflections collected); complet.- completeness (unique number of reflections collected/number of possible reflectionsx100); %>3- percentage of reflections collected with signal higher than 3xsigma

TABLE 3.3 HgI data quality (collected by diffractometer)

Cell dimensions a=54.2 Å, b=54.2 Å, c=113.1 Å, α=90°, β=90°, γ=120°

All data collected from 43 Å to 6.0 Å	606 reflections
Redundancy = 1	%>2 = 76%

Redundancy- number of observed reflections/unique number of reflections collected; %>2- percentage of reflections collected with signal higher than 3xsigma

TABLE 3.4 MMA anomalous data quality

Res.bins	complet.	%>2
∞ -6.6	58%	40%
6.6-4.5	65%	18%
4.5-3.7	70%	6.2%
3.7-3.2	56%	4.0%

Res.bins- Resolution bins; complet.- completeness (number of unique reflections collected with anomalous signal/number of unique reflections collectedx100); %>2- percentage of reflections with anomalous signal higher than 2xsigma

resolution and the symmetry residual (defined as $[\Sigma\Sigma(|I(h,i)-I(h)|)/\Sigma(I(h))]x100$ where I(h,i) is the scaled i^{th} observation of reflection h and I(h) is the merged final intensity of that reflection), were analysed as function of resolution.

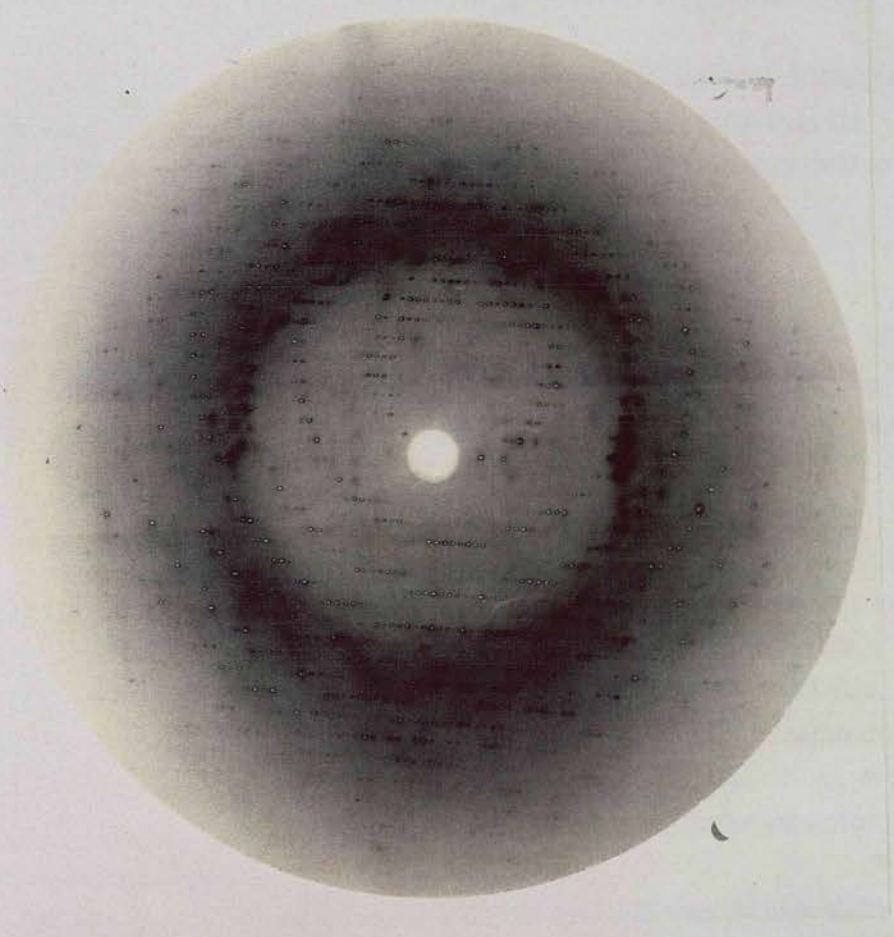
During the work with BlgZ several characteristics of this crystal form became apparent. One was the clear drop in intensity of the diffraction pattern at the higher angles and the almost absence of diffraction beyond the 3 Å resolution, as can be observed from the tables for the native and ascorbic acid co-crystals.

Another intrinsic problem in the crystal quality was the disorder observed. This is manifested by the wide shape of the diffraction spots and in the "stains" on photographs taken at the synchrotron source at Daresbury (FIG. 3.3), an effect attributed to thermal diffusion scattering.

The strategy for collection of good data by the rotation method depends on the variation of as many factors as possible (Evans, 1993). In effect this can be translated as the need for observation of each reflection and its symmetry equivalents as many times as possible during the collection process. Unfortunately, the data sets used for the

FIG. 3.3 Rotation photograph of BlgZ data collected in the SRS

An example of a 2.5° rotation photograph of BlgZ crystal, collected from station 7.2 at the SRS, Daresbury laboratory.



structure determination of BlgZ do not contain great amount of redundancy; the native data set has the best value for the average redundancy.

Both the MMA and the platinum derivatives exhibit a reduction in the resolution of diffraction and a general drop in the intensity of reflections. The deterioration of the derivative crystals may be explained by the low reactivity of the protein (under the crystallization conditions) with the heavy atom compounds and the need for long soaking times (see conditions in section 2.21). Despite the apparent low quality of the anomalous differences, since beyond the resolution of 4.5 Å only around 5% of the reflections have an amplitude higher than 2σ associated, they were essential for the phasing.

The data set collected on the diffractometer suffers from the required increase in time for the collection of reflections and it was only possible to observe diffraction to 6 Å resolution.

The native data set has been recollected at Daresbury SRS station 9.5 using a MAR image plate detector. Once processed, it is the intention to use these data for refinement of the structure as it will be a higher resolution, high redundancy data set.

3.2 Analysis of relative characteristics between data sets

Data sets from different crystals need to be put on the same scale for the extraction of structural and crystallographic information. This section will describe statistical aspects of the different comparisons done during the course of this work.

3.2.1 Scaling of data sets

The correct scaling of two data sets is of extreme importance. The production of phases involves comparison of native and heavy atom derivative structure factor amplitudes during the Patterson synthesis and phase triangle construction. As this comparison is done on an "one-to-one" amplitude basis, the importance of removing all the differences that are not due to changes in the crystal structure is clear. Blow, *et al.*, (1988) demonstrate that a change of 20% in the scaling factor affects the value for the phase calculated from a single derivative, by 180°.

The scaling and relative analysis of the different data sets was done initially with the program ANSC (CCP4 package), which calculated and applied an overall relative scaling and temperature factor to the derivative data set. Subsequently, it was considered that due to the rapid deterioration of the heavy atom derivatives during data collection and consequent introduction of systematic errors associated with low redundancy of those data sets, it would be advantageous to scale these reflection sets by local scaling as well. In reality, for the particular case of the HgI data set, which was

obtained from the diffractometer and thus collected and processed with a different strategy from the native data, it was considered advisable to apply a "finer" scaling procedure to try to account for the different systematic errors introduced. The local scaling procedure determines scaling factors by minimizing the differences between each data set, on a restricted number of neighbours around each particular reflection, and not the overall differences or the differences in resolution bins. If the number of reflections used for the scaling calculation is very small then any difference, including real differences, may be removed. The program used is called LOCAL and the number of reflections used in each zone was 25. The procedure was repeated until the residual stabilized.

The application and effect of the scaling was checked with plots of the scaling factors in resolution bins and in bins of Miller indices to monitor any systematic unaccounted behaviour. Nothing unusual was detected.

3.2.2 Analysis after scaling

ANSC performs an analysis of the differences between two data sets by determining the mean isomorphous difference and the mean fractional isomorphous difference as function of resolution. These analyses can give an estimate of the changes, and in particular of the non-isomorphous changes, which occur on adding the heavy atom compound to the crystal. The results after the the two scaling steps are presented in FIG. 3.4 and FIG. 3.5.

The overall mean isomorphous difference for MMA was 27%, for HgI 16%, for the platinum 21% and for the ascorbic data 26%. Both the MMA and platinum derivatives exhibit a peak at around 6 Å on the mean fractional isomorphous difference plot. This is characteristic of heavy atom substitution as it represents the substitution of the first solvent shell by a stronger scatterer. Better isomorphism is expected from the MMA than from the platinum derivative because the mean isomorphous difference decays more rapidly with resolution for the former and the mean fractional isomorphous difference does not increase as much. It should be pointed out that for a totally isomorphous derivative the mean isomorphous difference would drop with resolution due to the decrease in the scattering of the heavy atom at higher angles and to the increase in noise in the weaker high resolution reflections. The mean fractional difference should be stable with increasing resolution, indicating that the changes only result from the introduction of a very electron-dense atom and not from increased disorder at smaller resolution spacings.

The data for these two derivatives were truncated at 3.5 Å because the two indicators presented evidence of non-isomorphism beyond that. A cut-off of 30% for the mean fractional difference was introduced, and resulted in the removal of amplitudes beyond 3.5 Å for both derivative data sets. This in line with the 40%

FIG. 3.4 Mean isomorphous difference plot over resolution bins for the derivatives and"ascorbic" data sets

Mean isomorphous difference is defined as <|F-Fnat|>, where F and Fnat are the amplitudes of derivative and native data respectively. Ascorbic data (-•-•-), HgI data (——··—), MMA data (------), platinum data (— - —).

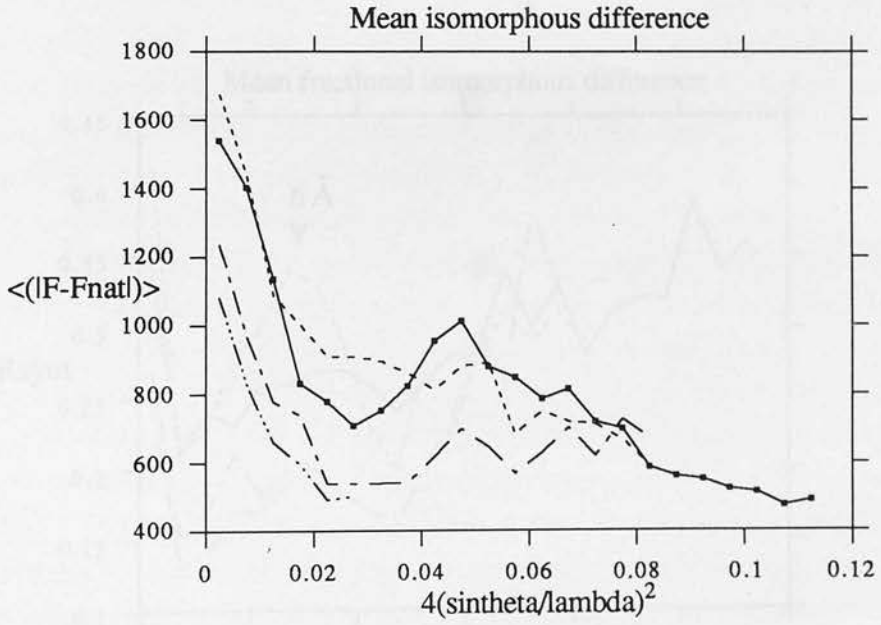
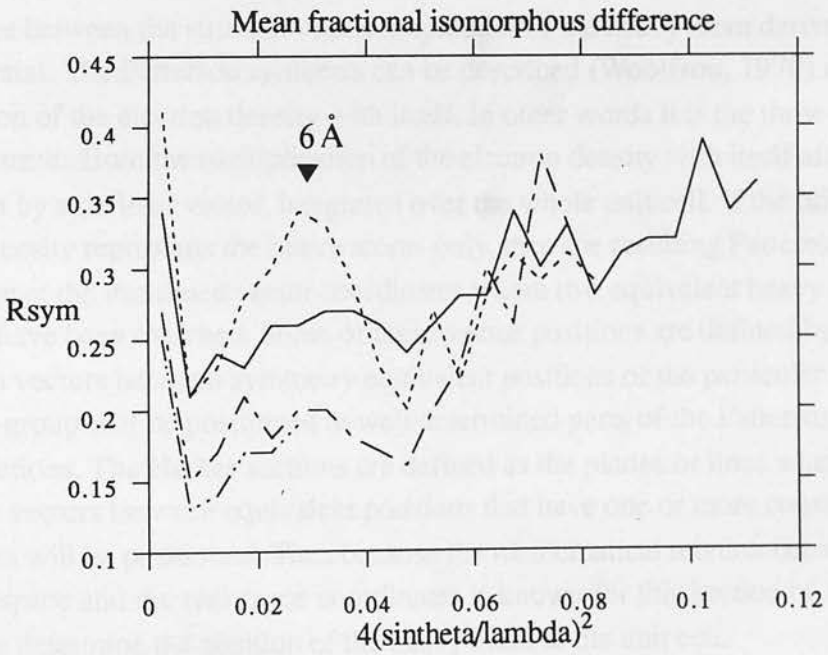


FIG. 3.5 Mean fractional isomorphous difference plot over resolution bins for the derivatives and "ascorbic" data sets

Mean fractional isomorphous difference or Rsym is defined as Σ(|F-Fnatl)/Σ(Fnat) where F and Fnat are the derivative and native amplitudes, respectively. Ascorbic data (——), HgI data (—···—), MMA data (------), platinum data (—-—).



theoretical change resulting from the totally isomorphous introduction of a mercury atom in a 24 kDa protein described by Blundell and Johnson, (1976).

The analysis of the low resolution derivative HgI does not reveal much except that modification of the protein is apparent.

The plots for ascorbic acid show that changes have occurred. A mean fractional isomorphous difference of 26% was unexpected for an experiment where the only anticipated modification was the binding of the ligand. However the changes are probably isomorphous since the mean diffference plot drops with resolution and the mean fractional plot increases slowly with resolution. This means that the ligand is bound to the molecule without provoking radical distortions.

3.3 Heavy atom positioning

The strategy for the determination of the heavy atom positions is usually based on solving one derivative, estimating protein phases from this derivative and then using these phases for difference Fourier syntheses to determine the positions of heavy atoms in other derivatives. The first step is usually achieved by Patterson synthesis of the differences between the structure factor amplitudes of the heavy atom derivative and native crystal. The Patterson synthesis can be described (Woolfson, 1970) as the convolution of the electron density with itself. In other words it is the three-dimensional map that results from the multiplication of the electron density with itself after translation by a defined vector, integrated over the whole unit cell. If the original electron density represents the heavy atoms only, then the resulting Patterson will be zero except at the translated vector coordinates where two equivalent heavy atom positions have been matched. Some of these vector positions are defined by the translation vectors between symmetry equivalent positions of the particular space group and a sub-group will be positioned in well determined parts of the Patterson space, the Harker sections. The Harker sections are defined as the planes or lines where the difference vectors between equivalent positions that have one or more constant coordinates will be positioned. Thus because the mathematical relation between the Patterson space and the real space coordinates is known for this section of space, it is possible to determine the position of the heavy atom in the unit cell.

The difference between the derivative and native structure factors amplitudes is a rough approximation to the heavy atom structure factors, in fact the Patterson synthesis with these terms will have contributions from the protein-protein and protein-heavy atom vectors (Blundell and Johnson, 1976). These extra contributions will be minor as long as there are enough terms and the heavy atom structure factors are small.

Difference Patterson syntheses were produced for the isomorphous differences of the three derivatives atoms and in the case of MMA, an anomalous difference Patterson was prepared too. The Harker sections w=1/3 of these maps are shown in

FIG. 3.6 to FIG. 3.9. The MMA isomorphous Patterson showed only the peaks due to the heavy atom site; the anomalous Patterson, although noisier, presented the same peaks as the isomorphous map demonstrating that the anomalous data would contribute to phasing. It is clear from the sections that the MMA and the HgI derivatives have been substituted in the same position. The platinum derivative produced a noisy map with a small peak confirming that it was weaker and probably more non-isomorphous than the other derivatives, as already expected from the ANSC analysis (see section 3.2.2). However, because the major peak position in the platinum derivative was different from any of the other derivative sites it would have an important contribution for the determination of phases.

3.3.1 Solution of the MMA Patterson function

The MMA derivative was automatically solved with the SHELX-86 package. This program determines atomic coordinates explaining all the Harker vectors present in a Patterson synthesis (Sheldrick, 1991). It produced one solution at the fractional coordinates x=0.855, y=0.482, z=0.238, equivalent to x=0.37, y=0.52, z=0.095.

The heavy atom parameters for MMA were refined with the program REFINE using FHLE refinement. FHLE refinement (Blundell and Johnson, 1976) minimizes the difference between the absolute values of the calculated heavy atom structure factor (a function of the heavy atom parameters) and the observed heavy atom structure factor. The latter is estimated from the structure factor amplitudes observed for the native and derivative crystals (Dodson, 1976):

for centric reflections $F_H = |F'_{PH} \pm F_P|$

for acentric reflections $F_H^2 \approx F'^2_{PH} + F_P^2 \pm 2F_P F'_{PH} [1 - 1/2(k\Delta ano/2F_P)^2]^{1/2}$

where F_P is the native structure factor amplitude, $F'_{PH} = Ke^{-Bs^{**2}}F_{PH}$, F_{PH} is the derivative structure factor amplitude, B and K are the overall temperature and scaling factors and $s = 2\sin\Theta/\lambda$, F_H the heavy atom structure factor contribution, Δano is the anomalous difference and k is the ratio between isomorphous and the anomalous heavy atom contributions. These expressions can both be split into two estimates, the F_{HLE} (heavy atom lower estimate) and F_{HUE} (heavy atom upper estimate) and for most reflections the lower estimate is the more correct.

The expression for the acentric reflections can be further simplified (Dodson, 1976) to:

$$F^2_{HLE} \approx \Delta^2 iso + (k\Delta ano/2)^2 \quad (1)$$

3.6 Harker section w=1/3 of MMA isomorphous Patterson

Only half of the section was calculated and the levels were contoured at one sigma intervals above the average.

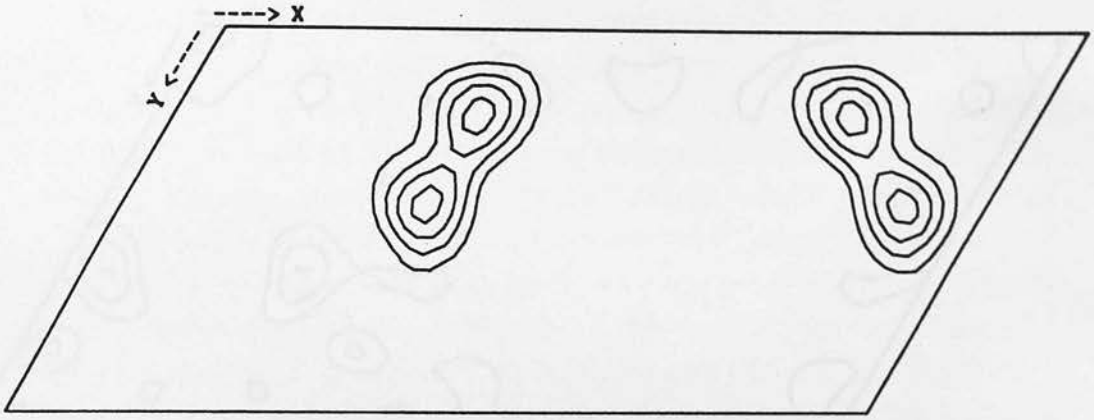


FIG. 3.7 Harker section w=1/3 of MMA anomalous Patterson

Only half of the section was calculated and the levels were contoured at one sigma intervals above the average.

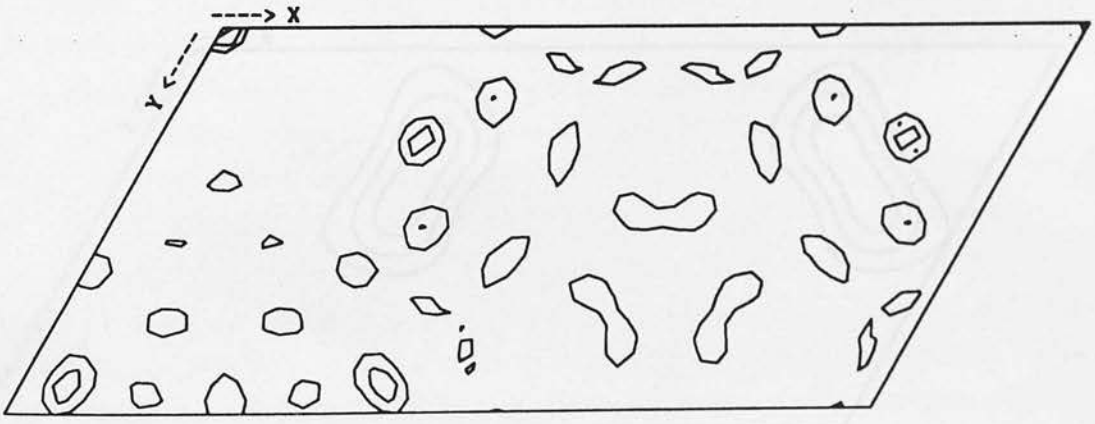


FIG. 3.8 Harker section w=1/3 of platinum derivative Patterson

Only half of the section was calculated and the levels were contoured at one sigma intervals above the average.

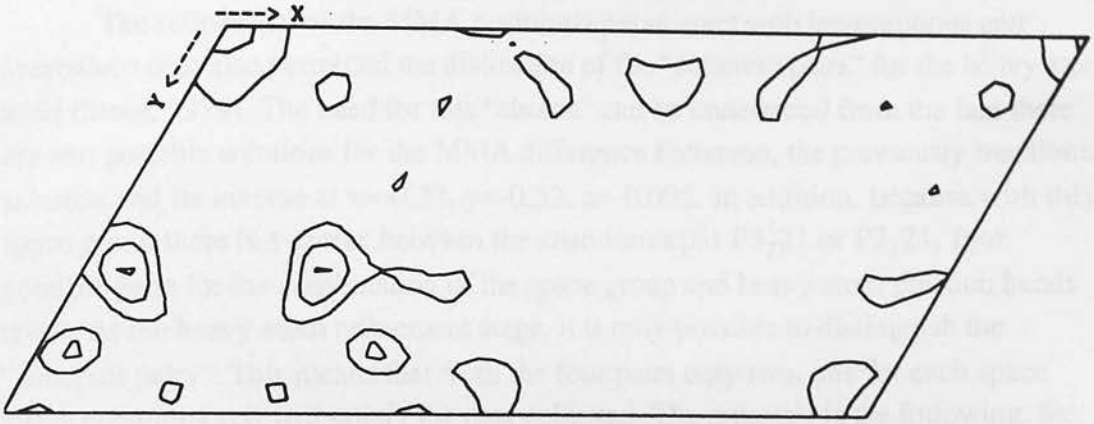
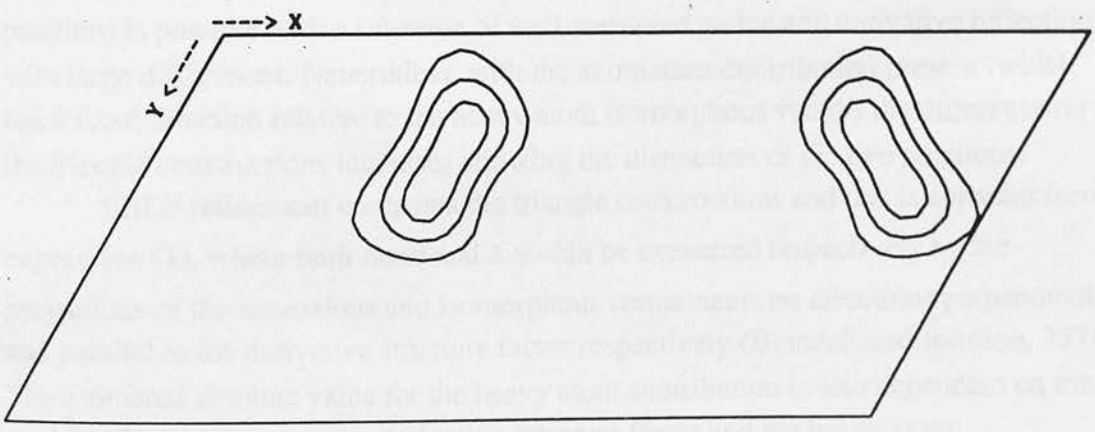


FIG. 3.9 Harker section w=1/3 of HgI Patterson

Only half of the section was calculated and the levels were contoured at one sigma intervals above the average.



where Δiso is the isomorphous difference.

This procedure is ideal for refinement of the heavy atom parameters of a single derivative (Blundell and Johnson, 1976) because the parameters are independent of the minimized function. The procedure works very securely with just centric reflections while with acentric the quality of the parameters depends on the accurate measurement of the anomalous differences. The relative F_{PH} to F_P temperature and scale factors are not very well refined nor are occupancy and atomic temperature factor parameters.

The refinement of the MMA positional parameters with isomorphous and anomalous data also permitted the distinction of the "coherent pairs" for the heavy atom hand (Stout, 1979). The need for this "choice" can be understood from the fact there are two possible solutions for the MMA difference Patterson, the previously mentioned solution and its inverse at x=-0.37, y=-0.52, z=-0.095. In addition, because with this space group there is a choice between the enantiomorphs $P3_221$ or $P3_121$, four possible pairs for the combination of the space group and heavy atom position hands result. At the heavy atom refinement stage, it is only possible to distinguish the "coherent pairs". This means that from the four pairs only two, one for each space group enantiomorph will satisfy the data collected. The rationale is the following: the absolute heavy atom contribution to the diffraction pattern protein crystal is the same either when positioned at x,y,z or -x,-y,-z. The direction, however, of the heavy atom vector changes totally, since the phase changes from positive to negative. The net result is that the phase triangle formed by the derivative, native and heavy atom structure factors is different for each position: a difference that is noticeable in the phase and absolute value of the derivative structure factor. However for most reflections the heavy atom contribution is very small when compared to the protein contribution which, with the added uncertainties resulting from the errors present on the observed terms, makes the two triangles quite similar. In other words, the derivative structure factor for one and the other cases are indistinguishable, especially when only the absolute value is considered. The question arises whether distinction between the two heavy atom positions is possible with a selection of well measured native and derivative reflections with large differences. Neverthless, with the anomalous contribution present (which has a fixed direction relative to the heavy atom isomorphous vector) the difference on the triangle constructions increases allowing the distinction of the two positions.

FHLE refinement compares the triangle constructions and this is apparent from expression (1), where both Δano and Δiso can be expressed respectively by the projections of the anomalous and isomorphous components on directions perpendicular and parallel to the derivative structure factor respectively (Blundell and Johnson, 1976). The estimated absolute value for the heavy atom contribution is also dependent on the angle difference between the derivative structure factor and the heavy atom isomorphous structure factor, which is different for each triangle construction. The

triangles will be the same for the pair composed by one hand of the space group and the positive coordinates of the heavy atom and the other pair composed by the other space group hand and the negative coordinates because the coordinate transformation between enantiomorphic space groups is related by an inversion centre too. Thus, for each space group there is one position that gives better agreement between the calculated and the observed heavy atom amplitude. For the wrong pair, the values of heavy atom occupancy and temperature factor and consequently the amplitude of the calculated heavy atom contribution can never account satisfactorly for the differences in the triangle construction.

For the MMA case the results after six cycles of refinement were the following:

	Space Group	x	y	z	R	g	corr
1	$P3_221$	0.37	0.52	0.095	55%	0.29	0.39
2	$P3_221$	0.63	0.48	0.91	75%	0.12	0.10
3	$P3_121$	0.37	0.52	0.095	67%	-0.09	0.05
4	$P3_121$	0.63	0.48	0.91	59%	0.30	0.33

where $R= \Sigma(F_{Hobs}-F_{Hcalc}/F_{Hobs})$, g is the gradient of $(<F_{Hcalc}^2>)^{1/2}$ against $|F_{Hobs}|$ and corr is a correlation factor between F_{Hobs} and F_{Hcalc}, the observed and calculated heavy atom amplitudes respectively. Both g and corr are equal to unity for the perfect case.

It is clear from the results that pairs 1 and 4 present the best statistics and were therefore considered as the real arrangement.

These positions are equivalent to those estimated for mercurial compounds used in previous work (Green, *et al.*, 1979; Monaco, *et al.*, 1987), where the chosen enantiomorph was $P3_221$. It must be noted that, Green, *et al.*, (1979) have used a non-conventional setting, with the crystallographic dyad that defines the origin coincident with the **b** axis, which is different from the one described in the International Tables, vol.A, (1983) for $P3_221$. In the work conducted by Monaco, *et al.*, (1987) the mercurial compound position coordinates are the same as described in (Green, *et al.*, 1979), which seems to indicate that the same non-conventional setting was used. The diagrams (Monaco, *et al.*, 1987) however, show the molecular packing in the conventional setting. This inconsistency is confusing as it could indicate that the heavy atom positions were determined with respect to a particular origin but the density was calculated with respect to another.

The parameters of the MMA derivative in $P3_221$ after FHLE refinement were:

fractional coordinates x,y,z:	0.379(2)	0.518(2)	0.094(1)
atomic temperature factor:	28.4		
occupancy:	7.1(3)		

3.3.2 Difference maps for the determination of the position of the other heavy atoms

Phases were determined from the MMA derivative with the program PHASE in the space group used by Green, *et al.*, (1979), $P3_221$. These phases were used to produce difference Fourier maps between all the derivatives and the native.

Observation of these maps allowed determination of the heavy atom coordinates. In the MMA case, it was also possible to check the peak shape of the already-determined position, so as to verify that no distortion indicative of a multiple site was present. No minor sites appeared in the MMA difference map. For the other derivatives, the positions of the major substitutions could be easily determined from the respective difference maps. The positions for the platinum derivative were x1=0.46, y1=0.2, z1=0.026 and x2=0.34, y2=0.72, z2=0.07, while for the HgI the most prominent position was x=0.37, y=0.52, z=0.09. This is the same position as that for the MMA derivative, while the top platinum position is very similar to the top solution described for the same compound in Green, *et al.*, (1979).

3.3.3 Choice of space group enantiomorph

The procedure described in Blundell and Johnson, (1976) was used to determine the exact enantiomorph. This procedure requires the existence of two heavy atom derivatives data sets, of which one will have measured anomalous differences. All the calculations must done in one of the possible space groups. Two sets of phases are calculated from a single derivative with anomalous measurements, the phases will include isomorphous and anomalous contributions; one set is prepared with the anomalous contribution unchanged and the other set with the signal of the anomalous differences inverted. With these phase sets, two difference Fourier syntheses are calculated for the the second derivative. The highest peaks in both maps indicate the heavy atom position of the second derivative site. The height will not be the same for these common peaks; if the higher peak is present in the map calculated from the phases with unchanged anomalous contribution then the enantiomorph chosen to prepare the maps is the correct one, if the higher peak is found in the map calculated with the inverted anomalous signal then the correct enantiomorph is the opposite of the one used for all these calculations.

The procedure was attempted with the phases calculated from the MMA derivative for both the platinum and HgI difference maps, in $P3_221$. The HgI difference maps showed no difference between each other, probably because the site of this derivative is common to MMA and so the image of the heavy atom used to produce the phases is present in the map.

For the platinum case, the map calculated with an unchanged anomalous signal

presented the highest peak at x=0.46, y=0.20, y=0.026 with a height of 1.59 units and an r.m.s. deviation of 0.15, while the inverted anomalous signal map had its highest peak placed in the same position with a value of 1.17 units and the same r.m.s. deviation. Sections of these two maps are shown on FIG. 3.10, and it is clear that the map with unchanged anomalous signal has the highest peak. This means that the correct enantiomorph is $P3_221$, also confirmed by molecular replacement (see section 3.8.4) and is in line with the choice described in the literature (Green, *et al.*, 1979; Monaco, *et al.*, 1987).

3.3.4 Heavy atom refinement and phase calculation

The positions of all the heavy atoms, defined in section 3.3.2, were refined with the phase refinement program, MLPHARE. This sort of refinement improves the heavy atom parameters by minimizing the difference between the absolute values of the observed and the calculated derivative structure factors (Blundell and Johnson, 1976). The calculated derivative structure factor amplitude is determined from the following expression:

$$F_{PHcalc} = | \mathbf{F}_{Pcalc} + \mathbf{F}_{Hcalc} |$$

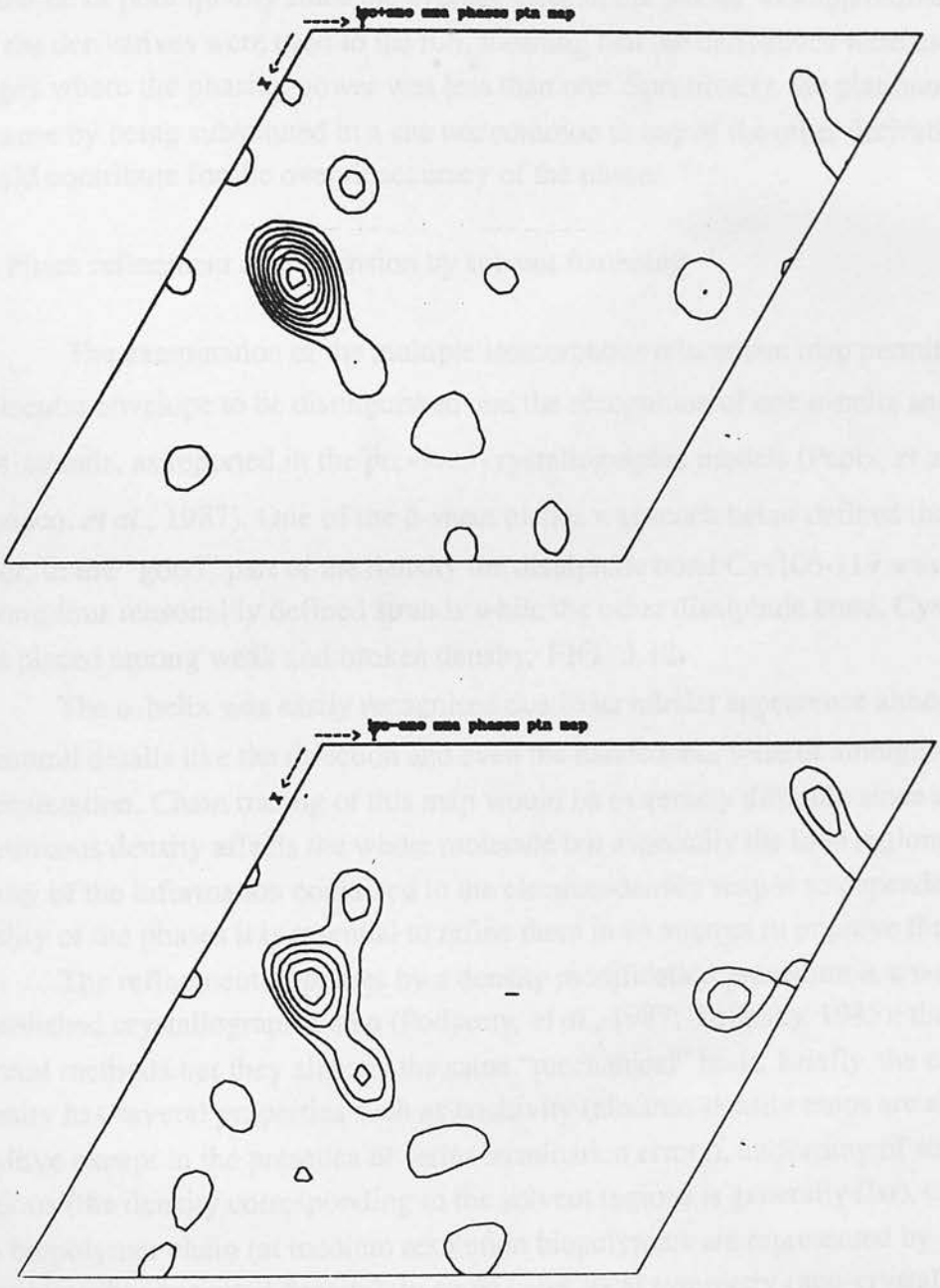
$\mathbf{F}_{Pcalc}$ and $\mathbf{F}_{Hcalc}$ are the vectors for the calculated protein and heavy atom contributions. This means that the difference between the phases of the protein and heavy atom vector has to be calculated in each cycle of refinement and that the method works correctly when several derivatives with different sites are refined together or when one derivative is being refined against outside phases, but not when the parameters of a single derivative (Dodson, 1976) or multiple derivatives with common sites are being refined. The procedure seems to detect incorrect positions, indicated by an occupancy close to zero, as long as there are well estimated positions being refined at the same time (Blundell and Johnson, 1976) and to be able to refine correctly the relative scale factors and temperature factors between the native and the derivative (Dodson, 1976).

The procedure was applied for all the positions described above, from which the minor site of the platinum derivative was rejected. Subsequently, phases were calculated and difference maps for all derivatives were inspected. Some possible extra sites were put into a few cycles of refinement but rejected.

The final positions and statistics are presented in TABLE 3.5. It is apparent that MMA is the derivative that is contributing the most to the phase calculation. This is demonstrated by the phasing power for both the isomorphous and anomalous contributions which are above 1 over all resolution ranges. The platinum derivative is the weaker one since it has the highest Rcullis and lowest phasing power, even at low

FIG. 3.10 Sections of platinum derivative difference map for the determination of the space group enantiomorph

Top: phases calculated with isomorphous plus positive anomalous contribution
Bottom: phases calculated with isomorphous plus negative anomalous contribution
Peaks were contoured at one sigma intervals above the average.



resolution. The HgI with a phasing power just over 1, is contributing reasonably to the lower resolution ranges.

The phases for BlgZ were calculated with MLPHARE for the resolution range 200 to 3.5 Å and the figure of merit for the final phases is plotted in resolution ranges in FIG. 3.11. This statistical parameter evaluates the sharpness of the phase probability distribution and can be approximated as the cosine of the error in phase angle for the reflection. The real quality of the phases can only be judged from the electron-density map but it was expected that with an average figure of merit value of 0.46 the map would be of poor quality since the average error in the phases was approximately 62°. All the derivatives were used to the full, meaning that the derivatives were used even in ranges where the phasing power was less than one. Specifically, the platinum was kept because by being substituted in a site not common to any of the other derivatives it would contribute for the overall accuracy of the phases.

3.4 Phase refinement and extension by solvent flattening

The examination of the multiple isomorphous relacement map permitted the molecular envelope to be distinguished and the recognition of one α-helix and a number of β-strands, as reported in the previous crystallographic models (Papiz, *et al.*, 1986; Monaco, *et al.*, 1987). One of the β-sheet planes was much better defined than the other; in the "good" part of the density the disulphide bond Cys106-119 was visible among four reasonably defined strands while the other disulphide bond, Cys66-160 was placed among weak and broken density, FIG. 3.12.

The α-helix was easily recognised due to its tubular appearence although structural details like the direction and even the handedness were of ambiguous interpretation. Chain tracing of this map would be extremely difficult, since a lack of continuous density affects the whole molecule but especially the loop regions. Since the clarity of the information contained in the electron-density map is so dependent on the quality of the phases it is essential to refine them in an attempt to improve the density.

The refinement of phases by a density modification procedure is a well established crystallographic step (Podjarny, *et al.*, 1987; Tulinsky, 1985); there are several methods but they all have the same "mechanical" basis. Briefly, the electron-density has several properties such as positivity (electron-density maps are always positive except in the presence of series termination errors), uniformity of solvent regions (the density corresponding to the solvent regions is generally flat), continuity of the biopolymer chain (at medium resolution biopolymers are represented by continuously connected density), in some cases local symmetry (non-crystallographic symmetry relates density within the molecular envelope) and others. If these characteristics are well understood then they can be explored for density improvement

TABLE 3.5 Refined heavy atom parameters and phasing statistics

MMA derivative

x=0.372	y=0.516		z=0.094		occ.=9.7 ano.occ.=6.0		bfac.=30.		
Res.bins	17.8	11.2	8.2	6.4	5.3	4.5	3.9	3.5	total
Phas.pow.									
Acentric	0.5	1.6	1.8	2.2	2.1	1.4	1.3	1.2	1.5
Centric	0.5	1.2	1.4	1.7	1.4	0.9	1.0	0.8	1.2
Anomalous	2.6	0.9	0.8	0.9	0.9	1.0	1.0	1.0	1.1
Rcullis	0.72	0.49	0.50	0.40	0.57	0.73	0.75	0.85	0.58

Platinum derivative

x=0.467	y=0.206		z=0.026		occ.=3.9		bfac.=30.		
Res.bins	17.8	11.2	8.2	6.4	5.3	4.5	3.9	3.5	total
Phas.pow.									
Acentric	0.4	0.8	0.8	0.9	1.0	0.6	0.6	0.6	0.7
Centric	0.2	0.6	0.7	0.7	0.8	0.5	0.5	0.6	0.6
Rcullis	0.93	0.68	0.65	0.62	0.76	0.89	1.06	0.96	0.78

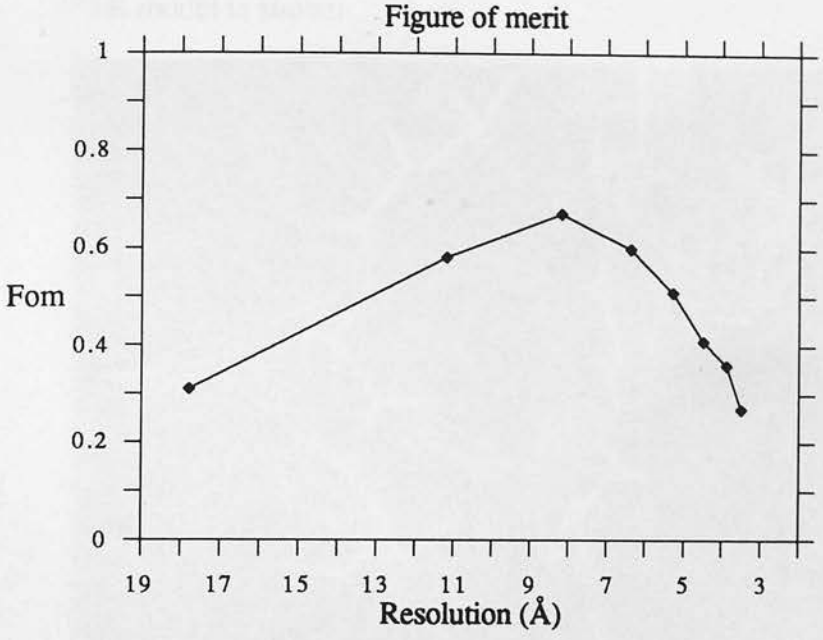
HgI derivative

x=0.375	y=0.529		z=0.093		occ.=4.0		bfac.=30.		
Res.bins	17.8	11.2	8.2	6.4	5.3	4.5	3.9	3.5	total
Phas.pow.									
Acentric	0.7	0.8	1.0	1.2	1.0	0.0	0.0	0.0	1.1
Centric	0.4	0.6	0.8	0.9	0.8	0.0	0.0	0.0	0.8
Rcullis	0.88	0.69	0.70	0.70	0.92	0.0	0.0	0.0	0.73

occ.- occupancy; ano.occ.- anomalous occupancy; bfac.- heavy atom temperature factor; Res.bins- Resolution bins in Å; Phas.pow.- Phasing power (<FHcalc>/<|FPHobs-FPHcalc|>; Rcullis- Cullis Rfactor (<|FPHobs-FPHcalc|>/<|FPHobs-FPobs|> for centric reflections only; FHcalc- heavy atom calculated structure factor amplitude; FPHobs- derivative observed structure factor amplitude; FPHcalc- derivative calculated structure factor amplitude; FPobs- native observed structure factor amplitude

FIG. 3.11 Plot of figure of merit over resolution bins

Fom - represents the average of figure of merit for each resolution bin.

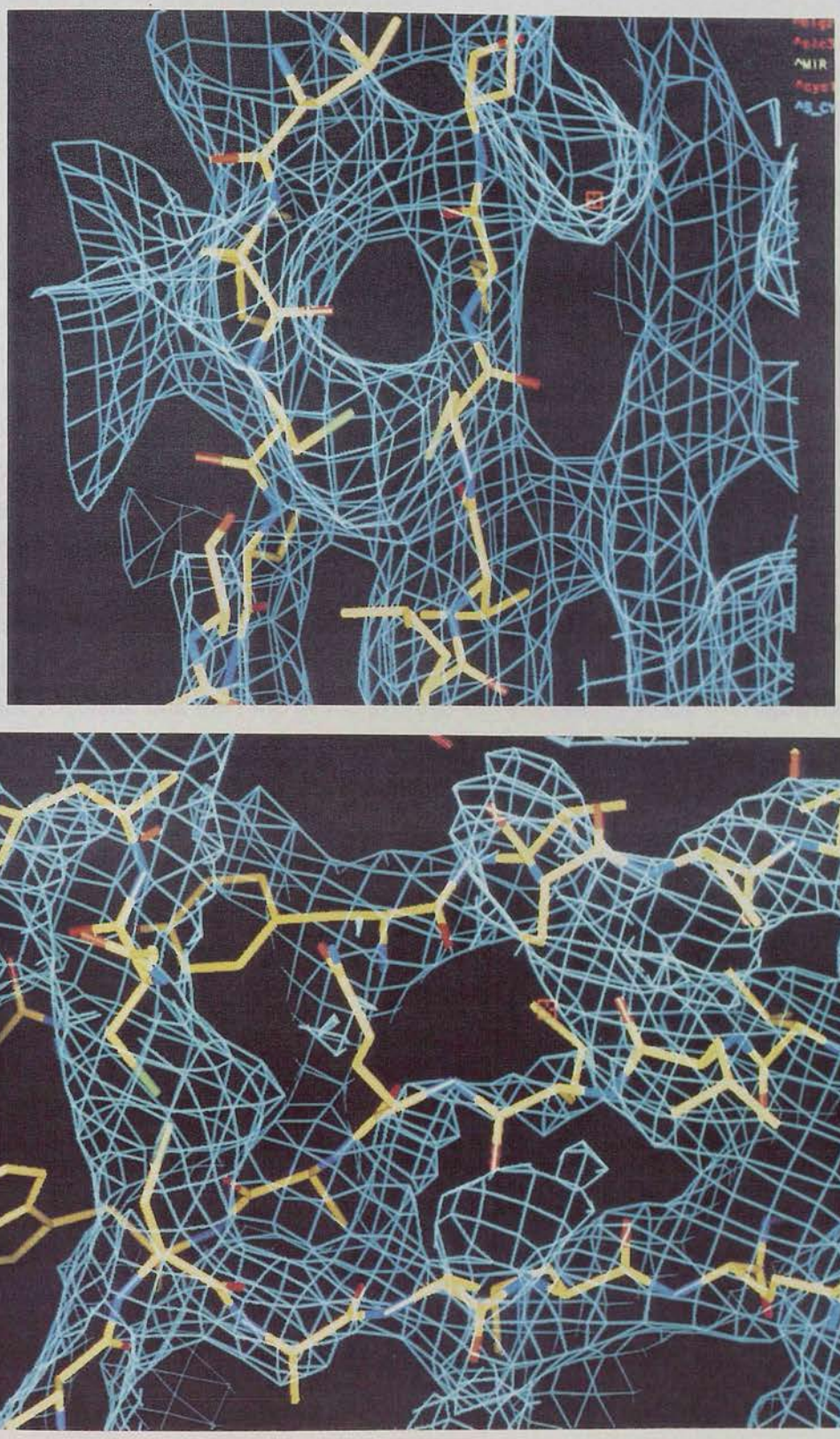


by constraining the overall density to a model of the chosen property. The method is more powerful the larger the change it induces and the larger the number of map grid points it affects (Podjarny, *et al.*, 1987). The general algorithm involves the calculation of an initial map with the highest quality phases possible, modification of the map, back transformation of the modified map and combination of the resulting information, calculated phases and in some cases calculated amplitudes, with the experimental data. A new map is calculated with the modified terms and the procedure is repeated until convergence is achieved.

Map quality is also related to the number of terms present in the Fourier transform, thus an increase in this number either at the same resolution as before or at a higher resolution will result in a more well-defined molecular picture. The density modification methods alter the initial density and introduce extra information of a general nature. During the back Fourier transform of the modified density, extra terms are calculated permiting the extension of phase information (Tulinsky, 1985). The errors in the procedure are filtered out as long as there is experimental information that can be incorporated into the extended terms.

FIG. 3.12 Two pictures of different parts of the density produced with the MIR phases

Top: view of the MIR density around Cys106-119
Bottom: view of the MIR density around Cys66-160
Final model is shown.



3.4.1 Phase refinement

The improvement of the density was performed by the application of solvent uniformity because it does not demand any interpretation of the existing density. The following lines will describe the method and the results obtained.

The procedure applied is designated as solvent flattening and is in fact a conjunction of two of the properties described above. The particular algorithm used (Wang, 1985; Leslie, 1988), flattens the solvent and applies a positivity constraint in the molecular region.

The different stages are the following, with the programs used at each stage enclose within brackets. The map was calculated with as many terms as possible (FFTKW). All the negative density present in this map was set to zero (TRUNCMAP), after which the map was back transformed (SFC) and all the structure factors at the resolution range used were calculated. The structure factors were multiplied by a weighting function (HKLWEIGHT) that corresponds in real space to the substitution of the value of the density at a particular grid point by a weighted average of the density within a given radius. This reciprocal space weighting is, in fact, only applied to the low resolution structure factors, at spacings larger than 5 Å (Leslie, 1988), or in other words, applied to the reflections that contain more information about the solvent. With the new structure factors, it was possible to determine the molecular boundaries because the electron density generated (FFTKW) was a "blurred" original map and defined a molecular mask. Thus with knowledge of the crystal solvent content, the solvent part of the cell is easily defined (ENVELOPE). To the density, a constant term is added (FLATMAP) that substitutes the missing F000 term. This added constant satisfies the requisite that the ratio of the average solvent density to the average protein density should be some defined value. After this addition, the density in the solvent regions is set to the average solvent density and the negative parts in the protein envelope are attenuated. The map thus modified is back transformed (SFC) and after scaling (SCALENEW), the calculated phases are combined (COMBINE) with the experimental ones. The procedure is repeated using the refined phases and the same envelope until no further modification is observed. A new envelope is created whenever it is considered that the existing one is not constraining the density anymore.

The essential parameter in this procedure is the solvent content, because it determines the number of the grid points affected by the modification. From Matthews, (1968) formula the solvent content for BlgZ was determined to be 42%, however for the application of the solvent flattening it was decided to use a conservative value of 32%, since it was observed that higher values removed the density corresponding to loops and other solvent exposed parts of the molecule. Another parameter observed to have some effect on the continuity of the density was the averaging radius for the determination of the "blurred" map or mask, with a smaller radius the mask was

detailed but loss of connectivity was observed in the weaker density present on the main-chain β-strands. The "improved" maps were determined using the combined phases and the observed amplitudes.

The evolution of the improvements was followed by observing the electron-density maps and by determining the residual between calculated and experimental amplitudes, the overall (and high resolution) average figures of merit and residual. The results are plotted in FIG. 3.13. During phase refinement only two envelopes were determined because the evolution from the first to the second mask was minor. Convergence was attained reasonably quickly as demonstrated by the values plotted. The map produced at the 4th cycle presented more connectivity than the original one but the differences between the fourth and the sixth maps were minute.

3.4.2 Phase extension

The extension of the phases by solvent flattening from 3.5 to 3.0 Å was attempted for two reasons. First, it has been reported that the refinement of the original phases continues as new terms are introduced (Zhang and Main, 1988) and second the determination of phases to at least 3.0 Å for β-sheet proteins can be important since beyond this resolution the "lateral fusion" of the strands disappears and the chain tracing is facilitated.

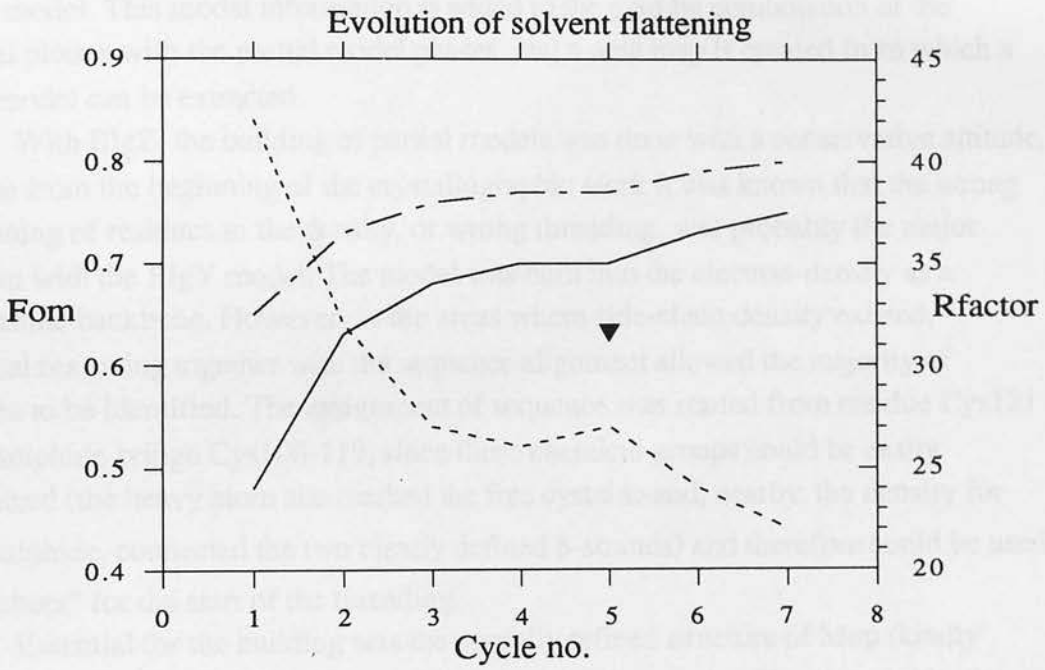
It has also been reported that the success of the extension is dependent on the completeness of the reflection set (Frederick, *et al.*, 1984; McClarin, *et al.*, 1986) and because the native data set was considered to be incomplete at 3 Å, data collected on a weak derivative was used instead. This derivative data set was >90% complete to 3.2 Å resolution and 82.5% complete in the range 3.2 to 3.0 Å and it had a mean isomorphous fractional difference with the native of 13%. It was not used for phase calculation because the unique heavy atom site present had very low occupancy.

The extension was done in the following way: substitution of the native amplitudes for the derivative ones while keeping the original and refined (see section 3.4.1) native phases, refinement of the phases at 3.5 Å to accommodate any difference between data sets, calculation of all the phases for reflections with amplitude present but no experimental phase, refinement of the new phases, extension and refinement of phases to 3.0 Å in three steps, 3.5 to 3.3 Å, 3.3 to 3.15 Å and 3.15 to 3.0 Å. The extension was done for one to two reciprocal space units at a time which meant that 300-400 reflections were added in each step. The evolution of the extension can be followed from the figure of merit and residual plotted on FIG. 3.14.

The map after this treatment showed major improvements, it was now possible to recognize the hand of the helix hand (right handed) (FIG. 3.15), and the β-sheets from the "bad side " (see section 3.4) of the molecule started to appear (FIG. 3.16). It

FIG. 3.13 Plot of phase refinement evolution

Fom is the average figure of merit, represented by (— · —) for all the data and (——) for the resolution range 3.5-4 Å. Rfactor is the residual between the calculated and observed amplitudes, represented by (--------). The black triangle marks the calculation of a new envelope.



has to be recognized that there were still regions of difficult interpretation and that the assignment of residues to the density in some parts of the map was difficult. However, because most of the main-chain could be followed, it was decided to initiate the building of the molecular model.

3.5 Model building

The strategy for model building was based on the need to improve the map further. For that, another method of map improvement, already mentioned in section 3.4 was applied. This method (Rice, *et al.*, 1988), constrains the density to a biopolymer continuity by introducing information from partial models, the constraint being the more effective the more atoms are included into the partial model. The "tricky step" with this procedure is the need for interpretation of the density to produce the partial model. This model information is added to the map by combination of the original phases with the partial model phases, and a new map is created from which a better model can be extracted.

With BlgZ, the building of partial models was done with a conservative attitude, because from the beginning of the crystallographic work it was known that the wrong positioning of residues in the density, or wrong threading, was probably the major problem with the BlgY model. The model was built into the electron-density as a polyalanine backbone. However, in the areas where side-chain density existed, chemical reasoning together with the sequence alignment allowed the majority of residues to be identified. The assignment of sequence was started from residue Cys121 and disulphide bridge Cys106-119, since these chemical groups could be easily recognized (the heavy atom site marked the free cysteine and, nearby, the density for the disulphide, connected the two clearly defined β-strands) and therefore could be used as "anchors" for the start of the threading.

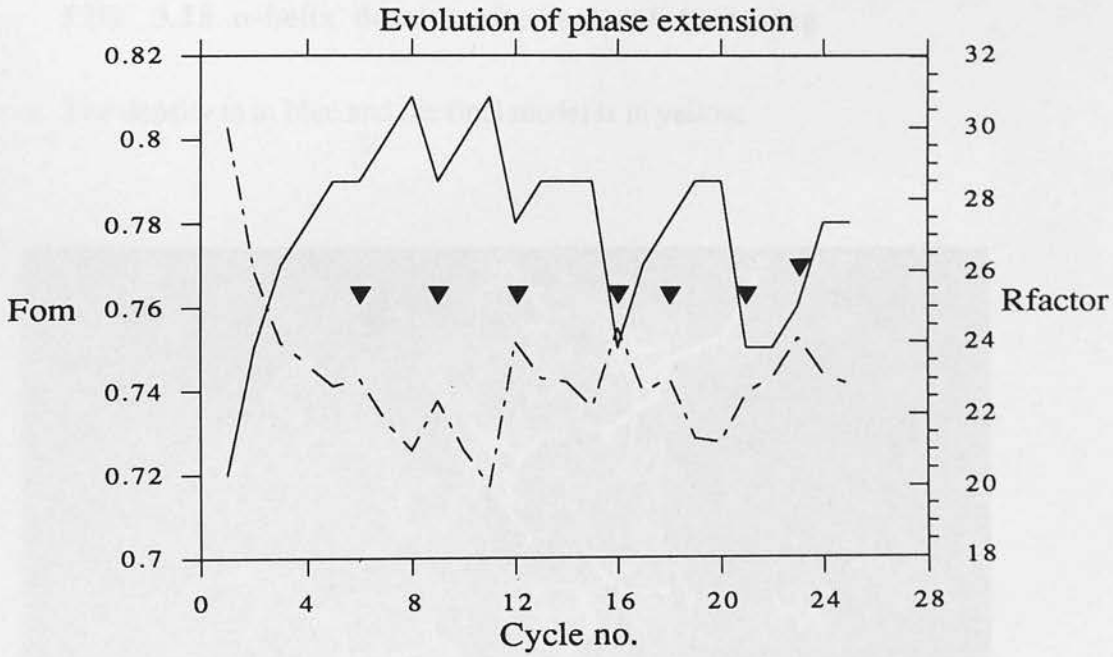
Essential for the building was the partially refined structure of Mup (kindly provided by Prof.A.C.T. North) superimposed with the density. To achieve this, the molecular replacement solution (see section 3.8) for Mup was used. This molecule was expected to be similar to Blg, from sequence alignment (Adams, 1992) and in fact, was seen, from the superimposition of the model into the density, to have many common three-dimensional features. Adding to these reasons was the fact that Mup, as an example of a well-built protein molecule, served as a teaching aid.

Model building proceeded through two partial model stages. The first model built consisted of 118 residues and 816 atoms, distributed in three segments:

first segment -	residue 21 to 48, all alanines
second segment -	residue 67 to 107, where 67 to 89 were alanines and 90 to 107 the Blg residues

FIG. 3.14 Plot of phase extension evolution

Fom is the average figure of merit, represented by (——). Rfactor is the residual between calculated and observed amplitudes, represented by (— - —). The triangles mark the cycles where new envelopes or phases were extended. Cycle 6 - new envelope and phases extended to 3.5 Å, cycle 9 - extension to 3.3 Å, cycle 12 - new envelope, cycle 16 extension to 3.15 Å, cycle 18 - new envelope, cycle 21 - extension to 3.0 Å, cycle 23 - new envelope.



third segment - residue 117 to 157, where 117 to 140 were Blg residues and the remaining alanines

The second and third segments were connected by the disulphide bond Cys106-119, forming the major segment.

The second model consisted of 131 residues and 1182 atoms distributed in four segments:

first segment - residue 16 to 49, built according to the sequence
second segment - residue 52 to 60, built as alanines
third segment - residues 67 to 75, built according to the sequence
fourth segment - residues 79 to 158, where 79-84 were built according to the sequence, 85-89 as alanines, 90-142 following the sequence, 143 as alanine

FIG. 3.15 α-helix density after solvent flattening

The density is in blue and the final model is in yellow.

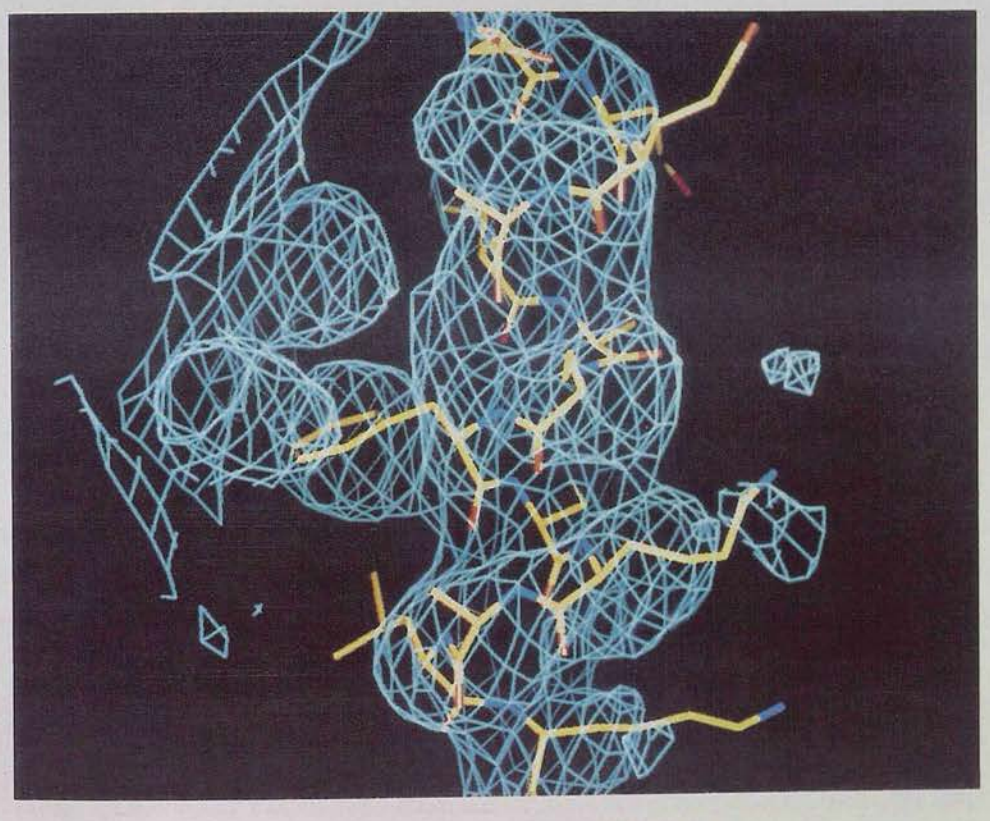
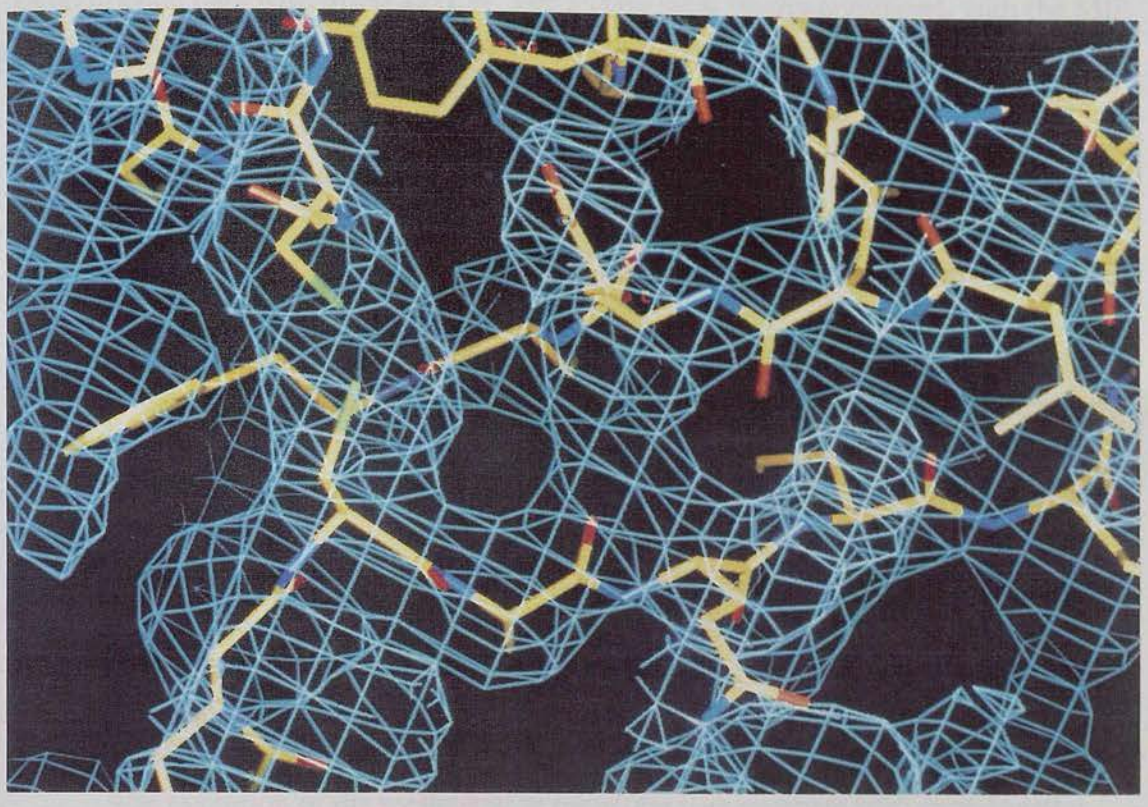


FIG. 3.16 Density of the strands on the "bad" side of the molecule after solvent flattening

The density is in blue and the final model is in yellow.



fifth segment - residues 145 to 159, where 145 and 146 were present as alanines, 147-155 were built according to the sequence and 156-159 as alanines

Each partial model was refined by a molecular dynamics procedure (X-PLOR) to improve the stereochemistry and fit to the X-ray data. The dynamics procedure consists of a search of the protein's conformation by simulating the effect of heating and slow cooling on the dynamics (direction and velocity) of each atom. The atoms should escape false local minima and following a gradient of energy, function by the stereochemical restrictions and agreements with the experimental data, attain an energy minima.

To remove any excessive influence of incorrect parts of the model on the combined phases, a temperature factor refinement was performed in X-PLOR after the

annealing procedure. The reason for applying this procedure is based on the concept that wrongly placed atoms would have higher temperature factors to account for the disagreement with the density. This results in weighting each residue's contribution to the partial phases and consequently to the combined phases.

The phases and calculated amplitudes were produced by SFC and scaled to the experimental data with SCALENEW. The phases were finally combined, with the MIR phases and with the solvent flattened phases, using the program COMBINE (all these programs, SFC, SCALENEW and COMBINE, are part of the CCP4 package).

It was found convenient when rebuilding the model to not only refer to both the MIR and solvent flattened combined with model information but also to the $2F_{obs}$-F_{calc} and F_{obs}-F_{calc} maps(where F_{obs} are the observed amplitude, F_{calc} the calculated amplitudes from the partial model and the phases obtained from the models) and the original maps, MIR and solvent flattened.

The evolution and details of the dynamics procedure are described in TABLE 3.6. The residual dropped between models, as expected from introducing more atoms into the models. In both models it was necessary to restrict the movement of the terminae of some of the fragments to avoid "flying" and subsequent steric clashes with other parts of the molecules, which would trigger the program to stop. During the minimization and dynamics procedure, the ionic side-chains were kept neutral, to avoid drastic movements of parts of the molecule due to ionic interactions. The maximum temperature of the slowcooling protocol was set at a low value for the first model (2000 K) and then increased to 3000 K for the second model, this way increasing the possiblity of releasing the molecule from local energy minima.

The evolution of the model and in particular its agreement with experimental data is probably better evaluated from the parameters measured during the phase combination procedure, especially through the comparison of observed and calculated phases since the latter are independent of the refinement done in the previous steps. In TABLE 3.7, the parameters relative to the comparison between the two partial models and the experimental data, amplitudes and phases (MIR and solvent flattened) are displayed. The changes were greater for the MIR data than for the solvent flattened data as expected, since the solvent flattening procedure refined the phase information. From the relative lack of closure and the phase change, it can be concluded that the influence of adding the model information is felt more strongly in the higher resolution ranges. Finally, it seems clear that the second model is an improvement over the first, because the overall relative lack of closure decreases from the first to the second model indicating that the building has evolved correctly.

3.6 Final model

The final model consisted of 158 residues and 1516 atoms, the first three

TABLE 3.6 Dynamics procedure details

Model 1 (constrained residues 21, 48, 67)

Refinement step	Details	Resolution	R-factor
1- Weight calculation	Wa = 45388	8-3 Å	50.8%
2- Minimization	160 cycles of minimization	8-3 Å	37.3%
3- Slow cooling	Max.Temp.= 2000 K no. of steps = 50 timestep = 0.0005 ps	8-3 Å	33.9%
4- Minimization	120 cycles of minimization	8-3 Å	33.3%
5- B-factor refinement	15 steps of refinement side-chain/main-chain	8-3 Å	32.3%

Model 2 (constrained 52, 60, 67, 75)

Refinement step	Details	Resolution	R-factor
1- Weight calculation	Wa = 54044	8-3 Å	48.9%
2- Minimization	160 cycles of minimization	8-3 Å	34.1%
3- Slow cooling	Max.Temp.= 3000 K no. of steps = 50 timestep = 0.0005 ps	8-3 Å	29.0%
4- Minimization	120 cycles of minimization	8-3 Å	28.1%
5- B-factor refinement	15 steps of refinement side-chain/main-chain	8-3 Å	26.1%

TABLE 3.7 Evolution of map improvement by partial model phase combination

MIR + model 1 combination to 3.5 Å

res.bins(Å)	r.lack clos.	no.refl.	old fm	new fm	abs.ph.ch.
50-7.8	0.8160	250	0.689	0.797	11.90
7.8-5.5	0.6112	445	0.577	0.780	23.80
5.5-4.5	0.4552	536	0.420	0.744	37.66
4.5-3.9	0.4850	606	0.322	0.718	47.96
3.9-3.5	0.5373	660	0.194	0.680	62.37
total	0.5882	2497	0.391	0.733	41.64

Solvent flattened + model 1 combination to 3.0 Å

res.bins(Å)	r.lack clos.	no.refl.	old fm	new fm	abs.ph.ch.
50-6.7	0.7698	395	0.856	0.867	8.37
6.7-4.7	0.5478	682	0.807	0.822	16.35
4.7-3.9	0.4540	806	0.768	0.796	21.71
3.9-3.4	0.5529	891	0.761	0.780	23.40
3.4-3.0	0.6634	842	0.740	0.747	26.15
total	0.5922	3616	0.777	0.793	20.69

MIR + model 2 combination to 3.5 Å

res.bins(Å)	r.lack clos.	no.refl.	old fm	new fm	abs.ph.ch.
50-7.8	0.7986	250	0.689	0.805	10.41
7.8-5.5	0.4559	445	0.577	0.818	29.74
5.5-4.5	0.3556	536	0.420	0.771	39.82
4.5-3.9	0.3692	606	0.322	0.758	55.48
3.9-3.5	0.4642	660	0.194	0.720	59.31
total	0.5017	2497	0.391	0.766	44.03

res.bins- resolution bins; r.lack clos.- relative lack of closure (relative lack of closure between calculated and observed structure factors); no.refl.- number of reflections in each bin; old fm- old average figure of merit (figure of merit before combination); new fm- new average figure of merit (figure of merit after combination); abs.ph.ch.- absolute phase change from observed to combined phase.

TABLE 3.7 Continuation

Solvent flattened + model 2 combination to 3.0 Å

res.bins(Å)	r.lack clos.	no.refl.	old fm	new fm	abs.ph.ch.
50-6.7	0.7267	395	0.856	0.863	8.94
6.7-4.7	0.4223	682	0.807	0.827	22.13
4.7-3.9	0.3536	806	0.768	0.798	28.74
3.9-3.4	0.4898	891	0.761	0.791	26.15
3.4-3.0	0.5882	842	0.740	0.754	30.61
total	0.5094	3616	0.777	0.799	25.13

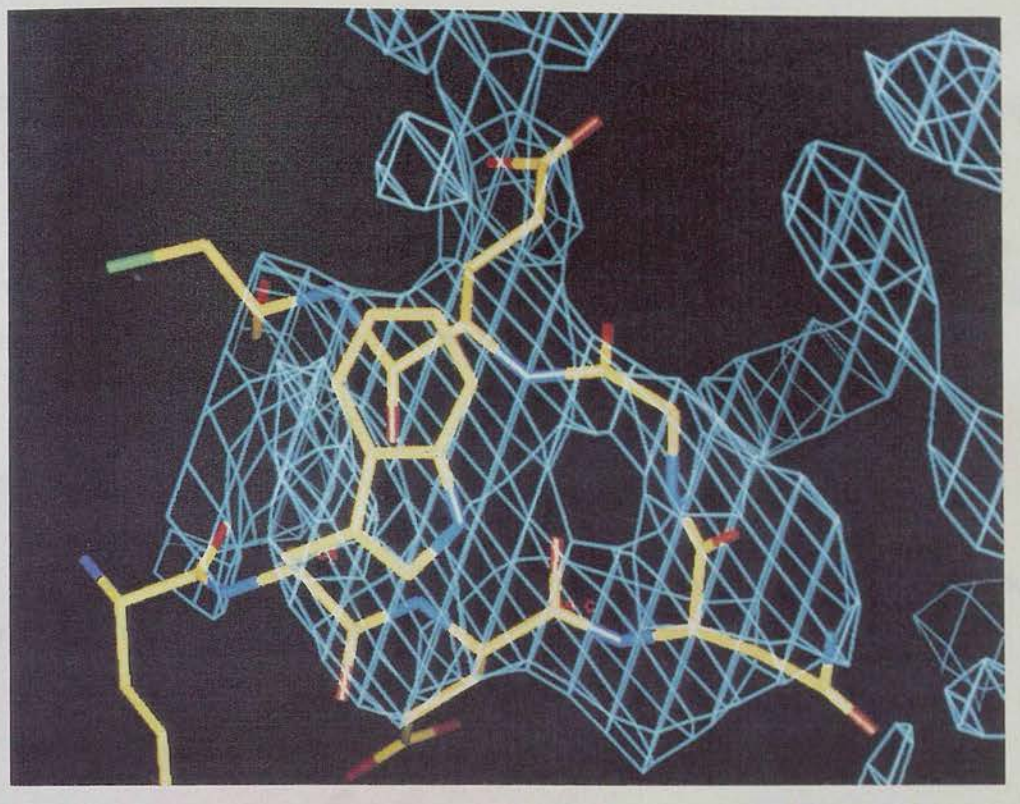
res.bins- resolution bins; r.lack clos.- relative lack of closure (relative lack of closure between calculated and observed structure factors); no.refl.- number of reflections in each bin; old fm- old average figure of merit (figure of merit before combination); new fm- new average figure of merit (figure of merit after combination); abs.ph.ch.- absolute phase change from observed to combined phase.

residues at the N-terminus and the last residue (C-terminus) had little or no density and therefore were not included. No stretches of alanines were left. This model was totally rebuilt from the second model. The rebuilding was done from Cys121 towards the C-terminus and towards the centre of the molecule at Cys66. Then, building was restarted from Trp19 towards the N-terminus and forward, towards Cys66. The two fragments formed in this way, met and closed at the loop between the third and fourth strands following the density from the MIR+2nd model combined map (FIG. 3.17), with no need for extra adjustments on the positions of residues previously positioned. This strategy was used because the density was weaker on the second and especially on the third β-strand. The model was then refined as before, with the results shown in TABLE 3.8. The residual attained is quite satisfactory for this resolution range, especially if compared with the value obtained for the BlgY model, positioned by molecular replacement (see section 3.8), after the same dynamics procedure. The final residual was 26.2% for 162 residues and 1548 atoms compared with 22.9% of the BlgZ model.

Bearing in mind that the model is far from refined, several analyses were performed. A geometry analysis using X-PLOR, which evaluates the deviation from ideal values for bond lengths, bond angles and dihedral angles, indicated that some regions of the molecule contained very poor geometry (FIG. 3.18). The program calculated the ratio between the empirical energy (function of bond lengths, bond angles and dihedral angles) and the value of the root mean square deviation for each residue. This value indicates the stretches of model with worst stereochemistry and a well refined structure typically presents values of 0-3 for this ratio. The residues that present

FIG. 3.17 Density for loop between the third and fourth strands

Density in blue and model in yellow.



problems are:

Ala26 - At the end of the first β-strand, for the moment this was built with an unusual bond angle formed by the N, Cα and C atoms.

Pro50 - Positioned in the loop between the second and third β-strand, the high energy value is due to unusual dihedral angle in the ring, N-Cα-C-Cγ.

Asp53 - Positioned after the loop between the second and third β-strand. It has an unusual bond angle, C-Cα-Cβ. The stretch from Thr49 to Asp53 has been difficult to model. The density for the main-chain was present but the exact position of the side-chains and therefore the overall aspect of the loop was not well determined.

Lys60 - Positioned at the end of the third β-strand, it presents two unusual dihedral angles, Cα-Cβ-Cγ-Cδ, and Cβ-Cγ-Cδ-Cε. Density was only visible in maps produced with combined phases.

Pro79 - Positioned in the loop between the fourth and fifth β-strand. Bad bond angle, C(79)-N(80)-Cα(80). The main-chain density for this loop was clearly seen but without much detail resulting in a difficult building of the residues.

Glu108 to Leu117 - All of these residues are positioned in a loop (between the seventh and eighth β-strands) for which not much density exists, it was just possible to guess the main-chain from the MIR map.

Thr125, Pro126 and Glu127 - Loop between eighth β-strand and α-helix. Density present but it was difficult to build.

Glu134, Lys135 and Phe136 - In the middle of the α-helix. It was badly built. The problem is simple to correct.

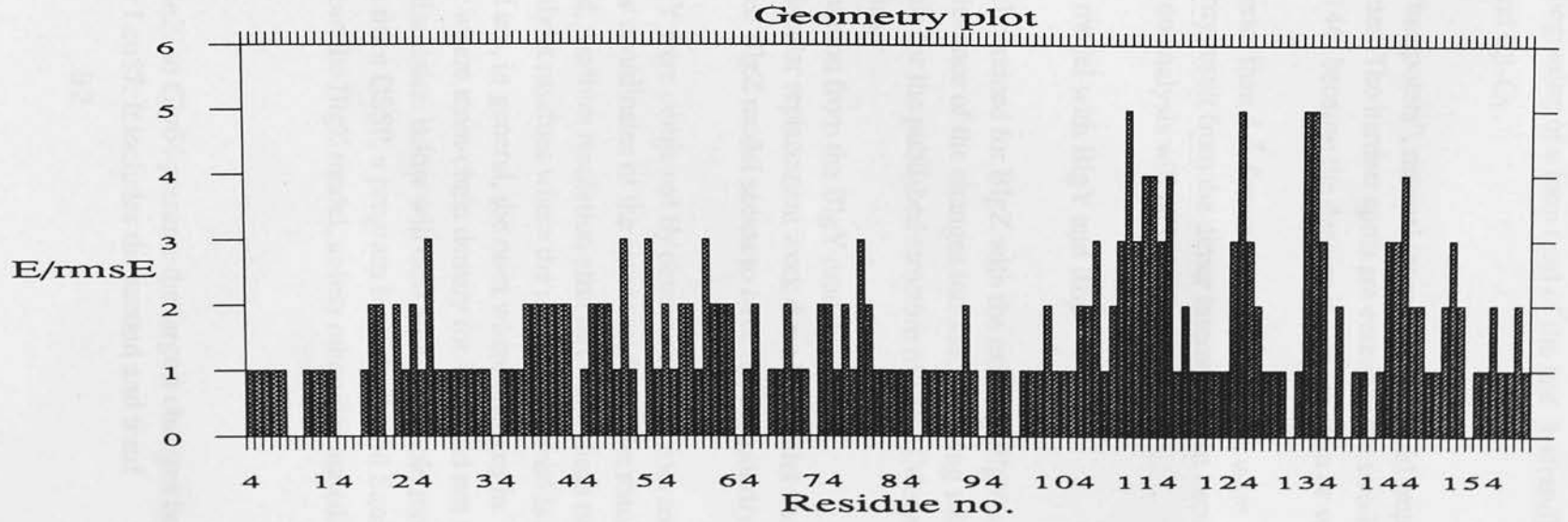
TABLE 3.8 Final model refinement

Final model

Refinement step	Details	Resolution	R-factor
1- Weight calculation	Wa = 65935	8-3 Å	44.2%
2- Minimization	100 cycles of minimization	8-3 Å	32.0%
3- Slow cooling	Max.Temp.= 4000 K no. of steps = 50 timestep = 0.0005 ps	8-3 Å	23.9%
4- Minimization	160 cycles of minimization	8-3 Å	22.9%

FIG. 3.18 Geometry plot for the final model.

An empirical energy was evaluated, which is function of bond length, bond angle and dihedral angle. The parameter plotted is the ratio between this energy and the root mean square deviation from ideal energy. A well refined structure has values between 0 and 3. Values were rounded up to the next unit and are plotted for the sequence number. Bad points are discussed in the text.



Pro144, Met145 and His146 - Positioned in the loop between the α-helix and the last β-strand. Not much density was present.

Asn152 - Positioned at the beginning of a loop (end of the last β-strand). It has two short bond distances, Cα-Cβ and Cβ-Cγ.

From the analysis of these "bad points", manual intervention should improve the geometry of most of these residues. The hardest spots are certainly the two loops (Glu108-Leu117) and (Pro144-His146) because the density is non-existent or very weak.

Crystal packing contacts shorter than 4 Å for non-hydrogen atoms were analysed using X-PLOR. The majority result from the dimer interaction but because changes will occur with refinement no analysis will be performed.

3.6.1 Comparison of the final BlgZ model with BlgY and Rbp

The comparison of the model obtained for BlgZ with the existing BlgY model was done in order to evaluate the relevance of the changes introduced during this work. No comparison to the model available for the published structure of BlgZ (Monaco, *et al.*, 1987) was done, the reasons being:

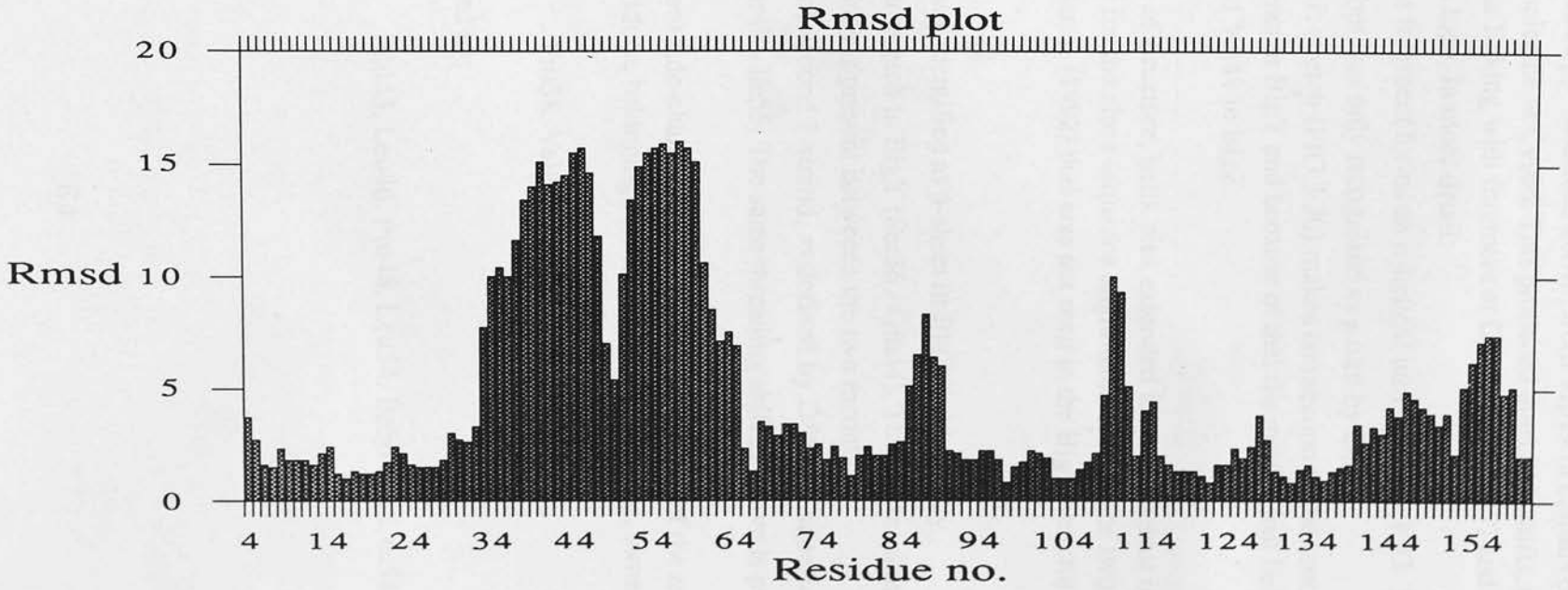
- the model contained information from the BlgY model.
- as demonstrated by the molecular replacement work done in parallel and described in section 3.8, the published BlgZ model seems to have little similarity with the structure here described.

The models of BlgZ and BlgY were compared by determining the root mean square difference (rmsd) between the coordinates of the same residues. The results are plotted in FIG. 3.19. For non-refined, medium resolution structures like these two it seems reasonable to only comment about residues where the main-chain rmsd is above 2.5 Å. The regions this way selected are, in general, the ones where changes in sequence threading where applied or where main-chain density for BlgZ did not agree with the existing BlgY model. The discussion below will deal with the whole model and will mention the analysis results from DSSP, a program by Kabsch and Sander, (1983). The residue numbering is from the BlgZ model, unless otherwise stated.

Thr4 - Asp28 - No major changes.

Ile29 - Gly64 - The fragment from Ile29 to Gly64 presents the largest changes between the two models, as much as 16 Å for Leu57. It includes the second and third

FIG. 3.19 Plot of the root mean square difference between the coordinates of BlgZ final model and BlgY model.



β-strands and three loops. The discrepancy was a result of the shortening of the loop between the third and fourth β-strands and the extension of the loop between the first and second β-strands. Because for the shortened loop the density was weak, building was done according to the strategy described above, from Trp19 (a conserved residue) towards the centre of the molecule at Cys66. This produced threading shifts, the extra long loop, the unambiguous linking with fragment at Cys66 (FIG. 3.17) and consequent shortening of a loop. In more detail:

Ile29 - Leu32 - This fragment forms an α-helical turn in BlgZ (FIG. 3.20), while in BlgY it is similar but was only recognized as a turn by DSSP.

Asp33 - Val41 - This stretch (FIG.3.20) makes connection to the second strand. It is much longer in BlgZ than in BlgY and because of that the residue just before the strand is Gln35 in BlgY and Val41 in BlgZ.

Interestingly, the same sort of structure, helix plus extended loop, is present in the Mup structure, and this accounts for the close sequence alignment between the two molecules described by Adams, (1992) that was not seen in the BlgY structure for this region.

Tyr42 - Thr49 - It was identified as β-sheet in BlgZ and is totally superimposable by a similar stretch in BlgY (Ser36 - Glu44). This means that a threading shift of five residues is present betweeen the two models.

Asp53 - Lys60 - This second β-strand, as defined by DSSP, is superimposed with the BlgY fragment Thr49 - Ile55. The same threading shift as above is present.

These shifts result in different side-chains being present in the pocket of the molecule. For BlgY the following residues, belonging to this part of the sequence, have their side-chain into the pocket:

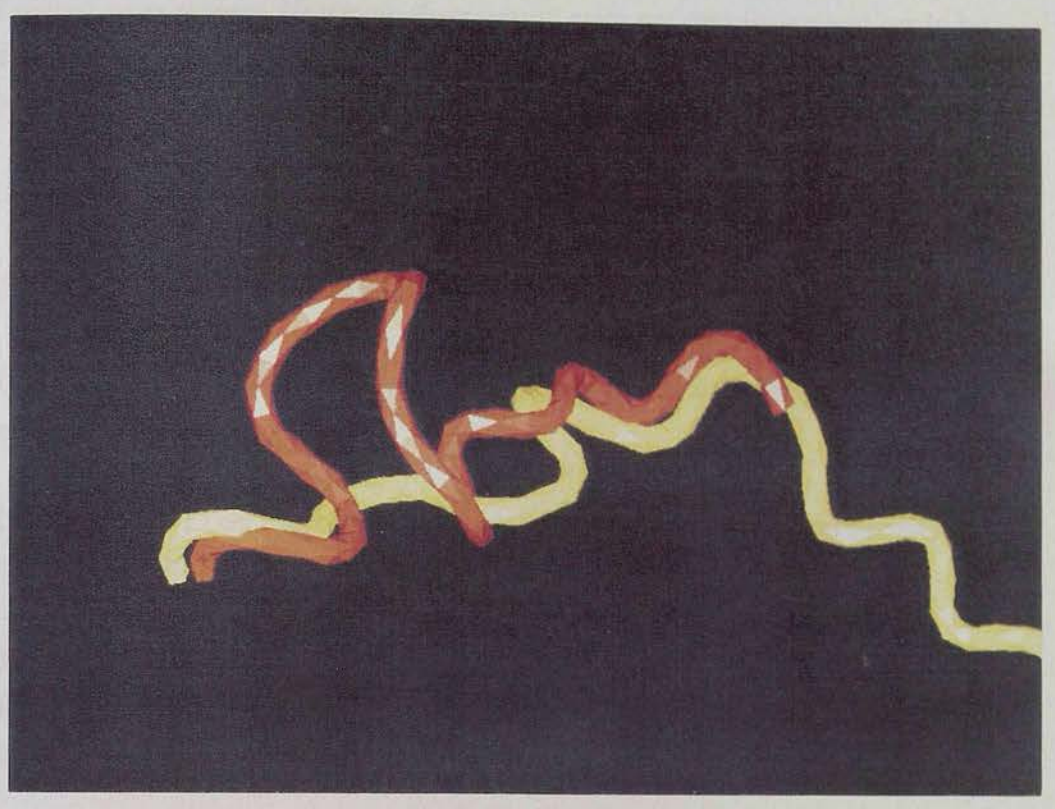
Pro38, Val41, Val43, Glu51 and Asp53.

In BlgZ they were substitued by:

Val43, Leu46, Pro48, Leu54, Ile56 and Leu58.

FIG. 3.20 Tube representation of new model and BlgY main-chain atoms between Ile29 and Arg40.

BlgZ new model in red, BlgY in yellow. The N-terminus of both fragments is towards the left of the picture.



It is clearly seen that in BlgY there are two charged residues present in the hydrophobic environment, while in BlgZ no such feature is observed. This results in an increase of the hydrophobicity of the pocket of BlgZ and confers some credibility to the changes between models.

Thr49 - Gly52 - Forms the small loop betwen the third and fourth strands which is equivalent to a slightly different shaped loop in BlgY (Glu45 - Pro48).

Trp61 - Glu65 - This loop (FIG. 3.21) is much smaller in BlgZ than in BlgY, where it is eleven residues long (Ile56 - Glu65). Part of this BlgY loop there are three solvent exposed residues Ile56, Leu57 and Leu58 which are not present in the same situation in BlgZ. In BlgZ, Trp61 is present in a small cleft closer to the "bulk" of the molecule, while in BlgY, it is quite exposed to the solvent. This change correlates with some spectroscopic data which suggest that the two tryptophans have similar environments (Mills, 1976). At the end of this loop the models are back in frame with each other.

Glu65 - Lys75 - In this stretch that includes a β-strand, Cys66 forms a disulphide bridge with Cys160 in both BlgZ and BlgY. The residues Ala67-Lys69 follow different paths such that in BlgZ it is Lys69 rather then Lys70 which has access to the binding pocket. This residue is important because a lysine has been implicated in an interaction with retinol (Horwitz and Heller, 1974). Cho, *et al.*, (1993) tried to establish the importance of the pocket and in particular of Lys70 for the binding of that ligand, by site-directed mutagenesis. However, their results were not conclusive and could be explained by having concentrated their efforts in the wrong residue. This stretch in BlgZ is out of frame with BlgY from Lys69 up to Lys75.

Thr76 - Asp85 - Similar loop and β-strand.

Ala86 - Glu89 - Loop without much density in BlgZ.

Asn90 - Glu108 - Includes two β-strands and one loop that suffered just small adjustments.

Asn109 - Leu117 - Loop that has changed place. In BlgZ there was just density enough to guess the main-chain.

Val118 - Glu127 - β-strand and loop, that includes one of the residues involved in the well defined disulphide bridge Cys106-119 and the thiol of Cys121. Only small adjustments were observed.

Val128 - Asp129 - Connection to α-helix. It was to decided make the connection in a slightly different way from BlgY. Not much density.

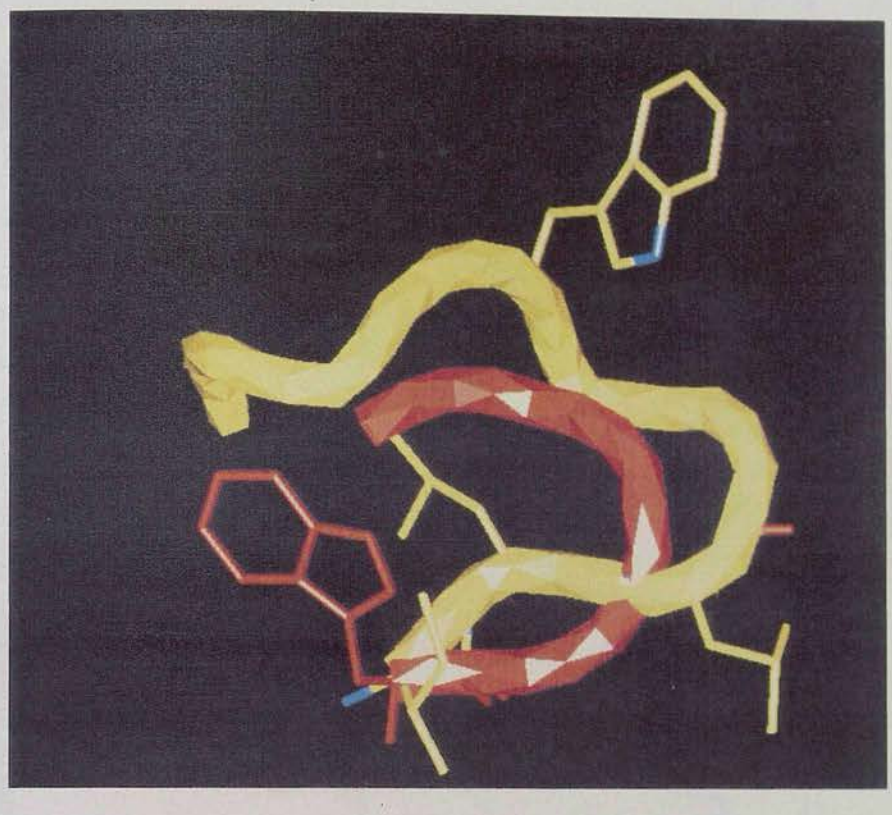
Asp130 - Ala142 - This stretch forms the long α-helix. There were no changes introduced except the addition of an extra turn to the end of the helix. This extra turn resulted from the inclusion of residues Leu140, Lys141 as part of the α-helix.

Leu143 - His146 - The loop is different in BlgZ from BlgY.

Ile147 - Phe151 - This is the last β-strand, it establishes the interaction with the other molecule in the dimer. The residues involved in the interaction have changed due to the two residue thread change towards the C-terminus.

FIG. 3.21 Loop Trp61-Glu65 in the final model

Tubular representation of main-chain atoms of BlgZ final model (red) -Trp61 to Glu65 - and BlgY (yellow) - Ile56 to Glu 65.



Asn152 - His161 - It includes a new loop in BlgZ, with a helical turn between Pro153 and Leu156. Cys160 forms the disulphide bridge Cys66-160 in both models.

The structures of BlgY, Rbp and BlgZ model were run through the program PROFILE, which evaluates and scores each residue according to its environment (Lüthy, *et al.*, 1992). The Rbp structure was analysed as a control of a refined lipocalin. The environments are evaluated in terms of the area of the residue that is buried, fraction of side-chain area covered by polar atoms and local secondary structure. The total score for the whole sequence is a function of the how consistent the sequence and fold are with each other, meaning that wrongly traced structures will score lower. The total score for BlgY was 27.14, for BlgZ final model 55.37 and for Rbp 48.09. By averaging the scores of 21 sequential residues and attributing this value to the central residue of that stretch it is possible to detect wrong parts of the structure. This was done for BlgY, BlgZ and Rbp and the result is presented in FIG. 3.22. For correct structures this value is usually between 0.2 and 0.5 and never close to zero or negative, just as demonstrated in the Rbp plot. The plot of BlgY shows two marked

FIG. 3.22 Plots of the profiles of Rbp, BlgY and BlgZ

Plot of BlgY and BlgZ final model profiles averaged over a window of 21residues. Continuous line BlgY. Broken line BlgZ.

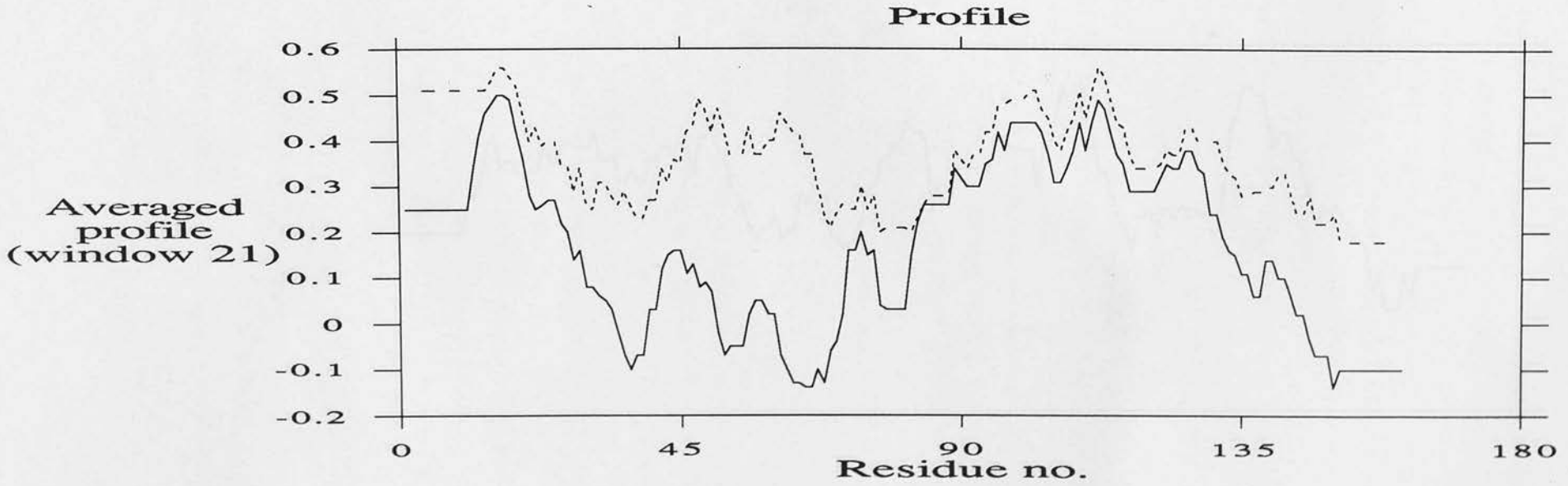
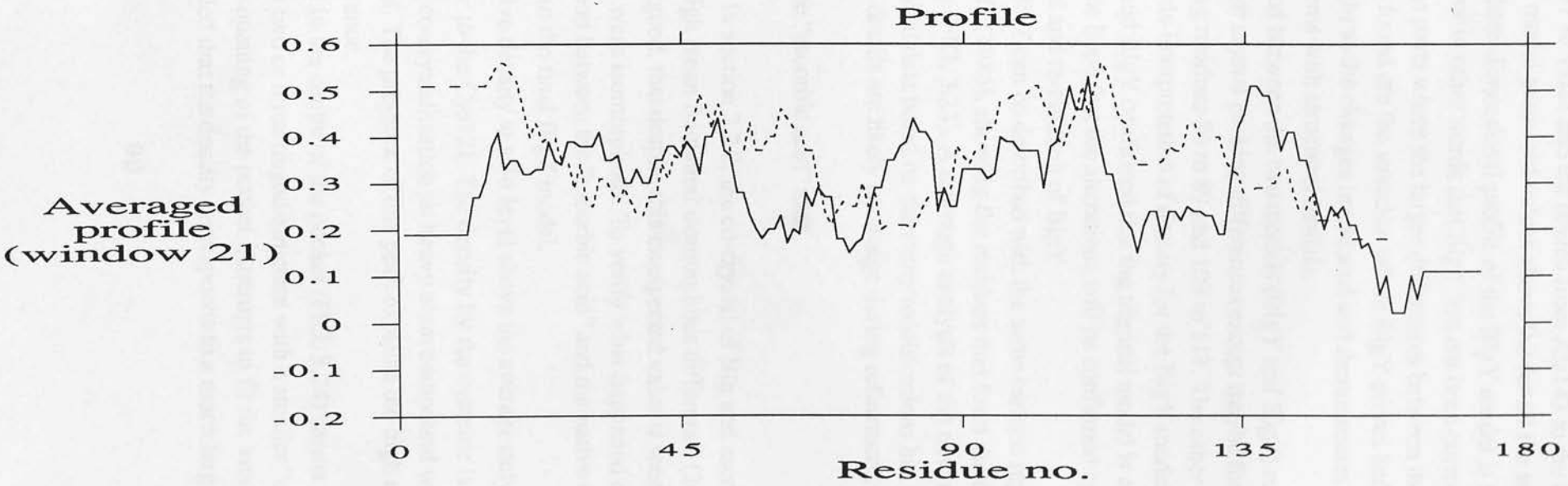


FIG. 3.22 Continuation

Plot of Rbp and BlgZ final model profiles averaged over a window of 21 residues. Continuous line Rbp. Broken line BlgZ.



stretches where the values are close to zero and even negative, from Ser30 to Ile71 plus the smaller stretch from Lys77 to Val81 and the stretch from Ala143 to the C-terminus. On the other hand, the BlgZ model presented values above 0.2 for all the sequence. These results mean that the three-dimensional profile of the BlgY model is not consistent with its sequence or in other words that BlgY has not been correctly built. Also interesting is the fact the parts where the larger differences between the two models, BlgY and BlgZ, are found are the stretches where BlgY scores badly. This observation gives more weight to the changes introduced and demonstrates the value of this program to detect problems with structural models.

The changes observed between the two models (BlgY and BlgZ) can not be attributed to conformational or crystal packing differences except maybe for the positioning of loops involving residues 86 to 89 and 109 to 117. The other changes were most probably due to mis-interpretation of density for the BlgY model. The calculated profiles for BlgZ and BlgY confirmed that the trigonal model is a better structure and gave credit to the hope that the alterations will be confirmed by the full refinement of the BlgZ model and re-building of BlgY.

The final model for BlgZ can be described with the same cartoon representation used for BlgY by Papiz, *et al.*, (1986), showing the residues that form the different secondary structure elements (FIG. 3.23). A thorough analysis of the new model and a full interpretation of biochemical data based on the many modifications has not yet been done at this stage since some details are likely to change during refinement.

3.7 Preliminary analysis of the "ascorbic acid" data

As already discussed in section 3.2.2, the co-crystal of Blg and ascorbic acid resulted in a data set with a high mean fractional isomorphous difference (26%) from the native data. It was then argued, that despite this unexpected value it seemed that the structures in the two crystals were isomorphous. To verify what happened during co-crystallization, a difference map between the "ascorbic acid" and the native data was produced with the phases from the final BlgZ model.

The map shows positive density at a 3σ level above the average only in the centre of the pocket and close to the Cys121. The density by the cysteine is round shaped indicating that during co-crystallization an heavy atom compound was present and reacted with the free-thiol. The presence of this peak explains the high mean fractional isomorphous difference.

The density observed in the centre of the pocket (FIG. 3.24) seems to belong to a long flat molecule, maybe a two or three ringed-structure with a smaller "detached" group positioned towards the opening of the pocket. Attempts to fit the ascorbic acid molecule to the density revealed that the density corresponds to a much larger molecule

FIG. 3.23 Cartoon of BlgY (Papiz, *et al.*, 1986) and BlgZ secondary structure elements

Squares represent the β-strands, circles represent α-helix. The numbering indicates the residues that constitute each secondary structure element. The definition of the elements in Blgz was done with DSSP.

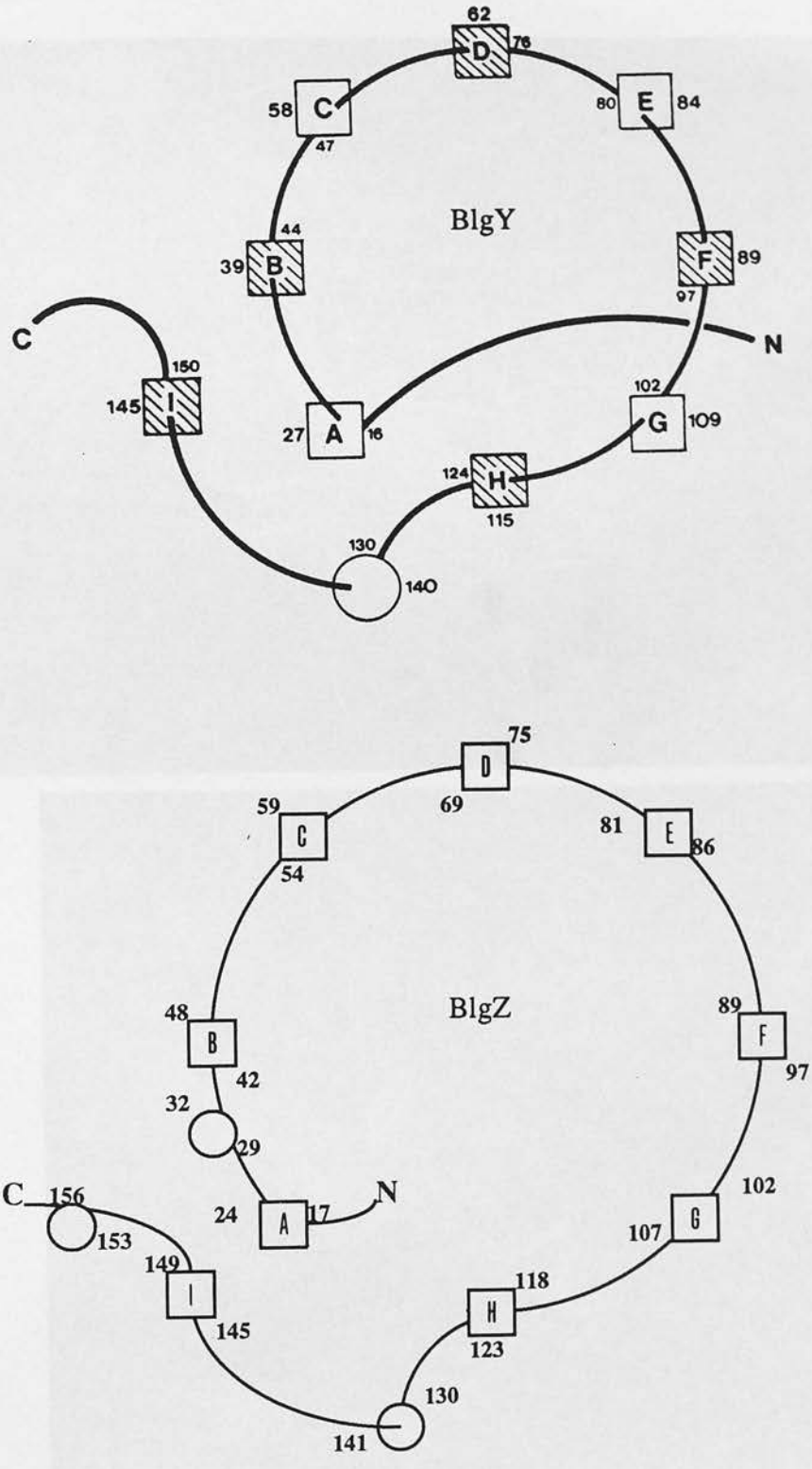
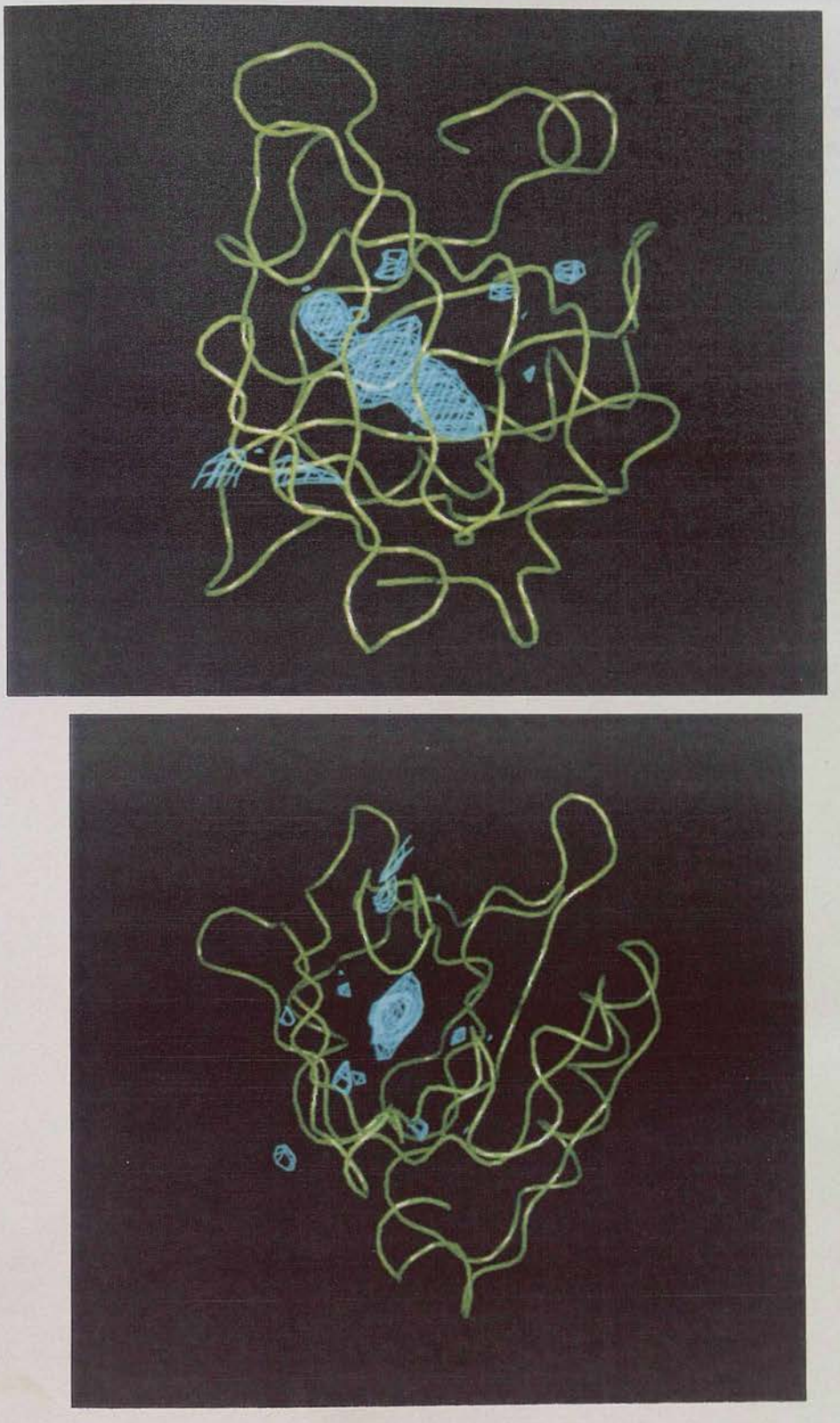


FIG. 3.24 Density found inside the pocket of Blg crystallized in the presence of ascorbic acid

In green, tubular represenation of BlgZ final model main-chain atoms. In blue the density found inside the pocket from a difference map between the ascorbic and the native data sets.
Top: view of density in between strands.
Bottom: view into the pocket.



than vitamin C. It seems improbable, as well, that this polar molecule will bind into the pocket especially since it became much more apolar due to the alterations on the threading discussed above. It seems more probable that, since the particular batch of protein was used only for this experiment, an apolar solvent used during the purification of the protein remained bound to the protein or that a contaminant with affinity for Blg was present in the co-crystallization experiment.

Despite the difficulties in determining the true nature of the ligand, the structural role of the pocket in Blg has now been established without doubt. As discussed in section 1.2, the fact that the protein binds small hydrophobic molecules has been long known but the location and environment of the binding site was never totally understood and contradictory results were leading to the idea that the pocket was not biologically functional in Blg. The finding of density has re-stated the importance of the pocket and the possibility of crystallographic analysis of the interactions between ligands and the protein.

Addendum

The calculation of difference maps for the *p*-nitrophenol (pnp) and palmitic acid co-crystals revealed density inside the pocket. Surprisingly, that density presented the same shape and was positioned in the same region of the pocket as the density observed for the ascorbic acid co-crystallization. Changes in the resolution range did not result in any alteration of the common charateristics of the three difference maps. It was verified that the information was not present as a "ghost" in the phases used, since no comparable density could be observed in a difference map calculated for a weak derivative (mean fractional isomorphous difference of 13%). It should be noted as well, that the mean fractional isomorphous difference for both the palmitic acid and pnp against the native was quite high, 26 and 30% respectively. It was concluded that changes have occurred either in the overall packing or in the protein conformation or both, which cannot be considered isomorphous. Therefore the basic principle for the use of the difference map is not valid and the density observed is a result of the overall alterations. For the full understanding of the modifications in the co-crystals, including the positioning of the ligand it will be necessary to refine the positioning of the whole molecule first and then refine the position of residues in the protein. For the moment it can only be concluded that modifications occur in the crystals upon co-crystallization and that the nature of molecular interactions of ligands to the binding pocket is still unknown.

3.8 Molecular replacement

In parallel to the structure solution by isomorphous replacement, molecular replacement was also undertaken to obtain approximate phases from BlgY (bovine beta-lactoglobulin - crystal form $B22_12$) and Mup (major urinary protein) models (see section 1.1). Mup was chosen since it was demonstrated to be one of the lipocalin family members most closely related to Blg (Ali, 1989; Adams, 1992). The model was kindly provided by A.C.T. North and had been refined to ~27% at 2.4 Å with X-PLOR.

The molecular replacement strategy consisted of both an "automatic" and a "manual" approach. The "automatic" molecular replacement was carried out mainly with the package X-PLOR using the independently determined BlgY model and the Mup model and is described in sections 3.8.1 and 3.8.2. The "manual" procedure was devised because it was foreseen that some difficulties would be encountered when looking for a solution using a traditional molecular replacement procedure with non-fully refined models (Tickle, 1992; Dauter, 1992) of proteins mainly constituted by β-sheet (Blundell and Tickle, 1985). The "manual" approach involved the use of the derivative heavy atom constellations in each of BlgY and BlgZ crystal forms as a basis for the determination of the rotation and translation parameters and will be discussed in more detail in section 3.8.5.

3.8.1 X-PLOR rotation function

For a predominantly β-sheet protein, Blundell and Tickle, (1985) recommended that a complete data set to roughly 3 Å be used. As the native data set were incomplete to this resolution, missing reflections were supplemented (see TABLE 3.9) from a data set of a very weak derivative (Riso 13%), shown subsequently by difference Fouriers 1 to be genuinely isomorphous with the native except for the presence of an heavy atom

with low occupancy.

The procedure is a real space Patterson search (Huber, 1985). In other words the Patterson function of the model, placed in a cell with P1 symmetry, and including around 70-80% of the model (Blow, 1985), is rotated over the Patterson map calculated from the measured intensities. The rotation is performed over a particular range of Eulerian angles defined by the asymmetric unit of rotation space and a correlation function between the maps is calculated at intervals. The origin peak is removed from the maps and only a set number of the strongest features is used. The correlation function calculated is the product function of the two maps.

Most of the recommended default values presented in the X-PLOR manual were used. To summarize, the top 6000 peaks in the rotation map were selected, the threshold of the peak height was set to 0.0, a 2.5° grid interval was used for the search. Data with spacings in the range 8 Å to 4 Å were used, where the low resolution limit was chosen to

TABLE 3.9 Completeness of the native data and supplemented data sets

Bins of resolution (Å)	native % complete	supplemented % complete
6.4	95.4	96.0
5.0	98.5	99.1
4.3	97.8	98.5
3.8	96.0	97.8
3.5	95.2	96.5
3.2	90.4	93.9
3.0	83.0	87.1

remove the solvent and associated salt features from the Patterson maps (Blow, 1985), while the high resolution limit was the maximum (see above) allowed by the restrictive computer memory available.

The maps were restricted to a maximum distance of 18 Å from the origin, a value considered enough to include most of the intra-molecular vectorsbut exclude most inter molecular vectors and a minimum distance of 5 Å chosen for the prupose of origin removal. The Lattman angle range for the search was established (Rao *et al.*, 1980) to be 0° to 120° for Θ-, 0° to 90.0° for Θ2 and 0° to 720° for Θ+.

Prior to the use of Mup as a search model, its coordinates were aligned with those of the BlgY model using the LSQ facility of the program O (Jones, *et al.*, 1991). This step was intended to facilitate the recognition of real solutions by comparing the

results from the two models.

The results for the rotation functions of both search models, BlgY and Mup were never clear since the top positions were not obviously separated from other solutions. The ten top positions for the two search models are presented in TABLE 3.10. It is apparent that the solutions are in the same area of space and it seems reasonable to assume that the real solution will be among the top 100.

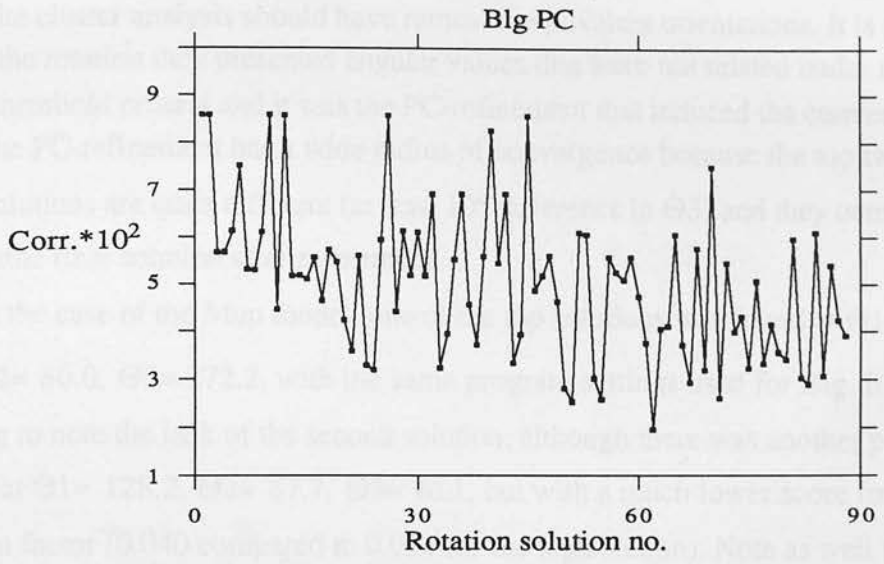
TABLE 3.10 Top rotation solutions for Mup and BlgY models

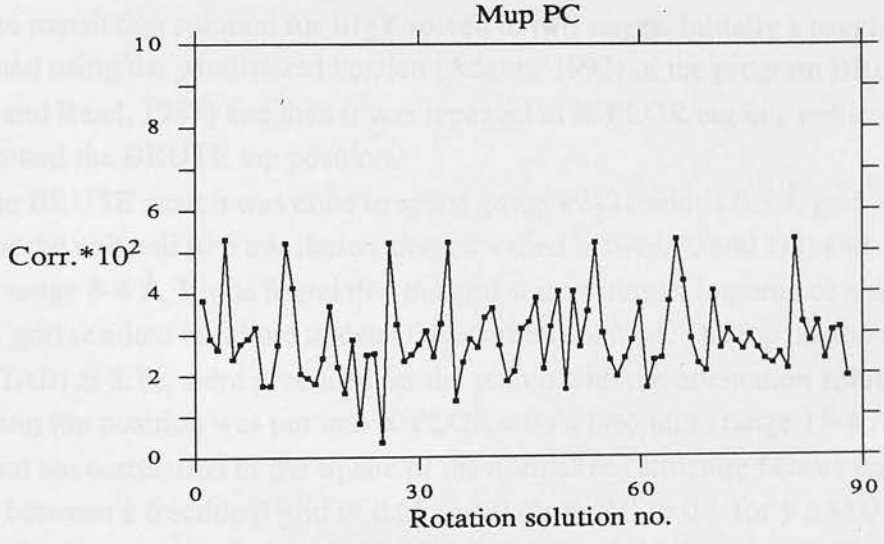
BlgY angles	Rotation correl. (Sigma=0.312)	Mup angles	Rotation correl. (Sigma=0.301)
314.3, 87.5, 270.7	3.615	143.7, 60.0, 58.70	3.395
313.2, 85.0, 260.7	3.349	142.6, 60.0, 117.6	3.393
314.1, 87.5, 281.3	3.349	144.4, 57.5, 87.1	3.305
313.8, 87.5, 292.0	3.430	314.4, 82.5, 275.6	3.303
316.2, 85.0, 237.4	3.354	142.9, 57.5, 142.9	3.280
312.9, 90.0, 251.1	3.335	145.7, 62.5, 35.7	3.262
313.5, 87.5, 302.1	3.260	146.2, 57.5, 125.3	3.236
133.5, 90.0, 93.5	3.243	146.9, 60.0, 46.9	3.236
310.0, 82.5, 213.3	3.226	145.8, 62.5, 105.8	3.212
133.7, 85.0, 36.2	3.213	142.8, 65.0, 80.4	3.186

The rotation was immediately followed by PC-refinement (Patterson correlation refinement) on all the peaks produced in the previous step. This procedure correlates the observed normalized structure factors and the normalized structure factors calculated from the rotated model. The latter are a function of the overall orientation parameters and of the translation and rotation parameters for selected rigid bodies. Thus the procedure allows for the filtering of a subset of orientations and it actually improves both the overall orientation parameters and packing relationships of different parts of the molecule. The range of resolution used was extended to 10-3.2 Å to try to use all the available information in determining the orientation. Because of the structural characteristics of the members of the lipocalin family as compact and single-domain proteins, no fragmentation of the molecules was conceivable and so the refinement was applied only to the overall rotational parameters. The results for the two search models are shown in FIG. 3.25.

FIG. 3.25 PC-refinement solutions for BlgY and Mup

Plot of the Patterson correlation value for BlgY and Mup after refinement of the rotational parameters. The correlation values were plotted according to the ranking of the rotation function solutions.





For BlgY, two solutions emerge having the same correlation factor of 0.0856:

a) $\Theta 1 = 312.1°$, $\Theta 2 = 82.0°$, $\Theta 3 = 268.1°$

b) $\Theta 1 = 132.0°$, $\Theta 2 = 98.0°$, $\Theta 3 = 31.8°$

The two solutions are in fact related by 180° rotation about the **y** axis of the trigonal space group, as determined by the ROTMAN facility in X-PLOR. This means that they are crystallographically related though the selection of the rotation asymmetric unit and the cluster analysis should have removed equivalent orientations. It is possible that after the rotation they presented angular values that were not related under the imposed threshold criteria and it was the PC-refinement that induced the convergence.

The PC-refinement has a wide radius of convergence because the top two rotation solutions are quite different (at least 10° difference in $\Theta 3$) and they converge into the same final solution after refinement.

In the case of the Mup model, one of the top solutions was found at $\Theta 1 = 313.7$, $\Theta 2 = 80.0$, $\Theta 3 = 272.2$, with the same program settings used for Blg. It is interesting to note the lack of the second solution, although there was another position emerging at $\Theta 1 = 128.2$, $\Theta 2 = 87.7$, $\Theta 3 = 40.1$, but with a much lower score for the correlation factor (0.040 compared to 0.053 for the top solution). Note as well that the scores for Mup are understandably lower than those obtained for BlgY since the sequence identity is reduced.

3.8.2 Translation solution

The translation solution for BlgY solved in two stages. Initially a rough solution was obtained using the parallelized version (Adams, 1992) of the program BRUTE (Fujinaga and Read, 1987) and then it was repeated in X-PLOR but in a reduced volume around the BRUTE top position.

The BRUTE search was done in space group $P3_221$ with a 0.5 Å grid spacing over half of the unit cell (the translation along **z** varied between 0 and 1/2) and resolution range 8-4 Å. It was found that the grid was of utmost importance since the use of 1 Å grid resulted in failure to detect the correct solution. The top positions shown in TABLE 3.11, were produced for the search with the orientation solution a). The resulting top position was put into X-PLOR with a resolution range 15-4 Å , a grid of 0.5 Å and the correlation of the square of the normalized structure factors was computed between a fractional grid of 0.65 to 1.0 for **x**, 0.0 to 0.1 for **y** and 0.10 to 0.25 for **z** for orientation a) and 0.0 to 0.15 for **x**, 0.0 to 0.15 for **y** an -0.10 to 0.10 for **z** in the case of orientation b). The results for the two orientation solutions are

TABLE 3.11 Top BRUTE translation positions for BlgY, solution a)

The grid on the sections is orthogonalized with **a** parallel to translation x and **c** parallel to translation z. z was calculated in 200 sections and x and y in 100 sections.
The starting position was -1,-1,-1.

Position (Maximum of section)	Translation correlation (Sigma=0.0613)
100, 2, 31	0.4593
92, 97, 27	0.4350
85, 96, 70	0.4238
59, 30, 98	0.4216
27, 64, 56	0.4147
51, 29, 41	0.4090
17, 51, 8	0.4087
17, 62, 21	0.4054
93, 98, 40	0.4020
53, 20, 37	0.4020

shown in TABLE 3.12. It is evident that the top solutions are very close to each other and differ only by fractions of an Ångstrom in **x** and **y**, while **z** is constant. This can be explained by the way the molecule is oriented in the cell, the **c** axis collinear with a fictitious molecular axis that goes through the pocket, such that slight adjustments in **x** and **y** compensate for the "breathing" of the β-strands.

The packing values, which correspond to the solvent content in the unit cell, were very sensitive to the validity of the solution since only the correct one gave values close to 42%, which is consistent with the calculations of Matthews' Vm (Matthews, 1968). The residual was also diagnostic, although not as evident, only dropping below 55% for the true solution.

The Mup rotation solution was submitted to an X-PLOR translation search consisting of four scans, each of 1/8 of the asymmetric unit (used because of demand on computer time and memory) on a 0.5 Å grid. The translation search produced the solutions shown in TABLE 3.13. The chosen solution is the first solution in the top table. It is not clear that the true solution is present because both the packing (39%) and the residual (54.8%) are not as good as for Blg. The packing function is clearly higher for the chosen solution but the correlation function is worse, an explanation could be that the packing function values are relative to the unit cell volume while the correlation function is probably scaled internally and so cannot be compared between different scans. However, by applying the rotation and translation parameters found for Mup to

TABLE 3.12 X-PLOR translation function solutions for BlgY orientations

The results are for the two orientations defined for BlgY. T is the translation correlation function value and P the packing function, which is equivalent to the solvent content of the unit cell. The positions are defined in Ångstroms as orthogonal coordinates, with **a** parallel to x translation and **c** parallel to z translation.

Orientation solution a)

Position	T[a]	P
53.037, 0.425, 16.905	0.3929	0.4141
52.798, 0.425, 16.905	0.3917	0.4132
52.920, 0.213, 16.905	0.3915	0.4134
52.681, 0.213, 16.905	0.3877	0.4130
53.160, 0.213, 16.905	0.3876	0.4130
52.914, 0.638, 16.905	0.3858	0.4145
53.276, 0.425, 16.905	0.3836	0.4145
53.153, 0.638, 16.905	0.3820	0.4146
52.558, 0.425, 16.905	0.3816	0.4124
52.804, 0.000, 16.905	0.3795	0.4135

a- Sigma of translation function =0.050

Orientation solution b)

Position	T[a]	P
0.123, 1.063, 1.753	0.3947	0.4132
0.245, 1.275, 1.753	0.3944	0.4134
0.368, 1.063, 1.753	0.3915	0.4130
0.000, 1.275, 1.753	0.3900	0.4140
0.491, 1.275, 1.753	0.3892	0.4115
0.245, 0.850, 1.753	0.3867	0.4131
0.368, 1.488, 1.753	0.3855	0.4138
-0.123, 1.063, 1.753	0.3854	0.4132
0.000, 0.850, 1.753	0.3845	0.4133
0.123, 1.488, 1.753	0.3830	0.4112

a- Sigma of translation function =0.063

TABLE 3.13 X-PLOR translation solutions for Mup

The two tables show the five top solutions of two 1/8 of the unit cell scans (see text for more details). The real solution is present in the top table. T is the translation correlation function value and P the packing function, which is equivalent to the solvent content of the unit cell. The positions are defined as Ångstroms of orthogonal coordinates, with **a** parallel to x translation and **c** parallel to z translation.

Search x= 0.75-1.0, y= 0.0-1.0, z= 0.0-0.5

Position	T[a]	P
26.152, 45.310, 16.100	0.3276	0.3858
26.032, 45.518, 16.100	0.3268	0.3859
25.911, 45.310, 16.100	0.3261	0.3853
25.432, 46.558, 30.691	0.3259	0.3144
52.312, 0.000, 30.691	0.3259	0.3147

a- Sigma of translation function =0.049

Search x= 0.50-0.75, y= 0.0-1.0, z= 0.0-0.5

Position	T[a]	P
26.073, 14.549, 12.075	0.3444	0.2829
26.314, 14.549, 12.075	0.3436	0.2825
26.193, 14.341, 12.075	0.3435	0.2823
25.952, 14.341, 12.075	0.3423	0.2823
26.194, 14.757, 12.075	0.3404	0.2835

a- Sigma of translation function =0.048

BlgY it was possible to verify that the "bad" residual and packing values are only caused by the low identity between the two proteins, indicating therefore that the molecular replacement has been solved for Mup too.

3.8.3 Rigid body refinement

The X-PLOR rigid body refinement routine was used with the entire molecule refined as a single rigid body.

Refinement of the positional and orientation parameters of the top translation

solution from BlgY rotation solution b) resulted in a drop of the residual from 51.9% to 50% with 30 cycles at 15-6 Å and then down to 49.2% with 50 cycles at 15-3.2 Å. The $2F_{obs}$-F_{calc} Fourier synthesis created from this model presents very reasonable density covering the whole protein. The same happened for solution a).

For Mup, rigid body refinement reduced the residual from 54.8% to 53.5% after 50 cycles at 15-3.2 Å.

3.8.4 Distinction between the space group enantiomorphs

The space group choice between $P3_221$ and $P3_121$ can be done at the stage of the translation function. When using a real space translation function the search vector is applied to the model and to its symmetry equivalents, thus the correlation depends on the position of all the molecules in the unit cell. It is then possible to distinguish a left from a right handed arrangement of the molecules since the correct arrangement will produce a higher correlation and a lower residual.

This was verified by doing the search in BRUTE with BlgY. The orientation matrix for solution a) was used and the search was done in each enantiomorphic space group. For $P3_221$ the top solution had a correlation value of 0.4593 and the residual was calculated to be 50.4%, while for $P3_121$ the top solution had a correlation value of 0.4481 and a residual of 53.1%. The choice is not as clear as in the case described by Baldwin, *et al.*, (1991) where the results with CORELS were 0.64 for the correlation factor, 44% for the residual in $P3_121$, while for the other enantiomorph the correlation was 0.428 and the residual 48.9%. However the choice of the enantiomorph $P3_221$ as the correct enantiomorph based on the translation function is consistent with the isomorphous replacement results (see section 3.3.3).

3.8.5 "Manual" procedure

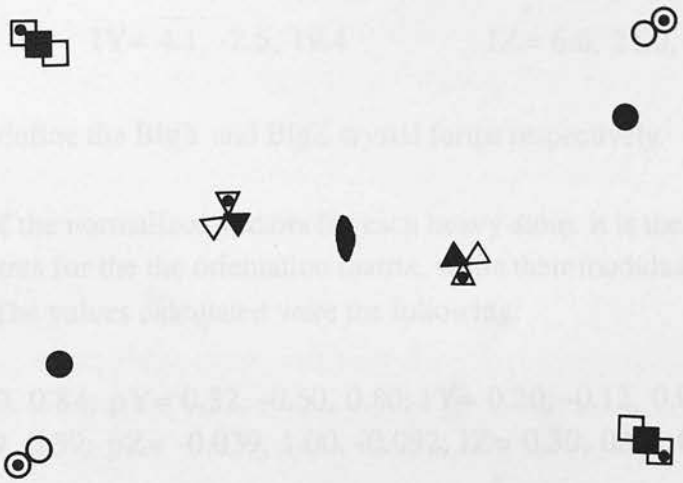
As previously discussed the manual approach was devised to surpass difficulties anticipated with a more traditional use of the molecular replacement technique.

The low resolution work described by Green, *et al.*, (1979) presents the idea that Blg is dimeric in both the trigonal and orthorhombic space groups. This idea was supported by the fact that the arrangement of the common heavy atoms in derivative crystals is identical along the crystallographic dyads, as shown in FIG. 3.26, and is consistent with the known dimeric form of the protein at neutral pH.

The "manual "approach involved the use of the coordinates of the three heavy atoms common to BlgY (orthorhombic) and BlgZ (trigonal) derivatives to determine the rotation matrix and the translation parameters. The assumption was made that the heavy atoms were bound to the same sites in both crystal forms, an assumption supported by

FIG. 3.26 Arrangement of BlgY and BlgZ heavy atoms around the crystallographic dyad (adapted from Green, *et al.*, 1979)

The figure represents a projection down the molecular dyad of the positions of the heavy atom compounds MMA (circle), HgI_4^{-2} (square) and $Pt(NO_2)_4$ (triangle) in Blg crystal lattices Z (open), Y (open with black dot) and X (filled). The dyad is represented by the ⬮.



their common arrangement around their repective dyad axes, as already described above, and by the similar reaction conditions (both pH and salt concentrations are very similar for the growth of these two crystal forms).

The mathematical procedure required the following steps:

1) Definition of a common reference system. This was achieved by aligning the **a** and **c** axes of both space groups and defining them as the x and z directions while the y direction was defined by a right hand orthogonal system. The base vectors of the coordinate system had the modulus of 1 Å.

2) Definition of the heavy atom positions or symmetry equivalents that correspond to one and the same protein molecule in each crystal form and determination of their positions relative to the new reference system. The positions of the heavy atoms of MMA, HgI_4^{-2} and $Pt(NO_2)_4$ were described for BlgY by Papiz, (1982), and by inspecting a plot with all the equivalent positions, three were chosen that appeared to have a reasonable arrangement around the 2-fold axis (the **c** axis in this case). The same was done in BlgZ for positions arranged around the crystallographic dyad at z=0. In this crystal form the positions of MMA and Pt $(NO_2)_4$ had been determined at a resolution of 3.5 Å, as described in section 3.3.4, while the position of HgI_4^{-2} was obtained from Green, *et al.*, (1979). Thus the vectors determined were:

for MMA	**MY**= 13.4, 2.6, 21.2	**MZ**= 6.1, 24.2, 10.5
for Pt $(NO_2)_4$	**PY**= 7.0, -10.9, 17.6	**PZ**= -0.9, 22.7, -2.1
for HgI_4^{-2}	**IY**= 4.1, -2.5, 19.4	**IZ**= 6.6, 21.0, 0.223

where **Y** and **Z** define the BlgY and BlgZ crystal forms respectively.

3) Calculation of the normalized vectors for each heavy atom. It is the direction of these vectors that matters for the the orientation matrix, while their modulus is a function of the translation. The values calculated were the following:

mY= 0.53, 0.10, 0.84; **pY**= 0.32, -0.50, 0.80; **iY**= 0.20, -0.12, 0.97
mZ= 0.22, 0.89, 0.39; **pZ**= -0.039, 1.00, -0.092; **iZ**= 0.30, 0.95, 0.010.

4) Solution of the [Y]R=[Z] equation, where $[Y]=[\mathbf{mY}, \mathbf{iY}, \mathbf{pY}]^T$ and $[Z]=[\mathbf{mZ}, \mathbf{iZ}, \mathbf{pZ}]^T$ are the matrices composed of the vectors defined in 3) which define the orientation of the molecule in each crystal form, while R is the rotation matrix that transforms one onto the other. The solution requires the calculation of $[Y]^{-1}$ and the final result for the rotation matrix R is:

$$R = \begin{vmatrix} -0.436 & 0.406 & 0.751 \\ 0.562 & -0.362 & 0.539 \\ 0.473 & 0.850 & -0.079 \end{vmatrix}$$

5) The differences between the rotated **MY**, **IY**, **PY** vectors and **MZ**, **IZ**, **PZ** define the translation vectors. In fact it was decided to use the center of gravity of the rotated three positions in Y. This resulted in **cog**(rotated)=3.6, 21.1, 2.7 and the distance to the center of gravity of the Z crystal form is

Δ**cog**= 0.3, 1.5, 0.2. As the necessary translation was very small, the translation vector used was (0.0, 1.5, 0.0).

6) the rotation was applied to the BlgY model using X-PLOR (in fact X-PLOR required the transpose of the matrix shown in 4)) and then the result was translated. Finally 60 steps of rigid body refinement at 15-6 Å dropped the residual from 58% to 50.6% (the residual was 52.3% when the range was 15-4 Å).

The resulting 2Fobs-Fcalc map was not satisfactory since the main chain density was frequently interrupted. However, the position of the mercury atom in relation to the free cysteine, Cys121, was quite acceptable which indicates that the molecule is approximately in the correct position and that the procedure for the positioning is self-consistent.

The orientated molecule was entered into a simulated annealing procedure to relieve any twists or alterations in the β-sheet structure from BlgY to BlgZ. The procedure was run in X-PLOR with the settings presented in the manual. The final residual value was 30% and the 2Fobs-Fcalc map was perfectly acceptable. The changes between the model before and after the simulated annealing can be described as "breathing" movements of all the β-strands. It was clear that no drastic movement of the model had occurred.

As a way of checking the reliability of the model after simulated annealing, the "manually" positioned molecule was randomly rotated over itself and put through a simulated annealing procedure. The center of gravity was determined with the use of the program COGCAL, (personal communication by Dr. P. Taylor), and the molecule was rotated around that point by 40° around **z** and 20° around the new **x**. After the procedure the residual was again 30%, but the empirical energy values were much higher and the molecule was visibly deformed. Furthermore the 2Fobs-Fcalc map was poor, showing , for example, no density for the disulphide bridge 106-119. This test lent confidence to the "manual" solution and the model produced from it .

3.8.6 Discussion

The work done with the molecular replacement techniques, although not used in the calculation of density, provided confirmation of the space group enantiomorph and of the suspicions that severe problems existed in the structure presented by Monaco, *et al.*, (1987).

The former work at medium resolution in the trigonal space group (Monaco, *et al.*,1987) concluded that the Blg molecule was not present as a dimer since the interactions among molecules related by 2-fold axes were found to be very distant and

no special interactions were observed between pairs of molecules. Some short contacts were described, "...the contacts are so tight that interpretation of the electron density map is not straightforward." (Monaco, *et al.*, 1987) and it was concluded that the protein was in some state of polymeric aggregation.

The molecular replacement solution obtained during this study shows a clear association between two monomers and that this association is related by a crystallographic dyad. It also shows that the dimer interactions appear to be generally the same in BlgY and BlgZ, being established by the β-strand closest to the C-terminus. In fact the overall packing of the molecules is entirely different from that described by Monaco, *et al.*, (1987), and no evidence for any kind of polymeric aggregation other than the described dimer formation, is observed. A comparison of the positioning in the unit cell of the present structure solution with that of Monaco, *et al.*, (1987) is illustrated in FIG. 3.27.

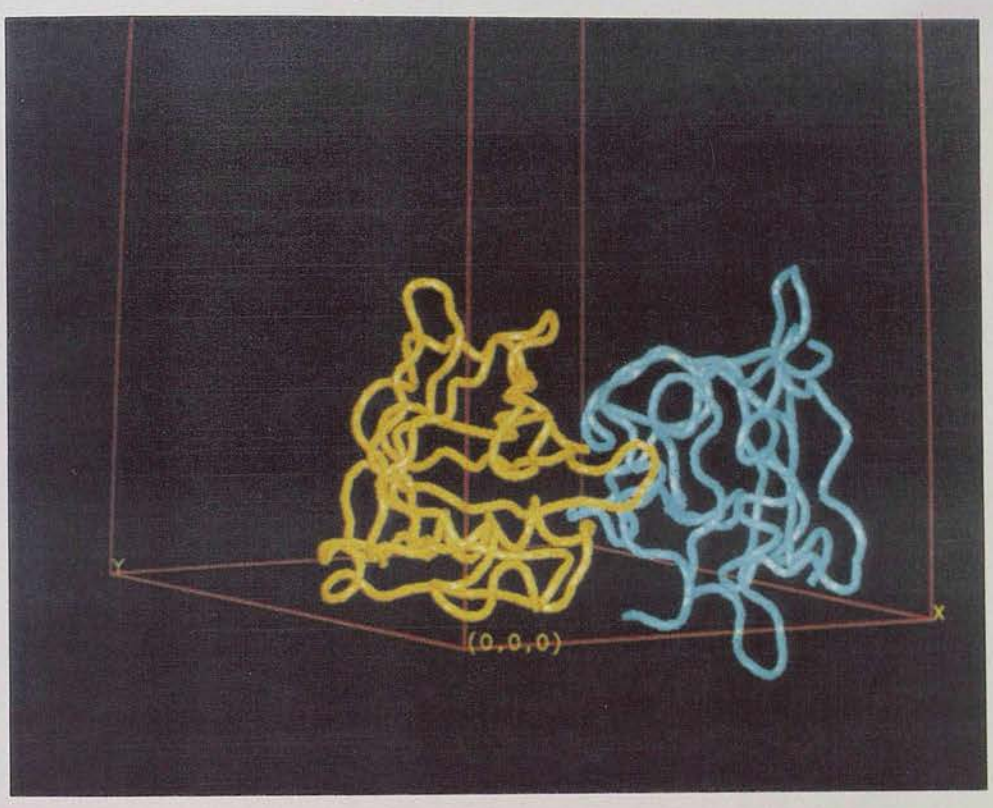
The comparison of the "manual" and automatic solutions reveals the power of the molecular replacement algorithms in X-PLOR, especially in their capacity to refine the orientation matrix. Despite the use of information from the heavy atoms and rigid body refinement, the "manual" solution only produced an acceptable map after simulated annealing. This can be explained by the small misorientation of the "manual" solution observed when compared with the automatic one. The "manually" produced solution is located in the same position as the "automatic" solution, however a slight but obvious difference in orientation is apparent between the two final models. It would be expected that rigid body refinement could ease this misorientation; instead there was a requirement for application of shifts to the each β-strand during the simulated annealing procedure to bring the resulting model closer to the automatic solution. Based on considerations of map quality the "automatic" solution was chosen as the better representation.

The translation function permitted the confirmation of $P3_221$ as the correct space group, since both the residual and the correlation values of the top solution are favourable to the search performed with this enantiomorph.

Finally, both the Blg and Mup models produced solutions for the molecular replacement procedure. The solution obtained with Mup is not as clear as the ones for Blg but it was still possible to recognize the correct from the wrong orientation and position parameters. It will be interesting to see if the results for molecular replacement are more obvious when the refined models are used. It must be emphasized that both the BlgY and Mup molecular replacement solutions were used as helplines in the building of the model of BlgZ, as discussed in section 3.5.

FIG. 3.27 Position of molecules in two studies of BlgZ crystal form

View of the $P3_221$ unit cell section closest to the origin. The unit cell is represented by red lines with the origin and both the x and y axes marked. The diagonal between the origin and the far corner on the background is the crystallographic dyad. In yellow, a tubular representation of BlgZ new model and in blue, the tubular representation of BlgZ from Monaco, *et al.*, (1987).



Chapter 4 Experiments involving apo D

4. Experiments involving apo D

The following sections will describe the various experiments done with apo D in order to gain insight into the biochemistry of this protein. It is important to re-state that most of the experiments are only preliminary, nothing more than "opening doors" into possibly relevant aspects of apo D. The different purification methods will be commented upon, followed by the structural and biological characterizations performed.

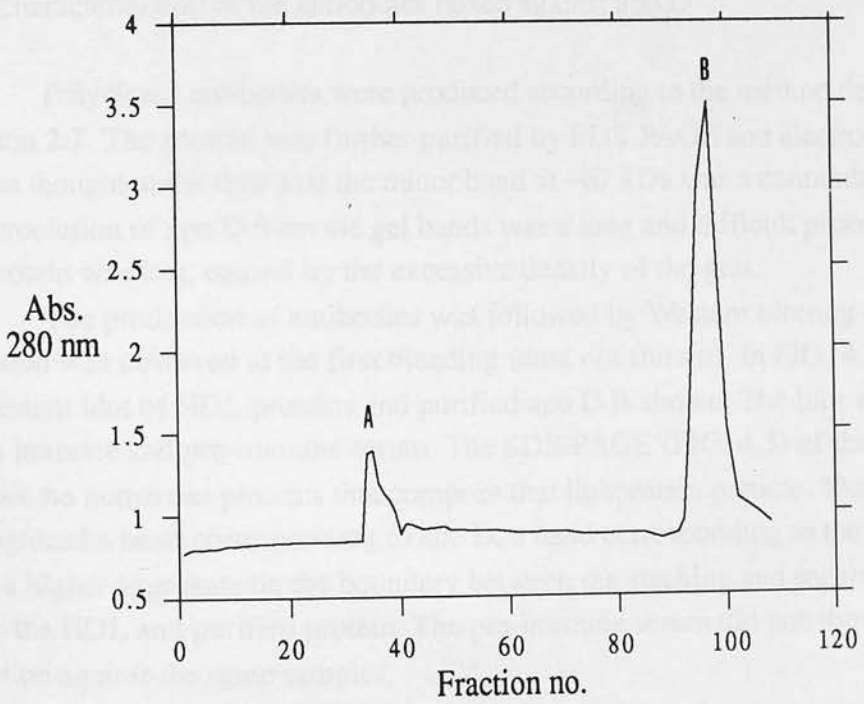
4.1 Purification of apo D from plasma

At the onset of this work the source of apo D was the plasma. The method of purification had been established since the detection and initial characterization of the protein (McConathy and Alaupovic, 1973). It was only necessary to scaleup the method (see section 2.1 and following) to purify larger amounts of protein for crystallization trials and general structural characterization.

A typical chromatography trace is shown on FIG. 4.1. Peak A was eluted with

FIG. 4.1 Chromatography trace from hydroxylapatite column

Peak A eluted with 1 mM K_2HPO_4, pH 8.0 (the loading buffer), peak B was eluted with 1M K_2HPO_4, pH 8.0.



the loading buffer (1 mM K_2HPO_4, pH 8.0), peak B was eluted with 1M K_2HPO_4, pH 8.0. The fractions of peak A and B were run in the SDS-PAGE shown on FIG. 4.2, peak A fractions contain two broad bands, while peak B fractions contain apo AI and all the other components of the HDL (see section 1.3.4.1). Apo D has been identified as the major broad band with a molecular weight at 30 kDa in peak A. The identification was confirmed (data not shown) by cross-reaction with antibodies against apo D, kindly provided by Dr. Kostner from Graz, Austria , and by confirming that the sequence (performed by Dr.C. Lopez-Otín and Mr.L.-M. Sanchez in Oviedo, Spain) of the first five residues was identical to the sequence reported in the literature (Drayna, *et al.*,1986). The band that migrates at 67 kDa was tentatively identified as one of the following: albumin, a contaminant from the plasma, LCAT (lecithin-cholesterol acyltransferase (see section 1.3.1), a contaminant from the HDL, or a dimer of apo D. The last possibility was shown to be the most probable by a series of facts:

a) no-cross reaction with antibodies raised against albumin (data not shown)

b) cross-reaction of the band with antibodies raised against protein electro-eluted from the major band (see below)

c) observation of the same extra high molecular weight, in gels of the protein purified from gross-cystic-disease fluid. This higher molecular weight band cross-reacted with antibodies raised against the plasma protein.

The yield of the procedure was 10-20%, since one liter of plasma, containing 60-100 mg of apo D (McConathy and Alaupovic, 1986), produced 10-20 mg of pure apo D.

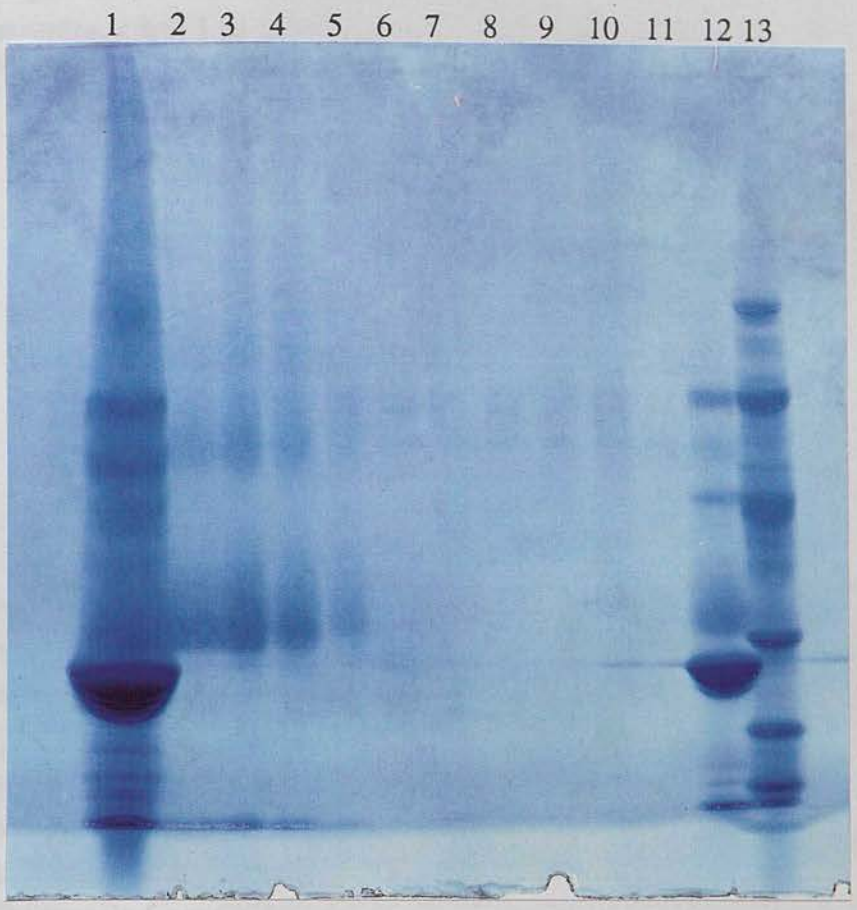
4.2 Characterization of the antibodies raised against apo D

Polyclonal antibodies were produced according to the method described in section 2.7. The protein was further purified by SDS-PAGE and electroelution because it was thought at the time that the minor band at ~67 kDa was a contaminant. The electroelution of apo D from the gel bands was a long and difficult process where a lot of protein was lost, caused by the excessive density of the gels.

The production of antibodies was followed by Western blotting and cross-reaction was observed at the first bleeding (data not shown). In FIG. 4.3, the result of a Western blot of HDL proteins and purified apo D is shown. The blot was incubated with immune and pre-immune serum. The SDS-PAGE (FIG. 4.3) of the HDL sample shows the numerous proteins that compose that lipoprotein particle. The immune serum recognized a band corresponding to apo D, a band corresponding to the dimer of apo D and a higher aggregate on the boundary between the stacking and separation gels, in both the HDL and purified protein. The pre-immune serum did not show any cross-reaction against the same samples.

FIG. 4.2 SDS-PAGE of fractions from hydroxylapatite column

12.5% SDS-PAGE. Fractions of peak A were run in lanes 2 to 6, fractions from peak B were run in lanes 10 to 12. Molecular weight markers were run in lane 13 and the loaded sample is present in lane 1. All other lanes were loaded with fractions between peak A and B. M.W. markers- 94, 67, 43, 30, 20, 14 kDa



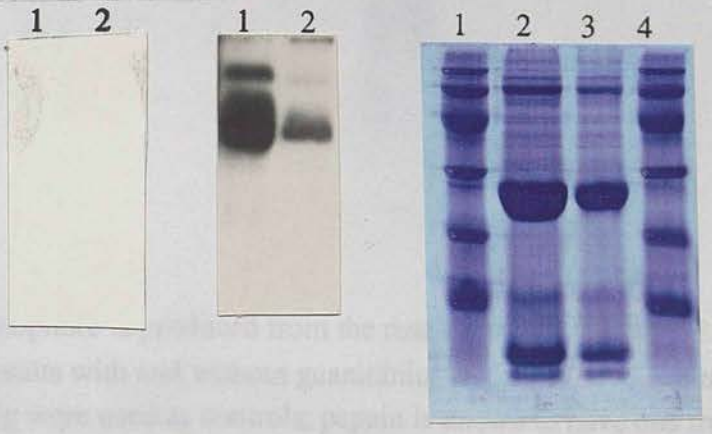
4.3 Apo D purification from gross-cystic disease fluid (GCDF)

The sequencing of the major protein (GCDF-24) from gross-cystic-disease fluid (GCDF) led to its identification as apo D (Balbín, *et al.*, 1990). Further, the fluid proved to be an excellent source of apo D since concentrations in the range of 10-50 mg/ml can be obtained.

The first attempt to isolate apo D from GCDF was done with the assistance of the workers in the laboratory of Dr.C. López-Otín in Oviedo, Spain, using the method described by Balbín, *et al.*, (1990) (section 2.3). Unfortunately, the preparative gel

FIG. 4.3 Western blot with antibodies raised against plasma apoD

Left: western blot of purified plasma apo D and HDL, incubated with pre-immune serum. Lane 1 - apo D purified from plasma, lane 2 - HDL. Middle: western blot of purified plasma apo D and HDL, incubated with immune serum. Lane 1 - apo D purified from plasma, lane 2 - HDL. Right: SDS-PAGE of purified HDL after delipidation. Lanes 2 and 3 - HDL samples, 2 and 1 µl respectively. Lanes 1 and 4 - molecular weight markers. 94, 67, 43, 30, 20, 14 kDa



filtration column used was old, not permitting good resolution of the peaks. Therefore the resultant apo D samples were heavily contaminated. Back in Edinburgh, the method previously used for the purification of apo D from plasma was slightly modified (section 2.2.2), and applied to the partially purified material. This resulted in an apparent total isolation of apo D from the other proteic components of cyst fluid. From this point onwards, apo D was routinely purified from GCDF by the single step hydroxylapatite column method.

The protein from GCDF migrates in SDS-PAGE at a lower molecular weight than the plasma protein (FIG. 4.4), which can be possibly explained by differences in glycosylation. The antibodies raised against plasma apo D cross-react against the GCDF apo D and its high molecular weight complex (data not shown).

Routinely, 10-12 mg of pure apo D were purified from 1-2 ml of GCDF.

4.4 Analysis of the cysteine residues

The five cysteines present in the GCDF apo D were characterized into three possible chemical states: free-thiol, disulphide bridge, other unknown oxidized state.

Ellman's reagent produces a chromophore allowing for the colorimetric determination of the number of free-thiols present in solution (see section 2.13.1). One

FIG. 4.4 SDS-PAGE of apo D from plasma and GCDF

12.5% SDS-PAGE. Lanes 1 - apo D from plasma, lane 2 - apo D from GCDF.

mole of chromophore is produced from the reaction with one mole of free-thiol.

The results with and without guanidinium chloride are presented in TABLE 4.1. Papain and Blg were used as controls; papain is known to have one free-thiol in the active site and one disulphide bond while Blg has one free-thiol and two disulphide bonds. It was shown that the free-thiol of Blg was not accessible to the reagent in the native conformation, a result consistent with the X-ray crystal structure that shows the Cys121 semi-buried under the α-helix (Papiz, *et al.*, 1986). On the other hand the cysteine in the active-site of papain was accessible to the reagent. Apo D behaves the same way as Blg, since no reaction was detected.

Under denaturing conditions papain still only showed one free-thiol accessible, while in Blg the free-thiol was now accessible to the reagent. Denatured apo D, however, did not seem to have any free-thiol, as no production of chromophore was detected. An increase in incubation time with denaturant did not alter the results.

An attempt to reduce the cysteines by incubating the proteins with DTT followed by extensive dialysis, before the colorimetric reaction, was not successful because the results were very dependent on the need for total removal of the reducing agent and the longer the dialysis time the greater the probability of the re-oxidation of the cysteines.

The number of disulphide bridges present in the proteins was determined by the method described by Thannhauser, *et al.*, (1984; 1987) (see section 2.13.2). By chemical modification of Ellman's reagent, 2-nitro-5-thiosulphobenzoate (NTSB) was obtained which when upon reaction with one mole of disulphide or one mole of free-

TABLE 4.1 Results of cysteine studies with Ellman's reagent

DTNB native conditions

	moles of protein	moles of chromophore	ratio
Papain	2 nmol	2.2 nmol	1.1
	4 nmol	4.4 nmol	1.1
	8 nmol	8.8 nmol	1.1
Apo D	4 nmol	0 nmol	0
	8 nmol	0 nmol	0
	10 nmol	0 nmol	0
Blg	4 nmol	0 nmol	0
	8 nmol	0 nmol	0
	8 nmol	0 nmol	0

DTNB denaturing conditions

	moles of protein	moles of chromophore	ratio
Papain	4 nmol	5.2 nmol	1.3
	8 nmol	9.6 nmol	1.2
Apo D	4 nmol	0 nmol	0
	8 nmol	0 nmol	0
	10 nmol	0 nmol	0
Blg	10 nmol	9.6 nmol	0.96
	10 nmol	9.6 nmol	0.96

ratio- number of moles of chromophore per mole of protein

thiol produces one mole of chromophore. The results are presented in TABLE 4.2. The two moles of chromophore produced for papain under denaturing conditions are due to the presence of one free-thiol, as demonstrated with Ellman's reagent, and one disulphide bridge. Apo D under denaturing conditions, presented two reactive groups,possibly corresponding to the two disulphide bonds. For Blg, under denaturing conditions, it was expected that three moles of chromophore would be produced per mole of protein, however an average of 2.5 reactive groups were determined. A possible explanation is indicated in the literature (Thannhauser, *et al.*, 1984; 1987), where the reactions that give rise to the chromophore are described as reversible at pH 9.5. Thus if the denaturant was not totally effective with Blg, the local concentration of

TABLE 4.2 Results of cysteine studies with NTSB

NTSB native conditions

	moles of protein	moles of chromophore	ratio
Blg	4 nmol	3.8 nmol	0.98
	6 nmol	6.1 nmol	1.0
Apo D	4 nmol	2.2 nmol	0.55
	6 nmol	3.2 nmol	0.53

NTSB denaturing conditions

	moles of protein	moles of chromophore	ratio
Papain	6 nmol	12.0 nmol	2
Apo D	4 nmol	7.3 nmol	1.8
	6 nmol	12.3 nmol	2.1
	10 nmol	21.0 nmol	2.1
Blg	4 nmol	9.4 nmol	2.4
	6 nmol	14.7 nmol	2.4
	12 nmol	30.1 nmol	2.6
	10 nmol	26.0 nmol	2.6

ratio- number of moles of chromophore per mole of protein

the cysteines that form one of the disulphides may have been high enough for the reactions not being totally drawn into the chromophore production. This averaged effect would result in the determination of half a mole of chromophore per mole of protein. Then the value 2.5 may include the free-thiol, detected above, one of the disulphide bonds and "half" of another disulphide bond.

In non-denaturing conditions, Blg produced one chromophore per mole of protein. The chromophore was being produced from the reaction with one of the disulphide bridges and not from the free-thiol, because the latter was not detected with Ellman's reagent. This is acceptable from the point of view of the three-dimensional structure since one of the disulphide bonds (Cys66-160) is exposed to the solvent. This conclusion when associated with the half chromophore detected under denaturing conditions leads to the argument that the "difficult" disulphide bridge (the one that is not totally detected) is the Cys106-119. This bond is situated between adjacent β-strands and so its stability is reinforced by the extensive hydrogen bonding characteristic of this

secondary structure element. For apo D under non-denaturing conditions, only half of a chromophore was detected. The existing three-dimensional model, based on the X-ray structure of insecticyanin (Ins) (Peitsch and Boguski, 1990), shows that the arrangement of the disulphide bonds is different from Blg. In apo D the disulphides are supposed to form between residues in the C- and N- terminae and residues situated in the middle of the sequence (Cys8-114 and Cys41-165). Both of these disulphide bridges are exposed to the solvent in the model. The result above, however, seems to indicate that one of the disulphides is more exposed than the other, even if not fully reactive.

The simple conclusion from these experiments is that the apo D has one cysteine not involved in disulphide bond and that under the conditions of purification used is not present as a free-thiol. Thus, in the crystallization of apo D from GCDF it would be advisable to include reducing agents during the purification and crystallization procedure so that this reactive group can be used for specific heavy-atom binding. It may happen of course, that the cysteine is oxidized *in vivo*, by being involved in thio-ester, thio-ether connections or simply is present as cysteic acid. Balbín, *et al.*, (1990) have sequenced apo D from GCDF and all cysteines were detected as cysteic acids; no mention is made of any unusual behaviour of the cysteines. Thus the possibility of the fifth cysteine sulfur being involved in covalent bonds with other molecules is very low. Just as a note, 67% of apo D in the plasma seems to be involved in inter-protein disulphide bonds with other apolipoproteins (Blanco-Vaca, *et al.*, 1990; 1992). In that work, none of the heterodimers was purified by hydroxylapatite column, which may explain why no such complexes were observed for apo D from GCDF during this work.

4.5 Protein / lipid interactions

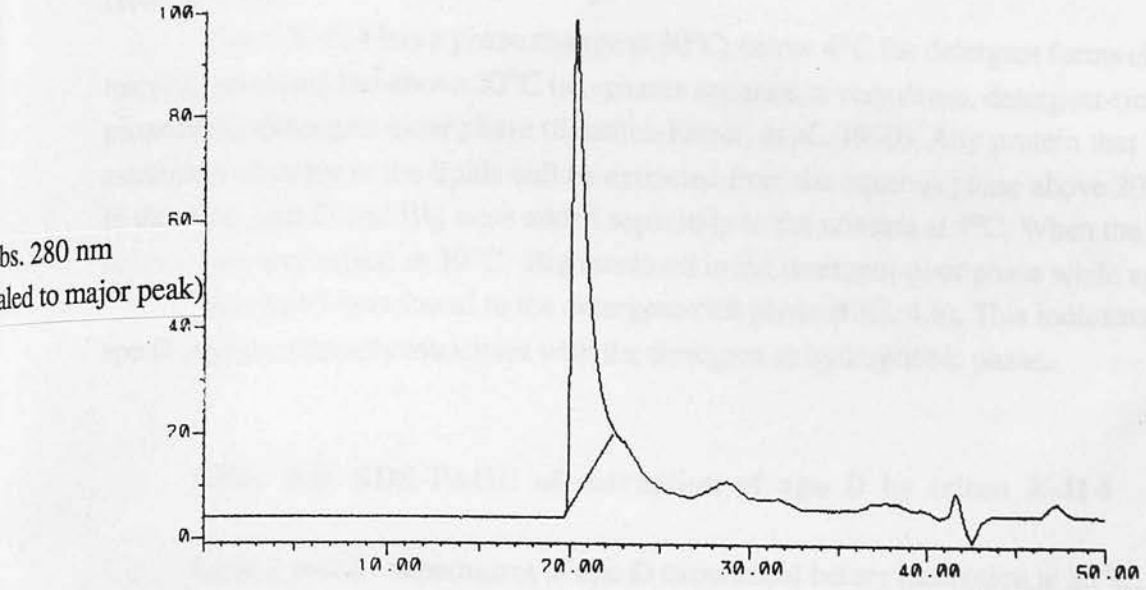
The interations of apo D with lipid vesicles have been previously observed in the formation of the lipoprotein particle, Lp-D, by disruption of HDL (McConathy and Alaupovic, 1973) and in the association of apo D to lipid vesicles described by Steyrer and Kostner, (1988).

Lipid vesicles, with DPPC and cholesterol, with or without apo D were prepared as described in section 2.14. Gel filtration trace of the vesicles (FIG. 4.5) prepared in the presence of apo D, is quite different from the one obtained from a sample with only apo D. In the sample with lipid vesicles plus protein, apo D emerged earlier than in the runs of protein only, as concluded from running the fractions from the peaks in SDS-PAGE. This seems to indicate that the protein was associated to the lipid vesicles.

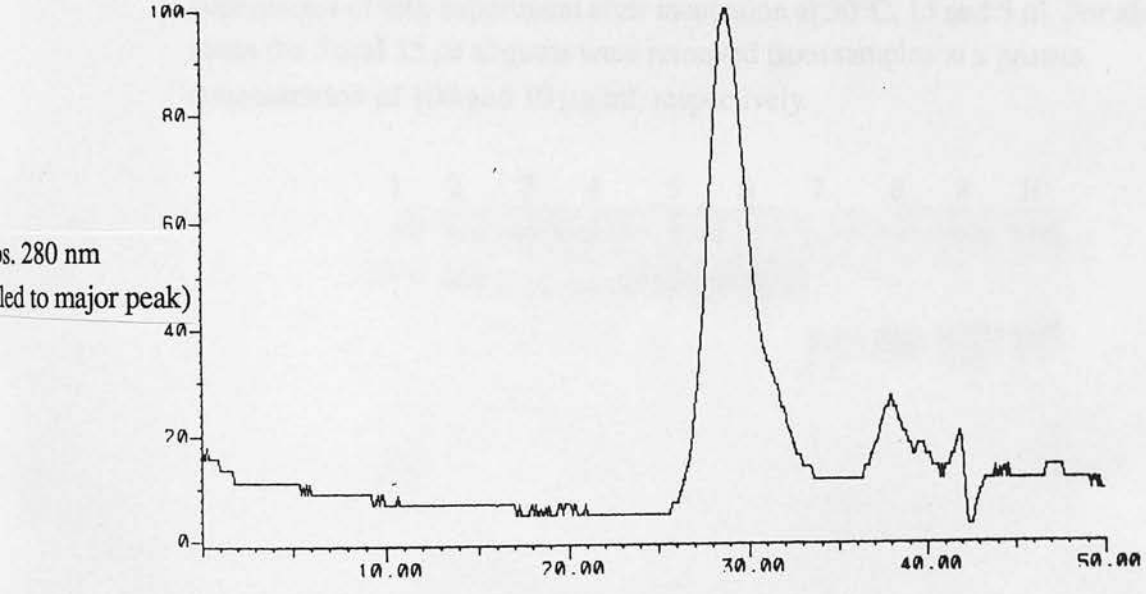
The procedure used for the preparation of the vesicles is ideal to obtain small homogeneous structures mainly because of the sonication step. However, it could be

FIG. 4.5 Gel filtration traces of the lipid/apo D interaction work

Top: trace sample containing of vesicles prepared with apo D
Bottom: trace of sample containing apo D only
For experimental conditions see sections 2.14 and 2.12



bs. 280 nm
aled to major peak)



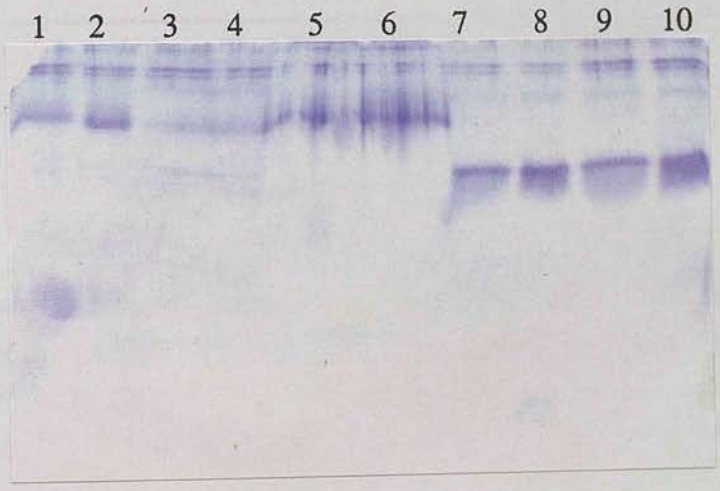
s. 280 nm
led to major peak)

considered that the formation of protein-lipid complexes was only caused by the denaturation of the protein during sonication, as the denatured protein would expose hydrophobic residues that would mediate the interaction with the lipids. To check that the association did not depend on denaturation, two different experiments were done:the protein was extracted from solution by triton X-114 as described in section 2.15, and the secondary structural elements were determined by circular dichroism (section 2.10).

Triton X-114 has a phase change at 20°C; below 4°C the detergent forms clear micellar solutions but above 20°C two phases separate, a very dense, detergent-rich phase and a detergent-poor phase (Sánchez-Ferrer, *et al.*, 1990). Any protein that associates strongly to the lipids will be extracted from the aqueous phase above 20°C. In this case, apo D and Blg were added separately to the mixture at 4°C. When the temperature was raised to 30°C, Blg remained in the detergent-poor phase while apo D was extracted and was found in the detergent-rich phase (FIG. 4.6). This indicates that apo D characteristically associates with the detergent or hydrophobic phase.

FIG. 4.6 SDS-PAGE of extraction of apo D by triton X-114

Lane 1 and 2 - supernatant of apo D experiment before incubation at 30°C, 15 and 5 µl. Lane 3 and 4 - supernatant of apo D experiment after incubation at 30°C, 15 and 5 µl. Lanes 5 and 6 - loaded with the pellet formed after incubation at 30°C of experiment with apo D. Lanes 6 and 7 - supernatant of Blg experiment before incubation at 30°C, 15 and 5 µl. Lanes 8 and 9 - supernatant of Blg experiment after incubation at 30°C, 15 and 5 µl. For all cases the 5 and 15 µl aliquots were removed from samples at a protein concentration of 100 and 10 µg/ml, respectively.



The analysis of apo D secondary structure when associated to lipids or free in solution was done with circular dichroism as described in section 2.10. The spectra of the protein without lipids, associated to DPPC and cholesterol vesicles and in the presence of 4% ethanol are presented in FIG. 4.7. The spectrum of lipid vesicles without protein showed negligible signal (data not shown) and no correction was applied to the protein-lipid complexes. By comparing the spectra it is visible that they are very similar over all the range. By analysing the spectra with the CONTIN algorithm (Provencher and Glöckner, 1981) the values for the secondary structure content were produced and the changes between native and the other two sets of conditions are presented on TABLE 4.3. The analysis at the 190 to 260 nm range provides, generally, only reliable secondary structure information on the amount of α-helix (Johnson,Jr.,1990). From the analysis of the spectra it is shown that apo D kept its integrity when associated to lipids with a small increase in the α-helix content. The apo D-ethanol mixture showed an increase in the amount of α-helix, as well. This was not observed for Blg (Dufour and Haertlé, 1990), which is supposed to have a similar structure, an increase of 3% in β-sheet was detected only at a 20% ethanol concentration while the α-helix content increased at 30% ethanol. For apo D, all the determined changes can, however, be attributed to increased noise level in the range 190-195 nm, at which the CONTIN algorithm is particular sensitive. It seems thus, based on these results alone, incorrect to withdraw conclusions about the efffect of polarity changes on the medium. It is clear, however, that the protein associates to lipid vesicles without denaturation.

TABLE 4.3 Changes in α-helix content relative to native apo D

conditions	α-helix change
pH 7.0, 4% ethanol, 0.3 mg/ml apo D, 0.02 cm cell path	+5±2.3%
pH 7.0, 0.165 mg/ml apo D (associated to lipid vesicles), 0.05 cm cell path	+3±2.3%

The DPPC and cholesterol vesicles with and without protein were observed in the electron-microscope as described in section 2.16. The samples without protein (FIG. 4.8) presented structures with an homogeneous globular aspect, with most structures showing a single bilayer although occasionally some presented multilayer

FIG. 4.7 Circular dichroism spectra of apo D, apo D associated to lipid vesicles and in the presence of 4% ethanol

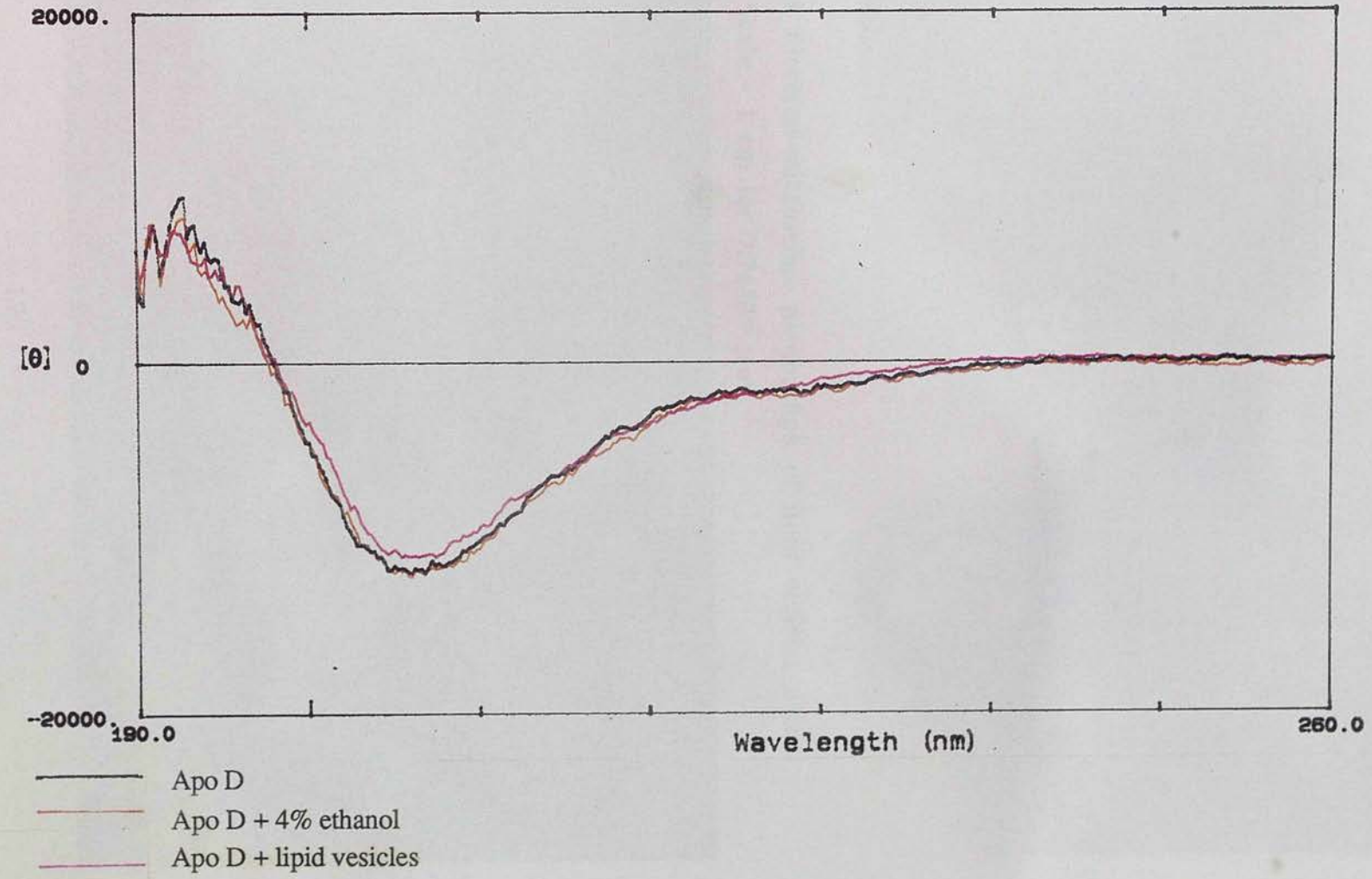


FIG. 4.8 Electron-microscope photograph of lipid vesicles without apo D. Scale - 1 cm for 0.26 μm.

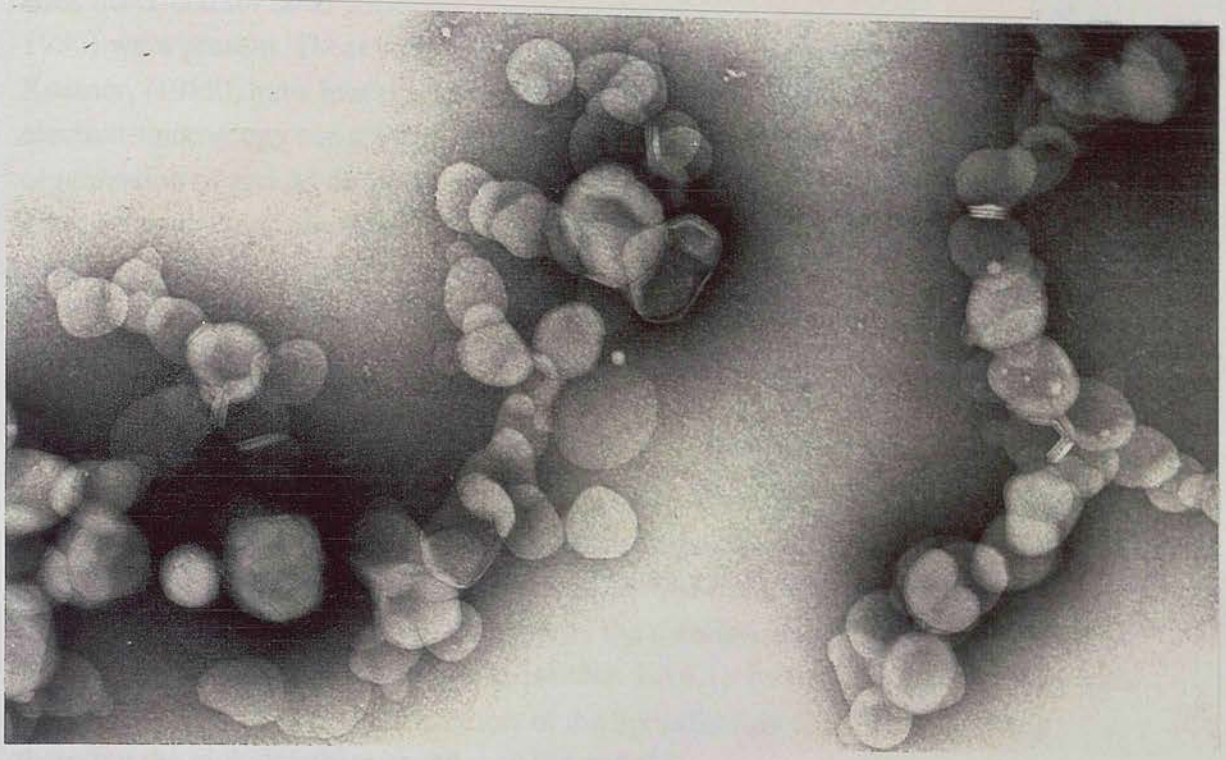
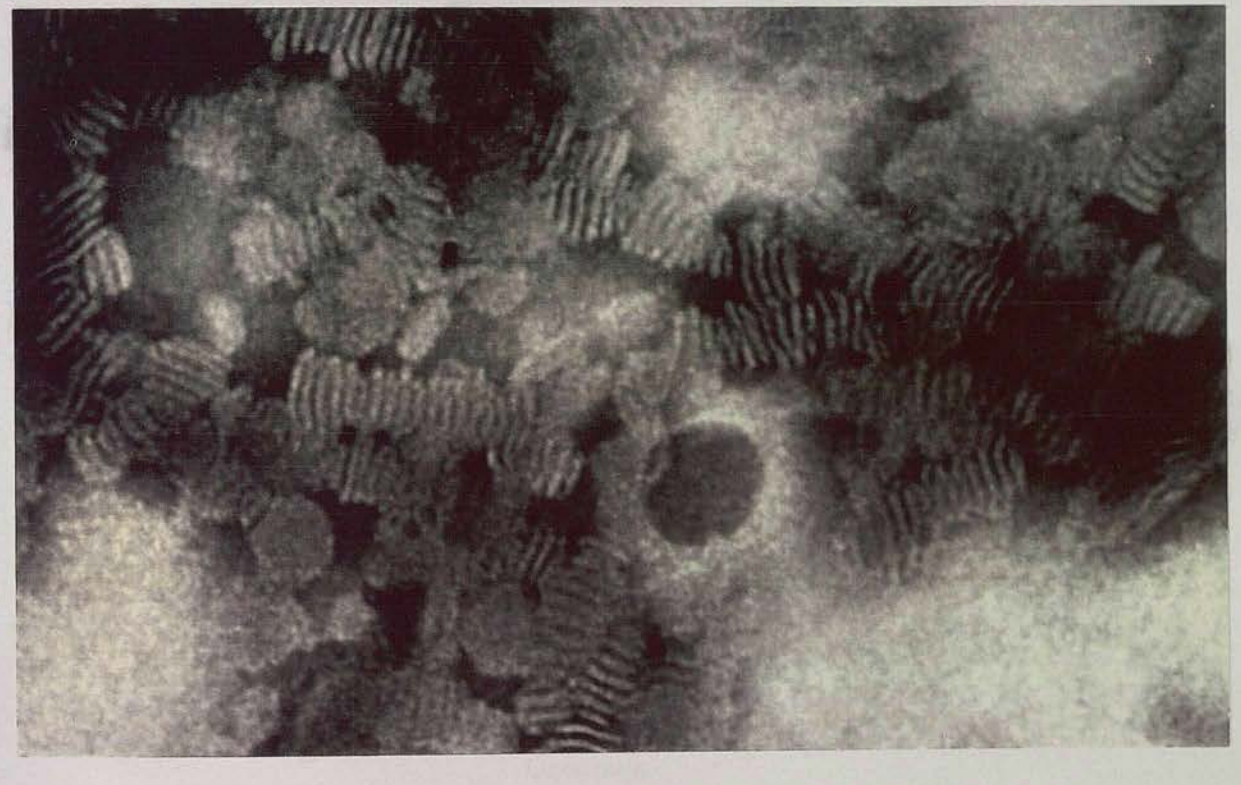


FIG. 4.9 Electron-microscope photograph of lipid vesicles with apo D. Scale - 1 cm for 7.7×10^{-2} μm.

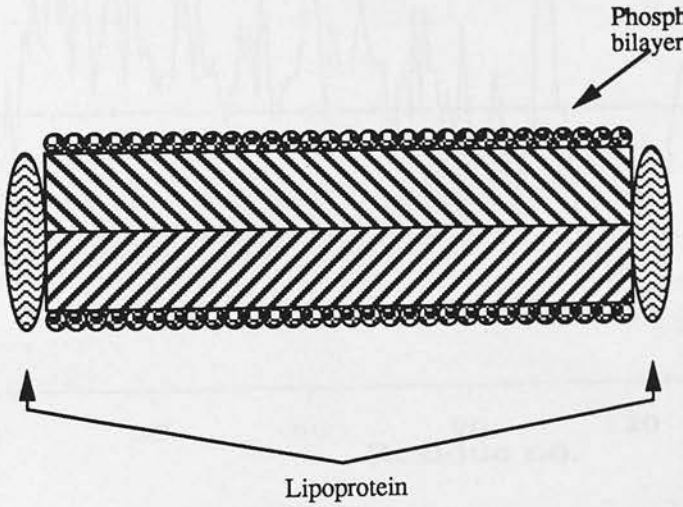


organization. The protein-lipid complex samples were observed (FIG. 4.9) after partial fractionation to homogeneous size by gel filtration. Discoidal structures just like the ones described for other apolipoproteins like apo AI (Tall, *et al.*, 1977; Guo, *et al.*, 1980) were present. These structures, already observed for apo D by Segrest and Kostner, (1988), have less of a globular aspect and under the conditions used for electron-microscopy can stack together. A model (FIG. 4.10) for the macromolecular organization of apo AI discoidal particles was presented by Tall, *et al.*, (1977) based on thermodynamic and electron-microscope considerations and was later confirmed by small angle X-ray scattering (Atkinson, *et al.*, 1980). In this model the discoidal particles are composed of a single bilayer arrangement of the lipid with protein ddisposed around the edge covering the hydrophobic tails of the phospholopid.

Apo AI is a typical apolipoprotein, its primary structure allows for a maximum flexibility and maximum interaction with the lipid, so as to stabilize the formation of the complex lipid particles (Segrest, *et al.*, 1974; Boguski, *et al.*, 1986). Free-apo AI has a variable α-helix content (around 54%) which increases to ~70% when associated with lipids in the discoidal complexes (Wald, *et al.*, 1990) and it is the existence of several stretches of amphipathic α-helix that permits the association with lipids. Apo D in the plasma is a minor component of HDL (section 1.3.4.1), forming tight complexes with apo AI and LCAT but is also a member of the lipocalin structural family, sharing therefore the globular characteristics of Blg and the

FIG. 4.10 Discoidal particles model

The particle here depicted corresponds to each of the "segments" that form the stacked structures of FIG. 4.9



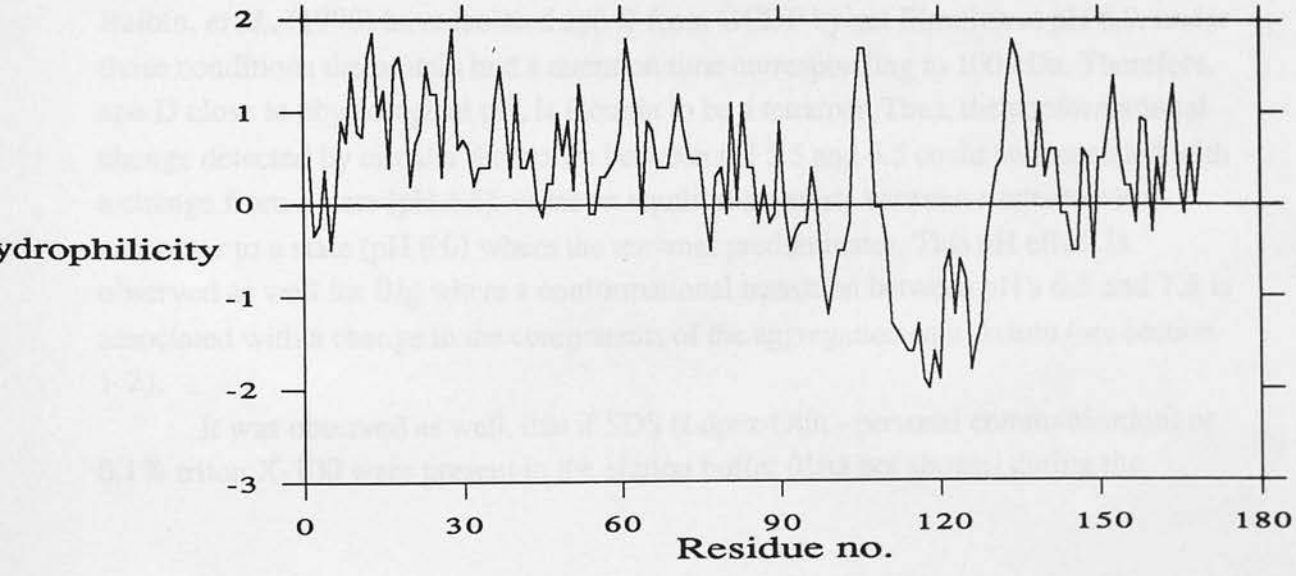
others. These proteins are very different from apolipoproteins; the free-energy of denaturation of the apolipoproteins is less than 4 kcal/mol while for Blg that value is 22.3 kcal/mole, which is a high value even when compared with other globular proteins like lysozyme (9 kcal/mol) and ribonuclease (16.3 kcal/mol) (Pace and Vanderburg, 1979; Boguski, *et al.*, 1986). This high value results from a rigid and compact structure, mainly β-sheet, as opposed to flexible apolipoprotein molecules which are mostly α-helix.

There are some other facts that may be related to the association of apo D to the lipid vesicles. A Kyte-Doolittle hydrophobicity plot (Kyte and Doolittle, 1982) of apo D shows that there is a stretch of roughly 10 residues (115-135) which is much more hydrophobic than any other part of the sequence (FIG. 4.11). Peitsch and Boguski, (1990) have noticed, in their apo D three-dimensional model, that those residues and others form a cluster of solvent exposed hydrophobic residues (Phe3, Leu5, Ile117, Ile118, Leu120, Phe121, Val123) in two closely positioned loops that are part of the "mouth" of the pocket. They have postulated that the interaction with the lipid particle is established by this arrangement of apolar residues.

The apparently contradictory behaviour of apo D, associating with lipid vesicles to form the same, albeit smaller (Steyrer and Kostner, 1988), discoidal particles as apo AI while remaining a water soluble, globular protein is certainly worth studying further.

FIG. 4.11 Kyte-Doolitle hydrophilicity plot for apo D

Hydrophobic parts of the molecule are represented when the tracing is below the zero line.



The pH-dependent conformation of apo D was analysed by circular dichroism and gel filtration.

The far UV (190-260 nm) circular dichroism spectra for apo D between pH's 4.5 and 8.5 (in one pH unit intervals) are shown in FIG. 4.12 . They are clearly very similar except for the pH 8.5 spectrum, but the application of the CONTIN algorithm (Provencher and Glöckner, 1981) applied to the spectra at pH 7.5 and 8.5 produced very similar values for the secondary structure elements content:

pH 7.5	15±1.6% α-helix, 57±3.1% β-sheet, 27±2.7% remainder
pH 8.5	14±3.6% α-helix, 60±3.7% β-sheet, 26±6.4% remainder

The α-helix content is generally the only reliable determination at this range (Johnson,Jr., 1990). The value of 15% α-helix content is in a agreement with what is expected from other members of the lipocalin family, for example, Rbp and Mup are constituted of 7% of α-helix, and in particular Ins, which is a closer "relative" of apo D, presents 12%.

The near UV spectra (260-320 nm) provide information about the aromatic side-chain environment in the protein. The spectra shown on FIG. 4.13, demonstrate a difference between pH's 4.5, 5.5, 8.5 and pH's 6.5, 7.5. This seems to indicate that there are conformational changes between pH 5.5-6.5 and pH 7.5-8.5. The conformational characteristics at the lower and higher range of pH analysed seem to be similar since the spectra are similar.

Changes at lower pH are visible by gel filtration too. The chromatograms shown on FIG. 4.14, were from FPLC runs of pure apo D at pH 5.5 and 7.5. It is clear that a lower molecular weight peak was present at pH 5.5 and not at pH 7.5. Balbín, *et al.*, (1990) have isolated apo D from GCDF by gel filtration at pH 6.0, under those conditions the protein had a retention time corresponding to 100 kDa. Therefore, apo D close to physiological pH, is thought to be a tetramer. Thus, the conformational change detected by circular dichroism between pH 5.5 and 6.5 could be associated with a change from a state (pH 5.5) where an equilibrium exists between a tetramer and monomer to a state (pH 6.0) where the tetramer predominates. This pH effect is observed as well for Blg where a conformational transition between pH's 6.5 and 7.5 is associated with a change in the components of the aggregation equilibrium (see section 1.2).

It was observed as well, that if SDS (López-Otín - personal communication) or 0.1% triton X-100 were present in the elution buffer (data not shown) during the

FIG. 4.12 Far UV circular dichroism spectra

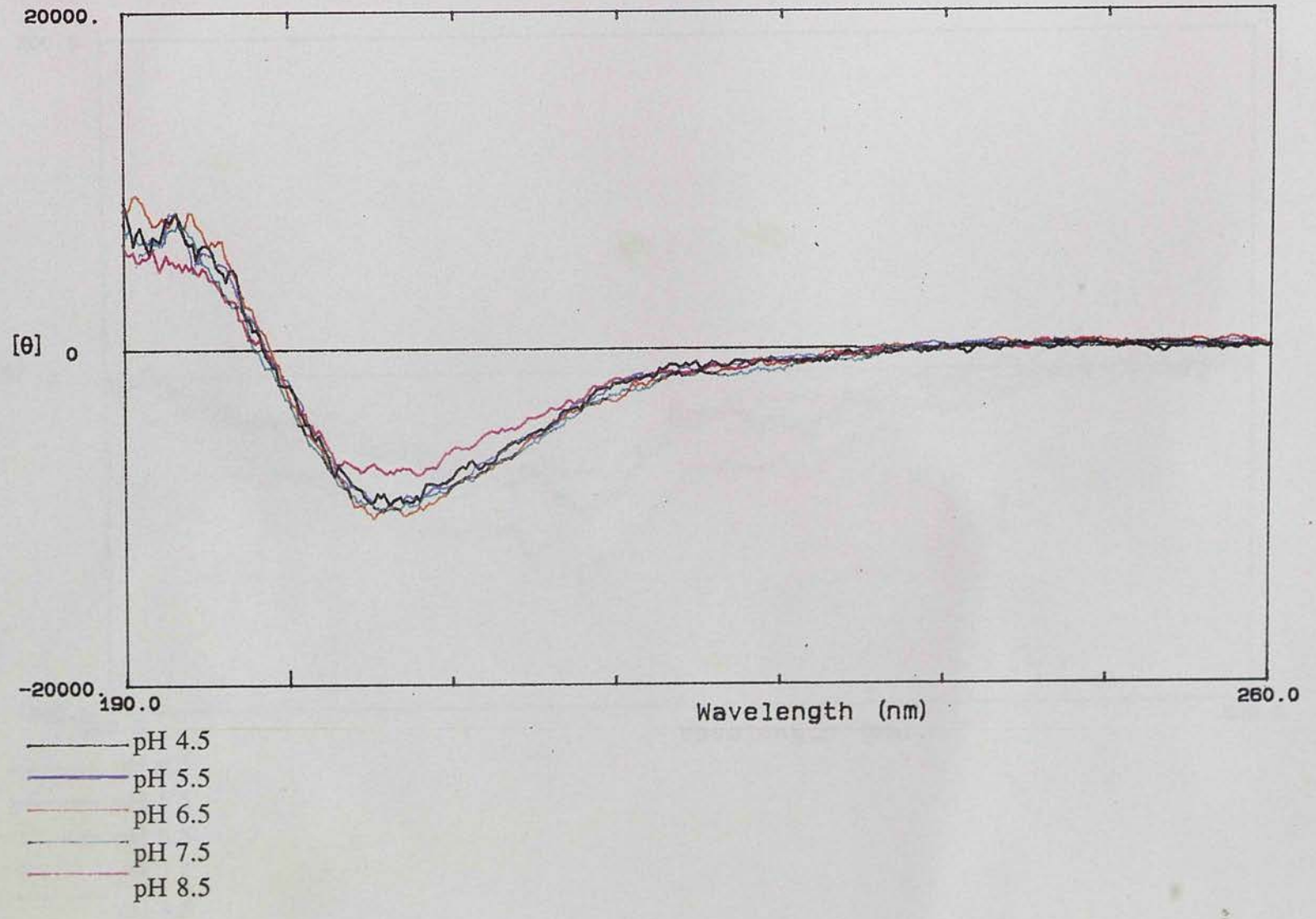


FIG. 4.13 Near UV circular dichroism spectra

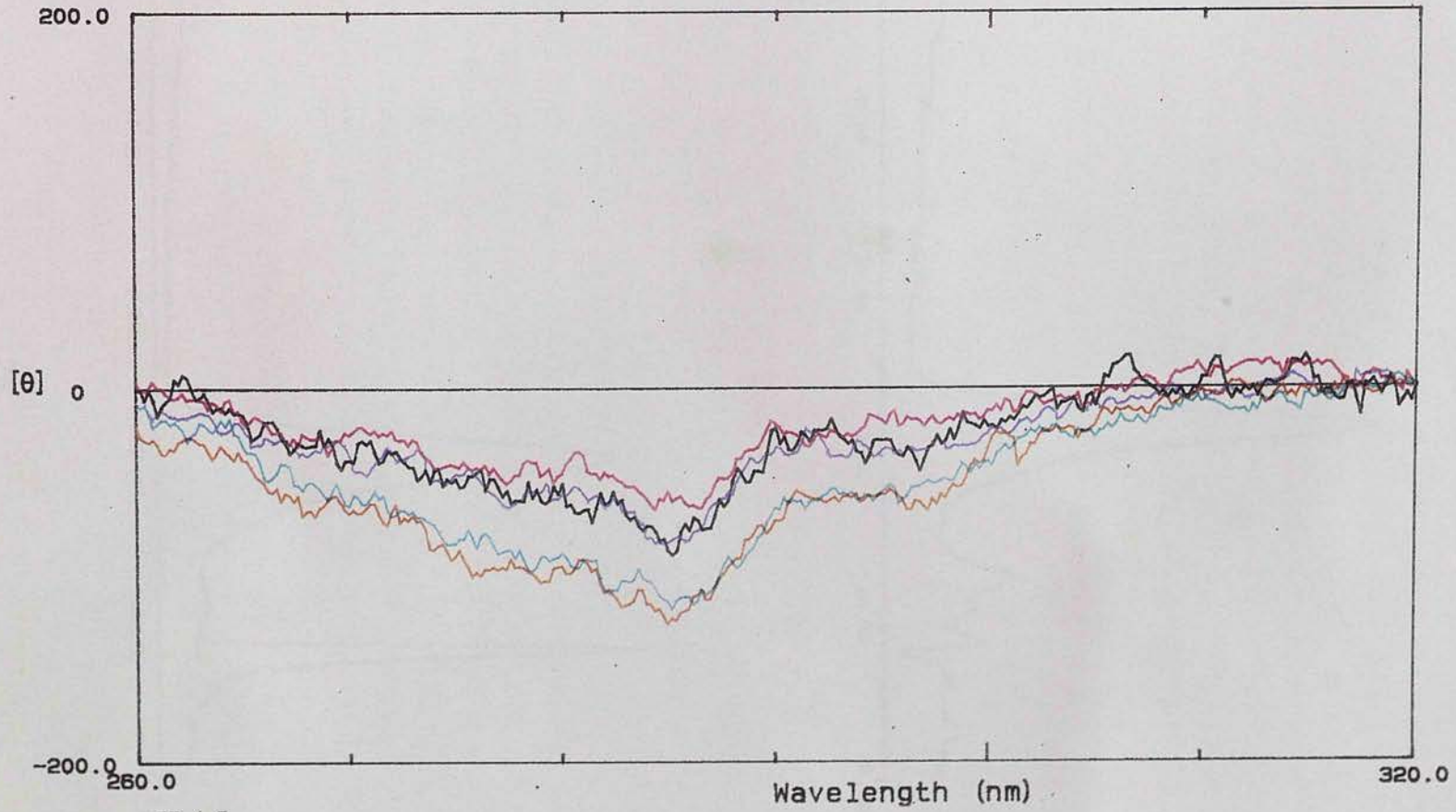
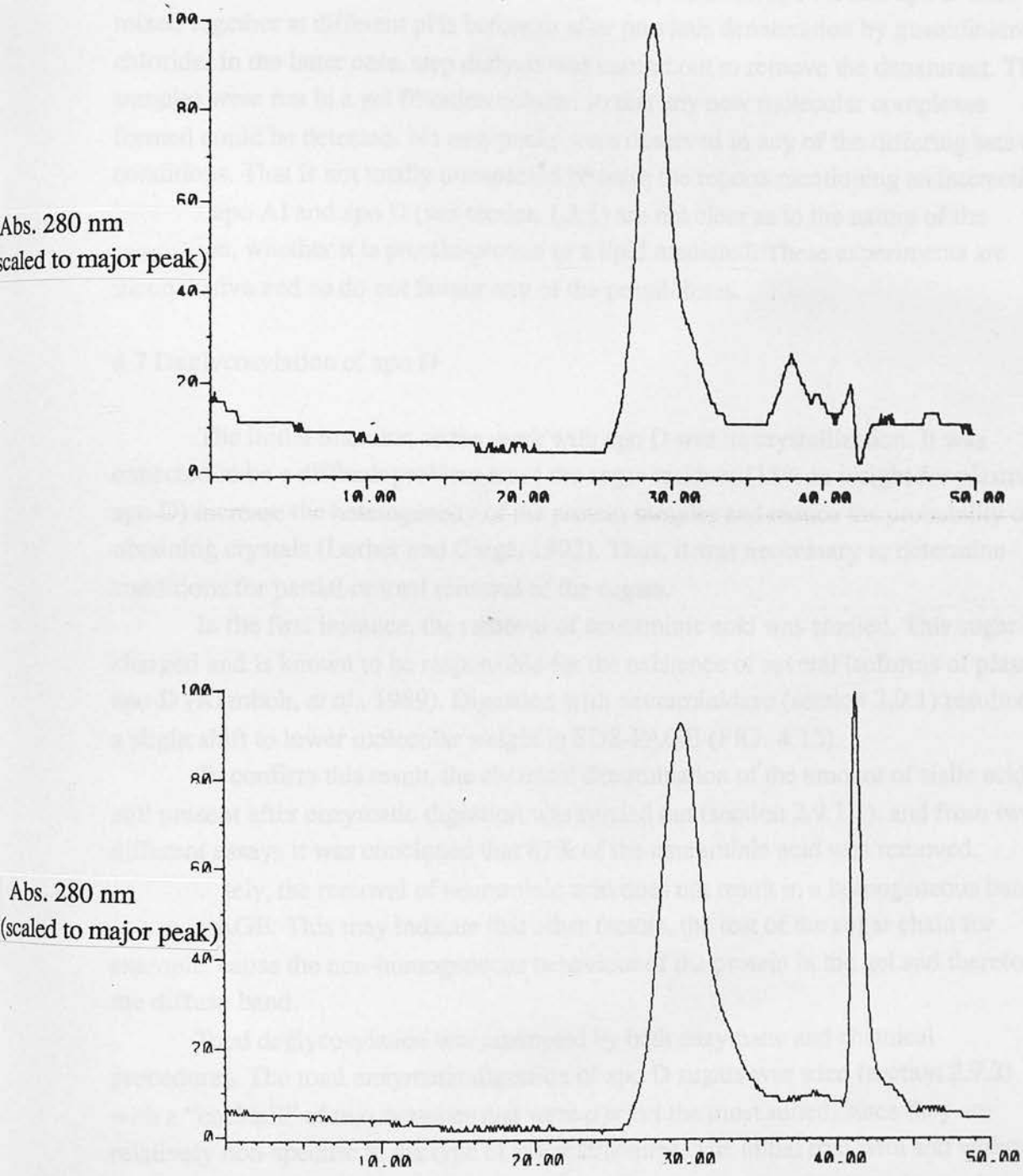


FIG. 4.14 Chromatograms of gel filtration experiments

Top: apo D run at pH 7.5

Bottom: apo D run at pH 5.5



purification of apo D by gel filtration at pH 6.0, the peak containing apo D emerged at lower molecular weight. This suggests that the interactions beween the subunits in the tetramer are dependent on an hydrophobic component too.

The close association reported for apo D and apo AI in the plasma (section 1.3.1), led to the attempts to form a complex between these two proteins without the presence of lipids. The possibility of forming a complex was explored by a series of experiments described in section 2.11. In these experiments apo AI and apo D were mixed together at different pHs before or after previous denaturation by guanidinium chloride. In the latter case, step dialysis was carried out to remove the denaturant. The samples were run in a gel filtration column so that any new molecular complexes formed could be detected. No new peaks were observed in any of the differing sets of conditions. That is not totally unexpected because the reports mentioning an interaction between apo AI and apo D (see section 1.3.1) are not clear as to the nature of the interaction, whether it is protein-protein or a lipid mediated. These experiments are inconclusive and so do not favour any of the possibilities.

4.7 Deglycosylation of apo D

The initial intention of the work with apo D was its crystallization. It was expected to be a difficult problem since the sugar residues (18% in weight for plasma apo D) increase the heterogeneity of the protein samples and reduce the probability of obtaining crystals (Lorber and Giegé, 1992). Thus, it was neccessary to determine conditions for partial or total removal of the sugars.

In the first instance, the removal of neuraminic acid was studied. This sugar is charged and is known to be responsible for the existence of several isoforms of plasma apo D (Kamboh, *et al.*, 1989). Digestion with neuraminidase (section 2.9.1) resulted in a slight shift to lower molecular weight in SDS-PAGE (FIG. 4.15).

To confirm this result, the chemical determination of the amount of sialic acid still present after enzymatic digestion was carried out (section 2.9.1.1), and from two different assays it was concluded that 87% of the neuraminic acid was removed. Unfortunately, the removal of neuraminic acid does not result in a homogeneous band in SDS-PAGE. This may indicate that other factors, the rest of the sugar chain for example, cause the non-homogeneous behaviour of the protein in the gel and therefore the diffuse band.

Total deglycosylation was attempted by both enzymatic and chemical procedures. The total enzymatic digestion of apo D sugars was tried (section 2.9.2) with a "cocktail" of two enzymes that were *a priori* the most suited, since they are relatively non-specific to the type of sugar structures. The initial trial with and without n-octylglucoside in the incubation mixture, resulted in a small migration change in SDS-PAGE (FIG. 4.16). The apo D band still presented a diffuse aspect and the

FIG. 4.15 SDS-PAGE of neuraminidase digested apo D

Lane 1 - non-digested apo D, lane 2 - apo D digested with neuraminidase.

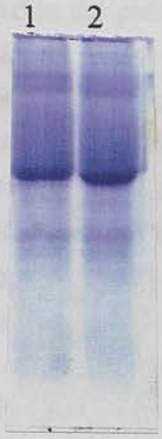
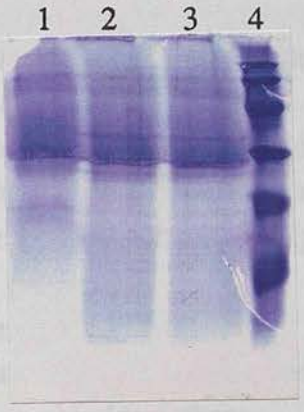


FIG. 4.16 Preliminary digestions with endoglycosidases

Lane 1- non-digested apo D, lane 2 - apo D digested with endoglycosidases, lane - 3 apo D digested with endoglycosidases in the presence of n-octylglucoside. Lane 4 - molecular weight markers. 94, 67, 43, 30, 20, 14 kDa



change in molecular weight was very small. Both facts indicate that the digestion is far from complete. In fact, it is not uncommon that steric hindrance will prevent effective action of the endoglycosidases and denaturation of the glycoprotein is recommended for a complete digestion (Biochemica Information, Boehringer Mannheim).

In the following trial of conditions, guanidinium chloride or SDS was present in the incubation mixture. All of the trials were done under reductive conditions by adding mercaptoethanol. Apo D was boiled in the presence of SDS, after which n-octylglucoside was added before the enzyme to prevent its denaturation by SDS(Biochemica Information, Boehringer Mannheim). After incubation with guanidinium chloride, n-octylglucoside was added in an attempt to keep the protein denatured (stabilizing the solvent exposed hydrophobic regions) while the denaturant was removed by dialysis. The SDS-PAGE (FIG. 4.17) of both samples shows an improvement in the deglycosylation relative to the conditions where no denaturation occurred, because the decrease in molecular weight was larger. Another interesting point is that the the diffuse aspect of the protein bands was reduced, as can be seen by comparing SDS or guanidinium chloride lanes with the controls. These facts point to a digestion of the carbohydrate moiety when denaturation of the protein occurs; however the digestion does not seem to be total because, even in the trial where the protein was boiled in the presence of SDS and mercaptoethanol, the apparent molecular weight of the protein is far from that expected from the amino acid sequence (19 kDa). From this set of experiments, guanidinium chloride emerged as a promising starting point of the search for better consitions.

On FIG. 4.18, an SDS-PAGE of several trials of different conditions is displayed. The conditions were:

with native apo D

lane 2) guanidinium chloride + n-octylglucoside
lane 3) guanidinium chloride + chaps
lane 4) guanidinium chloride + mercaptoethanol
lane 5) guanidinium chloride + n-octylglucoside + mercaptoethanol
lane 6) SDS + mercaptoethanol + n-octylglucoside

with apo D after neuraminidase digestion

lane 7) guanidinium chloride + n-octylglucoside
lane 8) guanidinium chloride + chaps
lane 9) guanidinium chloride + mercaptoethanol
lane 10) SDS + mercaptoethanol + n-octylglucoside

The conclusions from this set of conditions were: mercaptoethanol is essential for a successful deglycosylation because it facilitates the denaturation of the protein,

FIG. 4.17 SDS-PAGE of endoglycosidases digestion under denaturing conditions

Lane 1 - apo D incubated with endoglycosidases after denaturation with SDS and boiling and reduction with mercaptoethanol, lane 2 - apo D incubated with endoglycosidases after denaturation with guanidinium chloride and reduction with mercaptoethanol, lane 3 - apo D incubated with no denaturant but reduced, lane 4 - apo D incubated in the same conditions as lane 3 but without enzyme. Lane 5 - molecular weight markers. 94, 67, 43, 30, 20, 14 kDa

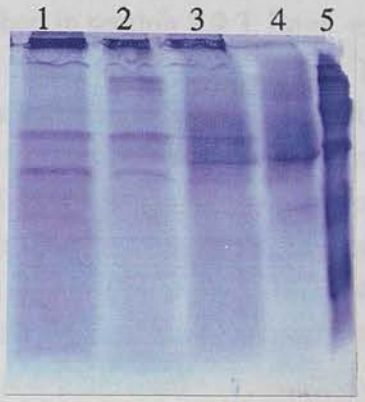
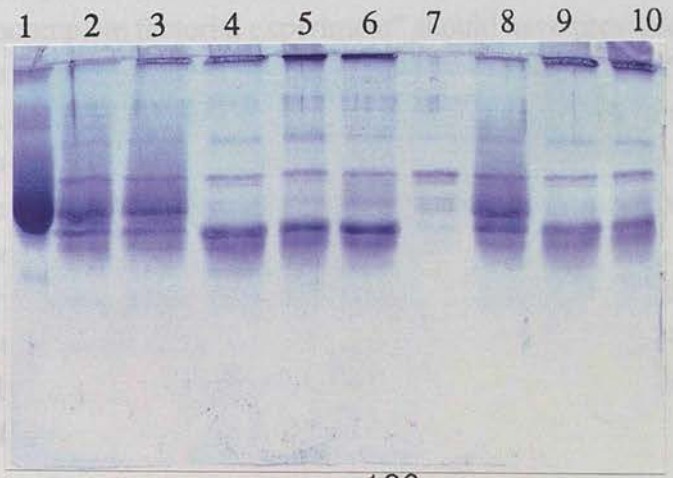


FIG. 4.18 SDS-PAGE of apo D deglycosylation

Lane 1 - non-digested apo D, lane 2 - apo D + Guanidinium chloride (GnCl)+ n-octylglucoside, lane 3 - apo D + GnCl + chaps, lane 4 - GnCl + mercaptoethanol (SH), lane 4 - apo D + GnCl + n-octylglucoside + SH, lane 6 - apo D + SDS + n-octylglucoside + SH, lane 7 - apo D pre-digested with neuraminidase (apo D N-) + GnCl +n-octylglucoside, lane 8 - apo D N- + GnCl + chaps, lane 9 - apo D N- + GnCl +SH, lane 10 - apo D N- + SDS + SH + n-octylglucoside.



detergents (besides stabilizing the enzyme) do not influence the outcome of the experiment and the previous removal of neuraminic acid does not influence the removal of the remaining sugars.

The need for extensive denaturation and reduction of apo D to obtain reasonable deglycosylation, raises the problem of refolding the protein after deglycosylation. Further, the enzymatic deglycosylation seems not to be complete since the digested protein never presents in SDS-PAGE, the molecular weight expected from the sequence. It most be noted, as well, that the high cost of the enzymes makes the procedure almost prohibitive.

The chemical approach was followed by using the trifluoromethane sulphonic acid (TFMS) method described in section 2.9.3. Edge, *et al.*, (1981) have discussed the influence of temperature on the deglycosylation by TFMS, particularly on the degradation of the proteic part of the molecule. Based on that report, trials at room temperature, 0°C and -10°C were setup and incubated for different times. The samples were run in SDS-PAGE and two of the gels are shown on FIG. 4.19. The higher the temperature, the earlier the protein was observed to degrade, as is concluded by comparing the 3 hour trial at room temperature with the 4 hour incubation at 0°C. Another fact observed was that no one set of conditions resulted in the formation of a single band. In fact, several populations were formed as seen by the many bands on the gels. One band was detected just below the 20 kDa standard that could have been the totally deglycosylated protein, but still the majority of the protein only showed a small molecular weight shift.

It is not clear if the total removal of sugars occurred or not, in either of the two methods. Another procedure, more specific than SDS-PAGE for the presence of sugars will have to be used to determine if deglycosylation is effective or not.

4.8 Apo D crystallization attempts

The methods described in section 2.18 were applied on either native apo D or neuraminidase digested apo D, after neuraminidase extraction (see section 2.9.1.2).

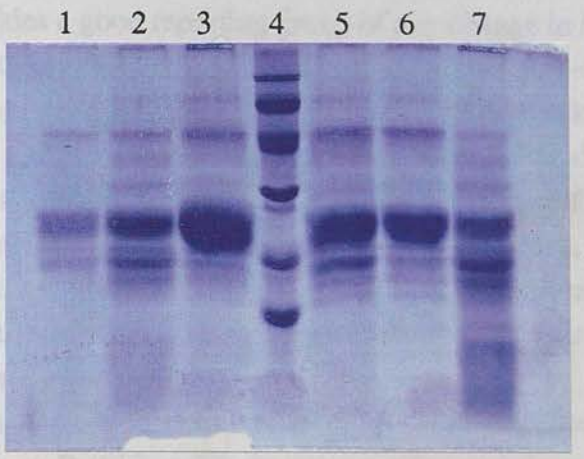
The "incomplete factorial experiment" should have provided information about the most favourable factors (type of precipitant, pH, importance of ions, *etc.*) for crystallization. The procedure is very dependent on the grading of the results and when no crystals appear it becomes "tea leaf reading" because distinctions between amorphous precipitate and crystalline precipitate, which are not always clear cut, have to be established.

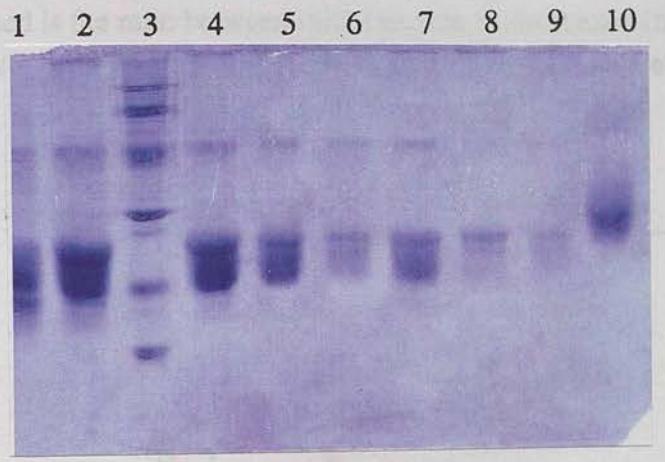
No crystals were observed with the "sparse matrix" method, either.

For both types of samples the results were discouraging, no crystals were grown probably because these samples are still very non-homogeneous as seen by the diffuse bands formed in SDS-PAGE.

FIG. 4.19 SDS-PAGE of chemical deglycosylation trials

Top gel: lanes 1 to 3 - apo D incubated at -10°C for 48, 25 and 8 hours respectively. Lanes 5 to 7 - apo D incubated at 0°C for 12, 8 and 6 hours, respectively. Lane 4 - molecular weight markers.
Bottom gel: lanes 1 and 2 - apo d incubated at 20°C for 3 and 1 hour, respectively. Lanes 4 to 6 - 5 µl of apo D incubated at 0°C for 4, 2 and 1 hours respectively. Lanes 7 to 9 - the same as 4 to 6 but 1 µl of solution. Lane 10 - non-incubated apo D. Lane 3 - molecular weight markers. 94, 67, 43, 30, 20, 14 kDa



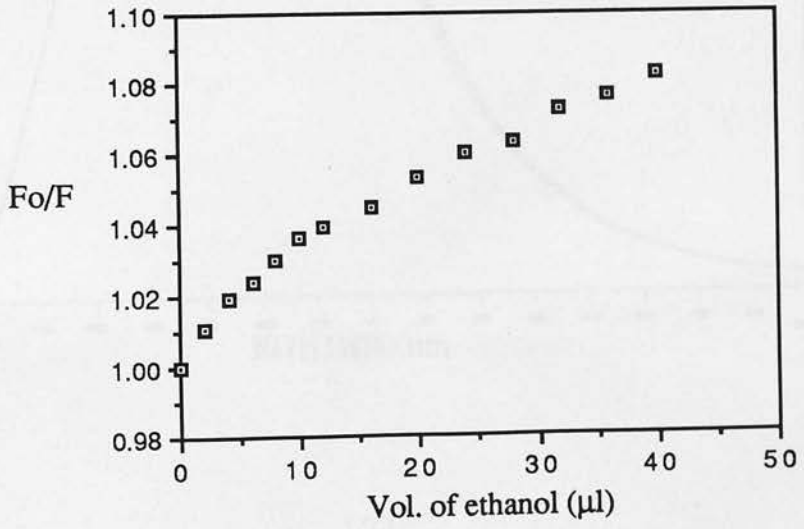


Apo D has long been thought to be a small hydrophobic molecule carrier (Chajek and Fielding, 1978). This idea was reinforced by its inclusion in the lipocalin family (Drayna, *et al.*, 1986) and finally confirmed by the discovery that the progesterone-binding protein from gross-cystic disease fluid was apo D (Balbín, *et al.*, 1990). The method chosen to evaluate the binding of molecules was fluorescence spectroscopy, in particular the detection of quenching of the protein fluorescence. This technique has been used for ligand-binding studies with many proteins (Ward, 1985) and in particular with Blg (see section 1.2). In Blg, the existence of one tryptophan in the pocket provides a good reporting group of any change in its environment. Human apo D has four tryptophans, of which, according to the three-dimensional model (Peitsch and Boguski, 1990), two are situated inside the pocket. Because of this, it was thought that the signal obtained would be sensitive to the binding of ligands. The other reason for choosing fluorescence quenching was the speed of each titration. Like most spectrophotometric phenomena (Bagshaw and Harris, 1987), fluorescence changes may be detected instantaneously, which is particularly important in cases where degradation of any component of the system is probable, like the oxidation of the ligands tested in this work.

The ethanol used to dissolve the ligands, was observed to quench the fluorescence of apo D (FIG. 4.20). In order to eliminate the possibility that the effect

FIG. 4.20 Ethanol quenching

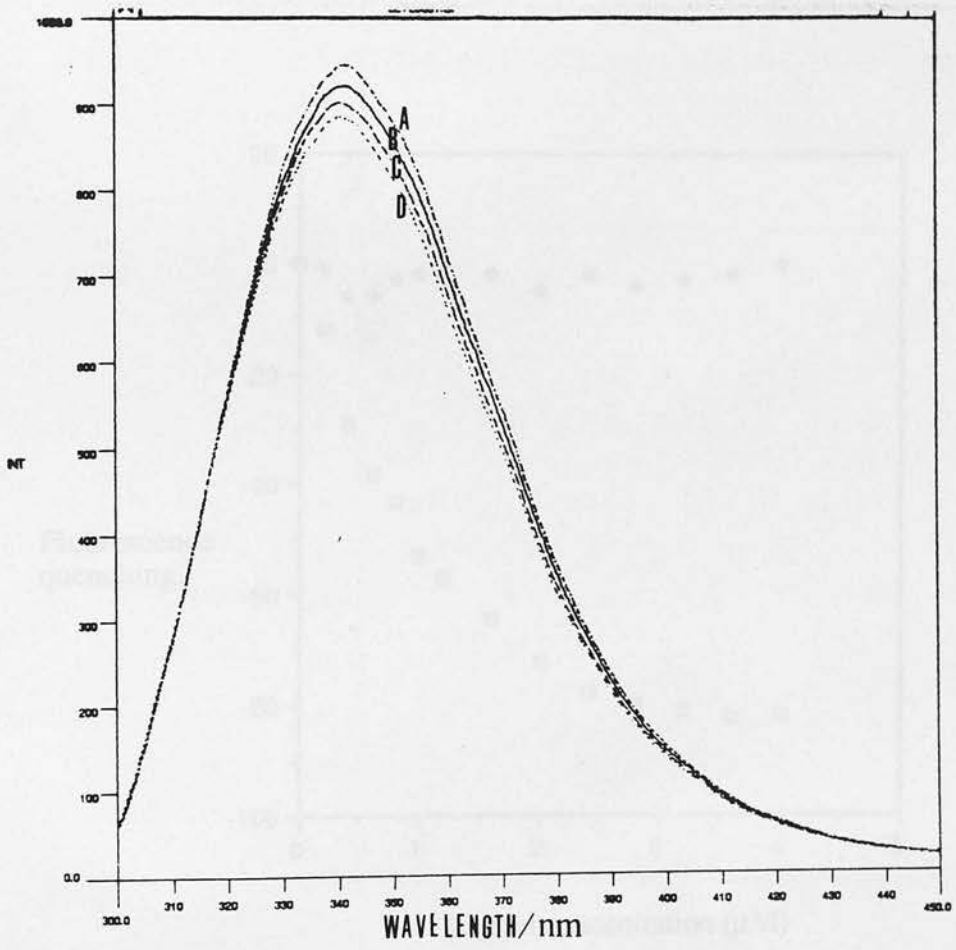
Plotted is the ratio between initial protein fluorescence (no ethanol present) and **fluorescence at each ethanol addition versus volume of ethanol added.**



was associated with protein conformational changes, during which the binding properties would be altered, several controls were made. The emission spectra at different ethanol concentrations were measured (FIG. 4.21), and no wavelength shift of the maximum of emission with increasing ethanol concentration was observed, indicating that no major conformation change was happening since the solvent-sheltered fluorescent groups were not being exposed. The absorbance spectra of apo D (240-350 nm), registered in the presence of ethanol within the concentration range 0-4%, did not reveal any changes besides a rise in the overall base line. The circular dichroism spectrum of apo D in the presence of 4% ethanol was registered between 190-260 nm and compared to the native spectrum (FIG. 4.7). The analysis of the spectra with the

FIG. 4.21 Emission spectra

Emission spectra A to D were recorded at 0, 12, 28 and 40 µl of added ethanol.

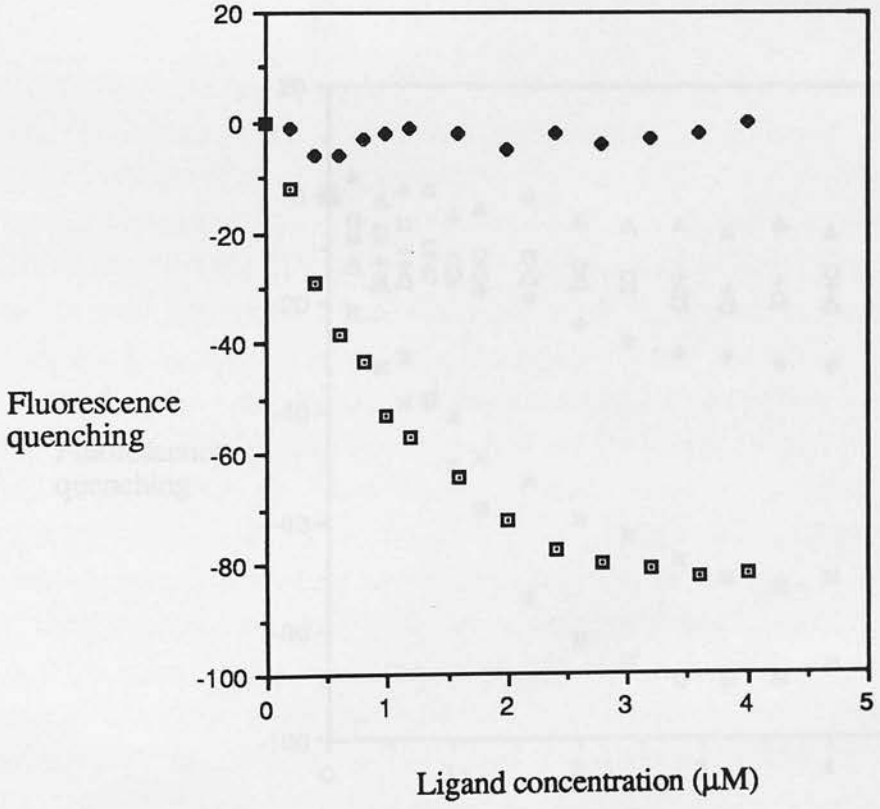


CONTIN algorithm produced a 5% increase in α-helix content. However, it was noticed that the spectra were superimposable except for the 190-200 nm range, where noise was clearly greater. Thus, it was pointed out (Ms.S. Kelly and Dr.N. Price from the CD Scottish facility at Stirling University -personnal communication) that the differences determined by the analysis algorithm were most certainly due to the noise increase in that range. All these facts indicate that no major conformational change occurs in the range of ethanol concentrations to which apo D was exposed. The ethanol effect was removed from the ligand titration curves by subtraction of the ethanol quenching from the overall ethanol/ligand quenching.

The titration curves for progesterone and cholesterol, after correction for ethanol and protein concentration, are presented in FIG. 4.22. These two molecules were used

FIG. 4.22 Cholesterol and progesterone titration curves

Cholesterol - black diamonds. Progesterone - dotted squares.

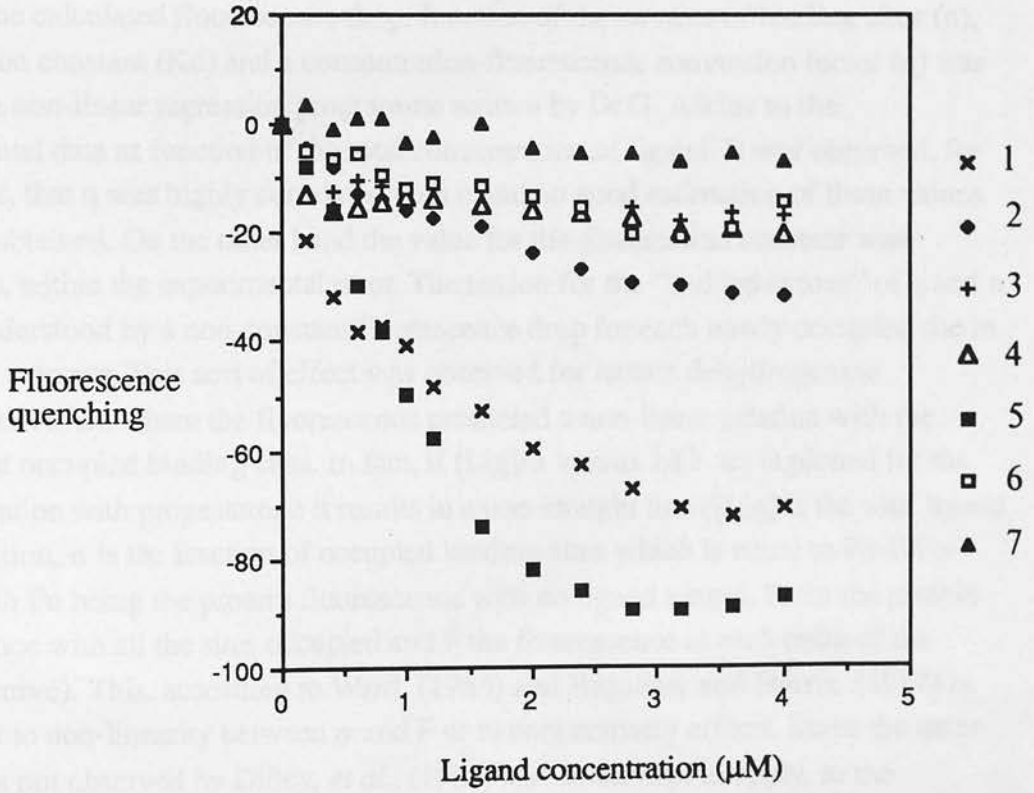


to test the method because it was known that progesterone (Lea, 1988; Dilley, *et al.*, 1990) was a strong ligand and cholesterol (Lea, 1988) was demonstrated not to bind to apo D. The marked difference between the titration curves, showing that progesterone-quenched protein fluorescence reaches a saturation point while cholesterol did not affect the fluorescence at all, demonstrate that the method was adequate for detection of binding.

By chance, (the molecule was available in the laboratory following an X-ray structure determination) EP092 was tested and shown to bind (FIG. 4.23). It is an analogue of prostaglandins (Wilson and Jones, 1985), an antagonist of thromboxane A2, and so led to the search among the prostaglandins and related compounds, for molecules with affinity for apo D. The titration curves of the few tested: prostaglandin E1, prostaglandin F2α, prostaglandin D2, arachidonic acid, 12-HETE and 5,15-diHETE (hydroxy and dihydroxyeicosatetranoic acid) are presented in FIG. 4.23. Only

FIG. 4.23 Prostaglandins and related compounds titration curves

1- EP092, 2- prostagl.D2, 3-prostagl.F2α, 4- prostagl.E1, 5-arachidonic acid, 6-12-hete, 7- 5,15-dihete



arachidonic acid and EP092 demostrated affinity for apo D as judged by the quenching of apo D fluorescence. Leukotriene D4 was tried too, but this compound has a strong absorption band at 280 nm and the emission spectra of the protein and small molecule are then superimposed. Not even with correction for the inner filter effect was it possible to separate the two signals. On the other other hand, the correction for the inner filter effect was necessary and successful in the case of EP092. In FIG. 4.24, the inner filter effects from several ligands are plotted from which the difference in the behaviour of EP092 from other ligands can be clearly seen.

A small series of fatty acids (arachidonic (C20:4), linoleic (C18:2), oleic (C18:1) and palmitic (C16:0)) and one phospholipid (dipalmitoyl phosphatidyl choline) were tested. Only arachidonic, as seen above, showed any affinity for apo D.

The analysis of the data for the extraction of association constants and number of binding sites per protein molecule was done with the help of Dr.G. Atkins, Biochemistry Department, University of Edinburgh. The simplest model used to explain the binding, where one molecule of protein binds an unknown number of molecules of ligand with the same affinity and in a random order, fitted the data. This same model was tested with the data extracted from FIG.6 in Dilley, *et al.*, (1990), where binding was determined by gel filtration separation of the bound from unbound ligand. The values obtained for number of binding of sites and association constant were very similar to the ones presented in that report, conferring validity to the chosen model.

The calculated fluorescence drop, function of the number of binding sites (n), dissociation constant (Kd) and a concentration-fluorescence conversion factor (q) was fitted by a non-linear regression programme written by Dr.G. Atkins to the experimental data as function of the total concentration of ligand. It was observed, for all ligands, that q was highly correlated with n and no good estimation of these values could be obtained. On the other hand the value for the dissociation constant was invariable, within the experimental error. The reason for the "bad behaviour" of q and n can be understood by a non-constant fluorescence drop for each newly occupied site in the apo D tetramer. This sort of effect was observed for lactate dehydrogenase (Holbrook, 1972), where the fluorescence presented a non-linear relation with the fraction of occupied binding sites. In fact, if [Lig]/α versus 1/(1- α) is plotted for the apo D titration with progesterone it results in a non-straight line ([Lig] is the total ligand concentration, α is the fraction of occupied binding sites which is equal to Fo-F/Fo-Fmin, with Fo being the protein fluorescence with no ligand bound, Fmin the protein fluorescence with all the sites occupied and F the fluorescence at each point of the titartion curve). This, according to Ward, (1985) and Bagshaw and Harris, (1987) is due either to non-linearity between α and F or to cooperativity effects. Since the latter effect was not observed by Dilley, *et al.*, (1988) and an attempt to apply, to the

fluorescence data, a model that would account for the cooperative effect was unsuccessful, then the non-linear fluorescence quenching effect must be present and be the reason for the non-determination of the number of binding sites.

The association constants determined for the ligands were the following:

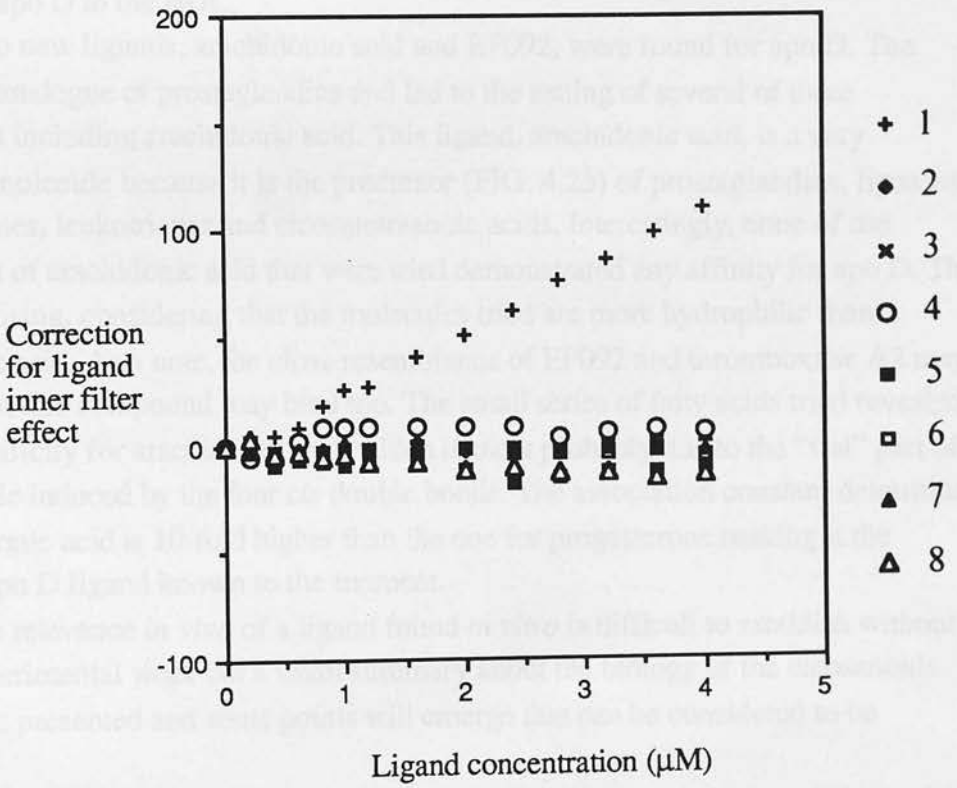
progesterone - $2.1\pm1\times10^6$ M^{-1} which is close to the value determined by Dilley, *et al.*, (1990) - 1.0×10^6 M^{-1} and Lea, (1988) - 1.3×10^6 M^{-1}.

EP092 - $1.2\pm1\times10^6$ M^{-1}

arachidonic acid - $4.1\pm4\times10^7$ M^{-1}

FIG. 4.24 Inner filter effect

1- EP092, 2- cholesterol, 3-prostagl.D2, 4-prostagl.F2α, 5- prostagl.E1, 6-progesterone, 7- arachidonic acid, 8- palmitate



4.9.1 Discussion

This study confirmed the affinity of apo D for progesterone, which was known to present the highest association constant among the steroids studied (see TABLE 1.6). Any of these steroids exist in the gross-cyst disease fluid at concentrations several fold higher than in plasma (Bradlow, *et al.*, 1981), but progesterone is present at a concentration (~$1.5x10^{-5}$ M) at which it will, probably, be the major ligand of the protein in the cyst fluid.

The first role proposed for apo D was that it was involved in cholesterol metabolism in the HDL (see section 1.3.4.1), where the protein would either be an activator of LCAT or a transfer protein of the cholesteryl-esters formed. Peitsch and Boguski, (1990) have shown some evidence that cholesterol has low affinity for apo D and proposed, based on the three-dimensional model then presented, that the affinity for cholesteryl-esters would be weak too. The work of Lea, (1988) showed that apo D, as GCDF-24, had very little affinity for cholesterol. This result has now been confirmed by fluorescence quenching. Unfortunately, it was not possible to study cholesteryl-esters due to their low solubility in the solvent and another method will have to be used to determine the affinity of apo D for these molecules and to understand fully the role of apo D in the HDL.

Two new ligands, arachidonic acid and EP092, were found for apo D. The latter is an analogue of prostaglandins and led to the testing of several of these compounds including arachidonic acid. This ligand, arachidonic acid, is a very important molecule because it is the precursor (FIG. 4.25) of prostaglandins, lipoxins, thromboxanes, leukotrienes and eicosatetraenoic acids. Interestingly, none of the metabolites of arachidonic acid that were tried demonstrated any affinity for apo D. This is not surprising, considering that the molecules tried are more hydrophilic than arachidonic acid. As a note, the close resemblance of EP092 and thromboxane A2 may indicate that this compound may bind too. The small series of fatty acids tried revealed a high specificity for arachidonic acid, which is most probably due to the "flat" part of the molecule induced by the four *cis* double bonds. The association constant determined for arachidonic acid is 10-fold higher than the one for progesterone making it the strongest apo D ligand known to the moment.

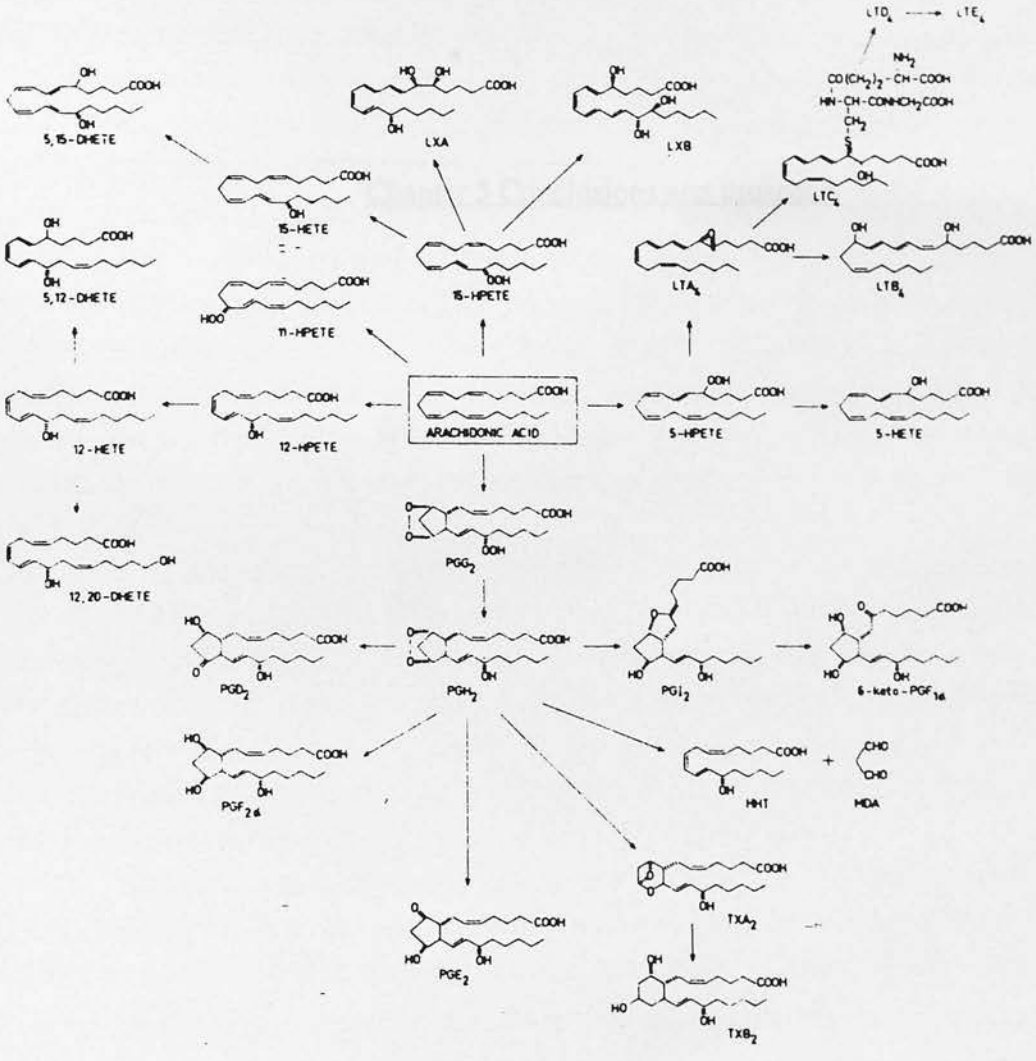
The relevance *in vivo* of a ligand found *in vitro* is difficult to establish without further experimental work but a small summary about the biology of the eicosanoids will now be presented and some points will emerge that can be considered to be relevant.

These compounds, the eicosanoids, are involved in a wide variety of biological phenomena, from platelet aggregation, sensitization to pain, involvement in inflammation, neurotropic effects, chemotactic action on leukocytes, and many others

(von Euler, 1988). The full understanding of their action is far from clear and is impaired due to their multiple, complex, and sometimes contradictory, action. However, it is accepted that these compounds are produced locally, throughout the organism and are not pre-stored (Deby, 1988) and, because of that, the major step-dependent factor for their production is the presence of the precursors, arachidonic acid, eicosapentanoic acid and homo-γ-linolenic acid. Arachidonic acid is produced from one of the essential fatty acids, linoleic acid, and is stored in the form of phospholipids (in any cellular membrane, and in lipoproteins), triglycerides (only in the renal medulla) and cholesteryl-ester (in lipoproteins). The release of arachidonic acid from these non-active forms depends mainly upon the regulation of the activity of phospholipases A2 and C on the membrane phospholipids, of LCAT (lecithin-cholesterol acyltransferase), and of a cholesterol-esterase (Behrman and Armstrong, 1969). It is through its association with LCAT (section 1.3.1), that apo D may have a role to play in the metabolism of arachidonic acid. LCAT is known to be able to act as a phospholipase and as an esterase separately (Kitabatake, *et al.*, 1979; Francone, *et al.*, 1993), but its usual role is a conjunction of the two activities by forming cholesteryl-esters from cholesterol and phosphatidylcholine. It is possible to envisage a role for apo D in the retention of arachidonic acid, avoiding further metabolisation and therefore resulting in a "fast" mobilization reserve. It is interesting too, that both apo D (see section 1.3) and the eicosanoids are involved in growth/regeneration processes. In fact, arachidonic acid metabolism has direct connections with cyclic adenosine monophosphate formation (Lagarde, 1988) in, for example, sensitization to pain, activity of nervous cells and cell proliferation. Apo D, on the other hand, is over-expressed during regeneration of peripherial rat nerve and during the non-growth stages of several cell lines, in particular cancerous cells. Thus, the presence of the two molecules can be related, apo D acting as a carrier of arachidonic acid to or from the "hot spot". In conclusion, the binding of arachidonic acid by apo D constitutes the establishment of a connection between the many apparently non-related biological situations where the protein is found.

FIG. 4.25 Prostaglandins and related eicosanoids

Diagram describing the metabolic conversion of these compounds.



Chapter 5 Conclusions and prospects

5. Conclusions and prospects

This chapter will repeat and summarize the observations and discussions already made in the previous sections. An attempt will be made to connect some aspects from different sections.

5.1 Blg

Bovine β-lactoglobulin was studied by crystallographic techniques, and the structure of the crystal form BlgZ (space group $P3_221$) was redetermined. The need for this redetermination arose from the inconsistencies between the medium resolution published structure (Monaco, *et al.*, 1987) and the low resolution study by Green, *et al.*, (1979), in particular the different packing of the protein molecules. By molecular replacement using the existing models of BlgY (model solved from crystal form in space group $B22_12$) and Mup it was possible to confirm the low resolution findings that the protein crystallizes as a dimer and not as stated by Monaco, *et al.*, (1987) in a high polymeric aggregation state. This led to the conclusion that the latter work was incorrectly approached which may have happened in the determination of the heavy atom positions using an unconventional origin choice and subsequent calculation of electron-density in the International Tables setting. Though improbable, it is possible that the crystals used in that work belong to yet another crystal form, different from the one studied here.

The re-solution of BlgZ by the heavy atom replacement method, produced a new molecular model that is different from any of the models available for Blg and in particular from the BlgY model, from which all the other models have been obtained. The phases determined from the derivatives data were refined and extended by solvent flattening so that model building could be initiated. Further improvement of the electron-density map was achieved by partial model phase combination before a model including 158 residues was finished.

This model was compared to the existing model of BlgY. It was clear that the overall folding was not changed but that sequence shifts or threading shifts were applied on two stretches of the structure. The first stretch includes the second and third β-strands and flanking loops (residues 29 to 65), where a five residue movement of the sequence towards the N-terminus took place. This resulted in the shortening of the loop between the third and fourth β-strands while extending the loop between the first and second β-strands. A similar change affected the last β-strand, where a two residue sequence shift towards the C-terminus was applied. This strand is the one responsible for the monomer-monomer interaction and so the residues responsible for the integrity of the dimer have been changed.

The quality of the new model can be evaluated by a series of factors. A lower residual of the new model when compared to the BlgY molecular replacement solution after a dynamics procedure (23% versus 26%) indicates a better fit of the new model to the data. The changes applied seem to be chemically reasonable like the movement of two charged side-chains (Asp53 and Glu51) from the inside (hydrophobic environment) to the outside of the pocket (a more solvent exposed position) and the sheltering of two of the three solvent exposed residues (Ile56, Leu57 and Leu57) that in BlgY are present in the loop between the third and fourth strands. The programme PROFILE confirmed the improvement in the new model since the plot of the averaged environment score does not present stretches below or close to zero which are observed for the BlgY model. The "bad" BlgY stretches revealed by that plot, map the stretches where the two models are different therefore conferring certainty to the changes applied. It must be emphasised that the differences between the models resulted from the incorrect interpretation of the BlgY density and not from conformational changes or different crystal packing constraints.

The structure still needs manual intervention at several points, in particular in some loops and the final refinement will take place in the near future, making use of a complete, high redundancy, 2.8 Å resolution image plate data set, collected at the Daresbury synchrotron. If the refinement procedes as expected, confirming the validiy of this model then it may be used to correct BlgY and BlgX and therefore make use of the very high resolution diffraction data, typically 2.0-1.9 Å, available for these two crystal forms.

The co-crystallization of Blg with ascorbic acid resulted in marked changes in the crystal structure, which have not yet been characterized. Future work will be based on the refinement of the model with the bound ligand, this will hopefully provide a clear picture of what residues are essential for binding and what conformational changes, if any, occur during binding. The extension of the same approach to other ligands will be important so that the relevance of the pocket and the possible existence of other binding sites can be evaluated.

5.2 Apo D

Apo D was purified both from plasma and gross-cystic disease fluid but the latter source was preferred for regular preparation of protein. The studies involving this protein covered several aspects of its biochemistry, in particular the analysis of cysteine residues, protein / lipid interactions, pH dependent conformational changes and ligand binding studies. Crystallization of native and partially carbohydrate-digested apo D was also attempted.

The analysis of the cysteines revealed that from the five residues present in apo D, four seem to be forming disulphide bonds, and the remaining one is oxidized either

as cysteic acid or, more improbably, as a thio-ester or thio-ether.

The association of apo D to lipid macro-structures was demonstrated for dipalmitoyl lecithin/cholesterol vesicles and Triton X-114. The observation by CD, that the content of α-helix does not increase very much upon lipid vesicle association clashes with the formation of discoidal particles which are typically the result of interaction between apolipoproteins and lipid vesicles. If apo D belongs to the lipocalin structural family then it will share their compact, mainly β-sheet three-dimensional characteristics which are quite different from the apolipoproteins, α-helix structures responsible for the stabilization of the lipid particles. The interaction between apo D and the lipid structures could be established by the hydrophobic peptide clearly seen in Kyte-Doolittle hydrophobicity plots. Peitsch and Boguski, (1990) have also postulated that this same peptide forms an hydrophobic loop close to the mouth of the pocket, interacting with the HDL and allowing an easy access of any ligand from the lipid moiety to the protein. There is as well, the possibility of having a cysteine forming a thio-ester bond with a fatty acid, like the one observed for ADP-ribosylation factor, which would be in line with the oxidized cysteine and the formation of apo D-lipid complexes. More experimental work is necessary to establish the nature of the apo D binding to lipid vesicles, in particular the determination of what part of the molecule establishes the hydrophobic contacts and how does the apo D structure fit the model for the discoidal particles.

The pH dependent conformational changes observed for apo D are similar to the ones observed for Blg, where the changes in CD spectra are associated with aggregation changes. But while in Blg these can be physiologically important because the protein when ingested as part of milk goes through a series of pH changes without denaturation, for apo D no immediate relevance can be seen for these conformational alterations since all the biological situations where apo D has been reported share the extracellular, controlled pH characteristic.

Finally, the ligand binding studies using fluorescence quenching have produced a new molecule which seems to present a higher affinity for apo D than any other so far. The importance of arachidonic acid for the biological role of apo D is evident as well, since it could provide a link between the different and apparently unrelated biological phenomena with which apo D is involved. Thus, apo D might, stop convertion of this fatty acid to its cholesterol-ester in the HDL, forming a fast mobilization reserve. At the cellular level, it might be used to excrete the excess of arachidonic acid and therefore stop its action as a precursor of prostaglandins and related compounds. These hypotheses have to be confirmed by determining the nature of the molecules that apo D binds when associated to LCAT (lecithin-cholesterol acyltransferase) and apo AI in HDL and by analysing the effects of the presence of apo D on the growth processes of some cancer cell lines.

Another approach for the understanding of many of these aspects would be to determine the three-dimensional structure of this protein. This would certainly provide information about the possible areas of interaction with lipids and would allow an examination of the pocket, determining the causes for a higher specificity for ligands than is found with Blg, for example. Unfortunately, it seems improbable that the protein will crystallize while the heterogeneity caused by the sugars has not been reduced. This can be achieved by cloning and expressing the protein in a organism that either does not insert sugars in proteins, like *E. coli*, or that adds carbohydrate chains which are easily removed or are short and homogeneous, like some yeast strains.

Bibliography

Adams, P. (1992) Structural studies and molecular modelling of Alpha-2u-globulin, *PhD*. thesis, University of Edinburgh

Åkerström, B. and Lögdberg (1990) TIBS 15:240-243

Albers, J.J., Cheng, M.C., Ewens, S.L., Tollefson, J.H. (1981) Atherosclerosis 39:395-409

Albers, J.J., Taggart, H.McA., Appleboum-Bowden, D., Haffner, S., Chestnut III, C.H., Hazzard, W.R. (1984) Bioch. Bioph. Acta 795:293-296

Ali, S. and Clark, J. (1988) J. Mol. Biol. 199:415-426

Ali, S. (1989) Structure and expression of the gene encoding ovine β-lactoglobulin, *PhD*. thesis, University of Edinburgh

Aschaffenburg, R., Green, D.W., Simmons, R.M. (1965) J. Mol. Biol. 13:194-201

Atkinson, D., Small, D.M., Shipley, G.G. (1980) Annals N.Y. Acad. of Sci. 348:284-298

Bagshaw, C.R. and Harris, D.A. (1987) Spectrophotometry and spectrofluorimetry - a practical approach, IRL press, Harris, D.A. and Bashford, C.L. (eds.) 91-113

Balbín, M., Freije, J.M.P., Fueyo, A., Sánchez, L.M., López-Otín, C. (1990) Biochem. J. 271:803-807

Balbín, M., Vizoso, F., Sánchez, L.M., Vents, R., Ruibal, A., Fueyo, A., López-Otín, C. (1991) Clinical Chem. 37:547-551

Baldwin, E.T., Weber, I.T., Charles, R.S., Xuan, J.-C., Appela, E., Yamada, M., Matsushima, K., Edwards, B.F.P., Clore, G.M., Gronenborn, A.M., Woldawer, A. (1991) Proc. Natl. Acad. Sci. U.S.A. 88:502-506

Banerjee, S., Katz, J., Levitz, M., Finlay, T.H. (1990) Annals N.Y. Acad. Sci. 586:204-212

Beeley, J.G. (1985) Glycoprotein and proteoglycan techniques - laboratory techniques

in Biochemistry and Molecular Biology 16:153-300

Behrman, H.R. and Armstrong, D.T. (1969) Endocrinology 85:474-480

Bennett, M. and Schmid, K. (1980) Proc. Natl. Acad, Sci. U.S.A. 77:6109-6113

Berman, P., Gray, P., Chen, E., Keyser, K., Ehrlich, D., Karten, H., LaCorbiere, M., Esch, F., Schubert, D. (1987) Cell 51:135-142

Betts, L., Frick, L., Wolfenden, R., Carter, Jr., C.W. (1989) J. Bio. Chem. 264:6737-6740

Birdsall, B., King, R.W., Whoeler, M.R., Lewis, Jr., C.A., Goode, S.R., Dunlap, R.B., Roberts, G.C.K. (1983) Anal. Biochem. 132:353-361

Blanco-Vaca, F., Via, D.P., Pownall, H.J., Massey, J.B. (1990) Arteriosclerosis 10:824a-825a

Blanco-Vaca, F., Via, D.P, Yang, C.-Y., Massey, J.B., Pownall, H. (1992) J. of Lipid Res. 33:1785-1796

Blow, D.M. (1985) Molecular replacement (Proc. of the Daresbury Study Weekend) Machin, P.A. (compiler) SERC, Daresbury laboratory 2-7

Blow, D.M., Henrick, K., Vrielink, A. (1988) Improving protein phases (Proc. of the Daresbury Study Weekend) Bailey, S. Dodson, E., Phillips, S. (compilers) SERC, Daresbury laboratory 32-38

Blundell, T.L. and Johnson, L.N. (1976) Protein crystallography, Academic Press

Blundell, T.L. and Tickle, I.J. (1985) Molecular replacement (Proc. of the Daresbury Study Weekend) Machin, P.A. (compiler) SERC, Daresbury laboratory 89-91

Böcskei, Z., Findlay, J.B.C., North, A.C.T., Phillips, S.E.V., Somer, W.S., Wright, C.E., Lionetti, C., Tirindelli, R., Cavaggioni, A. (1991) J. Mol. Biol. 218:699-701

Böcskei, Z., Groom, C.R., Flower, D.R., Wright, C.E., Phillips, S.E.V., Cavaggioni, A., Findlay, J.B.C., North, A.C.T. (1992) Nature 360:186-188

Boguski, M.S., Freeman, M., Elshourbagy, Taylor, J.M., Gordon, J.I. (1986) J.

Lipid Res. 27:1011-1034

Bojanovski, D., Alaupovic, P., McConathy, W.J., Kelly, J. (1980) FEBS Letts. 112:251-254

Bolognesi, M., Liberatori, J., Oberti, R., Ungaretti, L. (1979) J. Mol. Biol. 131:411-413

Bouma, M.-E., de Bandt, J.-P., Ayrault-Jarrie, M., Burdin, J., Verthier, N., Raisonnier, A. (1988) Scand. J. Gastroenterol. 23:477-483

Boyles, J.K., Zoellner, C.D., Anderson, L.J., Kosik, L.M., Pitas, R.E., Weisgraber, K.H., Hui, D.Y., Mahley, R.W., Gebicke-Haerter, P.J., Ignatius, M.J., Shooter, E.M. (1989) J. Clin. Invest. 83:1015-1031

Boyles, J.K., Notterpek, L.M., Anderson, L.J. (1990) J. Biol. Chem. 265:17805-17815

Bradlow, H.L., Rosenfeld, R.S., Kream, J., Fleisher, M., O'Connor, J., Schwartz, K. (1981) Cancer Res. 41:105-107

Brady, L. and Jian-Sheng, J. (1992) Molecular replacement (Proc. of the CCP4 Study Weekend) SERC, Daresbury laboratory, Dodson, E.J., Gover, S., Wolf, W. (compilers) 170-177

Braunitzer, G., Chen, R., Schrank, B., Stangl, A. (1972) Hoppe-Seyler's Z. Physiol. Chem. 354:867-878

Brooks, D.E., Means, A.R., Wright, E.J., Singh, S.P., Tiver, K.K. (1986) J. Biol. Chem. 261:4956-4861

Brünger, A.T. (1992) X-PLOR manual, version 3.0

Cancedda, F.D., Asaro, D., Molina, F., Cancedda, R., Caruso, C., Camardella, L., Negri, a., Ronchi, S. (1990) Biochem. Biophys. Res. Commun. 168:933-938

Carter, Jr., C.W. and Carter, C.W. (1979) J. Biol. Chem. 254:12219-12223

Carter, Jr., C.W., Baldwin, E.T., Frick, L. (1988) J. Crystal Growth 90:60-73

Chajek, T. and Fielding, C.J. (1978) Proc. Natl. Acad. Sci. U.S.A. 75:3445-3449

Chaplin, M.F. (1986) Carbohydrate analysis - a practical approach, IRL Press, Chaplin, M.F. and Kenedy, J.F. (eds.) 1-36

Chapman, MJ. (1986) Methods in enzymology, Segrest, J.P. and Albers, J.J. (eds.) 128:70-143

Cho, Y., Batt, C.A., Sawyer, L. (1993) Submitted

Clark, A.J., Clissold, P.M., Shaw, A., Beattie, P., Bishop, J. (1984) EMBO J. 3:1045-1052

Cooper, R., Eckely, D.M., Papaconstantinou, J. (1987) Biochemistry 26:5244-5250

Cowan, S.W., Newcomer, M.E., Jones, T.A. (1990) Proteins: structure, function and genetics 8:44-61

Creighton, T.E. (1989) Protein structure - a practical approach, IRL Press, Creighton, T.E. (ed.) 155-167

Cuatrecasas, P. and Illiano, G. (1971) Biochem. Biophys. Res. Commun. 44:178-184

Curry, M.D., McConathy, W.J., Alaupovic, P. (1977) Bioch. Bioph. Acta 49:232-241

Dauter, Z. (1992) Molecular replacement (Proc. of the CCP4 Study Weekend) SERC, Daresbury laboratory, Dodson, E.J., Gover, S., Wolf, W. (compilers) 163-169

Deby, C. (1988) Prostaglandins, Biology and Chemistry of prostaglandins and related eicosanoids, Churchill-Livingstone, Curtis-Prior, P.B. (ed.) 11-36

Dilley, W.G., Haagensen, D.E., Cox, C.E., Wells,Jr., S.A. (1990) Breast Cancer Res. and Treatment 16:253-260

Dodin, G., Andrieux, M., Alkabbani, H. (1990) Eur. J. Biochem. 193:697-700

Dodson, E.J. (1976) Crystallographic computing techniques, Ahmed, F.R. (ed.) Munksgaard, Copenhagen

Dogliotti, L., Orlandi, F., Caraci, P., Puligheddu, B., Torta, M., Angeli, A. (1990)

Annals N.Y. Acad. Sci. 586:17-28

Drayna, D., Fielding, C., McLean, J., Baer, B., Castro, G., Chen, Comstock, L., Henzel, W., Kohr, N., Rhee, L., Wion, K., Lawn, R. (1986) J. Biol. Chem. 261:16535-16539

Drayna, D.T., McLean, J.W., Wion, K.L., Trent, J.M., Drabkin, H.A., Lawn, R.M. (1987) DNA 6:199-204

Dufour, E. and Haertlé, T. (1990) Protein Engineering 4:185-190

Dufour, E., Marden, M.C., Haertlé, T. (1990) FEBS Letts. 277:223-226

Edge, A.S.B., Faltynek, C.R., Hof, L., Reichert,Jr., L.E., Weber, P. (1981) Anal. Biochem. 118:131-137

Erkelens, D.W. (1989) Postgraduate Medical J. 65:275-281

von Euler, U.S. (1988) Prostaglandins, Biology and Chemistry of prostaglandins and related eicosanoids, Churchill-Livingstone, Curtis-Prior, P.B. (ed.) 1-7

Evans, P. (1993) Data collection and processing (Proc. of the CCP4 Study Weekend), SERC, Daresbury laboratory (In press)

Farrel, H.M., Bede, M.J., Enyeart, J.A. (1987) J. Dairy Sci. 70:252-258

Fielding, P.E. and Fielding, C.J. (1980) Proc. Natl. Acad. Sci. U.S.A. 77:3327-3330

Francone, O.L., Evangelista, L., Fielding, C.J. (1993) Bioch. Bioph. Acta 1166:301-304

Frederick, C.A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobson, L., Wang, B.C., Greene, P., Boyer, H.W., Rosenberg, J.M. (1984) Nature 309:327-331

Fugate, R.D. and Song, P.-S. (1980) Bioch. Bioph. Acta 625:28-42

Fujinaga, M. and Read, R.J. (1987) J. Appl. Cryst. 20:517-521

Gill, S.C. and von Hippel, P.H. (1989) Anal. Biochem. 182:319-326

Godovac-Zimmerman, J. and Braunitzer, (1987) Biol. Chem. Hoppe-Seyler 366:431-434

Godovac-Zimmerman, J. (1988) TIBS 13:64-66

Gotto,Jr., A.M., Pownall, H.J., Havel, R.J. (1986) Methods in enzymology, Segrest, J.P. and Albers, J.J. (eds.) 28:3-41

Green, D.W., Aschaffenburg, R., Camerman, A., Coppola, J.C., Dunnill, P., Simmons, R.M., Komorowski, E.S., Sawyer, L., Turner, E.M.C., Woods, K,F. (1979) J. Mol. Biol. 131:375-397

Guo, L.S.S., Hamilton, R.L., Goerke, J., Weinstein, J.N., Havel, R.J. (1980) J. Lipid Res. 21:993-1003

Haagensen, C.D. (1971) Diseases of the breast, Saunders (ed.) 155-176

Haagensen,Jr., D.E., Mazoujian, G., Dilley, W.G., Pederson, C.E., Kister, S.J., Wells,Jr., S.A. (1979) J. Natl. Cancer Inst. 62:239-247

Haagensen, D.E., Stewart, P., Dilley, W.G., Wells, S.A. (1992) Breast Cancer Res. and Treatment 23:77-86

Haefliger, J.-A., Jenue, D., Stanley, K.K., Tschopp, J. (1987) Biochem. Biophys. Res. Comm. 149:750-754

Haffner, S.M., Applebaum-Bowden, D., Wahl, P.W., Hoover, J.J., Warnick, G.R., Albers, J.J., Hazzard, W.R. (1985) Arteriosclerosis 5:169-177

Hambling, S.G., McAlpine, A.S., Sawyer, L. (1992) Advanced dairy chemistry I, Elsevier, Fox, P.F. (ed.) 141-190

Handelmann, G.E., Boyles, J.K., Weisgraber, K.H., Mahley, R.W., Pitas, R.E. (1992) J. of Lipid Res. 33:1677-1688

Henzel, W.J., Rodriguez, H., Singer, A.G., Stults, J.T., Macrides, F., Agosta, W.C., Niall, H. (1988) J. Biol. Chem. 263:16682-16687

Holbrook, J.J. (1972) Biochem. J. 128:921-931

Holden, H.M., Rypniewski, W.R., Law, J.H., Rayment, I. (1987) EMBO J. 6:1565-1570

Hölmquist, L. (1989) J. of Bioch. and Bioph. methods 19:93-104

Hölmquist, L. (1990) Electrophoresis 11:93-94

Horwitz, J. and Heller, J. (1974) J. Biol. Chem. 249:4712-4719

Huber, R. (1985) Molecular replacement (Proc. of the Daresbury Study Weekend) Machin, P.A. (compiler), SERC, Daresbury laboratory 58-61

Huber, R., Schneider, M., Epp, O., Mayr, I., Messerschmidt, A., Pfulgrath, J., Kayser, H. (1987a) J. Mol. Biol. 195:423-434

Huber, R., Schneider, M., Mayr, I., Müller, R., Dentzmam, R., Suter, F., Zuber, H., Falk, H., Kayser, H. (1987b) J. Mol. Biol. 198:499-513

International tables (1983), IUC, Hahn, T. (ed.) Vol.A

Jancarik, J. and Kim, S.-H. (1991) J. Appl. Cryst. 24:409-411

Johnson,Jr., W.C. (1990) Proteins: structure, function and genetics 7:205-214

Jones, T.A., Zou, J.-Y., Cowan, S.W., Kjeldgaard, M. (1991) Acta Cryst. A47:110-119

Julkunen, M., Seppala, M., Janne, O.A. (1988) Proc. Natl. Acad. Sci. U.S.A. 85:8845-8849

Kabsch, W. (1988a) J. Appl. Cryst. 21:67-71

Kabsch, W. (1988b) J. Appl. Cryst. 21:916-924

Kabsch, W. and Sander, C. (1983) Biopolymers 22:2577-2637

Kamboh, M.I., Albers, J.J., Majurnder, P.P., Ferrel, R.E. (1989) Am. J. Human Genet. 45:147-154

Keen, J.N., Caceres, I., Eliopolous, E.E., Zagalsky, P.F., Findlay, P.F., (1990) Eur.

J. Biochem. 197:407-417

Kesner, L., Yu, W., Bradlow, H.L. (1990) Annals N.Y. Acad. Sci. 586:198-203

Kitakabe, K., Piran, U., Kamio, Y., Doi, Y., Nishida, T. (1979) Bioch. Bioph. Acta 573:145-154

Kostner, G. (1974) Scand. J. Clin. Lab. Invest. 33(supplement 137):19-21

Kostner, G. and Alaupovic, P. (1972) Biochemistry 11:3419-3428

Kyte, J. and Doolittle, R.R. (1982) J. Mol. Biol. 157:105-132

Labrie, F., Simard, J., Poulin, R., Hatton, A.-C., Labrie, C., Dauvois, S., Zhao, H., Petitclerc, L., Couet, J., Dumont, M., Haagensen,Jr., D.E. (1990) Annals N.Y. Acad. Sci. 586:174-187

Lagarde, M. (1988) Prostaglandins, Biology and Chemistry of prostaglandins and related eicosanoids, Churchill-Livingstone, Curtis-Prior, P.B. (ed.), 147-151

Laemmli, U.K. (1970) Nature 227:680-685

Lea, O. (1988) Steroids 52:337-338

Lee, K.-H., Wells, R.G., Beed, R.D. (1987) Science 235:1053-1056

Leslie, A.G.M. (1988) Improving protein phases (Proc. of the CCP4 Study Weekend) Bailey, S., Dodson, E., Phillips, S. (compilers) SERC, Daresbury laboratory 25-31

Lorber, B. and Giegé, R. (1992) Crystallization of nucleic acids and proteins - a practical approach, IRL Press Ducruix, A. and Giegé, R. (eds.) 19-45

Lovrien, R. and Anderson, W.F. (1969) Arch. Biochem. and Biophys. 131:139-144

Lüthy, R., Bowie, J.U., Eisenberg, D. (1992) Nature 356:83-85

Matthews, B.W. (1968) J. Mol. Biol. 33:491-493

Matthews, J.A., Frederick, C.A., Wang, B.C., Greene, P., Boyer, H.W., Grable, J., Rosenberg, J.M. (1986) Science 234:1526-1541

Mazoujian, G. (1990) American J. of Dermatopathology 12:452-457

Mazoujian, G. and Haagensen,Jr., D.E. (1990) Annals N.Y. Acad.Sci. 586:188-197

McAlpine, A.S. (1991) Structural and functional studies on bovine β-lactoglobulin, *PhD*. thesis, University of Edinburgh

McClarin, J.A., Frederick, C.A., Wang, B.C., Greene, P., Boyer, H.W., Grable, J., Rosenberg, J.M. (1986) Science 234:1526-1541

McConathy, W.J. and Alaupovic, P. (1973) FEBS Letts. 37:178-182

McConathy, W.J. and Alaupovic, P. (1976) Biochemistry 15:515-520

McConathy, W.J. and Alaupovic, P. (1986) Methods in enzymology, Segrest, J.P. and Albers, J.J. (eds.) 128:297-311

McKenzie, H.A. (1971) Milk proteins - II, Academic press, McKenzie, H.A. (ed.) 257-330

Mills, O.E. (1976) Bioch. Bioph. Acta 434:24-332

Mohammadzadeh, A., Feeney, R.E., smith, L.M. (1969) Bioch. Bioph. Acta 194:246-255

Molina, R., Filella, X., Herranz, M., Prats, M., Velasco, A., Zanon, G., Martinez-Osaba, M.J., Ballesta, A.M. (1990) Annals N.Y. Acad. Sci. 586:29-42

Monaco, H.L., Zanotti, G., Spadon, P., Bolognesi, M., Sawyer, L., Eliopoulos, E.E. (1987) J. Mol. Biol. 197:695-706

Morton, R.E. and Zilversmit, D.B. (1981) Bioch. Bioph. Acta 663:350-355

Nagata, A., Suzuki, Y., Igarashi, m., Eguchi, N., Toh, H., Urade, Y., Hayaishi, O, (1991) Proc. Natl. Acad. Sci. U.S.A. 88:4020-4024

Newcomer, M.E., Liljas, A., Sundelin, J., Rask, L., Peterson, P.A. (1984) J. Biol. Chem. 259:5230-5231

Newcomer, M.E., Jones, T.A., Åquist, J., Sundelin, J., Eriksson, U., Rask, L., Peterson, P.A. (1984) EMBO J. 3:1451-1454

North, A.C.T. (1989) Int. J. of Biol. Macromolecules 11:56-58

O'Neill, T.E. and Kinsella, J.E. (1987) J. Agric. Food Chem. 35:770-774

Osborne,Jr., J.C. (1986) Methods in enzymology, Segrest, J.P. and Albers, J.J. (eds.) 128:213-222

Pace, C.N. and Vanderburg, K.E. (1979) Biochemistry 19:288-292

Papiz, M. (1982) Crystallographic studies of beta-lactoglobulin at 2.8Å resolution, *PhD*. thesis, University of Edinburgh

Papiz, M.Z., Sawyer, L., Eliopoulos, E.E., North, A.C.T., Lindlay, J.B.C., Sivaprasadaro, R., Jones, T.A., Newcomer, M.E., Kraulis, P.J. (1986) Nature 324:383-385

Pearlman, W.H., Guériguian, J.L., Sawyer, M.E. (1973) J. Biol. Chem. 248:5736-5741

Peitsch, M.C. and Boguski, M.S. (1990) New Biologist 2:1-10

Pervaiz, S. and Brew, K. (1985) Science 228:335-337

Peterson, G.L. (1979) Anal. Biochem. 100:201-220

Pevsner, J., Reed, R.R., Feinstein, P.G., Suyder, S.H. (1988) FEBS Letts. 212:225-228

Podjarny, A.D., Bhat, T.N., Zwick, M. (1987) Ann. Rev. Biophys. Biophys. Chem. 16:351-373

Provencher, S.W. and Glöckner, J. (1981) Biochemistry 20:33-37

Provost, P.R., Weech, P.K., Tremblay, N.M., Marcel, Y.L., Rassart, E. (1990) J. Lipid Res. 31:2057-2065

Provost, P.R., Marcel, Y.L., Milne, R.W., Weech, P.K., Rassart, E. (1991) FEBS

Letts. 290:139-141

Rao, S.N., Jih, J.-H., Hartsuck, J.A. (1980) Acta Cryst. A36:878-884

Rask, L., Anundi, H., Peterson, P.A. (1979) FEBS Letts. 104:55-58

Redl, B., Holzfeind, Lottspeich, F. (1992) J. Biol. Chem. 267:20282-20287

Rice, D.W., Anderson, B.F., Baker, E.N. (1988) Improving protein phases (Proc. Of the Daresbury Study Weekend) Bailey, S., Dodson, E., Phillips, S. (compilers), SERC, Daresbury laboratory, 113-120

Riley, C.T., Barbean, B.K., Keim, P.S., Kezdy, F.J., Heinrikson, R.L., Law, J.H. (1984) J. Biol. Chem. 259:13159-13165

Robillard, K.A. and Wishnia, A. (1972) Biochemistry 11:3841-3845

Said, H.M., Ong, D.E., Singleton, J.L. (1989) Am. J. Clin. Nutr. 49:690-694

Sánchez, L.M., Freije, J.P., Merino, A.M., Vizoso, F., Foltmann, B., López-Ótin, C. (1992) J. Biol. Chem. 267:24725-24731

Sánchez-Ferrer, A., Bru, R., Garcia-Carmona, F. (1990) Anal. Biochem. 184:279-282

Sawyer, L. (1987) Nature 327:659

Sawyer, L., Papiz, M.Z., North, A.C.T., Eliopoulos, E.E. (1985) Biochem. Soc. Trans. 13:265-266

See, Y.P. and Jackowski, G. (1989) Protein structure - a practical approach, IRL press, Creighton, T.E. (ed.) 1-21

Segrest, J.P., Jackson, R.L., Morrisett, J.D., Gotto,Jr., A.M. (1974) FEBS Letts. 38:247-253

Sheldrick, G.M. (1991) Crystallographic computing 5, from chemistry to biology IUCR, Oxford university press, Moras, A.D., Podjarny, A.D., Thierry, J.C. (eds.) 145-157

Silva, J.S., Cox, C.E., Wells, S.A., Paull, D., Dilley, W.G., McCarty, K.S., Fetter,

B.F., Glaubitz, L.C., McCarty,Jr., K.S. (1982) Surgery 92:443-449

Simard, J., Dauvois, S., Haagensen, D.E., Levésque, C., Méraud, Y., Labrie, F. (1990) Endocrinology 126:3223-3231

Simard, J., Veilleux, R., de Lauvoit, Y., Haagensen, D.E., Labrie, F. (1991) Cancer Res. 51:4336-4341

Sizer, P.J.H., Miller, A., Watts, A. (1987) Biochemistry 26:5106-5113

Smith, K.M., Lawn, R.M., Wilcox, J.N. (1990) J. Lipid Res. 31:995-1004

Søreide, J.A., Lea, O., Kvinnsland, S. (1987) Breast Cancer Res. and Treatment 9:123-128

Søreide, J.A., Lea, O., Anda, O., Skarstein, A., Varhang, J.E., Kvinnsland, S. (1991a) Anticancer Res. 11:601-606

Søreide, J.A., Lea, O., Kvinnsland, S. (1991b) Anticancer Res. 11:1323-1326

Spector, A.A.and Fletcher, J.E. (1970) Lipids 5:403-411

Spence, A.M., Sheppard, P.C., Davie, J.R., Matuo, Y., Nishi, N., McKeehan, W.L., Dodd, J.G., Matusik, R.J. (1989) Proc. Natl. Acad. Sci. U.S.A. 86:7843-7847

Spreyer, P., Schaal, H., Kuhn, G., Rothe, T., Unterbeck, A., Olek, K., Müller, H.W. (1990) EMBO J. 9:2479-2484

Steyrer, E. and Kostner, G.M. (1988) Bioch. Bioph. Acta 958:484-491

Stout, C.D. (1979) Nature, 279:83-84

Sutor, F., Kayser, H., Zuber, H. (1988) Biol. Chem. Hoppe-Seyler 369:497-505

Tall, A.R., Small, D.M., Deckelbaum, R.J., Shipley, G.G. (1977) J. Biol. Chem. 252:4701-4711

Thannhauser, T.W., Konishi, Y., Sheraga, H.A. (1984) Anal. Biochem. 138:181-188

Thannhauser, T.W., Konishi, Y., Sheraga, H.A (1987) Methods in enzymology,

Jacoby, W.B. and griffith, O.W. (eds.) 143:115-118

Tickle, I.J. (1992) Molecular replacement (Proc. of the CCP4 Study Weekend) Dodson, E.J., Gover, S., Wolf, W. (compilers), SERC, Daresbury laboratory 20-32

Tilley, J.M.A. (1960) Dairy Science Abstract 22:111-125

Townend, R., Herskovits, T.T., Timasheff, S.N., Gorbunoff, M.J. (1969) Arch. Biochem. Biophys. 129:567-580

Tulinsky, A. (1985) Methods in enzymology, Whyckoff, H.W., Hirs, C.H.W., Timasheff, S.N. (eds.) 115:77-89

Unterman, R.D., Lynch, K.R., Nakhasi, H.L., Dolan, K.P., Hamilton, J.W., Cohn, D.V., Feigelson, P. (1981) Proc. Natl. Acad. Sci. U.S.A. 78:3478-3482

Veerkamp, J.H., Peeters, R.A., Maotman, R.G.H.G. (1991) Bioch. Bioph. Acta 1081:1-24

Wald, J.H., Krul, E.S., Jonas, A. (1990) J. Biol. Chem. 265:20037-20043

Walsh, M.T., Watzlawick, H., Putnam, F.W., Schmid, K., Brossmer, R. (1990) Biochemistry 29:6250-6252

Wang, B.-C. (1985) Methods in enzymology, Wyckoff, H.W., Hirs, C.H.W., Timasheff, S.N. (eds.) 115:90-112

Ward, L.D. (1985) Methods in enzymology, Hirs, C.H.W. and Timasheff, S.N. (eds.) 117:400-414

Wilson, N.H. and Jones, R.L. (1985) Thromboxane Leukotriene Res. 14:393-425

Wishnia, A, and Pinder, T.W. (1966) Biochemistry 5:1534-1542

Woolfson, M.M. (1970) An introduction to X-ray crystallography, Cambridge university press

Zhang, K.Y.J., Main, P. (1988) Improving protein phases (Proc. of the CCP4 Study Weekend), Bailey, S., Dodson, E., Phillips, S. (compilers), SERC, Daresbury laboratory, 57-64

Appendix 1

```
REMARK FILENAME="slowc_model3.pdb"
REMARK Written by O version 5.8, Thu Nov 26 09:29:04 MET 1992
REMARK Sun Mar 21 15:47:29 1993
REMARK DATE:20-Apr-93  04:01:51       created by user:
ATOM      1  CB  THR     4      20.781  20.073  15.153  1.00 15.00
ATOM      2  OG1 THR     4      20.263  19.851  13.804  1.00 15.00
ATOM      3  HG1 THR     4      20.920  20.101  13.138  1.00 15.00
ATOM      4  CG2 THR     4      22.105  19.297  15.407  1.00 15.00
ATOM      5  C   THR     4      18.497  20.549  16.225  1.00 15.00
ATOM      6  O   THR     4      17.513  20.165  15.583  1.00 15.00
ATOM      7  HT1 THR     4      19.689  17.471  16.096  1.00 15.00
ATOM      8  HT2 THR     4      18.986  18.443  14.915  1.00 15.00
ATOM      9  N   THR     4      19.159  18.343  15.941  1.00 15.00
ATOM     10  HT3 THR     4      18.185  18.283  16.312  1.00 15.00
ATOM     11  CA  THR     4      19.729  19.643  16.265  1.00 15.00
ATOM     12  N   GLN     5      18.517  21.715  16.878  1.00 15.00
ATOM     13  H   GLN     5      19.287  21.937  17.440  1.00 15.00
ATOM     14  CA  GLN     5      17.370  22.615  16.983  1.00 15.00
ATOM     15  CB  GLN     5      16.818  22.934  15.527  1.00 15.00
ATOM     16  CG  GLN     5      15.946  24.205  15.499  1.00 15.00
ATOM     17  CD  GLN     5      15.312  24.640  14.170  1.00 15.00
ATOM     18  OE1 GLN     5      14.093  24.861  14.058  1.00 15.00
ATOM     19  NE2 GLN     5      16.091  24.790  13.094  1.00 15.00
ATOM     20 HE21 GLN     5      17.045  24.587  13.151  1.00 15.00
ATOM     21 HE22 GLN     5      15.640  25.102  12.284  1.00 15.00
ATOM     22  C   GLN     5      16.329  21.961  17.928  1.00 15.00
ATOM     23  O   GLN     5      15.765  20.893  17.753  1.00 15.00
ATOM     24  N   THR     6      16.294  22.528  19.107  1.00 15.00
ATOM     25  H   THR     6      16.889  23.273  19.325  1.00 15.00
ATOM     26  CA  THR     6      15.385  22.164  20.179  1.00 15.00
ATOM     27  CB  THR     6      16.080  21.347  21.270  1.00 15.00
ATOM     28  OG1 THR     6      17.481  21.593  21.206  1.00 15.00
ATOM     29  HG1 THR     6      17.934  20.974  21.797  1.00 15.00
ATOM     30  CG2 THR     6      15.816  19.887  21.096  1.00 15.00
ATOM     31  C   THR     6      14.963  23.512  20.750  1.00 15.00
ATOM     32  O   THR     6      15.644  24.531  20.593  1.00 15.00
ATOM     33  N   MET     7      13.810  23.573  21.371  1.00 15.00
ATOM     34  H   MET     7      13.300  22.750  21.506  1.00 15.00
ATOM     35  CA  MET     7      13.267  24.822  21.861  1.00 15.00
ATOM     36  CB  MET     7      11.800  24.438  22.254  1.00 15.00
ATOM     37  CG  MET     7      10.759  25.515  22.519  1.00 15.00
ATOM     38  SD  MET     7      10.456  26.506  21.032  1.00 15.00
ATOM     39  CE  MET     7      11.601  27.865  21.129  1.00 15.00
ATOM     40  C   MET     7      14.109  25.463  22.991  1.00 15.00
ATOM     41  O   MET     7      15.139  24.961  23.448  1.00 15.00
ATOM     42  N   LYS     8      13.704  26.652  23.418  1.00 15.00
ATOM     43  H   LYS     8      12.954  27.068  22.967  1.00 15.00
ATOM     44  CA  LYS     8      14.252  27.297  24.591  1.00 15.00
ATOM     45  CB  LYS     8      14.380  28.822  24.438  1.00 15.00
ATOM     46  CG  LYS     8      14.905  29.598  25.671  1.00 15.00
ATOM     47  CD  LYS     8      14.855  31.128  25.512  1.00 15.00
ATOM     48  CE  LYS     8      15.857  31.567  24.416  1.00 15.00
ATOM     49  NZ  LYS     8      15.807  32.994  24.097  1.00 15.00
ATOM     50  HZ1 LYS     8      14.869  33.234  23.717  1.00 15.00
ATOM     51  HZ2 LYS     8      16.537  33.216  23.390  1.00 15.00
ATOM     52  HZ3 LYS     8      15.982  33.544  24.962  1.00 15.00
ATOM     53  C   LYS     8      13.133  26.984  25.550  1.00 15.00
ATOM     54  O   LYS     8      12.769  25.819  25.638  1.00 15.00
ATOM     55  N   GLY     9      12.456  27.922  26.212  1.00 15.00
ATOM     56  H   GLY     9      12.556  28.872  26.019  1.00 15.00
ATOM     57  CA  GLY     9      11.502  27.583  27.233  1.00 15.00
ATOM     58  C   GLY     9      10.253  27.114  26.565  1.00 15.00
ATOM     59  O   GLY     9       9.416  27.908  26.141  1.00 15.00
ATOM     60  N   LEU    10      10.191  25.819  26.359  1.00 15.00
ATOM     61  H   LEU    10      10.985  25.277  26.554  1.00 15.00
ATOM     62  CA  LEU    10       9.013  25.207  25.825  1.00 15.00
```

```
ATOM     63  CB  LEU    10       9.301  23.748  25.590  1.00 15.00
ATOM     64  CG  LEU    10       8.174  22.833  25.136  1.00 15.00
ATOM     65  CD1 LEU    10       7.865  23.012  23.662  1.00 15.00
ATOM     66  CD2 LEU    10       8.615  21.406  25.394  1.00 15.00
ATOM     67  C   LEU    10       8.040  25.417  26.973  1.00 15.00
ATOM     68  O   LEU    10       8.116  24.775  28.035  1.00 15.00
ATOM     69  N   ASP    11       7.156  26.386  26.800  1.00 15.00
ATOM     70  H   ASP    11       7.178  26.928  25.977  1.00 15.00
ATOM     71  CA  ASP    11       6.210  26.690  27.845  1.00 15.00
ATOM     72  CB  ASP    11       5.903  28.187  27.748  1.00 15.00
ATOM     73  CG  ASP    11       5.236  28.807  28.981  1.00 15.00
ATOM     74  OD1 ASP    11       4.003  28.914  28.966  1.00 15.00
ATOM     75  OD2 ASP    11       5.949  29.194  29.937  1.00 15.00
ATOM     76  C   ASP    11       5.047  25.796  27.502  1.00 15.00
ATOM     77  O   ASP    11       4.032  26.264  27.050  1.00 15.00
ATOM     78  N   ILE    12       5.171  24.493  27.673  1.00 15.00
ATOM     79  H   ILE    12       6.029  24.198  28.050  1.00 15.00
ATOM     80  CA  ILE    12       4.163  23.498  27.355  1.00 15.00
ATOM     81  CB  ILE    12       4.410  22.269  28.211  1.00 15.00
ATOM     82  CG2 ILE    12       3.293  21.279  28.155  1.00 15.00
ATOM     83  CG1 ILE    12       5.640  21.577  27.666  1.00 15.00
ATOM     84  CD1 ILE    12       6.431  20.745  28.737  1.00 15.00
ATOM     85  C   ILE    12       2.750  23.971  27.531  1.00 15.00
ATOM     86  O   ILE    12       2.007  23.884  26.576  1.00 15.00
ATOM     87  N   GLN    13       2.318  24.570  28.627  1.00 15.00
ATOM     88  H   GLN    13       2.944  24.732  29.362  1.00 15.00
ATOM     89  CA  GLN    13       0.929  25.015  28.736  1.00 15.00
ATOM     90  CB  GLN    13       0.715  25.588  30.182  1.00 15.00
ATOM     91  CG  GLN    13       0.412  24.492  31.292  1.00 15.00
ATOM     92  CD  GLN    13       1.573  23.805  32.040  1.00 15.00
ATOM     93  OE1 GLN    13       2.633  24.416  32.253  1.00 15.00
ATOM     94  NE2 GLN    13       1.449  22.543  32.474  1.00 15.00
ATOM     95 HE21 GLN    13       2.216  22.183  32.961  1.00 15.00
ATOM     96 HE22 GLN    13       0.617  22.043  32.326  1.00 15.00
ATOM     97  C   GLN    13       0.495  25.995  27.635  1.00 15.00
ATOM     98  O   GLN    13      -0.657  26.008  27.233  1.00 15.00
ATOM     99  N   LYS    14       1.395  26.735  27.033  1.00 15.00
ATOM    100  H   LYS    14       2.304  26.756  27.373  1.00 15.00
ATOM    101  CA  LYS    14       1.054  27.544  25.902  1.00 15.00
ATOM    102  CB  LYS    14       2.109  28.618  25.683  1.00 15.00
ATOM    103  CG  LYS    14       2.022  29.876  26.515  1.00 15.00
ATOM    104  CD  LYS    14       3.244  30.765  26.264  1.00 15.00
ATOM    105  CE  LYS    14       2.859  32.168  25.782  1.00 15.00
ATOM    106  NZ  LYS    14       2.504  32.154  24.378  1.00 15.00
ATOM    107  HZ1 LYS    14       3.343  31.922  23.814  1.00 15.00
ATOM    108  HZ2 LYS    14       2.141  33.086  24.093  1.00 15.00
ATOM    109  HZ3 LYS    14       1.774  31.431  24.223  1.00 15.00
ATOM    110  C   LYS    14       0.894  26.748  24.601  1.00 15.00
ATOM    111  O   LYS    14       0.467  27.385  23.632  1.00 15.00
ATOM    112  N   VAL    15       1.209  25.461  24.386  1.00 15.00
ATOM    113  H   VAL    15       1.599  24.929  25.103  1.00 15.00
ATOM    114  CA  VAL    15       0.934  24.850  23.078  1.00 15.00
ATOM    115  CB  VAL    15       2.117  23.975  22.464  1.00 15.00
ATOM    116  CG1 VAL    15       3.280  24.877  22.120  1.00 15.00
ATOM    117  CG2 VAL    15       2.495  22.822  23.383  1.00 15.00
ATOM    118  C   VAL    15      -0.280  23.928  23.102  1.00 15.00
ATOM    119  O   VAL    15      -0.447  23.032  22.267  1.00 15.00
ATOM    120  N   ALA    16      -1.139  24.101  24.089  1.00 15.00
ATOM    121  H   ALA    16      -1.016  24.854  24.701  1.00 15.00
ATOM    122  CA  ALA    16      -2.300  23.248  24.211  1.00 15.00
ATOM    123  CB  ALA    16      -2.824  23.440  25.633  1.00 15.00
ATOM    124  C   ALA    16      -3.403  23.510  23.155  1.00 15.00
ATOM    125  O   ALA    16      -3.381  24.480  22.372  1.00 15.00
ATOM    126  N   GLY    17      -4.385  22.593  23.118  1.00 15.00
ATOM    127  H   GLY    17      -4.237  21.752  23.604  1.00 15.00
ATOM    128  CA  GLY    17      -5.535  22.643  22.218  1.00 15.00
```

```
ATOM    129  C   GLY    17      -5.444  21.646  21.040  1.00 15.00
ATOM    130  O   GLY    17      -4.931  20.526  21.170  1.00 15.00
ATOM    131  N   THR    18      -5.943  22.098  19.871  1.00 15.00
ATOM    132  H   THR    18      -6.133  23.045  19.739  1.00 15.00
ATOM    133  CA  THR    18      -6.065  21.323  18.639  1.00 15.00
ATOM    134  CB  THR    18      -7.395  21.631  17.905  1.00 15.00
ATOM    135  OG1 THR    18      -7.601  23.047  17.990  1.00 15.00
ATOM    136  HG1 THR    18      -8.130  23.315  17.221  1.00 15.00
ATOM    137  CG2 THR    18      -8.573  20.886  18.502  1.00 15.00
ATOM    138  C   THR    18      -4.945  21.686  17.718  1.00 15.00
ATOM    139  O   THR    18      -4.590  22.861  17.634  1.00 15.00
ATOM    140  N   TRP    19      -4.449  20.717  16.979  1.00 15.00
ATOM    141  H   TRP    19      -4.866  19.826  16.979  1.00 15.00
ATOM    142  CA  TRP    19      -3.442  20.969  15.964  1.00 15.00
ATOM    143  CB  TRP    19      -2.102  20.459  16.426  1.00 15.00
ATOM    144  CG  TRP    19      -1.487  21.368  17.438  1.00 15.00
ATOM    145  CD2 TRP    19      -0.901  22.580  17.201  1.00 15.00
ATOM    146  CE2 TRP    19      -0.572  22.962  18.460  1.00 15.00
ATOM    147  CE3 TRP    19      -0.586  23.410  16.163  1.00 15.00
ATOM    148  CD1 TRP    19      -1.520  21.012  18.735  1.00 15.00
ATOM    149  NE1 TRP    19      -0.952  22.014  19.338  1.00 15.00
ATOM    150  HE1 TRP    19      -0.876  22.125  20.318  1.00 15.00
ATOM    151  CZ2 TRP    19       0.059  24.162  18.677  1.00 15.00
ATOM    152  CZ3 TRP    19       0.045  24.601  16.364  1.00 15.00
ATOM    153  CH2 TRP    19       0.364  24.971  17.625  1.00 15.00
ATOM    154  C   TRP    19      -3.828  20.231  14.696  1.00 15.00
ATOM    155  O   TRP    19      -4.536  19.224  14.813  1.00 15.00
ATOM    156  N   TYR    20      -3.374  20.689  13.515  1.00 15.00
ATOM    157  H   TYR    20      -2.839  21.497  13.525  1.00 15.00
ATOM    158  CA  TYR    20      -3.537  19.988  12.227  1.00 15.00
ATOM    159  CB  TYR    20      -4.314  20.845  11.164  1.00 15.00
ATOM    160  CG  TYR    20      -5.784  20.592  11.385  1.00 15.00
ATOM    161  CD1 TYR    20      -6.606  21.592  11.823  1.00 15.00
ATOM    162  CE1 TYR    20      -7.900  21.290  12.231  1.00 15.00
ATOM    163  CD2 TYR    20      -6.247  19.290  11.312  1.00 15.00
ATOM    164  CE2 TYR    20      -7.532  18.967  11.717  1.00 15.00
ATOM    165  CZ  TYR    20      -8.367  19.972  12.191  1.00 15.00
ATOM    166  OH  TYR    20      -9.635  19.633  12.691  1.00 15.00
ATOM    167  HH  TYR    20     -10.068  20.421  13.050  1.00 15.00
ATOM    168  C   TYR    20      -2.184  19.623  11.638  1.00 15.00
ATOM    169  O   TYR    20      -1.353  20.514  11.403  1.00 15.00
ATOM    170  N   SER    21      -1.893  18.340  11.423  1.00 15.00
ATOM    171  H   SER    21      -2.535  17.638  11.664  1.00 15.00
ATOM    172  CA  SER    21      -0.623  17.934  10.836  1.00 15.00
ATOM    173  CB  SER    21      -0.439  16.439  10.923  1.00 15.00
ATOM    174  OG  SER    21      -0.427  16.148  12.300  1.00 15.00
ATOM    175  HG  SER    21       0.478  15.848  12.485  1.00 15.00
ATOM    176  C   SER    21      -0.501  18.332   9.396  1.00 15.00
ATOM    177  O   SER    21      -0.527  17.509   8.504  1.00 15.00
ATOM    178  N   LEU    22      -0.297  19.600   9.132  1.00 15.00
ATOM    179  H   LEU    22      -0.120  20.224   9.868  1.00 15.00
ATOM    180  CA  LEU    22      -0.228  20.070   7.784  1.00 15.00
ATOM    181  CB  LEU    22      -0.199  21.575   7.925  1.00 15.00
ATOM    182  CG  LEU    22      -0.040  22.653   6.861  1.00 15.00
ATOM    183  CD1 LEU    22       1.369  23.141   6.833  1.00 15.00
ATOM    184  CD2 LEU    22      -0.570  22.121   5.543  1.00 15.00
ATOM    185  C   LEU    22       0.944  19.471   7.026  1.00 15.00
ATOM    186  O   LEU    22       0.881  19.566   5.817  1.00 15.00
ATOM    187  N   ALA    23       1.961  18.763   7.504  1.00 15.00
ATOM    188  H   ALA    23       1.896  18.379   8.399  1.00 15.00
ATOM    189  CA  ALA    23       3.099  18.354   6.679  1.00 15.00
ATOM    190  CB  ALA    23       4.073  19.525   6.651  1.00 15.00
ATOM    191  C   ALA    23       3.755  17.087   7.236  1.00 15.00
ATOM    192  O   ALA    23       3.142  16.590   8.163  1.00 15.00
ATOM    193  N   MET    24       4.880  16.450   6.878  1.00 15.00
ATOM    194  H   MET    24       5.362  16.762   6.081  1.00 15.00
```

```
ATOM    195  CA  MET    24       5.363  15.188   7.487  1.00 15.00
ATOM    196  CB  MET    24       4.458  13.986   7.280  1.00 15.00
ATOM    197  CG  MET    24       3.575  13.673   8.445  1.00 15.00
ATOM    198  SD  MET    24       2.244  12.505   8.107  1.00 15.00
ATOM    199  CE  MET    24       0.837  13.416   8.659  1.00 15.00
ATOM    200  C   MET    24       6.630  14.835   6.753  1.00 15.00
ATOM    201  O   MET    24       6.665  15.175   5.570  1.00 15.00
ATOM    202  N   ALA    25       7.654  14.148   7.259  1.00 15.00
ATOM    203  H   ALA    25       7.589  13.710   8.137  1.00 15.00
ATOM    204  CA  ALA    25       8.881  14.002   6.497  1.00 15.00
ATOM    205  CB  ALA    25       9.608  15.323   6.564  1.00 15.00
ATOM    206  C   ALA    25       9.830  12.899   6.931  1.00 15.00
ATOM    207  O   ALA    25       9.444  11.926   7.572  1.00 15.00
ATOM    208  N   ALA    26      11.082  13.115   6.505  1.00 15.00
ATOM    209  H   ALA    26      11.163  13.717   5.740  1.00 15.00
ATOM    210  CA  ALA    26      12.329  12.432   6.865  1.00 15.00
ATOM    211  CB  ALA    26      12.935  13.187   8.068  1.00 15.00
ATOM    212  C   ALA    26      12.512  10.933   7.128  1.00 15.00
ATOM    213  O   ALA    26      11.627  10.086   7.272  1.00 15.00
ATOM    214  N   SER    27      13.822  10.729   6.966  1.00 15.00
ATOM    215  H   SER    27      14.317  11.493   6.607  1.00 15.00
ATOM    216  CA  SER    27      14.571   9.493   7.103  1.00 15.00
ATOM    217  CB  SER    27      14.679   9.069   8.566  1.00 15.00
ATOM    218  OG  SER    27      15.388  10.041   9.323  1.00 15.00
ATOM    219  HG  SER    27      14.821  10.814   9.345  1.00 15.00
ATOM    220  C   SER    27      14.094   8.301   6.310  1.00 15.00
ATOM    221  O   SER    27      14.820   7.978   5.350  1.00 15.00
ATOM    222  N   ASP    28      12.950   7.639   6.524  1.00 15.00
ATOM    223  H   ASP    28      12.237   8.038   7.081  1.00 15.00
ATOM    224  CA  ASP    28      12.694   6.466   5.725  1.00 15.00
ATOM    225  CB  ASP    28      12.799   5.253   6.607  1.00 15.00
ATOM    226  CG  ASP    28      13.535   4.070   5.994  1.00 15.00
ATOM    227  OD1 ASP    28      12.968   3.379   5.155  1.00 15.00
ATOM    228  OD2 ASP    28      14.674   3.819   6.379  1.00 15.00
ATOM    229  C   ASP    28      11.362   6.507   5.031  1.00 15.00
ATOM    230  O   ASP    28      10.305   6.800   5.564  1.00 15.00
ATOM    231  N   ILE    29      11.501   6.204   3.765  1.00 15.00
ATOM    232  H   ILE    29      12.421   6.093   3.472  1.00 15.00
ATOM    233  CA  ILE    29      10.405   6.064   2.823  1.00 15.00
ATOM    234  CB  ILE    29      11.081   5.689   1.492  1.00 15.00
ATOM    235  CG2 ILE    29      11.960   4.441   1.729  1.00 15.00
ATOM    236  CG1 ILE    29      10.116   5.376   0.384  1.00 15.00
ATOM    237  CD1 ILE    29      11.007   5.005  -0.839  1.00 15.00
ATOM    238  C   ILE    29       9.317   5.069   3.232  1.00 15.00
ATOM    239  O   ILE    29       8.134   5.308   2.982  1.00 15.00
ATOM    240  N   SER    30       9.718   3.959   3.875  1.00 15.00
ATOM    241  H   SER    30      10.641   3.854   4.171  1.00 15.00
ATOM    242  CA  SER    30       8.798   2.902   4.233  1.00 15.00
ATOM    243  CB  SER    30       9.571   1.694   4.831  1.00 15.00
ATOM    244  OG  SER    30      10.652   2.033   5.681  1.00 15.00
ATOM    245  HG  SER    30      11.042   1.215   6.016  1.00 15.00
ATOM    246  C   SER    30       7.797   3.470   5.217  1.00 15.00
ATOM    247  O   SER    30       6.611   3.601   4.910  1.00 15.00
ATOM    248  N   LEU    31       8.390   3.997   6.300  1.00 15.00
ATOM    249  H   LEU    31       9.368   4.034   6.295  1.00 15.00
ATOM    250  CA  LEU    31       7.696   4.500   7.449  1.00 15.00
ATOM    251  CB  LEU    31       8.613   5.192   8.407  1.00 15.00
ATOM    252  CG  LEU    31       9.935   4.566   8.802  1.00 15.00
ATOM    253  CD1 LEU    31      10.460   5.451   9.953  1.00 15.00
ATOM    254  CD2 LEU    31       9.817   3.065   9.111  1.00 15.00
ATOM    255  C   LEU    31       6.616   5.468   7.120  1.00 15.00
ATOM    256  O   LEU    31       5.708   5.427   7.940  1.00 15.00
ATOM    257  N   LEU    32       6.591   6.278   6.066  1.00 15.00
ATOM    258  H   LEU    32       7.311   6.257   5.402  1.00 15.00
ATOM    259  CA  LEU    32       5.437   7.137   5.863  1.00 15.00
ATOM    260  CB  LEU    32       5.830   8.607   5.754  1.00 15.00
```

```
ATOM    261  CG  LEU    32       6.164   9.562   6.915  1.00 15.00
ATOM    262  CD1 LEU    32       7.428   9.247   7.698  1.00 15.00
ATOM    263  CD2 LEU    32       6.352  10.906   6.240  1.00 15.00
ATOM    264  C   LEU    32       4.612   6.824   4.615  1.00 15.00
ATOM    265  O   LEU    32       3.687   7.590   4.253  1.00 15.00
ATOM    266  N   ASP    33       4.905   5.738   3.877  1.00 15.00
ATOM    267  H   ASP    33       5.535   5.054   4.187  1.00 15.00
ATOM    268  CA  ASP    33       4.145   5.528   2.652  1.00 15.00
ATOM    269  CB  ASP    33       4.965   4.609   1.642  1.00 15.00
ATOM    270  CG  ASP    33       5.431   3.164   1.908  1.00 15.00
ATOM    271  OD1 ASP    33       6.631   2.972   2.088  1.00 15.00
ATOM    272  OD2 ASP    33       4.641   2.208   1.854  1.00 15.00
ATOM    273  C   ASP    33       2.744   4.972   2.948  1.00 15.00
ATOM    274  O   ASP    33       2.569   3.840   3.402  1.00 15.00
ATOM    275  N   ALA    34       1.734   5.851   2.802  1.00 15.00
ATOM    276  H   ALA    34       2.004   6.792   2.782  1.00 15.00
ATOM    277  CA  ALA    34       0.309   5.520   2.931  1.00 15.00
ATOM    278  CB  ALA    34      -0.103   4.344   2.089  1.00 15.00
ATOM    279  C   ALA    34      -0.217   5.182   4.316  1.00 15.00
ATOM    280  O   ALA    34       0.494   4.696   5.185  1.00 15.00
ATOM    281  N   GLN    35      -1.518   5.361   4.550  1.00 15.00
ATOM    282  H   GLN    35      -2.088   5.616   3.799  1.00 15.00
ATOM    283  CA  GLN    35      -2.114   5.232   5.874  1.00 15.00
ATOM    284  CB  GLN    35      -3.634   5.413   5.844  1.00 15.00
ATOM    285  CG  GLN    35      -4.428   6.473   5.028  1.00 15.00
ATOM    286  CD  GLN    35      -4.582   7.898   5.564  1.00 15.00
ATOM    287  OE1 GLN    35      -4.383   8.898   4.864  1.00 15.00
ATOM    288  NE2 GLN    35      -4.974   8.113   6.806  1.00 15.00
ATOM    289 HE21 GLN    35      -5.125   9.052   7.018  1.00 15.00
ATOM    290 HE22 GLN    35      -5.077   7.373   7.436  1.00 15.00
ATOM    291  C   GLN    35      -1.848   3.888   6.518  1.00 15.00
ATOM    292  O   GLN    35      -2.183   3.701   7.672  1.00 15.00
ATOM    293  N   SER    36      -1.304   2.916   5.809  1.00 15.00
ATOM    294  H   SER    36      -1.142   3.009   4.853  1.00 15.00
ATOM    295  CA  SER    36      -0.919   1.628   6.377  1.00 15.00
ATOM    296  CB  SER    36      -0.681   0.566   5.273  1.00 15.00
ATOM    297  OG  SER    36      -1.375   0.822   4.057  1.00 15.00
ATOM    298  HG  SER    36      -1.123   0.141   3.425  1.00 15.00
ATOM    299  C   SER    36       0.381   1.653   7.187  1.00 15.00
ATOM    300  O   SER    36       0.776   0.648   7.798  1.00 15.00
ATOM    301  N   ALA    37       1.120   2.748   7.077  1.00 15.00
ATOM    302  H   ALA    37       0.721   3.578   6.752  1.00 15.00
ATOM    303  CA  ALA    37       2.472   2.767   7.569  1.00 15.00
ATOM    304  CB  ALA    37       3.210   3.779   6.733  1.00 15.00
ATOM    305  C   ALA    37       2.652   3.058   9.049  1.00 15.00
ATOM    306  O   ALA    37       1.700   3.407   9.748  1.00 15.00
ATOM    307  N   PRO    38       3.863   2.870   9.581  1.00 15.00
ATOM    308  CD  PRO    38       4.818   1.876   9.127  1.00 15.00
ATOM    309  CA  PRO    38       4.271   3.347  10.882  1.00 15.00
ATOM    310  CB  PRO    38       5.731   2.992  10.938  1.00 15.00
ATOM    311  CG  PRO    38       5.690   1.617  10.345  1.00 15.00
ATOM    312  C   PRO    38       3.992   4.804  11.229  1.00 15.00
ATOM    313  O   PRO    38       2.855   5.125  11.586  1.00 15.00
ATOM    314  N   LEU    39       4.925   5.742  11.110  1.00 15.00
ATOM    315  H   LEU    39       5.720   5.599  10.559  1.00 15.00
ATOM    316  CA  LEU    39       4.619   7.039  11.619  1.00 15.00
ATOM    317  CB  LEU    39       5.943   7.585  11.991  1.00 15.00
ATOM    318  CG  LEU    39       6.506   7.020  13.262  1.00 15.00
ATOM    319  CD1 LEU    39       7.886   6.523  13.053  1.00 15.00
ATOM    320  CD2 LEU    39       6.488   8.099  14.302  1.00 15.00
ATOM    321  C   LEU    39       3.777   7.966  10.746  1.00 15.00
ATOM    322  O   LEU    39       3.960   9.191  10.663  1.00 15.00
ATOM    323  N   ARG    40       2.753   7.381  10.131  1.00 15.00
ATOM    324  H   ARG    40       2.534   6.436  10.263  1.00 15.00
ATOM    325  CA  ARG    40       1.804   8.159   9.408  1.00 15.00
ATOM    326  CB  ARG    40       1.444   7.535   8.104  1.00 15.00
```

```
ATOM    327  CG  ARG    40       0.978   8.622   7.126  1.00 15.00
ATOM    328  CD  ARG    40       0.871   8.086   5.680  1.00 15.00
ATOM    329  NE  ARG    40       0.842   9.136   4.678  1.00 15.00
ATOM    330  HE  ARG    40       1.499   9.861   4.728  1.00 15.00
ATOM    331  CZ  ARG    40      -0.042   9.137   3.689  1.00 15.00
ATOM    332 NH1  ARG    40      -0.962   8.213   3.510  1.00 15.00
ATOM    333 HH11 ARG    40      -1.596   8.291   2.741  1.00 15.00
ATOM    334 HH12 ARG    40      -1.049   7.453   4.153  1.00 15.00
ATOM    335 NH2  ARG    40      -0.013  10.127   2.841  1.00 15.00
ATOM    336 HH21 ARG    40       0.663  10.858   2.940  1.00 15.00
ATOM    337 HH22 ARG    40      -0.661  10.135   2.081  1.00 15.00
ATOM    338  C   ARG    40       0.627   8.085  10.344  1.00 15.00
ATOM    339  O   ARG    40      -0.315   7.299  10.169  1.00 15.00
ATOM    340  N   VAL    41       0.800   8.977  11.344  1.00 15.00
ATOM    341  H   VAL    41       1.552   9.602  11.257  1.00 15.00
ATOM    342  CA  VAL    41      -0.117   9.219  12.467  1.00 15.00
ATOM    343  CB  VAL    41       0.750   8.864  13.749  1.00 15.00
ATOM    344 CG1  VAL    41       1.666   9.995  14.150  1.00 15.00
ATOM    345 CG2  VAL    41      -0.168   8.431  14.844  1.00 15.00
ATOM    346  C   VAL    41      -0.576  10.685  12.283  1.00 15.00
ATOM    347  O   VAL    41       0.224  11.458  11.724  1.00 15.00
ATOM    348  N   TYR    42      -1.813  11.117  12.661  1.00 15.00
ATOM    349  H   TYR    42      -2.377  10.490  13.161  1.00 15.00
ATOM    350  CA  TYR    42      -2.384  12.498  12.416  1.00 15.00
ATOM    351  CB  TYR    42      -3.614  12.464  11.544  1.00 15.00
ATOM    352  CG  TYR    42      -3.362  11.765  10.247  1.00 15.00
ATOM    353 CD1  TYR    42      -4.191  10.728   9.924  1.00 15.00
ATOM    354 CE1  TYR    42      -3.997  10.093   8.728  1.00 15.00
ATOM    355 CD2  TYR    42      -2.341  12.189   9.423  1.00 15.00
ATOM    356 CE2  TYR    42      -2.159  11.556   8.229  1.00 15.00
ATOM    357  CZ  TYR    42      -2.994  10.515   7.894  1.00 15.00
ATOM    358  OH  TYR    42      -2.820   9.871   6.691  1.00 15.00
ATOM    359  HH  TYR    42      -3.508  10.139   6.067  1.00 15.00
ATOM    360  C   TYR    42      -2.820  13.341  13.629  1.00 15.00
ATOM    361  O   TYR    42      -3.761  12.966  14.339  1.00 15.00
ATOM    362  N   VAL    43      -2.220  14.496  13.897  1.00 15.00
ATOM    363  H   VAL    43      -1.753  14.987  13.188  1.00 15.00
ATOM    364  CA  VAL    43      -2.409  15.098  15.192  1.00 15.00
ATOM    365  CB  VAL    43      -1.080  15.989  15.447  1.00 15.00
ATOM    366 CG1  VAL    43      -1.124  16.771  16.758  1.00 15.00
ATOM    367 CG2  VAL    43       0.152  15.068  15.675  1.00 15.00
ATOM    368  C   VAL    43      -3.756  15.809  15.241  1.00 15.00
ATOM    369  O   VAL    43      -4.135  16.537  14.332  1.00 15.00
ATOM    370  N   GLU    44      -4.533  15.437  16.258  1.00 15.00
ATOM    371  H   GLU    44      -4.232  14.724  16.858  1.00 15.00
ATOM    372  CA  GLU    44      -5.828  16.042  16.462  1.00 15.00
ATOM    373  CB  GLU    44      -6.803  15.008  17.015  1.00 15.00
ATOM    374  CG  GLU    44      -8.282  15.479  16.878  1.00 15.00
ATOM    375  CD  GLU    44      -9.346  14.534  17.475  1.00 15.00
ATOM    376 OE1  GLU    44      -9.242  14.102  18.641  1.00 15.00
ATOM    377 OE2  GLU    44     -10.291  14.221  16.739  1.00 15.00
ATOM    378  C   GLU    44      -5.710  17.199  17.436  1.00 15.00
ATOM    379  O   GLU    44      -6.172  18.324  17.188  1.00 15.00
ATOM    380  N   GLU    45      -5.091  16.904  18.596  1.00 15.00
ATOM    381  H   GLU    45      -4.748  15.994  18.734  1.00 15.00
ATOM    382  CA  GLU    45      -4.865  17.870  19.668  1.00 15.00
ATOM    383  CB  GLU    45      -6.154  18.079  20.415  1.00 15.00
ATOM    384  CG  GLU    45      -6.360  17.294  21.704  1.00 15.00
ATOM    385  CD  GLU    45      -7.806  17.421  22.082  1.00 15.00
ATOM    386 OE1  GLU    45      -8.100  18.300  22.913  1.00 15.00
ATOM    387 OE2  GLU    45      -8.595  16.661  21.495  1.00 15.00
ATOM    388  C   GLU    45      -3.786  17.385  20.636  1.00 15.00
ATOM    389  O   GLU    45      -3.508  16.184  20.811  1.00 15.00
ATOM    390  N   LEU    46      -3.223  18.364  21.318  1.00 15.00
ATOM    391  H   LEU    46      -3.617  19.258  21.321  1.00 15.00
ATOM    392  CA  LEU    46      -2.088  18.155  22.193  1.00 15.00
```

```
ATOM    393  CB  LEU    46      -0.954  19.084  21.694  1.00 15.00
ATOM    394  CG  LEU    46       0.538  18.806  21.487  1.00 15.00
ATOM    395  CD1 LEU    46       0.791  17.718  20.434  1.00 15.00
ATOM    396  CD2 LEU    46       1.166  20.132  21.077  1.00 15.00
ATOM    397  C   LEU    46      -2.651  18.572  23.554  1.00 15.00
ATOM    398  O   LEU    46      -3.237  19.663  23.697  1.00 15.00
ATOM    399  N   LYS    47      -2.560  17.728  24.578  1.00 15.00
ATOM    400  H   LYS    47      -2.184  16.826  24.474  1.00 15.00
ATOM    401  CA  LYS    47      -3.004  18.149  25.872  1.00 15.00
ATOM    402  CB  LYS    47      -4.083  17.305  26.399  1.00 15.00
ATOM    403  CG  LYS    47      -5.284  17.010  25.530  1.00 15.00
ATOM    404  CD  LYS    47      -6.064  15.948  26.332  1.00 15.00
ATOM    405  CE  LYS    47      -5.253  14.657  26.755  1.00 15.00
ATOM    406  NZ  LYS    47      -5.628  14.179  28.092  1.00 15.00
ATOM    407  HZ1 LYS    47      -5.190  13.255  28.270  1.00 15.00
ATOM    408  HZ2 LYS    47      -6.662  14.085  28.144  1.00 15.00
ATOM    409  HZ3 LYS    47      -5.301  14.864  28.802  1.00 15.00
ATOM    410  C   LYS    47      -1.771  17.883  26.691  1.00 15.00
ATOM    411  O   LYS    47      -1.328  16.741  26.854  1.00 15.00
ATOM    412  N   PRO    48      -1.122  18.953  27.139  1.00 15.00
ATOM    413  CD  PRO    48      -1.229  20.293  26.595  1.00 15.00
ATOM    414  CA  PRO    48      -0.293  18.973  28.336  1.00 15.00
ATOM    415  CB  PRO    48       0.112  20.401  28.489  1.00 15.00
ATOM    416  CG  PRO    48      -0.999  21.165  27.830  1.00 15.00
ATOM    417  C   PRO    48      -1.049  18.464  29.556  1.00 15.00
ATOM    418  O   PRO    48      -2.293  18.416  29.538  1.00 15.00
ATOM    419  N   THR    49      -0.331  18.077  30.614  1.00 15.00
ATOM    420  H   THR    49       0.644  18.056  30.564  1.00 15.00
ATOM    421  CA  THR    49      -1.007  17.736  31.850  1.00 15.00
ATOM    422  CB  THR    49      -0.353  16.462  32.485  1.00 15.00
ATOM    423  OG1 THR    49       1.046  16.640  32.367  1.00 15.00
ATOM    424  HG1 THR    49       1.242  16.437  31.435  1.00 15.00
ATOM    425  CG2 THR    49      -0.775  15.151  31.845  1.00 15.00
ATOM    426  C   THR    49      -0.904  18.966  32.787  1.00 15.00
ATOM    427  O   THR    49      -0.136  19.919  32.543  1.00 15.00
ATOM    428  N   PRO    50      -1.653  18.996  33.904  1.00 15.00
ATOM    429  CD  PRO    50      -3.049  18.563  34.027  1.00 15.00
ATOM    430  CA  PRO    50      -1.241  19.581  35.186  1.00 15.00
ATOM    431  CB  PRO    50      -2.134  18.887  36.241  1.00 15.00
ATOM    432  CG  PRO    50      -3.027  17.909  35.443  1.00 15.00
ATOM    433  C   PRO    50       0.237  19.422  35.455  1.00 15.00
ATOM    434  O   PRO    50       0.926  20.405  35.660  1.00 15.00
ATOM    435  N   GLU    51       0.773  18.227  35.303  1.00 15.00
ATOM    436  H   GLU    51       0.235  17.518  34.908  1.00 15.00
ATOM    437  CA  GLU    51       2.176  18.016  35.582  1.00 15.00
ATOM    438  CB  GLU    51       2.374  16.561  36.056  1.00 15.00
ATOM    439  CG  GLU    51       1.274  15.540  35.753  1.00 15.00
ATOM    440  CD  GLU    51      -0.025  15.646  36.557  1.00 15.00
ATOM    441  OE1 GLU    51      -0.824  14.717  36.471  1.00 15.00
ATOM    442  OE2 GLU    51      -0.256  16.631  37.258  1.00 15.00
ATOM    443  C   GLU    51       3.101  18.328  34.401  1.00 15.00
ATOM    444  O   GLU    51       4.135  17.667  34.265  1.00 15.00
ATOM    445  N   GLY    52       2.790  19.283  33.504  1.00 15.00
ATOM    446  H   GLY    52       1.964  19.795  33.629  1.00 15.00
ATOM    447  CA  GLY    52       3.671  19.592  32.390  1.00 15.00
ATOM    448  C   GLY    52       4.004  18.450  31.382  1.00 15.00
ATOM    449  O   GLY    52       4.743  18.750  30.422  1.00 15.00
ATOM    450  N   ASP    53       3.563  17.162  31.500  1.00 15.00
ATOM    451  H   ASP    53       3.132  16.897  32.338  1.00 15.00
ATOM    452  CA  ASP    53       3.786  16.146  30.441  1.00 15.00
ATOM    453  CB  ASP    53       3.418  14.712  30.658  1.00 15.00
ATOM    454  CG  ASP    53       3.455  14.159  32.044  1.00 15.00
ATOM    455  OD1 ASP    53       2.401  13.622  32.427  1.00 15.00
ATOM    456  OD2 ASP    53       4.519  14.248  32.680  1.00 15.00
ATOM    457  C   ASP    53       2.846  16.348  29.275  1.00 15.00
ATOM    458  O   ASP    53       1.706  16.759  29.525  1.00 15.00
```

```
ATOM    459  N   LEU    54       3.231  16.013  28.049  1.00 15.00
ATOM    460  H   LEU    54       4.080  15.550  27.895  1.00 15.00
ATOM    461  CA  LEU    54       2.240  16.047  27.002  1.00 15.00
ATOM    462  CB  LEU    54       2.807  16.287  25.624  1.00 15.00
ATOM    463  CG  LEU    54       3.367  17.643  25.390  1.00 15.00
ATOM    464  CD1 LEU    54       3.970  17.681  24.003  1.00 15.00
ATOM    465  CD2 LEU    54       2.282  18.681  25.566  1.00 15.00
ATOM    466  C   LEU    54       1.648  14.649  26.987  1.00 15.00
ATOM    467  O   LEU    54       2.199  13.684  27.543  1.00 15.00
ATOM    468  N   GLU    55       0.467  14.610  26.373  1.00 15.00
ATOM    469  H   GLU    55       0.008  15.460  26.189  1.00 15.00
ATOM    470  CA  GLU    55      -0.264  13.417  25.984  1.00 15.00
ATOM    471  CB  GLU    55      -1.443  13.148  26.842  1.00 15.00
ATOM    472  CG  GLU    55      -1.263  12.023  27.814  1.00 15.00
ATOM    473  CD  GLU    55      -2.452  11.885  28.761  1.00 15.00
ATOM    474  OE1 GLU    55      -3.610  12.092  28.370  1.00 15.00
ATOM    475  OE2 GLU    55      -2.205  11.564  29.918  1.00 15.00
ATOM    476  C   GLU    55      -0.782  13.937  24.657  1.00 15.00
ATOM    477  O   GLU    55      -1.316  15.062  24.560  1.00 15.00
ATOM    478  N   ILE    56      -0.547  13.105  23.641  1.00 15.00
ATOM    479  H   ILE    56      -0.301  12.198  23.862  1.00 15.00
ATOM    480  CA  ILE    56      -0.825  13.389  22.240  1.00 15.00
ATOM    481  CB  ILE    56       0.303  12.812  21.307  1.00 15.00
ATOM    482  CG2 ILE    56       0.206  13.597  20.019  1.00 15.00
ATOM    483  CG1 ILE    56       1.733  12.975  21.810  1.00 15.00
ATOM    484  CD1 ILE    56       2.350  14.382  21.902  1.00 15.00
ATOM    485  C   ILE    56      -2.149  12.670  21.987  1.00 15.00
ATOM    486  O   ILE    56      -2.361  11.524  22.419  1.00 15.00
ATOM    487  N   LEU    57      -3.050  13.398  21.344  1.00 15.00
ATOM    488  H   LEU    57      -2.839  14.320  21.087  1.00 15.00
ATOM    489  CA  LEU    57      -4.340  12.882  21.005  1.00 15.00
ATOM    490  CB  LEU    57      -5.348  13.934  21.428  1.00 15.00
ATOM    491  CG  LEU    57      -6.631  13.612  22.236  1.00 15.00
ATOM    492  CD1 LEU    57      -7.577  12.736  21.387  1.00 15.00
ATOM    493  CD2 LEU    57      -6.249  12.955  23.582  1.00 15.00
ATOM    494  C   LEU    57      -4.262  12.653  19.501  1.00 15.00
ATOM    495  O   LEU    57      -4.251  13.539  18.628  1.00 15.00
ATOM    496  N   LEU    58      -4.273  11.377  19.211  1.00 15.00
ATOM    497  H   LEU    58      -4.479  10.707  19.898  1.00 15.00
ATOM    498  CA  LEU    58      -3.986  10.931  17.874  1.00 15.00
ATOM    499  CB  LEU    58      -2.657  10.278  17.974  1.00 15.00
ATOM    500  CG  LEU    58      -1.493  11.124  17.630  1.00 15.00
ATOM    501  CD1 LEU    58      -0.295  10.244  17.971  1.00 15.00
ATOM    502  CD2 LEU    58      -1.529  11.616  16.174  1.00 15.00
ATOM    503  C   LEU    58      -4.833  10.080  16.926  1.00 15.00
ATOM    504  O   LEU    58      -5.237   8.976  17.270  1.00 15.00
ATOM    505  N   GLN    59      -5.063  10.508  15.687  1.00 15.00
ATOM    506  H   GLN    59      -4.674  11.355  15.374  1.00 15.00
ATOM    507  CA  GLN    59      -5.718   9.640  14.741  1.00 15.00
ATOM    508  CB  GLN    59      -6.409  10.450  13.660  1.00 15.00
ATOM    509  CG  GLN    59      -7.436  11.555  13.995  1.00 15.00
ATOM    510  CD  GLN    59      -7.448  12.634  12.896  1.00 15.00
ATOM    511  OE1 GLN    59      -6.524  12.740  12.095  1.00 15.00
ATOM    512  NE2 GLN    59      -8.433  13.505  12.729  1.00 15.00
ATOM    513 HE21 GLN    59      -9.192  13.458  13.336  1.00 15.00
ATOM    514 HE22 GLN    59      -8.308  14.194  12.037  1.00 15.00
ATOM    515  C   GLN    59      -4.585   8.820  14.134  1.00 15.00
ATOM    516  O   GLN    59      -3.427   9.227  14.162  1.00 15.00
ATOM    517  N   LYS    60      -4.871   7.676  13.556  1.00 15.00
ATOM    518  H   LYS    60      -5.812   7.402  13.521  1.00 15.00
ATOM    519  CA  LYS    60      -3.921   6.758  12.928  1.00 15.00
ATOM    520  CB  LYS    60      -3.389   5.733  13.950  1.00 15.00
ATOM    521  CG  LYS    60      -1.881   5.388  13.943  1.00 15.00
ATOM    522  CD  LYS    60      -1.495   3.915  13.642  1.00 15.00
ATOM    523  CE  LYS    60      -0.513   3.960  12.446  1.00 15.00
ATOM    524  NZ  LYS    60      -1.025   4.730  11.306  1.00 15.00
```

```
ATOM    525  HZ1 LYS    60      -1.937   4.332  11.003  1.00 15.00
ATOM    526  HZ2 LYS    60      -0.344   4.673  10.522  1.00 15.00
ATOM    527  HZ3 LYS    60      -1.154   5.726  11.578  1.00 15.00
ATOM    528  C   LYS    60      -4.803   6.036  11.895  1.00 15.00
ATOM    529  O   LYS    60      -5.884   6.548  11.530  1.00 15.00
ATOM    530  N   TRP    61      -4.406   4.855  11.387  1.00 15.00
ATOM    531  H   TRP    61      -3.483   4.557  11.547  1.00 15.00
ATOM    532  CA  TRP    61      -5.204   4.029  10.455  1.00 15.00
ATOM    533  CB  TRP    61      -5.081   4.569   8.996  1.00 15.00
ATOM    534  CG  TRP    61      -6.091   4.021   7.994  1.00 15.00
ATOM    535  CD2 TRP    61      -7.263   4.610   7.611  1.00 15.00
ATOM    536  CE2 TRP    61      -7.781   3.676   6.740  1.00 15.00
ATOM    537  CE3 TRP    61      -7.974   5.761   7.862  1.00 15.00
ATOM    538  CD1 TRP    61      -5.905   2.820   7.381  1.00 15.00
ATOM    539  NE1 TRP    61      -6.952   2.644   6.624  1.00 15.00
ATOM    540  HE1 TRP    61      -7.074   1.870   6.035  1.00 15.00
ATOM    541  CZ2 TRP    61      -8.999   3.872   6.123  1.00 15.00
ATOM    542  CZ3 TRP    61      -9.201   5.966   7.244  1.00 15.00
ATOM    543  CH2 TRP    61      -9.717   5.027   6.376  1.00 15.00
ATOM    544  C   TRP    61      -4.516   2.663  10.605  1.00 15.00
ATOM    545  O   TRP    61      -3.568   2.297   9.894  1.00 15.00
ATOM    546  N   GLU    62      -4.883   1.890  11.620  1.00 15.00
ATOM    547  H   GLU    62      -5.692   2.100  12.134  1.00 15.00
ATOM    548  CA  GLU    62      -4.230   0.612  11.848  1.00 15.00
ATOM    549  CB  GLU    62      -3.500   0.539  13.204  1.00 15.00
ATOM    550  CG  GLU    62      -3.399  -0.816  13.988  1.00 15.00
ATOM    551  CD  GLU    62      -4.584  -1.199  14.897  1.00 15.00
ATOM    552  OE1 GLU    62      -4.879  -2.405  15.015  1.00 15.00
ATOM    553  OE2 GLU    62      -5.206  -0.280  15.473  1.00 15.00
ATOM    554  C   GLU    62      -5.423  -0.270  11.890  1.00 15.00
ATOM    555  O   GLU    62      -6.460   0.087  12.472  1.00 15.00
ATOM    556  N   ASN    63      -5.173  -1.358  11.163  1.00 15.00
ATOM    557  H   ASN    63      -4.331  -1.377  10.660  1.00 15.00
ATOM    558  CA  ASN    63      -6.046  -2.518  11.046  1.00 15.00
ATOM    559  CB  ASN    63      -5.334  -3.647  11.791  1.00 15.00
ATOM    560  CG  ASN    63      -5.166  -4.856  10.893  1.00 15.00
ATOM    561  OD1 ASN    63      -4.102  -5.051  10.295  1.00 15.00
ATOM    562  ND2 ASN    63      -6.215  -5.667  10.750  1.00 15.00
ATOM    563 HD21 ASN    63      -6.100  -6.454  10.184  1.00 15.00
ATOM    564 HD22 ASN    63      -7.048  -5.423  11.214  1.00 15.00
ATOM    565  C   ASN    63      -7.532  -2.481  11.452  1.00 15.00
ATOM    566  O   ASN    63      -8.115  -3.509  11.817  1.00 15.00
ATOM    567  N   GLY    64      -8.202  -1.334  11.321  1.00 15.00
ATOM    568  H   GLY    64      -7.727  -0.513  11.085  1.00 15.00
ATOM    569  CA  GLY    64      -9.593  -1.171  11.739  1.00 15.00
ATOM    570  C   GLY    64     -10.022   0.305  11.772  1.00 15.00
ATOM    571  O   GLY    64     -10.369   0.920  12.793  1.00 15.00
ATOM    572  N   GLU    65     -10.001   0.784  10.522  1.00 15.00
ATOM    573  H   GLU    65      -9.708   0.168   9.826  1.00 15.00
ATOM    574  CA  GLU    65     -10.344   2.130  10.124  1.00 15.00
ATOM    575  CB  GLU    65     -11.814   2.387  10.560  1.00 15.00
ATOM    576  CG  GLU    65     -12.943   1.558   9.894  1.00 15.00
ATOM    577  CD  GLU    65     -13.891   0.834  10.872  1.00 15.00
ATOM    578  OE1 GLU    65     -15.089   0.717  10.552  1.00 15.00
ATOM    579  OE2 GLU    65     -13.416   0.376  11.934  1.00 15.00
ATOM    580  C   GLU    65      -9.355   3.146  10.710  1.00 15.00
ATOM    581  O   GLU    65      -8.189   2.822  11.036  1.00 15.00
ATOM    582  N   CYS    66      -9.837   4.395  10.775  1.00 15.00
ATOM    583  H   CYS    66     -10.770   4.563  10.549  1.00 15.00
ATOM    584  CA  CYS    66      -9.096   5.545  11.243  1.00 15.00
ATOM    585  C   CYS    66      -9.175   5.348  12.752  1.00 15.00
ATOM    586  O   CYS    66     -10.057   5.867  13.433  1.00 15.00
ATOM    587  CB  CYS    66      -9.855   6.772  10.664  1.00 15.00
ATOM    588  SG  CYS    66      -9.080   8.426  10.519  1.00 15.00
ATOM    589  N   ALA    67      -8.315   4.501  13.290  1.00 15.00
ATOM    590  H   ALA    67      -7.700   4.019  12.701  1.00 15.00
```

```
ATOM    591  CA  ALA    67       -8.288   4.238  14.714  1.00 15.00
ATOM    592  CB  ALA    67       -7.522   2.916  14.956  1.00 15.00
ATOM    593  C   ALA    67       -7.666   5.336  15.588  1.00 15.00
ATOM    594  O   ALA    67       -6.487   5.637  15.421  1.00 15.00
ATOM    595  N   GLN    68       -8.368   5.992  16.514  1.00 15.00
ATOM    596  H   GLN    68       -9.306   5.751  16.620  1.00 15.00
ATOM    597  CA  GLN    68       -7.751   6.967  17.440  1.00 15.00
ATOM    598  CB  GLN    68       -8.761   7.536  18.444  1.00 15.00
ATOM    599  CG  GLN    68      -10.169   7.982  18.002  1.00 15.00
ATOM    600  CD  GLN    68      -11.251   7.980  19.122  1.00 15.00
ATOM    601  OE1 GLN    68      -10.966   7.933  20.334  1.00 15.00
ATOM    602  NE2 GLN    68      -12.550   8.031  18.766  1.00 15.00
ATOM    603 HE21 GLN    68      -12.787   8.063  17.818  1.00 15.00
ATOM    604 HE22 GLN    68      -13.193   8.040  19.504  1.00 15.00
ATOM    605  C   GLN    68       -6.623   6.325  18.294  1.00 15.00
ATOM    606  O   GLN    68       -6.601   5.107  18.504  1.00 15.00
ATOM    607  N   LYS    69       -5.719   7.099  18.895  1.00 15.00
ATOM    608  H   LYS    69       -5.822   8.069  18.818  1.00 15.00
ATOM    609  CA  LYS    69       -4.575   6.649  19.686  1.00 15.00
ATOM    610  CB  LYS    69       -3.300   6.629  18.852  1.00 15.00
ATOM    611  CG  LYS    69       -3.125   5.600  17.727  1.00 15.00
ATOM    612  CD  LYS    69       -2.911   4.174  18.241  1.00 15.00
ATOM    613  CE  LYS    69       -1.716   3.986  19.210  1.00 15.00
ATOM    614  NZ  LYS    69       -2.003   4.405  20.583  1.00 15.00
ATOM    615  HZ1 LYS    69       -2.847   3.906  20.930  1.00 15.00
ATOM    616  HZ2 LYS    69       -2.165   5.431  20.617  1.00 15.00
ATOM    617  HZ3 LYS    69       -1.189   4.171  21.187  1.00 15.00
ATOM    618  C   LYS    69       -4.372   7.682  20.792  1.00 15.00
ATOM    619  O   LYS    69       -4.661   8.849  20.525  1.00 15.00
ATOM    620  N   LYS    70       -3.957   7.391  22.038  1.00 15.00
ATOM    621  H   LYS    70       -3.847   6.453  22.291  1.00 15.00
ATOM    622  CA  LYS    70       -3.542   8.412  23.042  1.00 15.00
ATOM    623  CB  LYS    70       -4.325   8.367  24.367  1.00 15.00
ATOM    624  CG  LYS    70       -5.549   9.272  24.548  1.00 15.00
ATOM    625  CD  LYS    70       -5.652   9.523  26.074  1.00 15.00
ATOM    626  CE  LYS    70       -6.815  10.395  26.578  1.00 15.00
ATOM    627  NZ  LYS    70       -8.033   9.610  26.757  1.00 15.00
ATOM    628  HZ1 LYS    70       -8.305   9.178  25.852  1.00 15.00
ATOM    629  HZ2 LYS    70       -8.798  10.228  27.094  1.00 15.00
ATOM    630  HZ3 LYS    70       -7.857   8.866  27.461  1.00 15.00
ATOM    631  C   LYS    70       -2.074   8.076  23.401  1.00 15.00
ATOM    632  O   LYS    70       -1.815   7.005  23.970  1.00 15.00
ATOM    633  N   ILE    71       -1.080   8.895  23.019  1.00 15.00
ATOM    634  H   ILE    71       -1.310   9.778  22.655  1.00 15.00
ATOM    635  CA  ILE    71        0.320   8.581  23.277  1.00 15.00
ATOM    636  CB  ILE    71        1.309   8.915  22.140  1.00 15.00
ATOM    637  CG2 ILE    71        2.716   8.648  22.630  1.00 15.00
ATOM    638  CG1 ILE    71        1.064   8.073  20.918  1.00 15.00
ATOM    639  CD1 ILE    71        1.048   6.556  21.158  1.00 15.00
ATOM    640  C   ILE    71        0.636   9.510  24.392  1.00 15.00
ATOM    641  O   ILE    71        0.613  10.718  24.225  1.00 15.00
ATOM    642  N   ILE    72        0.911   8.952  25.545  1.00 15.00
ATOM    643  H   ILE    72        1.071   7.987  25.575  1.00 15.00
ATOM    644  CA  ILE    72        1.222   9.765  26.703  1.00 15.00
ATOM    645  CB  ILE    72        1.031   8.933  27.945  1.00 15.00
ATOM    646  CG2 ILE    72        1.160   9.851  29.134  1.00 15.00
ATOM    647  CG1 ILE    72       -0.317   8.224  27.900  1.00 15.00
ATOM    648  CD1 ILE    72       -0.584   7.112  28.931  1.00 15.00
ATOM    649  C   ILE    72        2.673  10.062  26.435  1.00 15.00
ATOM    650  O   ILE    72        3.390   9.152  26.000  1.00 15.00
ATOM    651  N   ALA    73        3.113  11.290  26.610  1.00 15.00
ATOM    652  H   ALA    73        2.545  11.974  27.008  1.00 15.00
ATOM    653  CA  ALA    73        4.498  11.592  26.306  1.00 15.00
ATOM    654  CB  ALA    73        4.575  12.626  25.199  1.00 15.00
ATOM    655  C   ALA    73        5.060  12.161  27.576  1.00 15.00
ATOM    656  O   ALA    73        4.867  13.361  27.833  1.00 15.00
```

```
ATOM    657  N   GLU    74       5.726  11.350  28.411  1.00 15.00
ATOM    658  H   GLU    74       6.070  10.481  28.109  1.00 15.00
ATOM    659  CA  GLU    74       6.163  11.930  29.668  1.00 15.00
ATOM    660  CB  GLU    74       6.279  10.911  30.827  1.00 15.00
ATOM    661  CG  GLU    74       6.492   9.420  30.587  1.00 15.00
ATOM    662  CD  GLU    74       5.251   8.640  30.145  1.00 15.00
ATOM    663  OE1 GLU    74       4.138   8.978  30.587  1.00 15.00
ATOM    664  OE2 GLU    74       5.418   7.670  29.383  1.00 15.00
ATOM    665  C   GLU    74       7.480  12.668  29.574  1.00 15.00
ATOM    666  O   GLU    74       8.475  12.285  28.963  1.00 15.00
ATOM    667  N   LYS    75       7.224  13.863  30.095  1.00 15.00
ATOM    668  H   LYS    75       6.303  14.002  30.388  1.00 15.00
ATOM    669  CA  LYS    75       8.121  14.982  30.272  1.00 15.00
ATOM    670  CB  LYS    75       7.387  15.831  31.312  1.00 15.00
ATOM    671  CG  LYS    75       8.046  16.942  32.123  1.00 15.00
ATOM    672  CD  LYS    75       8.488  18.077  31.197  1.00 15.00
ATOM    673  CE  LYS    75       8.271  19.439  31.886  1.00 15.00
ATOM    674  NZ  LYS    75       6.852  19.718  32.047  1.00 15.00
ATOM    675  HZ1 LYS    75       6.390  19.736  31.116  1.00 15.00
ATOM    676  HZ2 LYS    75       6.736  20.641  32.512  1.00 15.00
ATOM    677  HZ3 LYS    75       6.420  18.977  32.636  1.00 15.00
ATOM    678  C   LYS    75       9.519  14.533  30.660  1.00 15.00
ATOM    679  O   LYS    75       9.638  13.538  31.376  1.00 15.00
ATOM    680  N   THR    76      10.592  15.168  30.189  1.00 15.00
ATOM    681  H   THR    76      10.525  15.960  29.611  1.00 15.00
ATOM    682  CA  THR    76      11.931  14.739  30.568  1.00 15.00
ATOM    683  CB  THR    76      12.563  13.944  29.379  1.00 15.00
ATOM    684  OG1 THR    76      12.870  12.674  29.940  1.00 15.00
ATOM    685  HG1 THR    76      12.044  12.282  30.255  1.00 15.00
ATOM    686  CG2 THR    76      13.818  14.515  28.772  1.00 15.00
ATOM    687  C   THR    76      12.650  16.026  30.895  1.00 15.00
ATOM    688  O   THR    76      12.242  17.103  30.453  1.00 15.00
ATOM    689  N   LYS    77      13.714  15.893  31.676  1.00 15.00
ATOM    690  H   LYS    77      13.969  14.985  31.929  1.00 15.00
ATOM    691  CA  LYS    77      14.495  17.006  32.177  1.00 15.00
ATOM    692  CB  LYS    77      15.757  16.383  32.820  1.00 15.00
ATOM    693  CG  LYS    77      16.994  16.038  31.940  1.00 15.00
ATOM    694  CD  LYS    77      17.528  14.583  31.999  1.00 15.00
ATOM    695  CE  LYS    77      16.770  13.585  31.099  1.00 15.00
ATOM    696  NZ  LYS    77      15.355  13.480  31.440  1.00 15.00
ATOM    697  HZ1 LYS    77      14.899  14.397  31.273  1.00 15.00
ATOM    698  HZ2 LYS    77      15.249  13.212  32.439  1.00 15.00
ATOM    699  HZ3 LYS    77      14.909  12.760  30.837  1.00 15.00
ATOM    700  C   LYS    77      14.872  18.154  31.211  1.00 15.00
ATOM    701  O   LYS    77      15.346  19.208  31.656  1.00 15.00
ATOM    702  N   ILE    78      14.723  17.989  29.886  1.00 15.00
ATOM    703  H   ILE    78      14.199  17.238  29.547  1.00 15.00
ATOM    704  CA  ILE    78      15.051  19.010  28.908  1.00 15.00
ATOM    705  CB  ILE    78      15.849  18.488  27.700  1.00 15.00
ATOM    706  CG2 ILE    78      16.193  19.710  26.892  1.00 15.00
ATOM    707  CG1 ILE    78      17.144  17.823  28.021  1.00 15.00
ATOM    708  CD1 ILE    78      17.666  17.132  26.747  1.00 15.00
ATOM    709  C   ILE    78      13.664  19.345  28.389  1.00 15.00
ATOM    710  O   ILE    78      13.049  18.413  27.852  1.00 15.00
ATOM    711  N   PRO    79      13.118  20.571  28.477  1.00 15.00
ATOM    712  CD  PRO    79      13.477  21.559  29.464  1.00 15.00
ATOM    713  CA  PRO    79      11.992  21.025  27.668  1.00 15.00
ATOM    714  CB  PRO    79      11.723  22.411  28.192  1.00 15.00
ATOM    715  CG  PRO    79      12.181  22.318  29.612  1.00 15.00
ATOM    716  C   PRO    79      12.310  20.960  26.149  1.00 15.00
ATOM    717  O   PRO    79      13.267  21.540  25.613  1.00 15.00
ATOM    718  N   ALA    80      11.525  20.019  25.606  1.00 15.00
ATOM    719  H   ALA    80      11.017  19.489  26.258  1.00 15.00
ATOM    720  CA  ALA    80      11.364  19.544  24.242  1.00 15.00
ATOM    721  CB  ALA    80      12.484  19.888  23.267  1.00 15.00
ATOM    722  C   ALA    80      11.365  18.011  24.327  1.00 15.00
```

```
ATOM    723  O   ALA  80      10.400  17.409  23.869  1.00 15.00
ATOM    724  N   VAL  81      12.326  17.306  24.925  1.00 15.00
ATOM    725  H   VAL  81      12.890  17.738  25.600  1.00 15.00
ATOM    726  CA  VAL  81      12.363  15.856  24.808  1.00 15.00
ATOM    727  CB  VAL  81      13.823  15.409  25.000  1.00 15.00
ATOM    728  CG1 VAL  81      14.039  13.917  25.075  1.00 15.00
ATOM    729  CG2 VAL  81      14.546  15.888  23.765  1.00 15.00
ATOM    730  C   VAL  81      11.433  15.293  25.860  1.00 15.00
ATOM    731  O   VAL  81      11.426  15.801  27.002  1.00 15.00
ATOM    732  N   PHE  82      10.609  14.303  25.487  1.00 15.00
ATOM    733  H   PHE  82      10.631  13.984  24.557  1.00 15.00
ATOM    734  CA  PHE  82       9.711  13.645  26.427  1.00 15.00
ATOM    735  CB  PHE  82       8.257  14.028  26.173  1.00 15.00
ATOM    736  CG  PHE  82       7.807  15.490  26.342  1.00 15.00
ATOM    737  CD1 PHE  82       6.902  15.846  27.332  1.00 15.00
ATOM    738  CD2 PHE  82       8.237  16.479  25.488  1.00 15.00
ATOM    739  CE1 PHE  82       6.443  17.150  27.451  1.00 15.00
ATOM    740  CE2 PHE  82       7.783  17.778  25.605  1.00 15.00
ATOM    741  CZ  PHE  82       6.885  18.122  26.581  1.00 15.00
ATOM    742  C   PHE  82       9.913  12.170  26.110  1.00 15.00
ATOM    743  O   PHE  82      10.039  11.854  24.916  1.00 15.00
ATOM    744  N   LYS  83      10.002  11.218  27.059  1.00 15.00
ATOM    745  H   LYS  83       9.830  11.446  27.993  1.00 15.00
ATOM    746  CA  LYS  83      10.187   9.817  26.711  1.00 15.00
ATOM    747  CB  LYS  83      10.861   9.070  27.811  1.00 15.00
ATOM    748  CG  LYS  83      11.444   7.793  27.234  1.00 15.00
ATOM    749  CD  LYS  83      11.976   6.863  28.347  1.00 15.00
ATOM    750  CE  LYS  83      13.138   5.894  27.983  1.00 15.00
ATOM    751  NZ  LYS  83      12.837   4.911  26.940  1.00 15.00
ATOM    752  HZ1 LYS  83      13.696   4.366  26.726  1.00 15.00
ATOM    753  HZ2 LYS  83      12.512   5.402  26.083  1.00 15.00
ATOM    754  HZ3 LYS  83      12.091   4.269  27.276  1.00 15.00
ATOM    755  C   LYS  83       8.755   9.340  26.563  1.00 15.00
ATOM    756  O   LYS  83       7.915   9.678  27.419  1.00 15.00
ATOM    757  N   ILE  84       8.396   8.697  25.439  1.00 15.00
ATOM    758  H   ILE  84       9.044   8.634  24.719  1.00 15.00
ATOM    759  CA  ILE  84       7.020   8.272  25.253  1.00 15.00
ATOM    760  CB  ILE  84       6.427   8.887  23.950  1.00 15.00
ATOM    761  CG2 ILE  84       6.804  10.353  23.964  1.00 15.00
ATOM    762  CG1 ILE  84       6.945   8.279  22.661  1.00 15.00
ATOM    763  CD1 ILE  84       6.108   8.660  21.392  1.00 15.00
ATOM    764  C   ILE  84       6.715   6.778  25.245  1.00 15.00
ATOM    765  O   ILE  84       5.525   6.439  25.401  1.00 15.00
ATOM    766  N   ASP  85       7.731   5.871  25.097  1.00 15.00
ATOM    767  H   ASP  85       8.626   6.250  24.971  1.00 15.00
ATOM    768  CA  ASP  85       7.563   4.391  25.061  1.00 15.00
ATOM    769  CB  ASP  85       7.478   3.880  26.546  1.00 15.00
ATOM    770  CG  ASP  85       7.575   2.356  26.777  1.00 15.00
ATOM    771  OD1 ASP  85       8.522   1.756  26.267  1.00 15.00
ATOM    772  OD2 ASP  85       6.725   1.762  27.472  1.00 15.00
ATOM    773  C   ASP  85       6.328   3.907  24.218  1.00 15.00
ATOM    774  O   ASP  85       5.502   3.112  24.693  1.00 15.00
ATOM    775  N   ALA  86       6.162   4.303  22.933  1.00 15.00
ATOM    776  H   ALA  86       6.855   4.866  22.519  1.00 15.00
ATOM    777  CA  ALA  86       4.967   4.009  22.130  1.00 15.00
ATOM    778  CB  ALA  86       3.785   4.897  22.508  1.00 15.00
ATOM    779  C   ALA  86       5.149   4.222  20.629  1.00 15.00
ATOM    780  O   ALA  86       5.780   5.187  20.153  1.00 15.00
ATOM    781  N   LEU  87       4.561   3.253  19.895  1.00 15.00
ATOM    782  H   LEU  87       4.125   2.529  20.387  1.00 15.00
ATOM    783  CA  LEU  87       4.558   3.219  18.436  1.00 15.00
ATOM    784  CB  LEU  87       3.822   4.499  17.974  1.00 15.00
ATOM    785  CG  LEU  87       2.806   4.565  16.849  1.00 15.00
ATOM    786  CD1 LEU  87       2.449   6.036  16.682  1.00 15.00
ATOM    787  CD2 LEU  87       3.349   3.983  15.527  1.00 15.00
ATOM    788  C   LEU  87       6.011   3.141  17.938  1.00 15.00
```

```
ATOM    789  O   LEU   87       6.437   3.910  17.077  1.00 15.00
ATOM    790  N   ASN   88       6.800   2.222  18.543  1.00 15.00
ATOM    791  H   ASN   88       6.369   1.656  19.215  1.00 15.00
ATOM    792  CA  ASN   88       8.251   2.001  18.331  1.00 15.00
ATOM    793  CB  ASN   88       8.641   1.532  16.908  1.00 15.00
ATOM    794  CG  ASN   88       8.163   0.153  16.450  1.00 15.00
ATOM    795  OD1 ASN   88       6.984  -0.242  16.465  1.00 15.00
ATOM    796  ND2 ASN   88       9.100  -0.668  16.006  1.00 15.00
ATOM    797 HD21 ASN   88      10.031  -0.359  16.034  1.00 15.00
ATOM    798 HD22 ASN   88       8.815  -1.523  15.627  1.00 15.00
ATOM    799  C   ASN   88       9.075   3.263  18.572  1.00 15.00
ATOM    800  O   ASN   88      10.277   3.277  18.304  1.00 15.00
ATOM    801  N   GLU   89       8.498   4.333  19.131  1.00 15.00
ATOM    802  H   GLU   89       7.553   4.331  19.381  1.00 15.00
ATOM    803  CA  GLU   89       9.208   5.559  19.346  1.00 15.00
ATOM    804  CB  GLU   89       8.336   6.736  19.037  1.00 15.00
ATOM    805  CG  GLU   89       7.926   6.670  17.591  1.00 15.00
ATOM    806  CD  GLU   89       9.075   6.756  16.567  1.00 15.00
ATOM    807  OE1 GLU   89       9.427   7.879  16.186  1.00 15.00
ATOM    808  OE2 GLU   89       9.601   5.726  16.118  1.00 15.00
ATOM    809  C   GLU   89       9.596   5.638  20.774  1.00 15.00
ATOM    810  O   GLU   89       8.686   5.650  21.590  1.00 15.00
ATOM    811  N   ASN   90      10.884   5.618  21.142  1.00 15.00
ATOM    812  H   ASN   90      11.549   5.342  20.480  1.00 15.00
ATOM    813  CA  ASN   90      11.260   5.906  22.535  1.00 15.00
ATOM    814  CB  ASN   90      12.810   5.847  22.742  1.00 15.00
ATOM    815  CG  ASN   90      13.444   6.486  24.026  1.00 15.00
ATOM    816  OD1 ASN   90      13.244   7.669  24.329  1.00 15.00
ATOM    817  ND2 ASN   90      14.295   5.837  24.837  1.00 15.00
ATOM    818 HD21 ASN   90      14.545   4.916  24.634  1.00 15.00
ATOM    819 HD22 ASN   90      14.621   6.347  25.605  1.00 15.00
ATOM    820  C   ASN   90      10.767   7.308  22.945  1.00 15.00
ATOM    821  O   ASN   90      10.145   7.393  24.005  1.00 15.00
ATOM    822  N   LYS   91      10.919   8.355  22.105  1.00 15.00
ATOM    823  H   LYS   91      11.191   8.183  21.186  1.00 15.00
ATOM    824  CA  LYS   91      10.661   9.760  22.449  1.00 15.00
ATOM    825  CB  LYS   91      11.921  10.357  23.031  1.00 15.00
ATOM    826  CG  LYS   91      13.099  10.284  22.118  1.00 15.00
ATOM    827  CD  LYS   91      14.167  10.860  22.965  1.00 15.00
ATOM    828  CE  LYS   91      15.494  10.704  22.301  1.00 15.00
ATOM    829  NZ  LYS   91      15.557  11.490  21.088  1.00 15.00
ATOM    830  HZ1 LYS   91      15.386  12.492  21.303  1.00 15.00
ATOM    831  HZ2 LYS   91      14.829  11.146  20.430  1.00 15.00
ATOM    832  HZ3 LYS   91      16.499  11.380  20.659  1.00 15.00
ATOM    833  C   LYS   91      10.164  10.739  21.363  1.00 15.00
ATOM    834  O   LYS   91      10.299  10.492  20.160  1.00 15.00
ATOM    835  N   VAL   92       9.610  11.860  21.847  1.00 15.00
ATOM    836  H   VAL   92       9.690  11.992  22.819  1.00 15.00
ATOM    837  CA  VAL   92       9.051  12.994  21.116  1.00 15.00
ATOM    838  CB  VAL   92       7.621  13.290  21.642  1.00 15.00
ATOM    839  CG1 VAL   92       7.115  14.608  21.212  1.00 15.00
ATOM    840  CG2 VAL   92       6.638  12.455  20.954  1.00 15.00
ATOM    841  C   VAL   92      10.012  14.154  21.440  1.00 15.00
ATOM    842  O   VAL   92      10.723  14.105  22.451  1.00 15.00
ATOM    843  N   LEU   93      10.158  15.192  20.642  1.00 15.00
ATOM    844  H   LEU   93       9.800  15.153  19.728  1.00 15.00
ATOM    845  CA  LEU   93      10.917  16.349  21.040  1.00 15.00
ATOM    846  CB  LEU   93      12.434  16.080  20.881  1.00 15.00
ATOM    847  CG  LEU   93      13.018  15.446  19.633  1.00 15.00
ATOM    848  CD1 LEU   93      13.460  16.530  18.701  1.00 15.00
ATOM    849  CD2 LEU   93      14.284  14.694  19.929  1.00 15.00
ATOM    850  C   LEU   93      10.439  17.523  20.196  1.00 15.00
ATOM    851  O   LEU   93      10.414  17.516  18.938  1.00 15.00
ATOM    852  N   VAL   94       9.969  18.539  20.933  1.00 15.00
ATOM    853  H   VAL   94      10.044  18.495  21.907  1.00 15.00
ATOM    854  CA  VAL   94       9.391  19.697  20.265  1.00 15.00
```

```
ATOM    855  CB  VAL    94       8.541  20.498  21.277  1.00 15.00
ATOM    856  CG1 VAL    94       7.637  21.423  20.518  1.00 15.00
ATOM    857  CG2 VAL    94       7.634  19.597  22.084  1.00 15.00
ATOM    858  C   VAL    94      10.550  20.505  19.719  1.00 15.00
ATOM    859  O   VAL    94      11.371  20.860  20.576  1.00 15.00
ATOM    860  N   LEU    95      10.711  20.748  18.391  1.00 15.00
ATOM    861  H   LEU    95      10.035  20.449  17.749  1.00 15.00
ATOM    862  CA  LEU    95      11.858  21.544  17.918  1.00 15.00
ATOM    863  CB  LEU    95      12.326  21.321  16.483  1.00 15.00
ATOM    864  CG  LEU    95      12.696  20.008  15.820  1.00 15.00
ATOM    865  CD1 LEU    95      13.410  20.342  14.526  1.00 15.00
ATOM    866  CD2 LEU    95      13.661  19.191  16.622  1.00 15.00
ATOM    867  C   LEU    95      11.580  23.038  17.907  1.00 15.00
ATOM    868  O   LEU    95      12.472  23.803  18.273  1.00 15.00
ATOM    869  N   ASP    96      10.395  23.518  17.490  1.00 15.00
ATOM    870  H   ASP    96       9.680  22.894  17.237  1.00 15.00
ATOM    871  CA  ASP    96      10.145  24.964  17.370  1.00 15.00
ATOM    872  CB  ASP    96      10.944  25.417  16.103  1.00 15.00
ATOM    873  CG  ASP    96      10.867  26.839  15.576  1.00 15.00
ATOM    874  OD1 ASP    96      11.676  27.185  14.733  1.00 15.00
ATOM    875  OD2 ASP    96       9.994  27.600  15.956  1.00 15.00
ATOM    876  C   ASP    96       8.627  25.270  17.317  1.00 15.00
ATOM    877  O   ASP    96       7.872  24.556  16.673  1.00 15.00
ATOM    878  N   THR    97       8.120  26.328  17.939  1.00 15.00
ATOM    879  H   THR    97       8.720  27.003  18.324  1.00 15.00
ATOM    880  CA  THR    97       6.710  26.596  18.088  1.00 15.00
ATOM    881  CB  THR    97       6.225  25.688  19.258  1.00 15.00
ATOM    882  OG1 THR    97       4.829  25.919  19.341  1.00 15.00
ATOM    883  HG1 THR    97       4.385  25.457  18.617  1.00 15.00
ATOM    884  CG2 THR    97       6.865  25.939  20.630  1.00 15.00
ATOM    885  C   THR    97       6.538  28.091  18.364  1.00 15.00
ATOM    886  O   THR    97       7.442  28.720  18.944  1.00 15.00
ATOM    887  N   ASP    98       5.428  28.691  17.925  1.00 15.00
ATOM    888  H   ASP    98       4.790  28.167  17.396  1.00 15.00
ATOM    889  CA  ASP    98       5.152  30.081  18.245  1.00 15.00
ATOM    890  CB  ASP    98       5.268  30.930  16.993  1.00 15.00
ATOM    891  CG  ASP    98       4.101  30.905  16.022  1.00 15.00
ATOM    892  OD1 ASP    98       3.363  29.926  16.032  1.00 15.00
ATOM    893  OD2 ASP    98       3.926  31.860  15.255  1.00 15.00
ATOM    894  C   ASP    98       3.782  30.298  18.880  1.00 15.00
ATOM    895  O   ASP    98       3.234  31.404  18.928  1.00 15.00
ATOM    896  N   TYR    99       3.185  29.161  19.266  1.00 15.00
ATOM    897  H   TYR    99       3.638  28.315  19.070  1.00 15.00
ATOM    898  CA  TYR    99       1.937  29.039  20.020  1.00 15.00
ATOM    899  CB  TYR    99       2.040  29.858  21.372  1.00 15.00
ATOM    900  CG  TYR    99       3.309  29.517  22.156  1.00 15.00
ATOM    901  CD1 TYR    99       4.320  30.441  22.350  1.00 15.00
ATOM    902  CE1 TYR    99       5.499  30.048  22.969  1.00 15.00
ATOM    903  CD2 TYR    99       3.492  28.221  22.595  1.00 15.00
ATOM    904  CE2 TYR    99       4.667  27.819  23.211  1.00 15.00
ATOM    905  CZ  TYR    99       5.670  28.738  23.399  1.00 15.00
ATOM    906  OH  TYR    99       6.821  28.314  24.048  1.00 15.00
ATOM    907  HH  TYR    99       7.459  29.035  24.079  1.00 15.00
ATOM    908  C   TYR    99       0.709  29.464  19.233  1.00 15.00
ATOM    909  O   TYR    99      -0.147  28.631  18.974  1.00 15.00
ATOM    910  N   LYS   100       0.550  30.713  18.815  1.00 15.00
ATOM    911  H   LYS   100       1.242  31.372  19.039  1.00 15.00
ATOM    912  CA  LYS   100      -0.612  31.090  18.024  1.00 15.00
ATOM    913  CB  LYS   100      -0.520  32.596  17.688  1.00 15.00
ATOM    914  CG  LYS   100      -0.499  33.682  18.808  1.00 15.00
ATOM    915  CD  LYS   100       0.904  34.052  19.326  1.00 15.00
ATOM    916  CE  LYS   100       1.798  34.929  18.417  1.00 15.00
ATOM    917  NZ  LYS   100       3.204  34.864  18.846  1.00 15.00
ATOM    918  HZ1 LYS   100       3.283  35.204  19.826  1.00 15.00
ATOM    919  HZ2 LYS   100       3.788  35.459  18.225  1.00 15.00
ATOM    920  HZ3 LYS   100       3.535  33.880  18.792  1.00 15.00
```

```
ATOM    921  C   LYS  100      -0.760  30.276  16.709  1.00 15.00
ATOM    922  O   LYS  100      -1.799  29.693  16.402  1.00 15.00
ATOM    923  N   LYS  101       0.310  30.242  15.911  1.00 15.00
ATOM    924  H   LYS  101       1.138  30.644  16.232  1.00 15.00
ATOM    925  CA  LYS  101       0.263  29.578  14.619  1.00 15.00
ATOM    926  CB  LYS  101       0.686  30.640  13.639  1.00 15.00
ATOM    927  CG  LYS  101      -0.594  31.492  13.501  1.00 15.00
ATOM    928  CD  LYS  101      -1.722  30.628  12.838  1.00 15.00
ATOM    929  CE  LYS  101      -3.116  31.283  12.719  1.00 15.00
ATOM    930  NZ  LYS  101      -4.051  30.461  11.965  1.00 15.00
ATOM    931  HZ1 LYS  101      -4.152  29.534  12.426  1.00 15.00
ATOM    932  HZ2 LYS  101      -3.689  30.330  10.999  1.00 15.00
ATOM    933  HZ3 LYS  101      -4.975  30.936  11.929  1.00 15.00
ATOM    934  C   LYS  101       0.952  28.234  14.331  1.00 15.00
ATOM    935  O   LYS  101       0.209  27.251  14.182  1.00 15.00
ATOM    936  N   TYR  102       2.296  28.078  14.227  1.00 15.00
ATOM    937  H   TYR  102       2.854  28.872  14.270  1.00 15.00
ATOM    938  CA  TYR  102       2.940  26.779  14.002  1.00 15.00
ATOM    939  CB  TYR  102       4.027  27.036  12.976  1.00 15.00
ATOM    940  CG  TYR  102       5.255  27.845  13.362  1.00 15.00
ATOM    941  CD1 TYR  102       6.409  27.191  13.729  1.00 15.00
ATOM    942  CE1 TYR  102       7.588  27.877  13.903  1.00 15.00
ATOM    943  CD2 TYR  102       5.292  29.207  13.197  1.00 15.00
ATOM    944  CE2 TYR  102       6.475  29.902  13.368  1.00 15.00
ATOM    945  CZ  TYR  102       7.619  29.227  13.708  1.00 15.00
ATOM    946  OH  TYR  102       8.828  29.879  13.750  1.00 15.00
ATOM    947  HH  TYR  102       8.756  30.714  13.280  1.00 15.00
ATOM    948  C   TYR  102       3.499  25.919  15.159  1.00 15.00
ATOM    949  O   TYR  102       3.619  26.425  16.269  1.00 15.00
ATOM    950  N   LEU  103       3.871  24.639  15.021  1.00 15.00
ATOM    951  H   LEU  103       3.655  24.186  14.175  1.00 15.00
ATOM    952  CA  LEU  103       4.552  23.865  16.068  1.00 15.00
ATOM    953  CB  LEU  103       3.643  23.236  17.138  1.00 15.00
ATOM    954  CG  LEU  103       4.449  22.354  18.122  1.00 15.00
ATOM    955  CD1 LEU  103       4.327  22.835  19.556  1.00 15.00
ATOM    956  CD2 LEU  103       3.930  20.954  18.040  1.00 15.00
ATOM    957  C   LEU  103       5.280  22.693  15.432  1.00 15.00
ATOM    958  O   LEU  103       4.741  21.622  15.198  1.00 15.00
ATOM    959  N   LEU  104       6.531  22.872  15.105  1.00 15.00
ATOM    960  H   LEU  104       6.947  23.720  15.329  1.00 15.00
ATOM    961  CA  LEU  104       7.333  21.856  14.482  1.00 15.00
ATOM    962  CB  LEU  104       8.495  22.585  13.818  1.00 15.00
ATOM    963  CG  LEU  104       8.403  23.410  12.520  1.00 15.00
ATOM    964  CD1 LEU  104       6.993  23.753  12.071  1.00 15.00
ATOM    965  CD2 LEU  104       9.244  24.632  12.787  1.00 15.00
ATOM    966  C   LEU  104       7.809  20.860  15.536  1.00 15.00
ATOM    967  O   LEU  104       8.509  21.286  16.471  1.00 15.00
ATOM    968  N   PHE  105       7.480  19.570  15.476  1.00 15.00
ATOM    969  H   PHE  105       6.914  19.249  14.748  1.00 15.00
ATOM    970  CA  PHE  105       8.068  18.582  16.389  1.00 15.00
ATOM    971  CB  PHE  105       7.056  18.079  17.438  1.00 15.00
ATOM    972  CG  PHE  105       5.954  17.162  16.984  1.00 15.00
ATOM    973  CD1 PHE  105       6.110  15.812  17.105  1.00 15.00
ATOM    974  CD2 PHE  105       4.812  17.678  16.445  1.00 15.00
ATOM    975  CE1 PHE  105       5.115  14.968  16.679  1.00 15.00
ATOM    976  CE2 PHE  105       3.814  16.842  16.018  1.00 15.00
ATOM    977  CZ  PHE  105       3.969  15.486  16.134  1.00 15.00
ATOM    978  C   PHE  105       8.622  17.365  15.643  1.00 15.00
ATOM    979  O   PHE  105       8.352  17.197  14.440  1.00 15.00
ATOM    980  N   CYS  106       9.439  16.526  16.266  1.00 15.00
ATOM    981  H   CYS  106       9.736  16.684  17.189  1.00 15.00
ATOM    982  CA  CYS  106       9.840  15.321  15.590  1.00 15.00
ATOM    983  C   CYS  106       9.748  14.214  16.592  1.00 15.00
ATOM    984  O   CYS  106       9.673  14.444  17.782  1.00 15.00
ATOM    985  CB  CYS  106      11.254  15.403  15.065  1.00 15.00
ATOM    986  SG  CYS  106      11.528  16.457  13.597  1.00 15.00
```

```
ATOM    987  N   MET  107       9.626  12.999  16.160  1.00 15.00
ATOM    988  H   MET  107       9.637  12.817  15.202  1.00 15.00
ATOM    989  CA  MET  107       9.581  11.906  17.076  1.00 15.00
ATOM    990  CB  MET  107       8.426  11.005  16.675  1.00 15.00
ATOM    991  CG  MET  107       7.169  11.734  16.197  1.00 15.00
ATOM    992  SD  MET  107       5.719  11.740  17.269  1.00 15.00
ATOM    993  CE  MET  107       5.630   9.961  17.305  1.00 15.00
ATOM    994  C   MET  107      10.959  11.240  16.940  1.00 15.00
ATOM    995  O   MET  107      11.841  11.830  16.301  1.00 15.00
ATOM    996  N   GLU  108      11.161  10.011  17.460  1.00 15.00
ATOM    997  H   GLU  108      10.365   9.543  17.776  1.00 15.00
ATOM    998  CA  GLU  108      12.460   9.355  17.677  1.00 15.00
ATOM    999  CB  GLU  108      12.482   7.912  17.152  1.00 15.00
ATOM   1000  CG  GLU  108      12.448   6.845  18.232  1.00 15.00
ATOM   1001  CD  GLU  108      13.772   6.416  18.829  1.00 15.00
ATOM   1002  OE1 GLU  108      13.886   5.285  19.276  1.00 15.00
ATOM   1003  OE2 GLU  108      14.701   7.198  18.866  1.00 15.00
ATOM   1004  C   GLU  108      13.770   9.949  17.210  1.00 15.00
ATOM   1005  O   GLU  108      14.285   9.620  16.145  1.00 15.00
ATOM   1006  N   ASN  109      14.259  10.787  18.133  1.00 15.00
ATOM   1007  H   ASN  109      13.635  11.053  18.834  1.00 15.00
ATOM   1008  CA  ASN  109      15.558  11.411  18.111  1.00 15.00
ATOM   1009  CB  ASN  109      16.542  10.296  18.416  1.00 15.00
ATOM   1010  CG  ASN  109      17.945  10.677  17.986  1.00 15.00
ATOM   1011  OD1 ASN  109      18.566  11.671  18.414  1.00 15.00
ATOM   1012  ND2 ASN  109      18.354   9.909  16.986  1.00 15.00
ATOM   1013 HD21 ASN  109      17.722   9.241  16.668  1.00 15.00
ATOM   1014 HD22 ASN  109      19.248  10.019  16.608  1.00 15.00
ATOM   1015  C   ASN  109      16.011  12.244  16.906  1.00 15.00
ATOM   1016  O   ASN  109      16.412  11.705  15.867  1.00 15.00
ATOM   1017  N   SER  110      16.093  13.576  17.060  1.00 15.00
ATOM   1018  H   SER  110      15.896  13.958  17.935  1.00 15.00
ATOM   1019  CA  SER  110      16.530  14.456  15.987  1.00 15.00
ATOM   1020  CB  SER  110      15.826  15.764  16.269  1.00 15.00
ATOM   1021  OG  SER  110      16.179  16.842  15.449  1.00 15.00
ATOM   1022  HG  SER  110      15.387  17.303  15.136  1.00 15.00
ATOM   1023  C   SER  110      18.058  14.573  15.825  1.00 15.00
ATOM   1024  O   SER  110      18.593  15.692  15.655  1.00 15.00
ATOM   1025  N   ALA  111      18.697  13.366  15.896  1.00 15.00
ATOM   1026  H   ALA  111      18.132  12.578  16.018  1.00 15.00
ATOM   1027  CA  ALA  111      20.095  12.983  15.645  1.00 15.00
ATOM   1028  CB  ALA  111      21.074  14.154  15.745  1.00 15.00
ATOM   1029  C   ALA  111      20.823  11.870  16.429  1.00 15.00
ATOM   1030  O   ALA  111      20.864  10.731  15.955  1.00 15.00
ATOM   1031  N   GLU  112      21.286  12.116  17.670  1.00 15.00
ATOM   1032  H   GLU  112      20.746  12.743  18.190  1.00 15.00
ATOM   1033  CA  GLU  112      22.344  11.361  18.390  1.00 15.00
ATOM   1034  CB  GLU  112      22.522  12.222  19.688  1.00 15.00
ATOM   1035  CG  GLU  112      22.806  11.748  21.138  1.00 15.00
ATOM   1036  CD  GLU  112      24.175  11.175  21.530  1.00 15.00
ATOM   1037  OE1 GLU  112      24.269   9.956  21.744  1.00 15.00
ATOM   1038  OE2 GLU  112      25.124  11.959  21.684  1.00 15.00
ATOM   1039  C   GLU  112      22.557   9.841  18.717  1.00 15.00
ATOM   1040  O   GLU  112      21.774   9.147  19.381  1.00 15.00
ATOM   1041  N   PRO  113      23.750   9.329  18.327  1.00 15.00
ATOM   1042  CD  PRO  113      24.379   8.089  18.802  1.00 15.00
ATOM   1043  CA  PRO  113      24.671  10.024  17.429  1.00 15.00
ATOM   1044  CB  PRO  113      26.048   9.314  17.683  1.00 15.00
ATOM   1045  CG  PRO  113      25.893   8.451  18.917  1.00 15.00
ATOM   1046  C   PRO  113      24.070   9.899  16.006  1.00 15.00
ATOM   1047  O   PRO  113      23.282   8.978  15.750  1.00 15.00
ATOM   1048  N   GLU  114      24.398  10.906  15.169  1.00 15.00
ATOM   1049  H   GLU  114      25.102  11.510  15.465  1.00 15.00
ATOM   1050  CA  GLU  114      23.869  11.135  13.821  1.00 15.00
ATOM   1051  CB  GLU  114      24.972  11.900  12.987  1.00 15.00
ATOM   1052  CG  GLU  114      25.344  13.440  13.179  1.00 15.00
```

```
ATOM   1053  CD  GLU   114      26.522  13.980  14.065  1.00 15.00
ATOM   1054  OE1 GLU   114      27.081  13.274  14.924  1.00 15.00
ATOM   1055  OE2 GLU   114      26.885  15.153  13.887  1.00 15.00
ATOM   1056  C   GLU   114      23.300   9.980  12.951  1.00 15.00
ATOM   1057  O   GLU   114      24.047   9.169  12.379  1.00 15.00
ATOM   1058  N   GLN   115      21.936   9.947  12.947  1.00 15.00
ATOM   1059  H   GLN   115      21.503  10.478  13.658  1.00 15.00
ATOM   1060  CA  GLN   115      20.957   9.171  12.151  1.00 15.00
ATOM   1061  CB  GLN   115      21.294   7.635  12.152  1.00 15.00
ATOM   1062  CG  GLN   115      20.760   6.770  10.952  1.00 15.00
ATOM   1063  CD  GLN   115      19.352   6.099  11.001  1.00 15.00
ATOM   1064  OE1 GLN   115      18.353   6.476  10.373  1.00 15.00
ATOM   1065  NE2 GLN   115      19.167   5.070  11.808  1.00 15.00
ATOM   1066 HE21 GLN   115      19.903   4.779  12.382  1.00 15.00
ATOM   1067 HE22 GLN   115      18.289   4.634  11.785  1.00 15.00
ATOM   1068  C   GLN   115      19.574   9.455  12.861  1.00 15.00
ATOM   1069  O   GLN   115      19.447  10.425  13.649  1.00 15.00
ATOM   1070  N   SER   116      18.535   8.656  12.468  1.00 15.00
ATOM   1071  H   SER   116      18.624   8.241  11.593  1.00 15.00
ATOM   1072  CA  SER   116      17.171   8.510  13.024  1.00 15.00
ATOM   1073  CB  SER   116      17.300   8.229  14.557  1.00 15.00
ATOM   1074  OG  SER   116      18.244   7.231  15.007  1.00 15.00
ATOM   1075  HG  SER   116      19.054   7.708  15.234  1.00 15.00
ATOM   1076  C   SER   116      16.187   9.688  12.760  1.00 15.00
ATOM   1077  O   SER   116      16.734  10.787  12.646  1.00 15.00
ATOM   1078  N   LEU   117      14.830   9.515  12.540  1.00 15.00
ATOM   1079  H   LEU   117      14.503   8.604  12.425  1.00 15.00
ATOM   1080  CA  LEU   117      13.835  10.583  12.416  1.00 15.00
ATOM   1081  CB  LEU   117      14.289  11.725  11.648  1.00 15.00
ATOM   1082  CG  LEU   117      14.391  13.013  12.390  1.00 15.00
ATOM   1083  CD1 LEU   117      14.167  12.914  13.903  1.00 15.00
ATOM   1084  CD2 LEU   117      15.763  13.490  12.027  1.00 15.00
ATOM   1085  C   LEU   117      12.494  10.385  11.796  1.00 15.00
ATOM   1086  O   LEU   117      12.414   9.578  10.911  1.00 15.00
ATOM   1087  N   VAL   118      11.481  11.193  12.123  1.00 15.00
ATOM   1088  H   VAL   118      11.646  11.806  12.867  1.00 15.00
ATOM   1089  CA  VAL   118      10.150  11.193  11.491  1.00 15.00
ATOM   1090  CB  VAL   118       9.484   9.871  11.906  1.00 15.00
ATOM   1091  CG1 VAL   118       9.215   9.839  13.384  1.00 15.00
ATOM   1092  CG2 VAL   118       8.241   9.685  11.109  1.00 15.00
ATOM   1093  C   VAL   118       9.252  12.419  11.794  1.00 15.00
ATOM   1094  O   VAL   118       8.126  12.385  12.265  1.00 15.00
ATOM   1095  N   CYS   119       9.789  13.591  11.566  1.00 15.00
ATOM   1096  H   CYS   119      10.635  13.608  11.081  1.00 15.00
ATOM   1097  CA  CYS   119       9.154  14.864  11.884  1.00 15.00
ATOM   1098  C   CYS   119       7.817  15.152  11.227  1.00 15.00
ATOM   1099  O   CYS   119       7.464  14.532  10.230  1.00 15.00
ATOM   1100  CB  CYS   119      10.107  15.935  11.503  1.00 15.00
ATOM   1101  SG  CYS   119      11.759  15.374  11.932  1.00 15.00
ATOM   1102  N   GLN   120       7.091  16.137  11.725  1.00 15.00
ATOM   1103  H   GLN   120       7.473  16.670  12.461  1.00 15.00
ATOM   1104  CA  GLN   120       5.803  16.568  11.207  1.00 15.00
ATOM   1105  CB  GLN   120       4.613  16.029  11.948  1.00 15.00
ATOM   1106  CG  GLN   120       4.413  14.525  11.870  1.00 15.00
ATOM   1107  CD  GLN   120       3.000  14.141  12.235  1.00 15.00
ATOM   1108  OE1 GLN   120       2.133  15.012  12.351  1.00 15.00
ATOM   1109  NE2 GLN   120       2.719  12.861  12.400  1.00 15.00
ATOM   1110 HE21 GLN   120       3.452  12.214  12.333  1.00 15.00
ATOM   1111 HE22 GLN   120       1.777  12.612  12.507  1.00 15.00
ATOM   1112  C   GLN   120       5.790  18.045  11.440  1.00 15.00
ATOM   1113  O   GLN   120       6.810  18.609  11.811  1.00 15.00
ATOM   1114  N   CYS   121       4.682  18.718  11.271  1.00 15.00
ATOM   1115  H   CYS   121       3.846  18.267  11.033  1.00 15.00
ATOM   1116  CA  CYS   121       4.666  20.153  11.380  1.00 15.00
ATOM   1117  CB  CYS   121       5.284  20.754  10.198  1.00 15.00
ATOM   1118  SG  CYS   121       4.614  22.391   9.978  1.00 15.00
```

```
ATOM   1119  C   CYS  121       3.224  20.547  11.417  1.00 15.00
ATOM   1120  O   CYS  121       2.566  20.578  10.388  1.00 15.00
ATOM   1121  N   LEU  122       2.766  20.809  12.636  1.00 15.00
ATOM   1122  H   LEU  122       3.419  20.907  13.362  1.00 15.00
ATOM   1123  CA  LEU  122       1.384  21.077  12.949  1.00 15.00
ATOM   1124  CB  LEU  122       1.168  20.850  14.431  1.00 15.00
ATOM   1125  CG  LEU  122       1.587  19.524  15.040  1.00 15.00
ATOM   1126  CD1 LEU  122       1.243  19.388  16.495  1.00 15.00
ATOM   1127  CD2 LEU  122       0.838  18.469  14.311  1.00 15.00
ATOM   1128  C   LEU  122       1.020  22.489  12.606  1.00 15.00
ATOM   1129  O   LEU  122       1.880  23.320  12.334  1.00 15.00
ATOM   1130  N   VAL  123      -0.271  22.767  12.673  1.00 15.00
ATOM   1131  H   VAL  123      -0.913  22.032  12.735  1.00 15.00
ATOM   1132  CA  VAL  123      -0.799  24.111  12.531  1.00 15.00
ATOM   1133  CB  VAL  123      -0.965  24.295  10.990  1.00 15.00
ATOM   1134  CG1 VAL  123      -2.330  24.880  10.567  1.00 15.00
ATOM   1135  CG2 VAL  123       0.191  25.193  10.576  1.00 15.00
ATOM   1136  C   VAL  123      -2.089  24.260  13.384  1.00 15.00
ATOM   1137  O   VAL  123      -2.884  23.317  13.588  1.00 15.00
ATOM   1138  N   ARG  124      -2.271  25.408  14.046  1.00 15.00
ATOM   1139  H   ARG  124      -1.580  26.108  14.014  1.00 15.00
ATOM   1140  CA  ARG  124      -3.501  25.693  14.784  1.00 15.00
ATOM   1141  CB  ARG  124      -3.353  26.904  15.708  1.00 15.00
ATOM   1142  CG  ARG  124      -4.705  27.477  16.203  1.00 15.00
ATOM   1143  CD  ARG  124      -4.957  28.959  15.791  1.00 15.00
ATOM   1144  NE  ARG  124      -6.329  29.425  16.021  1.00 15.00
ATOM   1145  HE  ARG  124      -6.981  28.824  16.437  1.00 15.00
ATOM   1146  CZ  ARG  124      -6.707  30.663  15.664  1.00 15.00
ATOM   1147  NH1 ARG  124      -5.858  31.541  15.085  1.00 15.00
ATOM   1148 HH11 ARG  124      -4.911  31.281  14.900  1.00 15.00
ATOM   1149 HH12 ARG  124      -6.181  32.455  14.838  1.00 15.00
ATOM   1150  NH2 ARG  124      -7.976  31.016  15.877  1.00 15.00
ATOM   1151 HH21 ARG  124      -8.612  30.365  16.291  1.00 15.00
ATOM   1152 HH22 ARG  124      -8.288  31.933  15.628  1.00 15.00
ATOM   1153  C   ARG  124      -4.435  26.050  13.650  1.00 15.00
ATOM   1154  O   ARG  124      -4.278  27.057  12.953  1.00 15.00
ATOM   1155  N   THR  125      -5.383  25.148  13.475  1.00 15.00
ATOM   1156  H   THR  125      -5.397  24.393  14.091  1.00 15.00
ATOM   1157  CA  THR  125      -6.407  25.254  12.444  1.00 15.00
ATOM   1158  CB  THR  125      -7.675  26.018  13.119  1.00 15.00
ATOM   1159  OG1 THR  125      -7.343  26.691  14.367  1.00 15.00
ATOM   1160  HG1 THR  125      -6.668  27.350  14.141  1.00 15.00
ATOM   1161  CG2 THR  125      -8.812  24.948  13.359  1.00 15.00
ATOM   1162  C   THR  125      -6.094  25.820  11.038  1.00 15.00
ATOM   1163  O   THR  125      -5.698  26.970  10.891  1.00 15.00
ATOM   1164  N   PRO  126      -6.497  25.004  10.019  1.00 15.00
ATOM   1165  CD  PRO  126      -7.741  24.216  10.053  1.00 15.00
ATOM   1166  CA  PRO  126      -5.996  24.918   8.637  1.00 15.00
ATOM   1167  CB  PRO  126      -7.019  23.962   8.016  1.00 15.00
ATOM   1168  CG  PRO  126      -8.308  24.322   8.679  1.00 15.00
ATOM   1169  C   PRO  126      -5.604  26.045   7.649  1.00 15.00
ATOM   1170  O   PRO  126      -5.019  27.055   8.020  1.00 15.00
ATOM   1171  N   GLU  127      -5.961  25.948   6.352  1.00 15.00
ATOM   1172  H   GLU  127      -6.715  25.380   6.122  1.00 15.00
ATOM   1173  CA  GLU  127      -5.324  26.650   5.251  1.00 15.00
ATOM   1174  CB  GLU  127      -5.264  28.138   5.359  1.00 15.00
ATOM   1175  CG  GLU  127      -6.442  28.936   4.854  1.00 15.00
ATOM   1176  CD  GLU  127      -6.313  30.385   5.326  1.00 15.00
ATOM   1177  OE1 GLU  127      -6.840  30.728   6.416  1.00 15.00
ATOM   1178  OE2 GLU  127      -5.644  31.140   4.594  1.00 15.00
ATOM   1179  C   GLU  127      -3.905  26.156   5.360  1.00 15.00
ATOM   1180  O   GLU  127      -3.648  25.120   5.977  1.00 15.00
ATOM   1181  N   VAL  128      -2.928  26.812   4.795  1.00 15.00
ATOM   1182  H   VAL  128      -3.030  27.745   4.508  1.00 15.00
ATOM   1183  CA  VAL  128      -1.596  26.276   4.731  1.00 15.00
ATOM   1184  CB  VAL  128      -1.519  25.738   3.260  1.00 15.00
```

```
ATOM   1185  CG1 VAL   128      -0.111  25.816   2.774  1.00 15.00
ATOM   1186  CG2 VAL   128      -1.931  24.267   3.150  1.00 15.00
ATOM   1187  C   VAL   128      -0.797  27.520   5.099  1.00 15.00
ATOM   1188  O   VAL   128      -1.172  28.677   4.839  1.00 15.00
ATOM   1189  N   ASP   129       0.310  27.302   5.785  1.00 15.00
ATOM   1190  H   ASP   129       0.602  26.387   5.976  1.00 15.00
ATOM   1191  CA  ASP   129       1.137  28.438   6.188  1.00 15.00
ATOM   1192  CB  ASP   129       1.480  28.347   7.731  1.00 15.00
ATOM   1193  CG  ASP   129       2.044  29.629   8.369  1.00 15.00
ATOM   1194  OD1 ASP   129       2.803  30.340   7.707  1.00 15.00
ATOM   1195  OD2 ASP   129       1.716  29.937   9.520  1.00 15.00
ATOM   1196  C   ASP   129       2.396  28.301   5.313  1.00 15.00
ATOM   1197  O   ASP   129       3.116  27.313   5.542  1.00 15.00
ATOM   1198  N   ASP   130       2.756  29.145   4.317  1.00 15.00
ATOM   1199  H   ASP   130       2.158  29.859   4.013  1.00 15.00
ATOM   1200  CA  ASP   130       3.980  28.822   3.633  1.00 15.00
ATOM   1201  CB  ASP   130       4.129  29.471   2.256  1.00 15.00
ATOM   1202  CG  ASP   130       5.058  28.568   1.402  1.00 15.00
ATOM   1203  OD1 ASP   130       6.269  28.813   1.345  1.00 15.00
ATOM   1204  OD2 ASP   130       4.598  27.574   0.818  1.00 15.00
ATOM   1205  C   ASP   130       5.105  29.277   4.518  1.00 15.00
ATOM   1206  O   ASP   130       6.146  28.620   4.510  1.00 15.00
ATOM   1207  N   GLU   131       4.904  30.290   5.373  1.00 15.00
ATOM   1208  H   GLU   131       4.067  30.789   5.346  1.00 15.00
ATOM   1209  CA  GLU   131       5.959  30.694   6.297  1.00 15.00
ATOM   1210  CB  GLU   131       5.555  31.987   7.072  1.00 15.00
ATOM   1211  CG  GLU   131       5.622  33.333   6.262  1.00 15.00
ATOM   1212  CD  GLU   131       5.242  34.718   6.901  1.00 15.00
ATOM   1213  OE1 GLU   131       4.526  35.468   6.210  1.00 15.00
ATOM   1214  OE2 GLU   131       5.672  35.102   8.018  1.00 15.00
ATOM   1215  C   GLU   131       6.248  29.555   7.287  1.00 15.00
ATOM   1216  O   GLU   131       7.257  29.603   7.961  1.00 15.00
ATOM   1217  N   ALA   132       5.439  28.498   7.462  1.00 15.00
ATOM   1218  H   ALA   132       4.617  28.448   6.943  1.00 15.00
ATOM   1219  CA  ALA   132       5.788  27.369   8.317  1.00 15.00
ATOM   1220  CB  ALA   132       4.617  26.887   9.167  1.00 15.00
ATOM   1221  C   ALA   132       6.175  26.221   7.410  1.00 15.00
ATOM   1222  O   ALA   132       7.265  25.690   7.513  1.00 15.00
ATOM   1223  N   LEU   133       5.328  25.839   6.457  1.00 15.00
ATOM   1224  H   LEU   133       4.477  26.307   6.390  1.00 15.00
ATOM   1225  CA  LEU   133       5.564  24.751   5.520  1.00 15.00
ATOM   1226  CB  LEU   133       4.507  24.761   4.453  1.00 15.00
ATOM   1227  CG  LEU   133       3.619  23.552   4.415  1.00 15.00
ATOM   1228  CD1 LEU   133       2.881  23.579   3.103  1.00 15.00
ATOM   1229  CD2 LEU   133       4.397  22.254   4.467  1.00 15.00
ATOM   1230  C   LEU   133       6.921  24.706   4.814  1.00 15.00
ATOM   1231  O   LEU   133       7.479  23.652   4.427  1.00 15.00
ATOM   1232  N   GLU   134       7.503  25.874   4.610  1.00 15.00
ATOM   1233  H   GLU   134       7.059  26.733   4.790  1.00 15.00
ATOM   1234  CA  GLU   134       8.809  25.853   4.023  1.00 15.00
ATOM   1235  CB  GLU   134       8.561  26.302   2.506  1.00 15.00
ATOM   1236  CG  GLU   134       7.391  25.727   1.544  1.00 15.00
ATOM   1237  CD  GLU   134       7.244  24.219   1.105  1.00 15.00
ATOM   1238  OE1 GLU   134       6.177  23.588   1.301  1.00 15.00
ATOM   1239  OE2 GLU   134       8.185  23.670   0.510  1.00 15.00
ATOM   1240  C   GLU   134       9.618  26.767   4.987  1.00 15.00
ATOM   1241  O   GLU   134      10.079  27.839   4.628  1.00 15.00
ATOM   1242  N   LYS   135       9.635  26.348   6.286  1.00 15.00
ATOM   1243  H   LYS   135       8.951  25.671   6.454  1.00 15.00
ATOM   1244  CA  LYS   135      10.403  26.798   7.511  1.00 15.00
ATOM   1245  CB  LYS   135       9.634  27.544   8.618  1.00 15.00
ATOM   1246  CG  LYS   135      10.409  27.839   9.930  1.00 15.00
ATOM   1247  CD  LYS   135       9.641  28.787  10.831  1.00 15.00
ATOM   1248  CE  LYS   135       9.524  30.151  10.125  1.00 15.00
ATOM   1249  NZ  LYS   135       8.376  30.944  10.549  1.00 15.00
ATOM   1250  HZ1 LYS   135       8.426  31.108  11.575  1.00 15.00
```

```
ATOM   1251  HZ2 LYS   135       7.499  30.431  10.324  1.00 15.00
ATOM   1252  HZ3 LYS   135       8.381  31.856  10.048  1.00 15.00
ATOM   1253  C   LYS   135      10.786  25.507   8.203  1.00 15.00
ATOM   1254  O   LYS   135      11.795  25.351   8.873  1.00 15.00
ATOM   1255  N   PHE   136       9.783  24.645   8.080  1.00 15.00
ATOM   1256  H   PHE   136       8.903  25.057   8.016  1.00 15.00
ATOM   1257  CA  PHE   136       9.773  23.212   8.213  1.00 15.00
ATOM   1258  CB  PHE   136       8.565  22.639   7.513  1.00 15.00
ATOM   1259  CG  PHE   136       8.304  21.158   7.645  1.00 15.00
ATOM   1260  CD1 PHE   136       8.537  20.502   8.827  1.00 15.00
ATOM   1261  CD2 PHE   136       7.738  20.519   6.589  1.00 15.00
ATOM   1262  CE1 PHE   136       8.181  19.196   8.952  1.00 15.00
ATOM   1263  CE2 PHE   136       7.386  19.210   6.722  1.00 15.00
ATOM   1264  CZ  PHE   136       7.602  18.552   7.891  1.00 15.00
ATOM   1265  C   PHE   136      10.995  22.842   7.425  1.00 15.00
ATOM   1266  O   PHE   136      11.927  22.406   8.079  1.00 15.00
ATOM   1267  N   ASP   137      11.052  23.077   6.096  1.00 15.00
ATOM   1268  H   ASP   137      10.235  23.366   5.641  1.00 15.00
ATOM   1269  CA  ASP   137      12.223  22.788   5.294  1.00 15.00
ATOM   1270  CB  ASP   137      12.145  23.518   3.995  1.00 15.00
ATOM   1271  CG  ASP   137      11.092  22.926   3.047  1.00 15.00
ATOM   1272  OD1 ASP   137      10.366  21.983   3.447  1.00 15.00
ATOM   1273  OD2 ASP   137      11.021  23.409   1.898  1.00 15.00
ATOM   1274  C   ASP   137      13.473  23.183   6.001  1.00 15.00
ATOM   1275  O   ASP   137      14.298  22.309   6.203  1.00 15.00
ATOM   1276  N   LYS   138      13.527  24.407   6.533  1.00 15.00
ATOM   1277  H   LYS   138      12.771  25.013   6.440  1.00 15.00
ATOM   1278  CA  LYS   138      14.691  24.850   7.304  1.00 15.00
ATOM   1279  CB  LYS   138      14.510  26.305   7.847  1.00 15.00
ATOM   1280  CG  LYS   138      15.561  27.357   7.337  1.00 15.00
ATOM   1281  CD  LYS   138      15.696  27.582   5.782  1.00 15.00
ATOM   1282  CE  LYS   138      14.608  28.372   5.003  1.00 15.00
ATOM   1283  NZ  LYS   138      14.487  29.772   5.426  1.00 15.00
ATOM   1284  HZ1 LYS   138      13.778  30.250   4.833  1.00 15.00
ATOM   1285  HZ2 LYS   138      15.406  30.247   5.322  1.00 15.00
ATOM   1286  HZ3 LYS   138      14.188  29.806   6.421  1.00 15.00
ATOM   1287  C   LYS   138      15.036  23.947   8.476  1.00 15.00
ATOM   1288  O   LYS   138      16.159  23.473   8.452  1.00 15.00
ATOM   1289  N   ALA   139      14.227  23.639   9.491  1.00 15.00
ATOM   1290  H   ALA   139      13.317  24.006   9.484  1.00 15.00
ATOM   1291  CA  ALA   139      14.642  22.729  10.564  1.00 15.00
ATOM   1292  CB  ALA   139      13.514  22.455  11.533  1.00 15.00
ATOM   1293  C   ALA   139      15.049  21.394   9.982  1.00 15.00
ATOM   1294  O   ALA   139      16.154  20.874  10.113  1.00 15.00
ATOM   1295  N   LEU   140      14.153  20.939   9.149  1.00 15.00
ATOM   1296  H   LEU   140      13.409  21.520   8.893  1.00 15.00
ATOM   1297  CA  LEU   140      14.284  19.665   8.508  1.00 15.00
ATOM   1298  CB  LEU   140      12.940  19.379   7.829  1.00 15.00
ATOM   1299  CG  LEU   140      12.611  18.181   6.972  1.00 15.00
ATOM   1300  CD1 LEU   140      12.715  16.870   7.708  1.00 15.00
ATOM   1301  CD2 LEU   140      11.179  18.363   6.536  1.00 15.00
ATOM   1302  C   LEU   140      15.446  19.626   7.530  1.00 15.00
ATOM   1303  O   LEU   140      15.765  18.530   7.066  1.00 15.00
ATOM   1304  N   LYS   141      16.176  20.705   7.279  1.00 15.00
ATOM   1305  H   LYS   141      16.049  21.537   7.778  1.00 15.00
ATOM   1306  CA  LYS   141      17.119  20.677   6.172  1.00 15.00
ATOM   1307  CB  LYS   141      17.701  22.066   5.915  1.00 15.00
ATOM   1308  CG  LYS   141      17.994  22.402   4.451  1.00 15.00
ATOM   1309  CD  LYS   141      18.449  23.883   4.426  1.00 15.00
ATOM   1310  CE  LYS   141      19.187  24.360   3.142  1.00 15.00
ATOM   1311  NZ  LYS   141      19.335  25.815   3.148  1.00 15.00
ATOM   1312  HZ1 LYS   141      19.905  26.102   3.969  1.00 15.00
ATOM   1313  HZ2 LYS   141      18.396  26.258   3.205  1.00 15.00
ATOM   1314  HZ3 LYS   141      19.811  26.118   2.274  1.00 15.00
ATOM   1315  C   LYS   141      18.259  19.708   6.355  1.00 15.00
ATOM   1316  O   LYS   141      18.705  19.167   5.346  1.00 15.00
```

```
ATOM   1317  N   ALA  142      18.728  19.404   7.565  1.00 15.00
ATOM   1318  H   ALA  142      18.265  19.754   8.355  1.00 15.00
ATOM   1319  CA  ALA  142      19.871  18.490   7.690  1.00 15.00
ATOM   1320  CB  ALA  142      20.969  19.079   8.578  1.00 15.00
ATOM   1321  C   ALA  142      19.541  17.133   8.272  1.00 15.00
ATOM   1322  O   ALA  142      20.429  16.370   8.641  1.00 15.00
ATOM   1323  N   LEU  143      18.291  16.772   8.431  1.00 15.00
ATOM   1324  H   LEU  143      17.576  17.264   7.972  1.00 15.00
ATOM   1325  CA  LEU  143      18.040  15.449   8.949  1.00 15.00
ATOM   1326  CB  LEU  143      16.718  15.566   9.679  1.00 15.00
ATOM   1327  CG  LEU  143      16.644  16.739  10.633  1.00 15.00
ATOM   1328  CD1 LEU  143      15.343  16.783  11.353  1.00 15.00
ATOM   1329  CD2 LEU  143      17.680  16.590  11.666  1.00 15.00
ATOM   1330  C   LEU  143      18.055  14.507   7.718  1.00 15.00
ATOM   1331  O   LEU  143      18.152  14.982   6.585  1.00 15.00
ATOM   1332  N   PRO  144      18.015  13.182   7.762  1.00 15.00
ATOM   1333  CD  PRO  144      18.489  12.372   8.858  1.00 15.00
ATOM   1334  CA  PRO  144      17.668  12.383   6.605  1.00 15.00
ATOM   1335  CB  PRO  144      17.646  11.014   7.175  1.00 15.00
ATOM   1336  CG  PRO  144      18.799  11.109   8.109  1.00 15.00
ATOM   1337  C   PRO  144      16.418  12.728   5.766  1.00 15.00
ATOM   1338  O   PRO  144      15.345  13.283   6.071  1.00 15.00
ATOM   1339  N   MET  145      16.660  12.077   4.658  1.00 15.00
ATOM   1340  H   MET  145      17.403  11.441   4.653  1.00 15.00
ATOM   1341  CA  MET  145      15.913  12.192   3.437  1.00 15.00
ATOM   1342  CB  MET  145      16.997  11.803   2.324  1.00 15.00
ATOM   1343  CG  MET  145      17.748  10.422   2.344  1.00 15.00
ATOM   1344  SD  MET  145      16.788   8.877   2.482  1.00 15.00
ATOM   1345  CE  MET  145      17.473   8.257   3.994  1.00 15.00
ATOM   1346  C   MET  145      14.600  11.414   3.356  1.00 15.00
ATOM   1347  O   MET  145      13.811  11.276   4.277  1.00 15.00
ATOM   1348  N   HIS  146      14.435  10.930   2.129  1.00 15.00
ATOM   1349  H   HIS  146      14.977  11.305   1.408  1.00 15.00
ATOM   1350  CA  HIS  146      13.447  10.055   1.595  1.00 15.00
ATOM   1351  CB  HIS  146      13.252   8.754   2.357  1.00 15.00
ATOM   1352  CG  HIS  146      13.960   7.529   1.723  1.00 15.00
ATOM   1353  CD2 HIS  146      13.681   6.995   0.481  1.00 15.00
ATOM   1354  ND1 HIS  146      14.863   6.715   2.285  1.00 15.00
ATOM   1355  HD1 HIS  146      15.374   6.922   3.108  1.00 15.00
ATOM   1356  CE1 HIS  146      15.105   5.731   1.456  1.00 15.00
ATOM   1357  NE2 HIS  146      14.384   5.902   0.373  1.00 15.00
ATOM   1358  HE2 HIS  146      14.248   5.230  -0.339  1.00 15.00
ATOM   1359  C   HIS  146      12.125  10.659   1.489  1.00 15.00
ATOM   1360  O   HIS  146      11.731  10.627   0.331  1.00 15.00
ATOM   1361  N   ILE  147      11.445  11.246   2.456  1.00 15.00
ATOM   1362  H   ILE  147      11.817  11.473   3.338  1.00 15.00
ATOM   1363  CA  ILE  147      10.079  11.586   2.099  1.00 15.00
ATOM   1364  CB  ILE  147       9.298  10.300   2.547  1.00 15.00
ATOM   1365  CG2 ILE  147       9.229  10.325   4.058  1.00 15.00
ATOM   1366  CG1 ILE  147       7.938  10.169   1.866  1.00 15.00
ATOM   1367  CD1 ILE  147       7.422   8.705   1.768  1.00 15.00
ATOM   1368  C   ILE  147       9.487  12.900   2.569  1.00 15.00
ATOM   1369  O   ILE  147       9.788  13.277   3.704  1.00 15.00
ATOM   1370  N   ARG  148       8.738  13.656   1.749  1.00 15.00
ATOM   1371  H   ARG  148       8.580  13.364   0.830  1.00 15.00
ATOM   1372  CA  ARG  148       7.959  14.772   2.271  1.00 15.00
ATOM   1373  CB  ARG  148       8.386  16.120   1.771  1.00 15.00
ATOM   1374  CG  ARG  148       9.263  16.675   2.840  1.00 15.00
ATOM   1375  CD  ARG  148       9.586  18.124   2.635  1.00 15.00
ATOM   1376  NE  ARG  148       8.376  18.927   2.579  1.00 15.00
ATOM   1377  HE  ARG  148       7.585  18.648   3.087  1.00 15.00
ATOM   1378  CZ  ARG  148       8.332  20.048   1.847  1.00 15.00
ATOM   1379  NH1 ARG  148       9.411  20.436   1.147  1.00 15.00
ATOM   1380 HH11 ARG  148      10.251  19.895   1.161  1.00 15.00
ATOM   1381 HH12 ARG  148       9.372  21.272   0.600  1.00 15.00
ATOM   1382  NH2 ARG  148       7.214  20.806   1.835  1.00 15.00
```

```
ATOM   1383 HH21 ARG   148       6.414  20.528   2.367  1.00 15.00
ATOM   1384 HH22 ARG   148       7.184  21.638   1.280  1.00 15.00
ATOM   1385  C   ARG   148       6.524  14.602   1.850  1.00 15.00
ATOM   1386  O   ARG   148       6.258  14.149   0.740  1.00 15.00
ATOM   1387  N   LEU   149       5.617  14.896   2.767  1.00 15.00
ATOM   1388  H   LEU   149       5.914  15.197   3.654  1.00 15.00
ATOM   1389  CA  LEU   149       4.184  14.879   2.518  1.00 15.00
ATOM   1390  CB  LEU   149       3.508  13.839   3.364  1.00 15.00
ATOM   1391  CG  LEU   149       3.419  12.367   3.026  1.00 15.00
ATOM   1392  CD1 LEU   149       2.500  12.294   1.853  1.00 15.00
ATOM   1393  CD2 LEU   149       4.775  11.721   2.828  1.00 15.00
ATOM   1394  C   LEU   149       3.630  16.247   2.910  1.00 15.00
ATOM   1395  O   LEU   149       4.114  16.807   3.879  1.00 15.00
ATOM   1396  N   SER   150       2.705  16.894   2.239  1.00 15.00
ATOM   1397  H   SER   150       2.402  16.507   1.391  1.00 15.00
ATOM   1398  CA  SER   150       2.121  18.163   2.650  1.00 15.00
ATOM   1399  CB  SER   150       2.576  19.350   1.784  1.00 15.00
ATOM   1400  OG  SER   150       3.964  19.672   1.690  1.00 15.00
ATOM   1401  HG  SER   150       4.323  19.737   2.581  1.00 15.00
ATOM   1402  C   SER   150       0.654  17.911   2.339  1.00 15.00
ATOM   1403  O   SER   150       0.420  17.462   1.216  1.00 15.00
ATOM   1404  N   PHE   151      -0.400  18.068   3.120  1.00 15.00
ATOM   1405  H   PHE   151      -0.300  18.502   3.998  1.00 15.00
ATOM   1406  CA  PHE   151      -1.718  17.802   2.599  1.00 15.00
ATOM   1407  CB  PHE   151      -2.417  16.881   3.570  1.00 15.00
ATOM   1408  CG  PHE   151      -1.826  15.517   3.974  1.00 15.00
ATOM   1409  CD1 PHE   151      -0.708  15.408   4.778  1.00 15.00
ATOM   1410  CD2 PHE   151      -2.469  14.342   3.628  1.00 15.00
ATOM   1411  CE1 PHE   151      -0.279  14.163   5.209  1.00 15.00
ATOM   1412  CE2 PHE   151      -2.035  13.099   4.060  1.00 15.00
ATOM   1413  CZ  PHE   151      -0.935  13.004   4.856  1.00 15.00
ATOM   1414  C   PHE   151      -2.414  19.168   2.421  1.00 15.00
ATOM   1415  O   PHE   151      -1.820  20.193   2.752  1.00 15.00
ATOM   1416  N   ASN   152      -3.641  19.337   1.918  1.00 15.00
ATOM   1417  H   ASN   152      -4.265  18.590   1.941  1.00 15.00
ATOM   1418  CA  ASN   152      -4.180  20.683   1.607  1.00 15.00
ATOM   1419  CB  ASN   152      -4.461  20.882   0.052  1.00 15.00
ATOM   1420  CG  ASN   152      -3.544  21.792  -0.858  1.00 15.00
ATOM   1421  OD1 ASN   152      -2.796  21.339  -1.750  1.00 15.00
ATOM   1422  ND2 ASN   152      -3.568  23.128  -0.797  1.00 15.00
ATOM   1423 HD21 ASN   152      -4.169  23.588  -0.184  1.00 15.00
ATOM   1424 HD22 ASN   152      -2.948  23.578  -1.405  1.00 15.00
ATOM   1425  C   ASN   152      -5.505  20.840   2.359  1.00 15.00
ATOM   1426  O   ASN   152      -6.170  19.827   2.566  1.00 15.00
ATOM   1427  N   PRO   153      -6.020  21.998   2.789  1.00 15.00
ATOM   1428  CD  PRO   153      -5.643  23.307   2.302  1.00 15.00
ATOM   1429  CA  PRO   153      -7.019  22.127   3.839  1.00 15.00
ATOM   1430  CB  PRO   153      -7.151  23.596   4.029  1.00 15.00
ATOM   1431  CG  PRO   153      -6.832  24.151   2.668  1.00 15.00
ATOM   1432  C   PRO   153      -8.368  21.455   3.640  1.00 15.00
ATOM   1433  O   PRO   153      -9.022  21.049   4.611  1.00 15.00
ATOM   1434  N   THR   154      -8.823  21.269   2.403  1.00 15.00
ATOM   1435  H   THR   154      -8.318  21.549   1.615  1.00 15.00
ATOM   1436  CA  THR   154     -10.109  20.592   2.165  1.00 15.00
ATOM   1437  CB  THR   154     -10.585  20.761   0.670  1.00 15.00
ATOM   1438  OG1 THR   154      -9.569  20.184  -0.174  1.00 15.00
ATOM   1439  HG1 THR   154      -9.952  20.027  -1.046  1.00 15.00
ATOM   1440  CG2 THR   154     -10.794  22.258   0.260  1.00 15.00
ATOM   1441  C   THR   154      -9.894  19.134   2.472  1.00 15.00
ATOM   1442  O   THR   154     -10.825  18.356   2.577  1.00 15.00
ATOM   1443  N   GLN   155      -8.615  18.799   2.526  1.00 15.00
ATOM   1444  H   GLN   155      -7.969  19.428   2.174  1.00 15.00
ATOM   1445  CA  GLN   155      -8.189  17.489   2.869  1.00 15.00
ATOM   1446  CB  GLN   155      -6.922  17.219   2.189  1.00 15.00
ATOM   1447  CG  GLN   155      -6.543  15.807   2.218  1.00 15.00
ATOM   1448  CD  GLN   155      -5.177  15.641   1.624  1.00 15.00
```

```
ATOM   1449  OE1 GLN   155      -4.635  14.536   1.684  1.00 15.00
ATOM   1450  NE2 GLN   155      -4.583  16.649   0.982  1.00 15.00
ATOM   1451 HE21 GLN   155      -5.120  17.418   0.696  1.00 15.00
ATOM   1452 HE22 GLN   155      -3.616  16.611   0.895  1.00 15.00
ATOM   1453  C   GLN   155      -7.993  17.404   4.345  1.00 15.00
ATOM   1454  O   GLN   155      -8.398  16.419   4.966  1.00 15.00
ATOM   1455  N   LEU   156      -7.353  18.473   4.844  1.00 15.00
ATOM   1456  H   LEU   156      -7.111  19.184   4.225  1.00 15.00
ATOM   1457  CA  LEU   156      -6.986  18.625   6.266  1.00 15.00
ATOM   1458  CB  LEU   156      -6.253  19.993   6.470  1.00 15.00
ATOM   1459  CG  LEU   156      -4.811  20.079   7.011  1.00 15.00
ATOM   1460  CD1 LEU   156      -3.938  18.911   6.562  1.00 15.00
ATOM   1461  CD2 LEU   156      -4.223  21.381   6.519  1.00 15.00
ATOM   1462  C   LEU   156      -8.181  18.523   7.202  1.00 15.00
ATOM   1463  O   LEU   156      -8.075  17.906   8.264  1.00 15.00
ATOM   1464  N   GLU   157      -9.339  19.100   6.884  1.00 15.00
ATOM   1465  H   GLU   157      -9.444  19.634   6.069  1.00 15.00
ATOM   1466  CA  GLU   157     -10.420  18.892   7.802  1.00 15.00
ATOM   1467  CB  GLU   157     -11.353  19.993   7.652  1.00 15.00
ATOM   1468  CG  GLU   157     -11.200  20.608   8.995  1.00 15.00
ATOM   1469  CD  GLU   157     -12.064  21.831   9.038  1.00 15.00
ATOM   1470  OE1 GLU   157     -13.165  21.747   9.576  1.00 15.00
ATOM   1471  OE2 GLU   157     -11.643  22.861   8.499  1.00 15.00
ATOM   1472  C   GLU   157     -11.167  17.582   7.727  1.00 15.00
ATOM   1473  O   GLU   157     -12.062  17.358   8.544  1.00 15.00
ATOM   1474  N   GLU   158     -10.887  16.684   6.791  1.00 15.00
ATOM   1475  H   GLU   158     -10.067  16.762   6.262  1.00 15.00
ATOM   1476  CA  GLU   158     -11.604  15.416   6.792  1.00 15.00
ATOM   1477  CB  GLU   158     -11.605  14.817   5.363  1.00 15.00
ATOM   1478  CG  GLU   158     -12.528  15.643   4.401  1.00 15.00
ATOM   1479  CD  GLU   158     -12.592  15.275   2.885  1.00 15.00
ATOM   1480  OE1 GLU   158     -11.551  15.172   2.201  1.00 15.00
ATOM   1481  OE2 GLU   158     -13.702  15.107   2.351  1.00 15.00
ATOM   1482  C   GLU   158     -10.800  14.581   7.814  1.00 15.00
ATOM   1483  O   GLU   158      -9.766  14.999   8.285  1.00 15.00
ATOM   1484  N   GLN   159     -11.347  13.450   8.303  1.00 15.00
ATOM   1485  H   GLN   159     -12.224  13.174   7.977  1.00 15.00
ATOM   1486  CA  GLN   159     -10.650  12.599   9.287  1.00 15.00
ATOM   1487  CB  GLN   159     -11.612  11.538   9.778  1.00 15.00
ATOM   1488  CG  GLN   159     -11.244  10.673  10.954  1.00 15.00
ATOM   1489  CD  GLN   159     -12.196  10.903  12.120  1.00 15.00
ATOM   1490  OE1 GLN   159     -13.198  11.630  11.997  1.00 15.00
ATOM   1491  NE2 GLN   159     -11.913  10.328  13.297  1.00 15.00
ATOM   1492 HE21 GLN   159     -11.125   9.752  13.373  1.00 15.00
ATOM   1493 HE22 GLN   159     -12.547  10.503  14.023  1.00 15.00
ATOM   1494  C   GLN   159      -9.473  11.937   8.570  1.00 15.00
ATOM   1495  O   GLN   159      -9.475  11.911   7.334  1.00 15.00
ATOM   1496  N   CYS   160      -8.528  11.320   9.274  1.00 15.00
ATOM   1497  H   CYS   160      -8.590  11.285  10.247  1.00 15.00
ATOM   1498  CA  CYS   160      -7.362  10.730   8.664  1.00 15.00
ATOM   1499  C   CYS   160      -6.659  11.407   7.490  1.00 15.00
ATOM   1500  O   CYS   160      -6.406  10.744   6.469  1.00 15.00
ATOM   1501  CB  CYS   160      -7.691   9.340   8.267  1.00 15.00
ATOM   1502  SG  CYS   160      -7.274   8.307   9.668  1.00 15.00
ATOM   1503  N   HIS   161      -6.279  12.703   7.597  1.00 15.00
ATOM   1504  H   HIS   161      -6.393  13.202   8.436  1.00 15.00
ATOM   1505  CA  HIS   161      -5.544  13.469   6.575  1.00 15.00
ATOM   1506  CB  HIS   161      -6.538  14.094   5.602  1.00 15.00
ATOM   1507  CG  HIS   161      -7.040  13.098   4.562  1.00 15.00
ATOM   1508  CD2 HIS   161      -8.174  12.317   4.699  1.00 15.00
ATOM   1509  ND1 HIS   161      -6.517  12.805   3.366  1.00 15.00
ATOM   1510  HD1 HIS   161      -5.639  13.113   3.041  1.00 15.00
ATOM   1511  CE1 HIS   161      -7.294  11.896   2.798  1.00 15.00
ATOM   1512  NE2 HIS   161      -8.292  11.601   3.610  1.00 15.00
ATOM   1513  HE2 HIS   161      -8.936  10.872   3.496  1.00 15.00
ATOM   1514  C   HIS   161      -4.703  14.544   7.279  1.00 15.00
```

```
ATOM   1515  O1  HIS   161      -4.194  14.204   8.338  1.00 15.00
ATOM   1516  O2  HIS   161      -4.547  15.678   6.814  1.00 15.00
```