INVESTIGATION OF THE SUBSTRATE BINDING SITES OF ELASTASE

by

VILASBEN K. SHAH

A Thesis submitted in partial fulfillment of the requirements for the award of the degree of Doctor of Philosophy of the Council for National Academic Awards

Sponsoring Establishment :

Napier College of Commerce and Technology
Colinton Road
Edinburgh
EH10 5DT

Collaborating Establishment :

Roche Products Ltd
PO Box 8
Welwyn Garden City
Hertfordshire
AL7 3AY

JUNE 1986

I declare that this thesis was composed by me, that the work of which it is a record was done by me, and that the thesis has not been accepted in any previous application for a degree. All the sources of information are acknowledged.

V.K. Shah.

VILASBEN K. SHAH

INVESTIGATION OF THE SUBSTRATE BINDING SITES OF ELASTASE

ABSTRACT

Studies of pancreatic elastase were undertaken to examine substrate binding and of human leukocyte elastase to try to gain insight into the disease, pulmonary emphysema. Porcine pancreatic elastase was inhibited by various inhibitors : p-toluene sulphonyl fluoride, phenylmethane sulphonyl fluoride, 1-dimethylamino-naphthalene-5-sulphonyl fluoride or chloride and p-nitrophenyl-anthranilate. The enzyme was only fully inhibited by the first two inhibitors which were then treated under alkaline conditions to convert the serine-195, $-CH_2OH$ to $=CH_2$ to form the anhydroelastase. The enzyme was shown to be totally inactive but X-ray diffraction data to 2.5Å resolution from one preparation showed no changes at the active site. Further studies were carried out to modify the serine-195 to cysteine by thiolation of the inhibited enzyme and anhydroelastase but proved unsuccessful. Binding studies were carried out with the prepared hexapeptide substrate, H.Pro-Ala-Pro-Ala-Lys-Phe.OH and tetrapeptide inhibitors, Ac-Pro-Ala-Pro-Ala.OH (1), Ac-Pro-Ala-Pro-Alaninal (2) and TFA-Pro-Ala-Pro-Alaninal (3) using anhydroelastase and native elastase. Inhibition with inhibitors (2) and (3) showed TFA-peptide-aldehyde bound tighter than Ac-peptide-aldehyde with K_i for (2) = 6.1 µM and for (3) = 8.6 µM. A TFA-peptide-chloromethylketone was also examined. It showed anomalous binding with the TFA group bound to the active site region whilst the rest of the peptide bound in parallel mode with the main chain residues 214-216. Data on substrate (1) was collected but no binding around the active site region was observed.

Small crystals of human leukocyte elastase were obtained by the technique of vapour diffusion, however sealed tube X-ray experiments and using the sychroton revealed no diffraction pattern. Moreover, studies were done on the enzyme to remove the sugar content using various glycosidases of which almond emulsin showed some conclusive results.

In addition, the crystal structures of two small molecules containing the heavy atoms ruthenium and platinium were solved by conventional techniques of X-ray crystallography.

ACKNOWLEDGEMENT

It gives me great pleasure to acknowledge my indebtedness to all my friends and colleagues who have given invaluable help in the work for this thesis.

To Dr L. Sawyer for the supervision and advice throughout the course of this work. He has given me a great deal of help, encouragement and for the careful reading of the entire thesis, which cut out innumerable outrages against the English language.

Dr W.H. Johnson, Dr N.A. Roberts and all the colleagues at Roche Products for their help and advice.

To Dr G.H.W. Milburn who encouraged me throughout.

To Dr F. Körber and Mr E. Eliopoulos of the University of Leeds for their help in collecting data and assisting me with the Evans and Sutherland graphics system.

To Dr M. Paziz and Dr J. Raftery for the help with the computing.

To Napier College, the University of Edinburgh and the Edinburgh Regional Computing Centre for allowing me to use their facilities.

I am most grateful to the SERC, Roche Products and Napier College for their financial support.

To Marlyn Clark for her excellent typing and to my brother for assisting me with the figures.

Finally, I take all the responsibility for any errors which remain.

ABBREVIATIONS

SGPA	Streptomyces griseus protease A
SGPB	Streptomyces griseus protease B
PPE	Porcine pancreatic elastase
HLE	Human leukocyte elastase
DFP	Di-isopropyl-phosphofluoridate
Tosyl	p-Toulene sulphonyl
SH	Thiol
Ac	Acetyl
kcat	Rate constant for the reaction
K_m	Michaelis constant
K_i	Inhibition constant
CMK	Chloromethyl Ketone
N.M.R.	Nuclear Magnetic Resonance
Z	Benzyloxycarbonyl
ONp	p-Nitrophenyl
TFA	Trifluoroacetyl
Ph	Phenyl
kosbd	Observed first order rate constant
[I]	Inhibitor concentration
PMSF	Phenylmethane Sulphonyl Fluoride
Dansyl	1-dimethylamino-naphthalene-5-sulphonyl
NPA	p-Nitrophenol anthranilate
NBA	N-t-Butyloxycarbonyl-L-alaninate-p-nitrophenyl ester
B.A.E.E.	N-α-Benzoyl-L-arginine ethyl ester
A.T.E.E.	N-Acetyl-L-tyrosine ethyl ester
Ellman's reagent	5,5' -dithio-bis-(2-nitrobenzoic acid)
Tris	2-amino-2-hydroxy-methyl-propane-1,3-diol
PMSE	Phenylmethane Sulphonyl Elastase
BOC	Tert-Butyloxycarbonyl
-OBz	Benzyl ester
O^tBu	Tert-Butyl ester
DCCI	Dicyclohexylcarbodiimide
NHS	N-hydroxysuccinimide

ABBREVIATIONS (Contd.)

NEM	N-ethyl morpholine
THF	Tetrahydrofuran
DCM	Dichloromethane
Pd/C	5% palladium on charcoal
DME	1,2-dimethoxy ethane
DMF	N,N'-dimethyl formamide
E	Enzyme
I	Inhibitor
[ES]	Enzyme-Substrate complex
[EI]	Enzyme-Inhibitor complex
V_{max}	Maximum velocity obtained at high substrate concentration
[S]	Substrate concentration
Suc	Succinyl
MeO-Suc	Methyl Succinyl
$\underline{S}$	Scattering vector as $(s/\lambda - s_o/\lambda)$
$\rho(\underline{r})$	Electron density in a small unit of volume dv at a position r
$F_{(\underline{S})}$	Structure factor
B	Thermal motion of atoms
$\bar{U}^2$	Mean displacement of atoms along normal to the reflecting plane
$F_{(hkl)}$	Structure factor at the reciprocal lattice point hkl
F(hkl)	Structure factor amplitude
α(hkl)	Phase
I(hkl)	Intensity at the reciprocal lattice point hkl
P(UVW)	Patterson function
ΔF	$= \lvert F_o \rvert - \lvert F_c \rvert$
$\lvert F_o \rvert$	The scaled observed structure factor amplitude
$\lvert F_c \rvert$	The calculated structure factor amplitude
F_{PH}	The structure factor for the heavy atom isomorphous derivative
F_P	The structure factor for the native protein
F_H	The structure factor for the heavy atom

ABBREVIATIONS (Contd.)

R	Residual index
M.I.R.	Multiple Isomorphous Replacement
F_{PL}	The structure factor for the protein-ligand complex
F_{L}	The structure factor for the ligand
W	Weighting factor
L	Lorentz factor
P	Polarization factor
C	Scale factor
F_{HLE}	Heavy atom lower estimate

CONTENTS

CHAPTER THREE

PEPTIDE SYNTHESIS

CHAPTER FOUR

KINETIC STUDIES

CONTENTS (Contd.)

CHAPTER FIVE

HUMAN LEUKOCYTE ELASTASE

CHAPTER SIX

X-RAY CRYSTALLOGRAPHY

CHAPTER SEVEN

INTENSITY-DATA COLLECTION AND DATA REDUCTION

CONTENTS (Contd).

CHAPTER EIGHT

CHAPTER NINE

CHAPTER ONE

INTRODUCTION

1.1 HISTORY AND OCCURRENCE

The serine proteases, e.g. chymotrypsin, trypsin, elastase, thrombin, plasmin and subtilisin, are a class of proteolytic enzymes characterised by the presence of uniquely reactive serine residue. They are of extremely widespread occurrence and their function is to alter or decompose other proteins into fragments by hydrolysis of peptide bonds[1]. The two main groups most widely studied, are the trypsin family and the subtilisin family[2].

Members of the trypsin family (e.g. trypsin, chymotrypsin and elastase) occur in the vertebrates, phyla of higher organisms and also in prokaryotes, while the subtilisin family has been identified mainly among the bacilli[2]. Both families have their catalytically important functional groups arranged in the same geometrical pattern, but have entirely different overall three-dimensional structures.

The serine proteases that are involved in blood clotting are plasmin and thrombin which are present in the serum of all vertebrates[1]. Another group of serine proteases which exist in bacteria, Streptomyces griseus protease A and B (SGPA and SGPB) are regarded as possible evolutionary precursors of the mammalian serine proteases. They are isolated from the commercially available extracellular bacterial culture filtrate, pronase[3-5]. More recently, a member closely related to the mammalian serine proteases of animal origin, kallikreins, has been investigated[6,7]. They have the ability to release a vasodepressor decapeptide called, kallidin or Lys-bradykinin from kininogens[8].

The serine proteases that exhibits trypsin-like specificity are porcine kallikrein, thrombin and plasmin. Porcine kallikrein splits peptide substrates preferentially on the C-terminal side of internal arginyl and lysyl residue. Plasmin is also like trypsin and breaks down the fibrin

matrix by splitting all the peptide bonds following arginine or lysine units. Thrombin, on the other hand, is much more selective and splits only four bonds in fibrinogen, between arginine and glycine whereas trypsin would reduce the same molecule to 150 fragments[1]. SGPA, SGPB and subtilisin all closely resembles chymotrypsin specificity, but subtilisin will also accept non-aromatic a polar side-chains and even charged side-chains.

The pancreas of all mammals investigated, including man, contains a specific elastase (EC 3.4.21.11) which was first investigated in 1949 by Balo and Banga[9]. The work described in this thesis has concentrated mainly with the porcine pancreatic elastase (PPE) and human leukocyte elastase, (HLE), (discussed later in chapter five).

PPE has been the focus of most attention, chiefly because it is easily obtained in high yield. In 1956, Lewis et al[10] were first to acquire the pure enzyme from porcine pancreas. PPE is one of the members of the serine proteinases (Hartley 1960)[11] which originates in the pancreas as proelastase and is secreted into the duodenum by the acinar tissue of the pancreas. The zymogen, is then activated by trypsin, during which a small peptide of determined sequence[12] is removed from the N-terminal amino group valine-16 (the chymotrypsinogen-A numbering scheme will be used throughout to describe the position of amino-acid residues in PPE[13], thus facilitating comparison with chymotrypsin, trypsin and other homologous enzymes). This enables the molecule to adopt its active conformation where valine-16 forms a salt bridge with asparate-194.

The activated enzyme, elastase, catalyzes the decomposition of proteins into fragmentary peptides, which are reduced further into individual amino-acids by trypsin and chymotrypsin amongst others. These amino-acids are then absorbed by the intestine and carried to the sites of protein synthesis.

PPE was shown to digest a wide variety of substrates[10] but is distinguished from the other pancreatic endopeptidases by its ability to digest elastin, the elastic fibrous protein of connective tissue[21]. It's specificity is for substrates with uncharged non-aromatic side-chains[22] unlike those of trypsin and chymotrypsin whose preferred substrates are basic and aromatic side-chains respectively.

PPE is inhibited by di-isopropyl-phosphofluoridate (DFP) in a similar way to trypsin and chymotrypsin and a comparison of the pancreatic serine proteases, SGPA, SGPB, porcine kallikrein, plasmin and thrombin (only the B chain shows homology) shows that they all have similar sequences around the active centre, i.e. Gly-Asp-Ser*-Gly-Gly, and it is the serine that reacts with DFP[23,24]. In the case of subtilisin, Gly-Thr-Ser*-Met-Ala sequence is characteristic around the active centre serine residue in this group of bacterial proteins.

PPE consists of a single polypeptide chain of 240 amino-acid residues, with a molecular weight of 25,900[13], which is homologous to trypsin and the B and C chains of chymotrypsin. It also contains a disulphide-bridged "histidine loop" sequence similar to that in trypsin and chymotrypsin[25].

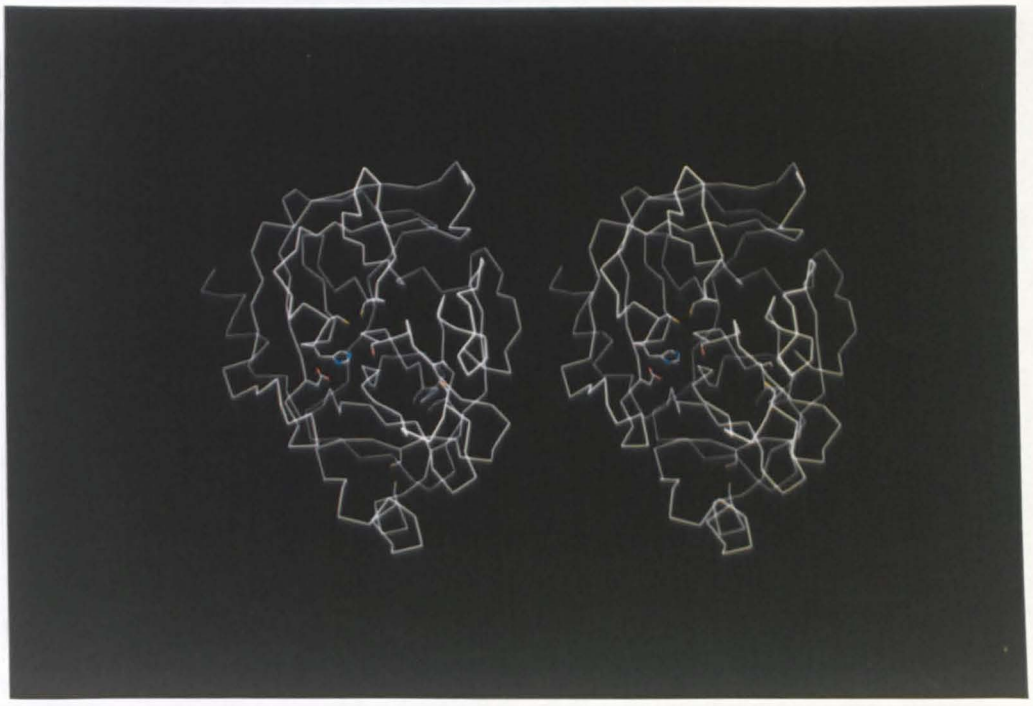
A general analysis carried out on the degree of homology by Shotton et al[13] between chymotrypsin and trypsin and elastase showed between 35 per cent and 45 per cent identity in amino-acid sequence, even though their specificities are different. They noticed that the amino-acid homologies are greater around the residues which are important for catalysis i.e., serine-195, histidine-57 and asparate-102 than elsewhere. Also, analysis of the distribution of homologous residues, in terms of per cent chemical similarity[a] between the internal residues and external

positions showed 71 per cent internal residues and only 19 per cent external ones.

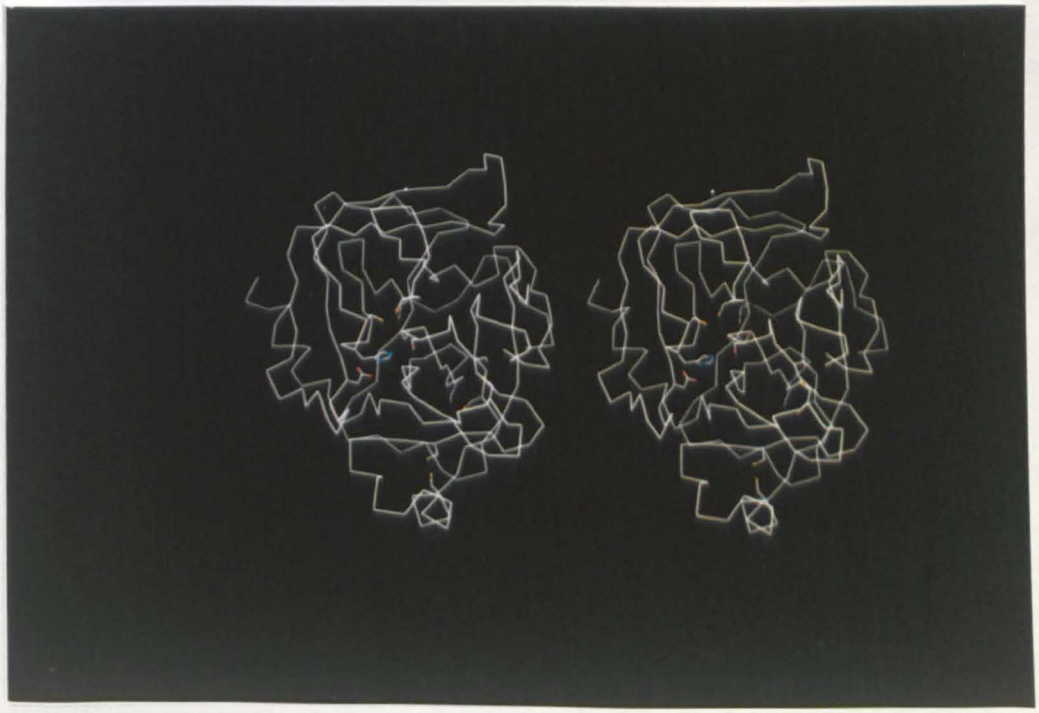
[a]Chemical similarity is defined as : Arg = Lys; Asp = Glu; Asn = Gln; Asp = Asn; Glu = Gln; Ser = Thr; Val = Ile; Ile = Leu; Tyr = Phe = Trp.

Figure 1.0 shows the similarity in the tertiary structure between the various serine proteases, with the residues which are important for catalysis, i.e., serine-195, histidine-57 and asparate-102. Also the cysteine's in the molecule are shown.

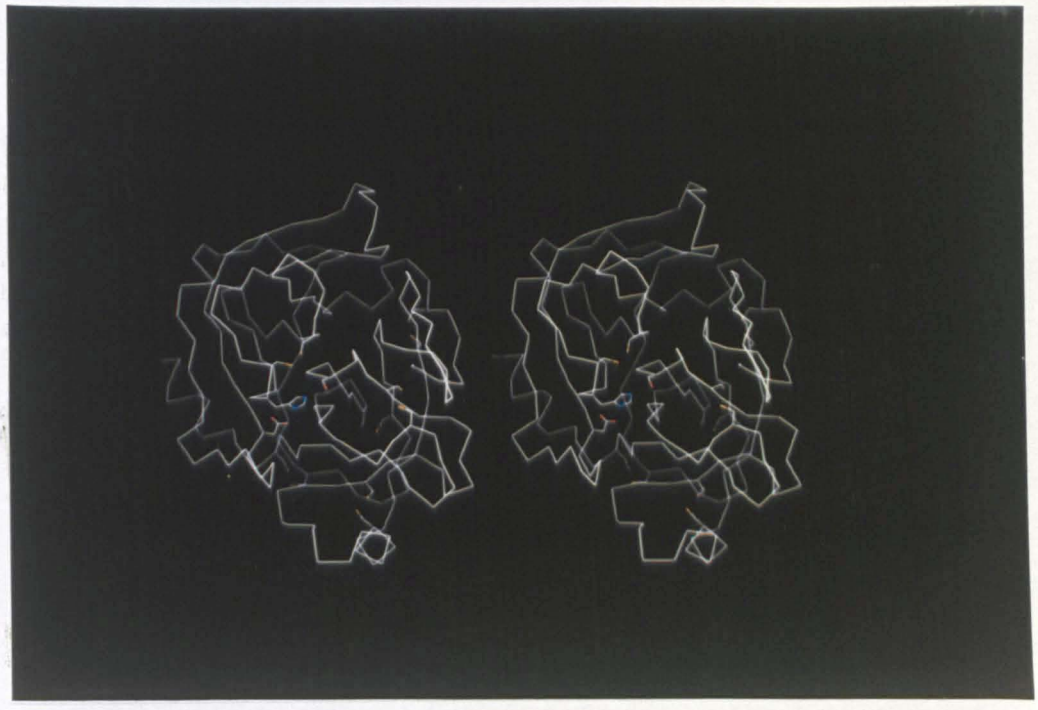
(1) PPE



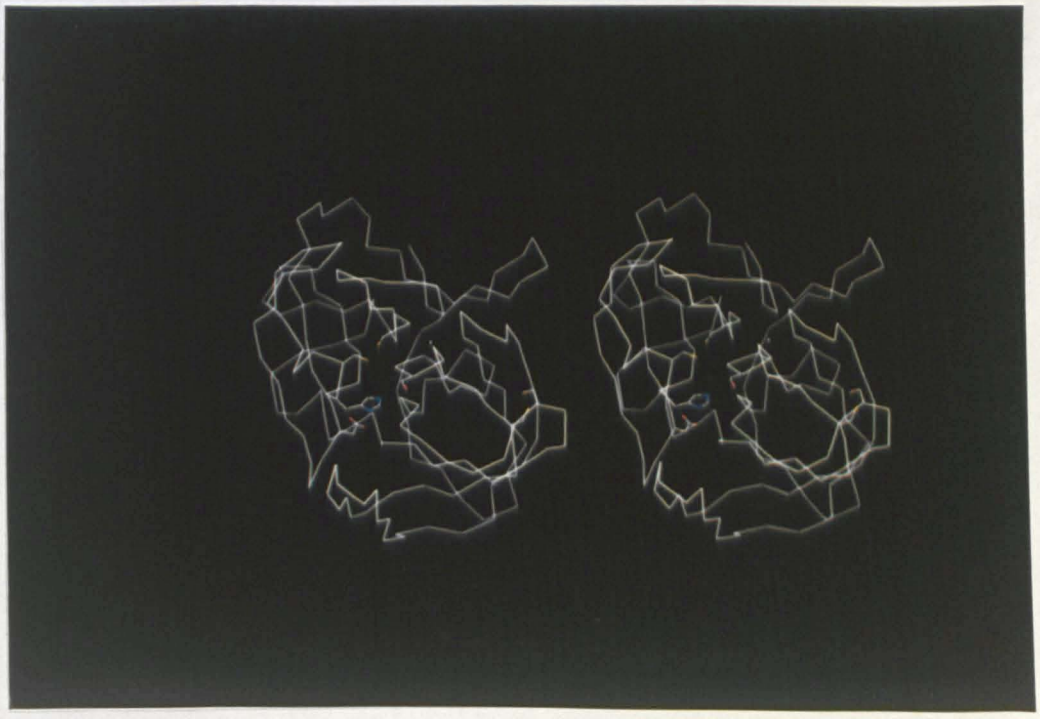
(2) CHYMOTRYPSIN



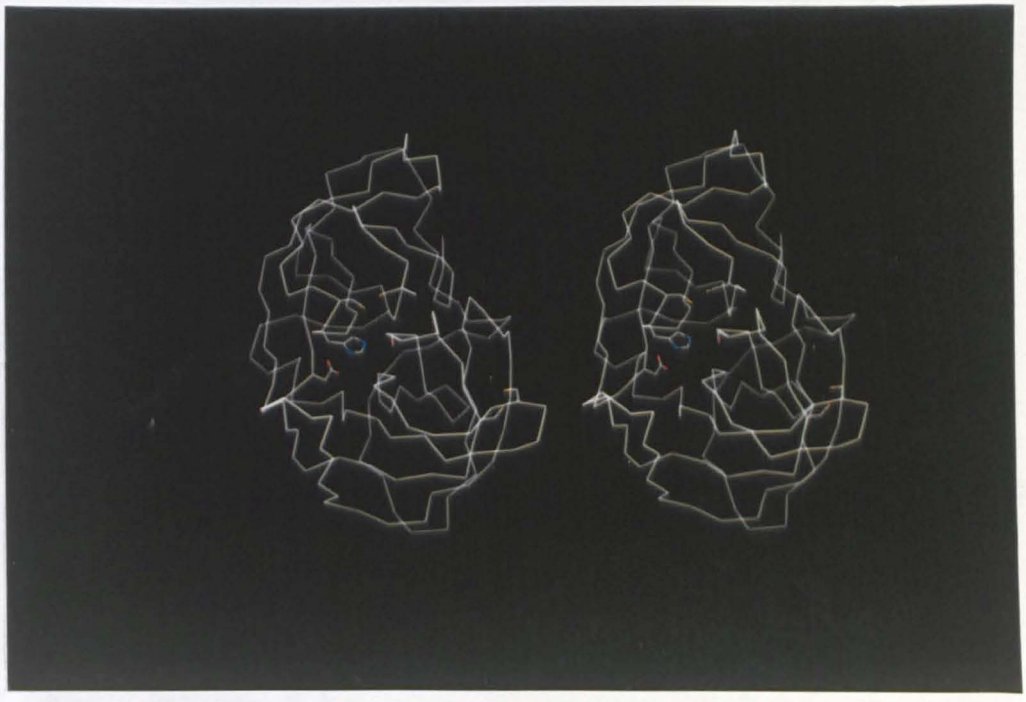
(3) TRYPSIN



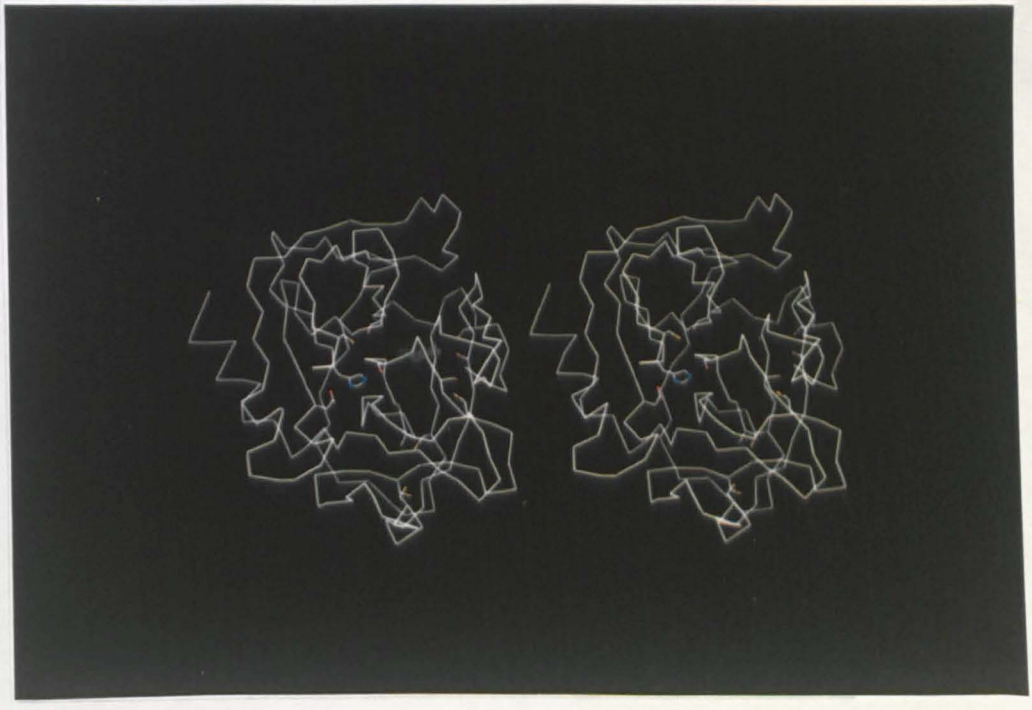
(4) SGPA



(5) SGPB



(6) PORCINE PANCREATIC KALLIKREIN



(7) α-LYTIC PROTEASE

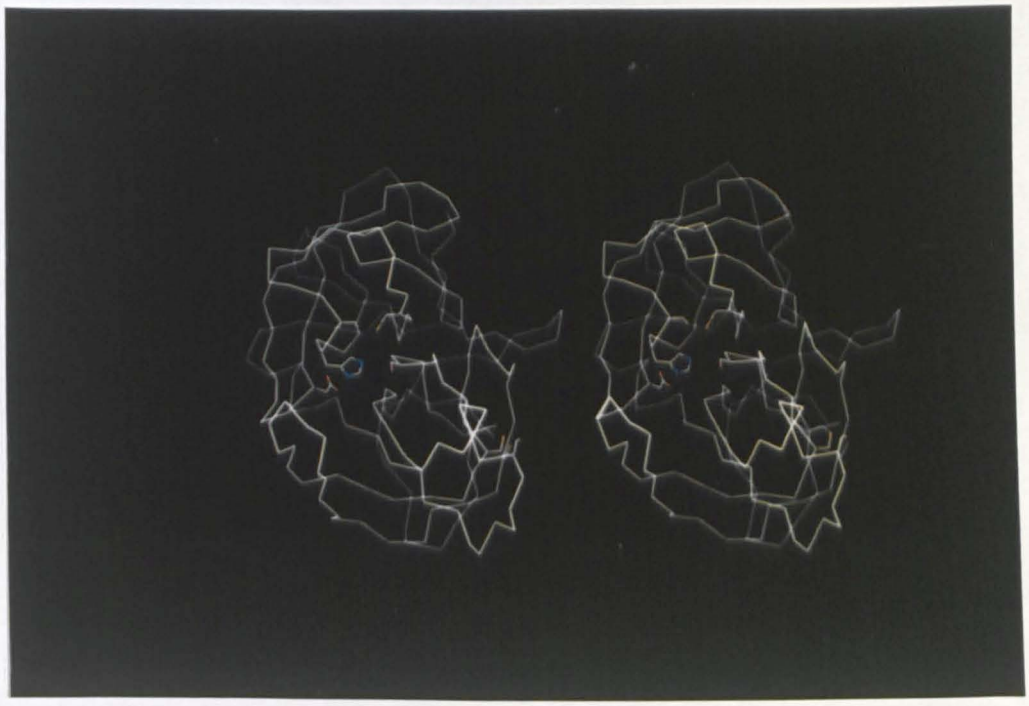


FIGURE 1.0

1.2 PURIFICATION

PPE is usually isolated from an activated extract of porcine pancreas, commercially available as trypsin 1-300 (Nutritional Biochemical Corporation). Shotton[26] has described the complete isolation procedure of the enzyme from the acetone powder.

The steps involved are : (1) crude PPE is extracted from trypsin 1-300 with 0.1M sodium acetate buffer pH 4.5[10]; (2) precipitate from the extract by 45% saturation of ammonium sulphate[10]; (3) the precipitate is dialysed against distilled water and the euglobulin precipitate collected[10]. (4) Batch adsorption of impurities on DEAE-Sephadex at pH 8.8[27]; is followed (5) by dialysis against 1mM acetic acid and freeze-drying; (6) The first batch crystallisation with 0.1M sodium sulphate buffered at pH 5.0 with 0.01M sodium acetate; is followed (7) by dialysis against 1mM acetic acid and freeze-drying; (8) re-crystallisation is as in steps 6 and 7.

PPE is always associated with some trypsin and chymotrypsin activity. The degree of contamination by trypsin and chymotrypsin for the above procedure has been shown by Shotton[28] to be less than 0.01% and 0.04% respectively. The yield obtained is about 2.7g of twice-crystallised PPE per 500g of trypsin 1-300. This procedure has so far given the purest PPE reported.

1.3 STABILITY

In comparison with the other two pancreatic proteases, PPE seems more stable. The study reported by Lewis et al[10] shows when PPE is incubated for 24 hours between pH 4 to 12 at 5° C, it does not lose its activity. However, below pH 4.0 rapid inactivation occurs. This has been shown by Gertler and Hofmann[29] and Wasi and Hofmann[30] to be related

to the pH-dependent conformation change of the PPE molecule (pK_a 3.55). The stability of trypsin and chymotrypsin, in comparison with PPE is greater at pH 3.0.

Concentrations of up to 50 mg/ml of PPE were found to be readily soluble in water or dilute salt solutions between pH 4.0 and 10.5[26]. At 2°C solutions of PPE are stable for prolonged periods below pH 6.0 and also reasonably stable at higher pH's. However, if incubated at room temperature at or near its optimum pH of 8.8, PPE was found to undergo autolysis to mixture of peptides[28].

Crystals of PPE in 1.2M sodium sulphate buffered at pH 5.0 with 0.01M sodium acetate are stable indefinitely at room temperature and freeze-dried PPE is stable for long periods at -10°C[28].

1.4 PHYSICAL PROPERTIES

The PPE molecule consists of a single polypeptide chain of 240 amino-acid residues containing four disulphide bridges. The amino-acid sequence is shown in table 1.1. From the amino-acid composition[13], the molecular weight of PPE was calculated at 25,900 which compared very well with the figure of 25,000 estimated from ultracentrifugal analysis[10]. The other physical properties are summarised in table 1.2

TABLE 1.1

THE AMINO-ACID SEQUENCE ALIGNMENT OF PPE[13] WITH THOSE OF BOVINE CHYMOTRYPSIN A (BCA)[14], BOVINE CHYMOTRYPSIN B (BCB)[15], PORCINE PANCREATIC TRYPSIN (PPTY)[6], SGPA[15], SGPB[15], PORCINE PANCREATIC KALLIKREIN (PPKK)[16], CRAB COLLAGENASE (COLL)[17], AND THE LIGHT CHAINS OF HUMAN COMPLEMENT COMPONENT $C\bar{1}r$[18], AND $C\bar{1}s$[19]

	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
PPE	V	V	G	G	T	E	A	Q	R	N	S	W	P	S	Q	I	S	L	Q	Y
BCA	I	V	N	G	E	E	A	V	P	G	S	W	P	W	Q	V	S	L	Q	D
BCB	I	V	N	G	E	D	A	V	P	G	S	W	P	W	Q	V	S	L	Q	D
PPTY	I	V	G	G	Y	T	C	A	A	N	S	V	P	Y	Q	V	S	L	N	-
SGPA	I	A	G	G	-	-	-	-	-	-	-	-	-	E	A	I	T	T	G	-
SGPB	I	S	G	G	-	-	-	-	-	-	-	-	-	D	A	I	Y	S	S	-
PPKK	I	I	G	G	R	E	C	E	K	N	S	H	P	W	Q	V	A	I	Y	-
COLL	I	V	G	G	V	E	A	V	P	N	S	W	P	H	Q	A	A	L	F	I
$C\bar{1}r$	I	I	G	G	Q	K	A	K	M	G	N	F	P	W	Q	V	F	-	T	N
$C\bar{1}s$	I	I	G	G	S	D	A	D	I	K	N	F	P	W	Q	V	F	F	D	N

TABLE 1.1 (Contd.

	36	36A	36B	36C	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52
PPE	R	S	G	S	S	W	A	H	T	C	G	G	T	L	I	R	Q	N	W	V
BCA	K	-	-	-	T	G	F	H	F	C	G	G	S	L	I	N	E	N	W	V
BCB	S	-	-	-	T	G	F	H	F	C	G	G	S	L	I	S	E	D	W	V
PPTY	-	-	-	-	S	G	S	H	F	C	G	G	S	L	I	N	S	Q	W	V
SGPA	-	-	-	-	-	-	G	S	R	C	S	L	G	F	N	V	V	A	H	A
SGPB	-	-	-	-	-	-	T	G	R	C	S	L	G	F	N	V	T	Y	Y	F
PPKK	Y	-	-	-	H	S	S	F	Q	C	G	G	V	L	V	N	P	K	W	V
COLL	D	-	-	-	D	M	Y	-	F	C	G	G	S	L	I	S	P	E	W	I
$C\bar{1}r$	I	-	-	-	H	G	R	-	-	G	G	G	A	L	L	G	D	R	W	I
$C\bar{1}s$	P	-	-	-	-	-	-	-	W	A	G	G	A	L	I	N	E	Y	W	V

TABLE 1.1 (Contd.)

	53	54	55	56	57	58	59	60	61	62	63	64	65	65A	66	67	68	69	70	71
PPE	M	T	A	A	H	C	V	D	R	E	L	T	F	R	V	V	V	G	E	H
BCA	V	T	A	A	H	C	G	V	T	T	S	D	V	-	V	V	A	G	E	F
BCB	V	T	A	A	H	C	G	V	T	T	S	D	V	-	V	V	A	G	E	F
PPTY	V	S	A	A	H	C	Y	K	S	R	I	Q	V	R	L	-	-	G	E	H
SGPA	L	T	A	G	H	C	T	N	-	I	S	A	-	S	W	S	-	-	-	-
SGPB	L	T	A	G	H	C	T	D	-	G	A	T	-	T	W	W	-	-	-	-
PPKK	L	T	A	A	H	C	K	N	D	N	Y	E	V	W	L	-	-	G	R	H
COLL	L	T	A	A	H	C	M	D	G	A	G	F	V	D	V	V	V	G	G	A
$C\bar{1}r$	L	T	A	A	H	T	L	Y	A	S	L	D	V	-	P	L	G	H	T	N
$C\bar{1}s$	L	T	A	A	H	V	V	E	-	-	-	-	M	-	Y	V	G	S	T	S

TABLE 1.1 (Contd.)

	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91
PPE	N	L	N	Q	N	N	G	T	E	Q	Y	V	G	V	Q	K	I	V	V	H
BCA	D	Q	G	S	S	S	E	K	I	Q	K	L	K	I	A	K	V	F	K	N
BCB	D	Q	G	L	E	T	E	D	T	Q	V	L	K	I	G	K	V	F	K	N
PPTY	N	I	D	V	L	E	G	N	E	Q	F	I	N	A	A	K	I	I	T	H
SGPA	-	-	-	-	-	-	-	-	-	-	-	-	-	I	G	T	R	T	G	T
SGPB	-	-	-	-	-	A	N	S	A	R	T	T	V	L	G	T	T	S	G	S
PPKK	N	I	K	E	D	E	N	T	A	Q	F	F	G	V	T	A	D	F	P	H
COLL	N	I	R	-	E	D	E	A	T	Q	V	T	I	S	T	D	F	T	V	H
$\overline{C1r}$	V	-	-	-	-	-	-	-	-	E	E	L	-	-	M	K	L	G	N	H
$\overline{C1s}$	V	Q	T	S	R	L	A	K	S	K	M	L	T	P	E	H	V	F	I	H

TABLE 1.1 (Contd.)

	92	93	94	95	96	97	98	99	99A	99B	100	101	102	103	104	105	106	107	108	109
PPE	P	Y	W	N	T	D	D	V	A	A	G	Y	D	I	A	L	L	R	L	A
BCA	S	K	Y	N	S	L	T	I	-	-	N	N	D	I	T	L	L	K	L	S
BCB	P	K	F	S	I	L	T	V	-	-	R	N	D	I	T	L	L	K	L	A
PPTY	P	N	F	N	G	N	T	L	-	-	D	N	D	I	M	L	I	K	L	S
SGPA	-	S	F	-	-	-	-	-	P	-	N	N	D	Y	G	I	I	R	H	S
SGPB	-	S	F	-	-	-	-	-	P	-	N	N	D	Y	G	I	V	R	Y	T
PPKK	P	G	F	N	G	K	D	Y	-	-	S	H	D	L	M	L	L	R	L	Q
COLL	E	N	Y	N	S	F	V	I	-	-	S	N	D	I	A	V	V	R	L	P
$\overline{C1r}$	P	I	R	E	S	Y	N	F	-	-	E	G	D	I	A	L	L	E	L	E
$\overline{C1s}$	P	G	W	F	D	-	-	-	-	-	-	N	D	I	A	L	V	R	L	K

TABLE 1.1 (Contd.)

	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129
PPE	Q	S	V	T	L	N	S	Y	V	Q	L	G	V	L	P	R	A	G	T	I
BCA	T	A	A	S	F	S	Q	T	V	S	A	V	C	L	P	S	A	S	D	D
BCB	T	P	A	Q	F	S	E	T	V	S	A	V	C	L	P	S	A	D	E	D
PPTY	S	P	A	T	L	N	S	R	V	A	T	V	S	L	P	R	-	S	C	A
SGPA	N	P	A	A	A	N	G	R	V	Y	L	Y	Q	D	I	T	T	A	G	N
SGPB	N	T	T	I	P	K	D	G	T	V	G	G	Q	D	I	T	S	A	A	N
PPKK	S	P	A	K	I	T	D	A	V	K	V	L	Q	L	P	T	-	Q	E	P
COLL	V	P	V	T	L	T	A	A	I	A	T	V	G	L	P	S	V	G	V	-
$C\bar{1}r$	N	S	V	T	L	G	P	N	L	L	P	I	C	L	P	D	N	D	T	F
$C\bar{1}s$	D	P	V	K	M	G	P	T	V	S	P	I	C	L	P	G	T	S	S	D

TABLE 1.1 (Contd.)

	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148
PPE	L	A	N	N	S	P	C	Y	I	T	G	W	G	L	T	R	-	T	N
BCA	F	A	A	G	T	T	C	V	T	T	G	W	G	L	T	R	Y	T	N
BCB	F	P	A	G	M	L	C	A	T	T	G	W	G	K	T	K	Y	N	A
PPTY	A	-	A	G	T	E	C	L	I	S	G	W	G	N	T	K	S	S	G
SGPA	A	F	V	G	Q	A	V	Q	R	S	G	S	T	T	-	-	-	-	-
SGPB	A	T	V	G	M	A	V	T	R	R	G	S	T	T	-	-	-	-	-
PPKK	E	L	L	G	S	T	C	E	A	S	G	W	G	S	I	E	P	P	B
COLL	-	-	-	G	T	V	V	T	P	T	G	W	G	L	S	-	D	S	A
$C\bar{1}r$	Y	-	-	-	D	L	G	L	V	S	G	P	G	V	-	-	M	E	E
$C\bar{1}s$	Y	N	L	G	D	L	G	L	I	S	G	W	G	R	-	-	T	E	K

TABLE 1.1 (Contd.)

	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167
PPE	G	Q	L	A	Q	T	L	Q	Q	A	Y	L	P	T	V	D	Y	A	I
BCA	A	N	T	P	D	R	L	Q	Q	A	S	L	P	L	L	S	N	T	N
BCB	L	K	T	P	D	K	L	Q	Q	A	T	L	P	I	V	S	N	T	D
PPTY	S	S	Y	P	S	L	L	Q	C	L	K	A	P	V	L	S	D	S	S
SGPA	-	-	-	-	-	-	-	G	L	R	S	G	S	V	T	G	L	N	A
SGPB	-	-	-	-	-	-	-	G	T	H	S	G	S	V	T	A	L	N	A
PPKK	F	E	F	P	D	E	I	Q	C	V	Q	L	T	L	L	Q	N	T	F
COLL	L	G	I	S	D	V	L	R	Q	V	D	V	P	I	M	-	N	S	D
$C\bar{1}r$	K	-	I	A	R	D	L	R	F	V	R	L	P	V	A	N	P	Q	A
$C\bar{1}s$	R	D	R	A	V	R	L	K	A	A	R	L	P	V	A	P	L	R	K

TABLE 1.1 (Contd.)

	168	169	170	170A	170B	171	172	173	174	175	176	177	178	179	180	181	182	183	184
PPE	C	S	S	S	S	Y	W	G	S	T	V	K	N	S	M	V	C	A	G
BCA	C	K	K	-	-	Y	W	G	T	K	I	K	D	A	M	I	C	A	G
BCB	C	R	K	-	-	Y	W	G	S	R	V	T	D	V	M	I	C	A	G
PPTY	C	K	S	-	A	Y	-	P	G	Q	I	T	G	N	M	I	C	V	G
SGPA	T	V	N	-	-	Y	G	S	S	G	I	V	Y	G	M	I	Q	T	N
SGPB	T	V	N	-	-	Y	G	G	G	D	V	V	Y	G	M	I	R	T	N
PPKK	C	A	B	-	A	H	-	P	B	K	V	T	E	S	M	L	C	A	G
COLL	C	D	-	A	V	Y	-	G	I	-	V	T	D	G	N	I	C	I	D
$C\bar{1}r$	C	E	N	-	-	W	L	R	G	K	N	R	Q	N	M	F	C	A	G
$C\bar{1}s$	C	K	E	-	-	Y	K	V	E	K	P	T	P	N	M	I	C	A	G

TABLE 1.1 (Contd.)

	185	186	187	188	188A	189	190	191	192	193	194	195	196	197	198	199	200	201
PPE	G	N	G	V	R	S	G	C	Q	G	D	S	G	G	P	L	H	C
BCA	A	S	G	V	-	S	S	C	M	G	D	S	G	G	P	L	V	C
BCB	A	S	G	V	-	S	S	C	M	G	D	S	G	G	P	L	V	C
PPTY	L	E	G	G	K	D	S	C	Q	G	D	S	G	G	P	V	V	C
SGPA	-	-	-	-	-	-	V	C	A	G	D	S	G	G	S	L	F	A
SGPB	-	-	-	-	-	-	V	C	A	G	D	S	G	G	P	L	Y	S
PPKK	L	P	G	G	K	D	T	C	M	G	D	S	G	G	P	L	I	C
COLL	S	T	G	G	K	G	T	C	D	G	D	S	G	G	P	L	N	Y
$C\bar{1}r$	H	P	S	L	K	D	A	C	Q	G	D	S	G	G	V	F	V	R
$C\bar{1}s$	G	E	K	G	M	D	S	C	K	G	D	S	G	G	A	F	V	Q

TABLE 1.1 (Contd.)

	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	217A	218	219
PPE	L	V	N	G	Q	Y	A	V	H	G	V	T	S	F	V	S	R	L	G
BCA	K	K	N	G	A	W	T	L	V	G	I	V	S	W	G	S	–	S	T
BCB	Q	K	N	G	A	W	T	L	A	G	I	V	S	W	G	S	–	S	T
PPTY	N	–	–	–	–	G	Q	L	Q	G	I	V	S	W	G	Y	–	–	G
SGPA	G	–	–	–	–	S	T	A	L	G	L	T	S	G	G	S	–	G	N
SGPB	G	–	–	–	–	T	R	A	I	G	L	T	S	G	G	S	–	C	N
PPKK	N	–	–	–	–	G	M	W	Q	G	I	T	S	W	G	H	–	T	P
COLL	–	–	D	G	–	–	L	T	Y	G	I	T	S	F	G	S	–	–	G
$C\bar{1}r$	D	P	N	R	–	W	V	A	T	G	I	V	S	W	G	–	–	I	G
$C\bar{1}s$	D	P	N	G	K	F	Y	A	A	G	L	V	S	W	G	–	–	P	Q

TABLE 1.1 (Contd.)

	220	221	221A	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237
PPE	C	N	V	T	R	K	P	T	V	F	T	R	V	S	A	Y	I	S	W
BCA	C	S	–	T	S	T	P	G	V	Y	A	R	V	T	A	L	V	N	W
BCB	C	S	–	T	S	T	P	A	V	Y	A	R	V	T	A	L	M	P	W
PPTY	C	A	Q	K	N	K	P	G	V	Y	T	K	V	C	N	Y	V	N	W
SGPA	C	R	–	T	G	G	T	T	F	Y	Q	P	V	T	E	A	L	A	Y
SGPB	C	S	–	S	G	G	T	T	F	F	Q	P	V	T	E	A	L	A	Y
PPKK	C	G	S	A	N	K	P	S	I	Y	T	K	L	I	F	Y	L	D	W
COLL	C	–	E	A	G	Y	P	D	A	F	T	R	V	T	Y	F	L	D	W
$C\overline{1r}$	C	S	–	R	–	G	Y	G	F	Y	T	K	V	L	N	Y	V	D	W
$C\overline{1s}$	C	G	–	–	–	T	Y	G	L	Y	T	R	V	K	N	Y	V	D	W

TABLE 1.1 (Contd.)

	238	239	240	241	242	243	244	245
PPE	I	N	N	V	I	A	S	N
BCA	V	Q	Q	T	L	A	A	A
BCB	V	Q	E	T	L	A	A	N
PPTY	I	Q	Q	T	L	A	A	N
SGPA	G	A	T	V	L	-	-	-
SGPB	G	V	S	V	Y	-	-	-
PPKK	I	B	B	T	I	T	E	N
COLL	I	Q	Q	T	I	-	T	P
$C\overline{1r}$	I	K	K	E	M	E	E	E
$C\overline{1s}$	I	M	K	T	M	Q	E	N

NOTE :

The numbering is that of chymotrypsinogen A[14], with insertions in the sequences of the other enzymes denoted as 36A, 36B, etc. The deletions are denoted by broken lines. Those residues that are chemically identical in all ten protein sequences are enclosed in solid lines. The residues that are doubly underlined are those that are topologically equivalent in the serine proteases[a]. The single-letter code for the amino-acids is used in this table[20]. PPKK, COLL, $C\overline{1r}$ and $C\overline{1s}$ were only aligned in sequences.

TABLE 1.2

PHYSICAL PROPERTIES OF PPE

Parameter	
$E^{1\%}_{1cm}$	20.2[26]
Isoelectric point	9.5[2]
$S_{20,w}$	2.6[2]
$D_{20,w}$	9.5[2]
Partial specific volume	0.726[26]
Frictional ratio	1.2[2]

1.5 THREE-DIMENSIONAL STRUCTURE OF PPE

Before any X-ray structure determination of PPE, Hartley[31] in 1970 built a hypothetical model of the enzyme, into the three-dimensional structure of chymotrypsin, determined in 1967[32]. This work was based on the assumption that the tertiary structures were very similar, since the covalent structure of PPE showed it to be homologous with trypsin and chymotrypsin.

However, shortly after this model, the three-dimensional structure of PPE was determined at 3.5 Å resolution from inhibited PPE rather than the native[33-35].

The X-ray diffraction data were collected on four isomorphous crystalline derivatives; p-toluene sulphonyl (tosyl) elastase, p-chloromercurisulphonyl elastase, in each case the active centre serine residue is blocked, and on two uranyl soaked crystals of these derivatives. The crystals crystallised were orthorhombic space group $P2_12_12_1$ with unit cell dimensions a = 51.5 Å, b = 58.0 Å, and c = 75.5 Å. This corresponded to a unit cell of four molecules, with one molecule per asymmetric unit[34].

They used tosyl elastase as the "parent" structure, rather than using the native PPE, in order to exploit high isomorphism in the structure determination, and provide an active site label. The three-dimensional structure resembled the hypothetical model of Hartley[31] and showed that the single polypeptide chain of PPE was folded into two distinct halves to form a compact globular molecule of dimensions 55 x 40 x 38 Å.

The X-ray structure showed the C-terminal part as an α-helix and the rest of the molecule was composed of six anti-parallel loops of β-sheet folded into two barrel-like domains. The two continuous chain domains, residues 27-127 and 128-230, were stabilised by one and three disulphide bridges respectively, and each had one of the catalytically essential residues, i.e., histidine-57 and serine-195[13,35]

Ghelis et al[36] investigated the role of the domains in the refolding of PPE. They were able to show that the fragment 126-245 corresponded to one of the two domains and was able to refold independently. Moreover, this fragment, prepared by limited proteolysis was reduced in strong denaturing solvent to give the theoretical titre of six free thiol (SH) groups. They were not successful in obtaining a intact fragment, 16-125, to study unfolding and folding of this domain but managed to identify the disulphide content consisting of two SH groups.

Shotton et al[35] observed that the organisation of the N-terminal half of the molecule resembled that of the C-terminal half of the molecule when traced to the middle of the chain. This similarity had been inferred from the hypothetical model of PPE. The structure of native PPE was determined at 3.5 Å resolution by difference electron density maps between tosyl and native PPE. The structures were found to be essentially similar except the orientation of side-chain of histidine-57.

Shotton et al[37] and Sawyer et al[38] have improved the structure of tosyl elastase to 2.5 Å resolution. At this resolution all the amino-acids could be identified and were found to be identical with the chemically determined sequence[a]. Also some twenty-five water molecules bound internally and a large number of external water molecules were identified. The internal water molecules play a major role in stabilizing the active conformation of the enzyme. The interior of trypsin and chymotrypsin showed identical positions for many of these water molecules. By comparison, with the bacterial serine proteases, show only two water molecules structurally equivalent to internal water molecules with pancreatic serine proteases[39].

THE SPECIFICITY OF ENZYME AND THE NATURE OF ITS SUBSTRATE/ INHIBITOR BINDING SITES

1.6 THE SPECIFICITY OF PANCREATIC SERINE PROTEASES

The proteolytic enzymes, trypsin, chymotrypsin and elastase being members of the serine protease family hydrolyse peptide and synthetic ester substrates by the acylenzyme mechanism (figure 1.1).

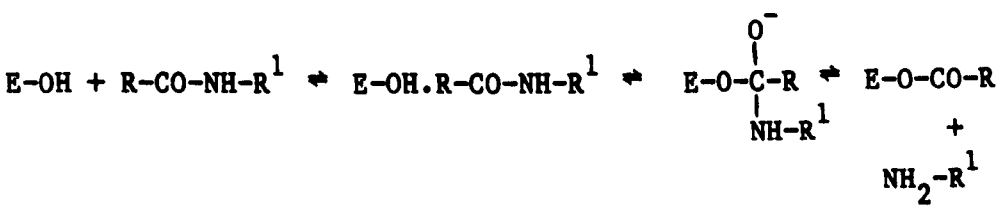


FIGURE 1.1

However there are major differences in the specificity of the substrates[40].

In the case of chymotrypsin, cleavage occurs on the C-terminal side of large hydrophobic side-chains, e.g. phenylalanine, tryptophan and tyrosine while with trypsin,

lysine and arginine are required. PPE requires small hydrophobic side-chains such as those of alanine. This major difference is due to the residues surrounding the binding pocket in the active site.

In chymotrypsin[41], the binding site is formed of a deep hydrophobic pocket involving glycine-216, glycine-226 and serine-189, which not only binds substrates but has the ability to orient them correctly for hydrolysis. In trypsin, the serine-189 in chymotrypsin is replaced by asparate and its negatively charged carboxylate forms a salt linkage with positively charged groups. That is, with the ammonium or guanidinium at the end of the side-chains of lysine or arginine. However, in the case of PPE both the flanking glycines are replaced by the bulky amino-acids valine-216 and threonine-226. This restricts the entry of any large side-chains of substrate/inhibitors into the pocket. Hence, only small side-chains are capable of efficient binding. The enzyme-substrate binding difference between chymotrypsin and PPE is shown schematically in figure 1.2[41].

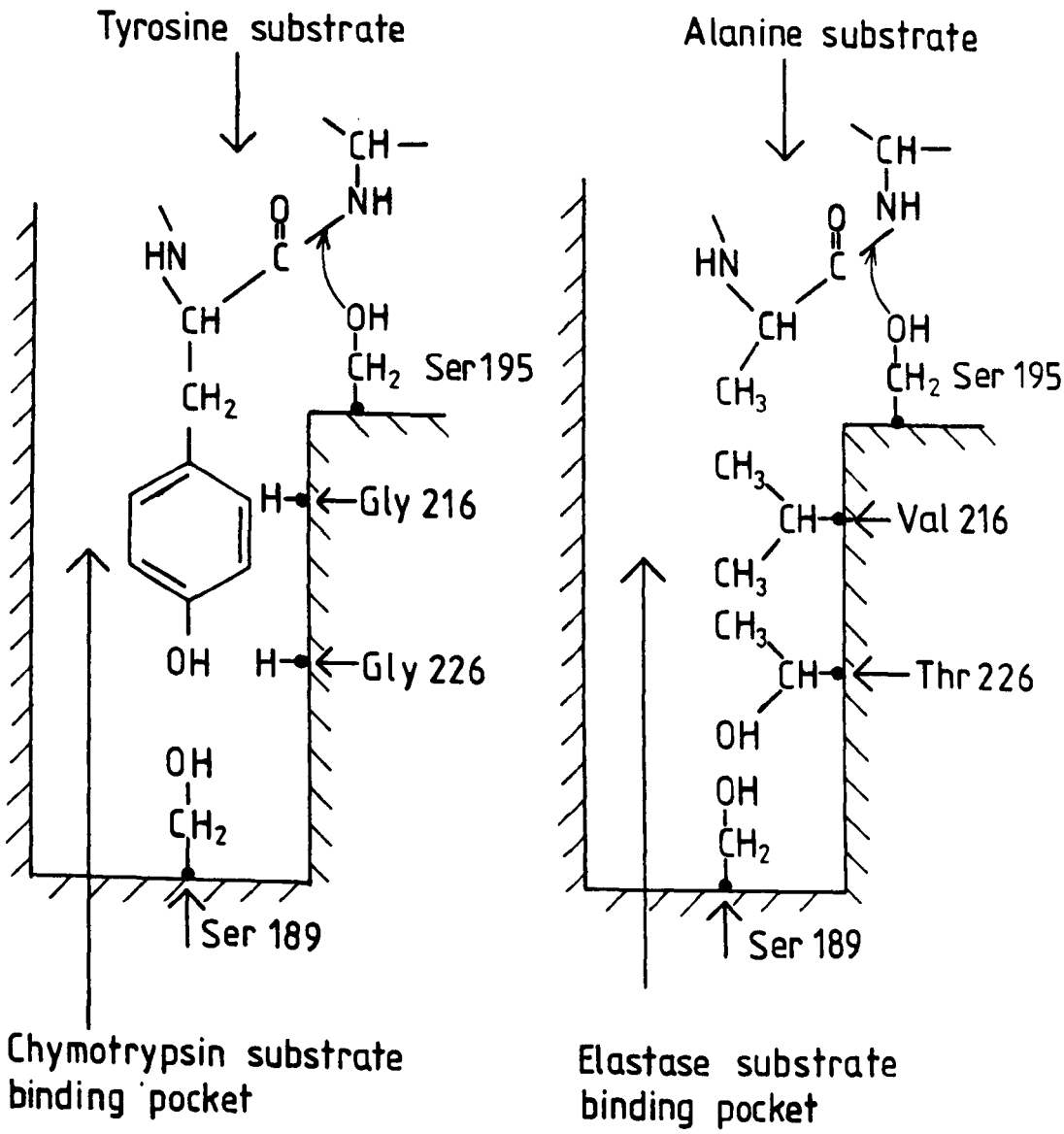
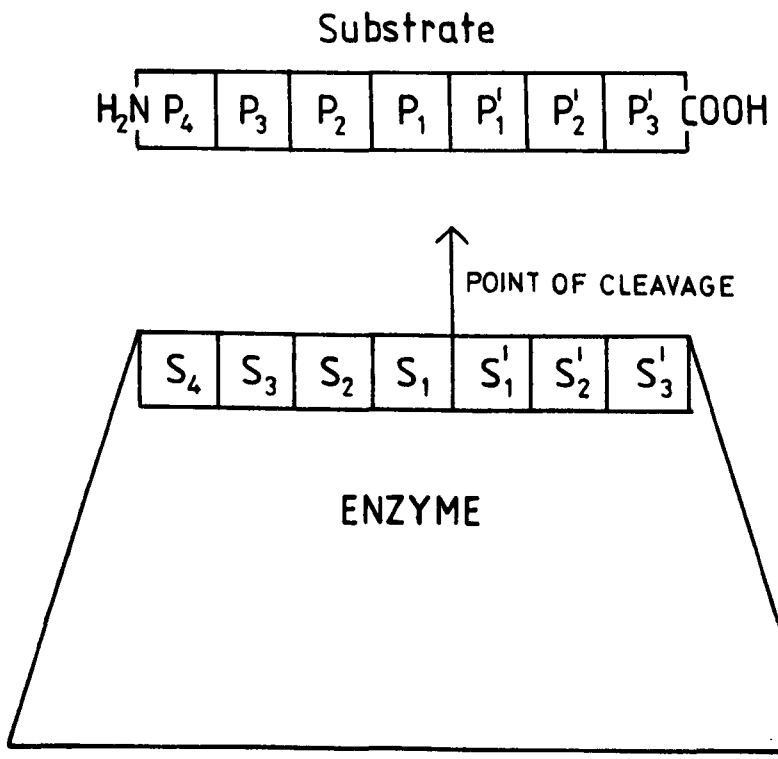

Tyrosine substrate
Alanine substrate
Ser195
Gly 216
Gly 226
Ser 189
Val 216
Thr 226
Ser 189
Chymotrypsin substrate binding pocket
Elastase substrate binding pocket

FIGURE 1.2

1.7 EVIDENCE THAT SUBSTRATE BINDING TO PPE REQUIRES AN EXTENDED BINDING SITE

The substrate binding site of PPE has been shown by Atlas and Berger[42,43] to be composed of at least seven subsites : five subsites towards the N-terminal end and at least two subsites towards the C-terminal end of the substrate. These subsites are usually referred to by the nomenclature introduced by Schecter and Berger[44]. The active site is divided into two categories : those subsites that are occupied by the segment of the substrate which is towards the N-terminus from the scissile bond are donoted by S_1, S_2 and S_3 etc., whereas those extending towards the C-terminus are S_1', S_2' and S_3'.



Schematic representation of enzyme-substrate complex

FIGURE 1·3

Also, studies on the binding sites have been carried out by several groups and have concluded that firstly, at least seven subsites are involved in the binding of substrate. Secondly, the catalytic efficiency is increased considerably with the extension in the chain length of the substrates and inhibitors[45-47]. This is emphasised in table 1.3 showing substrates and inhibitors with different leaving groups.

TABLE 1.3

INFLUENCE OF CHAIN LENGTH OF SUBSTRATE AND INHIBITORS ON PPE.

Peptide	$kcat/K_m$ (M^{-1} Sec^{-1})	k_2/K_i^3 (M^{-1} Sec^{-1})	K_i (mM)
ESTER SUBSTRATES[46,a]			
Ac-Ala-OMe	4.9×10^1		
Ac-Ala-Ala-OMe	2.2×10^3		
Ac-Ala-Ala-Ala-OMe	3.0×10^5		
Ac-Ala-Ala-Ala-Ala-OMe	1.8×10^6		
AMIDE SUBSTRATES AND INHIBITORS[46,a]			
$Ac\text{-}Ala\text{-}NH_2$			1.5×10^2
$Ac\text{-}Ala\text{-}Ala\text{-}NH_2$			5.0×10^1
$Ac\text{-}Ala\text{-}Ala\text{-}Ala\text{-}NH_2$	1.5×10^1		3.0
$Ac\text{-}Ala\text{-}Ala\text{-}Ala\text{-}Ala\text{-}NH_2$	2.5×10^3		
PEPTIDE CHLOROMETHYLKETONE (CMK)[47,b]			
Ac-Pro-Ala-CMK		5.0×10^{-2}	
Ac-Pro-Ala-Pro-CMK		3.5×10^1	
Ac-Pro-Ala-Pro-Ala-CMK		2.6×10^2	

[a] pH 9.0 and 37° C

[b] pH 6.5 and 25° C

This evidence has also been supported by the crystallographic studies of Shotton et al[48] on the inhibitors Ac-Ala_3, Ala_3 and Ac-Pro-Ala-Pro-Ala.OH. These inhibitors were all demonstrated to bind at the major site of PPE with full occupancy at subsites, i.e., S_{321}, S_{432}, and S_{54321}. (The abbreviation suggested by Thompson et al[49] will be used for a binding mode of a peptide to the subsites for instance to S_4, S_3, S_2 and S_1 by the symbol S_{4321}). Also, Hughes et al[50] and Renaud et al[51] have reported extended binding sites towards the C-terminal end. Nuclear Magnetic Resonance (N.M.R.) studies by Dimicoli et al[52] on oligopeptide inhibitor have added further evidence.

1.8 THE SPECIFICITY OF PPE

The specificity of PPE was first studied by Naughton and Sanger[22] towards the A and B chains of oxidized insulin. They reported that PPE cleaved peptide bonds adjacent to neutral aliphatic amino-acids. Later studies by Sampath-Narayanan et al[53] observed a narrower specificity toward peptide bonds. It was shown that the specificity was mainly directed toward the bonds involving alanine-8 and serine-12 in the A chain and alanine-14, valine-18 and glycine-23 in the B chain.

The specificity of PPE has also been studied using synthetic substrates[54,55]. Bender et al[54] reported their results, table 1.4 using p-nitrophenyl esters of benzyloxycarbonyl-L-amino-acids. From table 1.4, they found the preferred residue was alanine whereas proline or tyrosine esters were virtually unhydrolysed. These results have also been confirmed by others[56,57]. Hence it can be concluded that subsite S_1 is specific for alanine.

TABLE 1.4

THE SPECIFICITY OF SUBSITE S_1 in PPE

Substrate	kcat/K_m (M^{-1} Sec^{-1})
Z-Ala-ONp	1.8×10^5
Z-Gly-ONp	1.5×10^4
Z-Val-ONp	2.2×10^3
Z-Leu-ONp	2.9×10^4
Z-Ile-ONp	2.9×10^2
Z-Pro-ONp	2.0×10^1
Z-Tyr-ONp	8.0×10^1

Z = Benzyloxycarbonyl

ONp = p-Nitrophenyl

This subsite has also been shown to be rather stereo-specific, when an oligopeptide amide with a D-alanine instead of L-alanine at P_1 was not hydrolysed by PPE[46], that is, Ac-Pro-Ala-Pro-D-Ala-NH_2 K_i = 48mM while Ac-Pro-Ala-Pro-Ala-NH_2 K_m = 3.9mM. Shotton et al[49] have also concluded that to facilitate acylation, the methyl group of P_1 alanine must be sandwiched between the side-chains of valine-216 and glutamine-192. This enables the P_1-P_1' scissile bond to orient favourably relative to the position of serine-195, facilitating acylation.

However, subsite S_1 shows a different preference in the case of trifluoroacetyl (TFA) peptides, which have shown to be very potent PPE inhibitors. N.M.R. studies and kinetic measurement[52,58-60] demonstrated that the TFA group forces the TFA peptides to bind in a special mode to the active center of the enzyme.

The site of binding of the TFA group to PPE in solution is unique and is in the vicinity of the subsite S_1 of the enzyme[61]. X-ray study by Hughes et al[50] of a crystallised complex of PPE with CF_3CO-Lys-Ala-NH-Ph-p-CF_3 (Ph = Phenyl) showed the CF_3CO group bound at S_1 and the rest of the peptide forming a parallel sheet with residues 214-216.

Subsite S_2 is able to bind proline or alanine equally effectively as the P_2 residue in substrates or inhibitors, table 1.5[62,63]. However, subsite S_2 is found to depend on the residue at subsite S_4, since this subsite, S_4, is also important for the binding process. Therefore, when alanine in P_2 was replaced by proline, in the tetrapeptide substrates and inhibitors, $kcat/K_m$ and kobsd/[I] increased by a factor of 2 and 4[62-64], table 1.5. Hence, the subsites S_2 and S_4 bind a proline residue satisfactorily. Proline in S_4 is shown to bind efficiently at these subsites[48] and also the proline side-chain forms a good hydrophobic contact with the side-chain of leucine-151 and its carbonyl oxygen is able to form a hydrogen bond with the $N^{\delta 1}$ nitrogen atom of histidine-40 or with the main chain nitrogen atom of threonine-41[48].

TABLE 1.5

EFFECT OF ALANINE OR PROLINE RESIDUES IN POSITION P_2 AND THE EFFECT OF SUBSTITUTION OF ALANINE IN P_2 RESIDUE IN TETRAPEPTIDE SUBSTRATE OR INHIBITOR FOR PROLINE FOR PPE

Peptide P_4 P_3 P_2 P_1	K_m (mM)	K_i (mM)	kobsd/[I] ($M^{-1}sec^{-1}$)	kcat/K_m ($M^{-1}sec^{-1}$)
Ac-Ala-Ala-Ala-Ala-NH_2 [62,c]	2.9			
Ac-Ala-Ala-Pro-Ala-NH_2 [62,c]	2.1			
Ac-Pro-Ala-Ala-Ala-NH_2 [62,c]	3.9		9.8	
Ac-Pro-Ala-Pro-Ala-NH_2 [63,c]	3.9		38.0	
Ac-Ala-Ala-Ala-CMK [64,d]		1.5		
Ac-Ala-Pro-Ala-CMK [64,d]		1.2		
Ac-Ala-Ala-Ala-Ala-CMK [64,e]				1.2×10^3
Ac-Ala-Ala-Pro-Ala-CMK [64,e]				2.2×10^3

[c] pH 9.0 and 37° C
[d] pH 6.5 and 30° C
[e] pH 5.0 and 30° C

Subsite S_3 has been shown to bind only alanine. When proline is the P_3 residue no binding occurs. However their analogues with alanine show reactivity.

Renaud et al[51] have investigated the specificity of subsite S_1' and S_2' using trifluoroacetyl dipeptide anilides. They carried out mapping of these subsites for PPE, since Hughes et al[50] had already studied, its complex with CF_3CO-Lys-Ala-NH-Ph-p-CF_3 at a resolution of 2.5 Å.

Subsite S_1' is found to prefer large P_1' residues compared to the other subsites, (table 1.6). They observed that substitution of alanine by lysine in P_1' leads to a 6-fold

decrease of the Michaelis constant. This is also in agreement with their previous result shown by enzymatic and N.M.R. studies[61].

TABLE 1.6

SPECIFICITY OF SUBSITES S_1' AND S_2'

Peptide P_1 P_1' P_2'	K_i (M)
CF_3CO-Ala-Ala-NH-Ph-p-CF_3	9.0×10^{-8}
CF_3CO-Lys-Ala-NH-Ph-p-CF_3	2.5×10^{-8}
CF_3CO-Lys-Ala-NH-Ph-p-$CH(CH_3)_2$	5.0×10^{-8}
CF_3CO-Lys-Phe-NH-Ph-p-$CH(CH_3)_2$	4.4×10^{-7}

Renaud et al[51] explained the high affinity of the Lys-containing peptide on the basis of the examination of the atomic coordinates obtained by Hughes et al[50] with CF_3CO-Lys-Ala-NH-Ph-p-CF_3.PPE complex. It was due to a favourable van der waals interaction of the δ-CH_2 group of lysine with the γ-CH_3 group of valine-99 which may not exist or be weaker in alanine containing peptides.

Subsite S_2', shows a marked specificity for alanine based on results of Hughes et al[50] and Renaud et al[51]. From table 1.6, it was found that substitution of alanine by phenylalanine for the inhibitor caused a 10-fold increase in K_i. X-ray results of CF_3CO-Lys-Ala-NH-Ph-p-CF_3 showed that C_β of the alanine was very far from the centre of phenylalanine-215 (9.4 Å) and was directed away from the ring suggesting that a strong ring interaction between phenylalanine-215 and the phenylalanine residue of the peptide CF_3CO-Lys-Phe-NH-Ph-p-$CH(CH_3)_2$ was most unlikely. Furthermore C_δ of lysine in S_1' appeared to interact with the side-chain of valine-99 rather than the side-chains of S_2'.

Hence the overall conclusion for the preference of subsite is :- S_1 has preference for two different residues, that is, alanine in the case of acetyl peptides or for TFA group in the case of TFA peptides; S_2 and S_4 for proline; S_3 for alanine; S'_1 for lysine and S'_2 for alanine.

1.9 X-RAY DIFFRACTION STUDIES ON THE BINDING OF SUBSTRATE/ INHIBITORS

PPE might be expected to exhibit similar substrate binding sites to those of the other members of the family. However, low resolution studies (3.5 Å) by Shotton et al[48] have shown that the binding sites for substrates are apparently different to those of chymotrypsin and trypsin[65,66].

Thus, there is an apparent confusion as to the exact mode of substrate binding in PPE between Shotton and other workers. Further confusion has arisen from Hughes et al[50] involving binding of a trifluoroacetyl dipeptide inhibitor with PPE.

The mode of binding will be referred to the subsites assigned to trypsin, chymotrypsin and SGPA[39]. X-ray studies on bacterial serine proteases have shown that in the immediate vicinity of the active site, there is strong homology with the pancreatic enzymes[67]. The binding sites in SGPA which form anti-parallel β-sheet contacts between the peptide group and the main chain residues 214-216 have been referred to P_1, P_2 and P_3 i.e., to the N-terminal side of the scissile bond (figure 1.4). Similar β-sheet binding arrangements are found in complexes of chymotrypsin and trypsin. X-ray diffraction studies on various peptides have shown the C-terminal side of the scissile bond in SGPA was found close to the active site i.e., P'_1, P'_2 etc.

Shotton et al[48] indicated the binding mode of di, tri and tetrapeptides substrate were towards the subsite P_1 in SGPA

ACTIVE-SITE OF PPE WITH RESIDUES 214-216
IN ORANGE

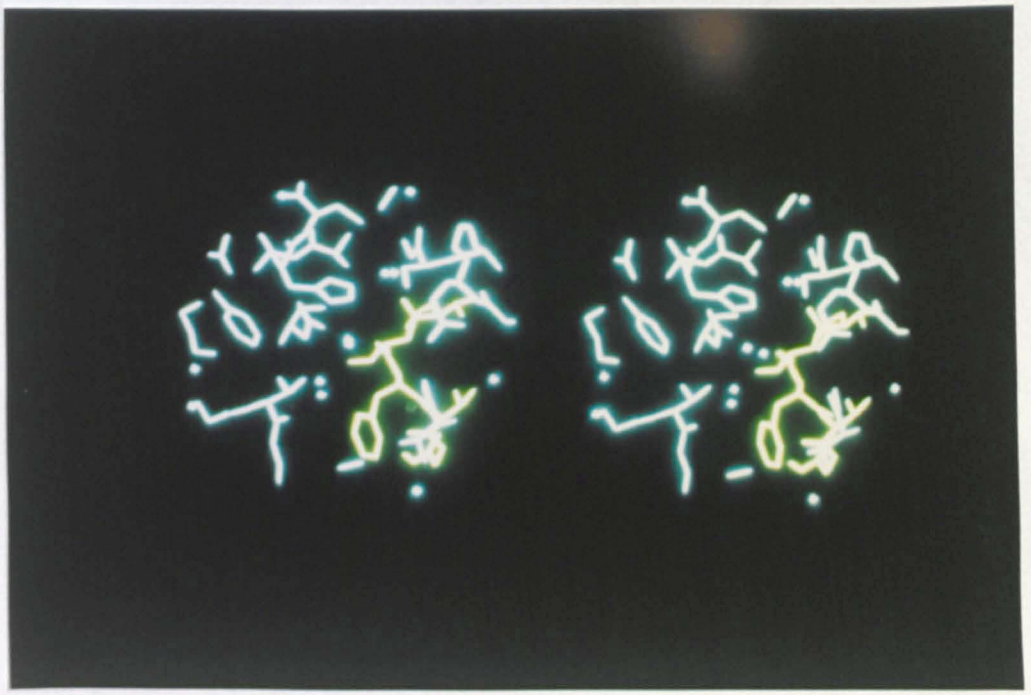


FIGURE 1.4

where as has been said, the substrate binding in the other enzymes appeared to be associated with the enzyme residue 214, 215 and 216. Also, the substrate residues for PPE were oriented very differently in the direction of the main chain of threonine-41 and glutamine-192.

However, a similar mode of binding to SGPA was observed by Hughes et al[50] (CF_3CO-Lys-Ala-NH-Ph-p-CF_3) but resulted in further confusion. Firstly, the result obtained contrast sharply with those reported by Shotton et al[48]. Secondly, it was found that the TFA group was found at S_1 compared to similar group, N-acetyl bound at S_4. Thirdly, Hughes et al[50] observed a similar mode of binding in the same vicinity as SGPA but the peptide chain portion was associated in a parallel pleated-sheet fashion with the protein chain (he describes these subsites as S'_1, S'_2 etc).

Further evidence for the TFA binding has been obtained from ^{19}F N.M.R. spectra of TFA-oligopeptide and TFA-oligopeptide anilide inhibitor molecules (Dimicoli et al 1980)[61]. Also by enzymatic and N.M.R. studies (Dimicoli et al 1976)[52].

The binding sites observed by Sawyer[67] (unpublished reported by Johnson et al) on X-ray diffraction studies at 3.5 Å resolution of two inhibitors were as that observed by Hughes et al[50]. The binding mode of these inhibitors was anti-parallel along the protein chain in the vicinity of residues 214-216 with the C-terminal group between phenylalanine-215 and glutamine-192 and the N-terminal groups in the region between tryptophan-172 and threonine-179 (binding of the inhibitor at subsites S_1, S_2 and S_3). This is in the same manner with SGPA, chymotrypsin etc.

Thus, there is still uncertainty as to the exact physical mode of binding of substrate/inhibitors to PPE. In order to map out conclusively the substrate binding site (both to the

N-terminal and C-terminal side of the scissile bond i.e., P_1-P_1' sites) evidence from a variety of techniques is required.

1.10 TECHNICAL APPROACHES TO THE INVESTIGATION OF BINDING SITES

In order to describe the catalytic mechanism of PPE in detail, it is necessary to correlate evidence from a variety of techniques. Thus, kinetic data from solution studies with a variety of substrates and inhibitors need to be correlated with structural data showing detailed binding of these molecules to the enzyme. The main structural technique available is X-ray crystallography, but being a time average technique does not allow one to examine the catalytic event at regular intervals throughout the turnover.

Several possible strategies are available which approximate the ideal experiment. One such is to cool the crystal down to low temperatures where the turnover is essentially stopped completely and by careful control of temperature, to allow the reaction to proceed by discrete steps until the product is formed. Various workers[69-71] are thought to be investigating the structure of PPE at high resolution under a variety of conditions of temperature and pH but no results have been published.

An alternative approach can be used where substrate analogues for the enzyme allow room temperature mapping of the binding site and this approach has already yielded excellent results with many enzymes. These inhibitors can either be naturally occurring or synthetically produced. The naturally occurring inhibitors are classified into two categories i.e., inhibitors from animal tissues, turkey ovomucoid and chicken ovoinhibitor[72] and the serum inhibitors, α-1-antitrypsin[73] and α_2-macroglobulin[74]. Also small size natural inhibitors, such as elastinal[75,76], have been investigated.

Turkey ovomucoid and chicken ovoinhibitor (molecular weight of 28,000[77] and 46,500[78]) inhibit PPE with a 1:1 stoichiometry. They both inhibit the esterolytic, proteolytic and elastolytic activities of PPE[72]. The complexes formed are stable and the inhibitor totally inhibits the activity of PPE.

Gertler et al[72] have shown that chicken ovoinhibitor has three independent sites for trypsin, chymotrypsin and PPE, while turkey ovomucoid has only two binding sites, one for trypsin and the other for PPE and chymotrypsin.

The other category, the serum inhibitors like α-1-antitrypsin and α_2-macroglobulin are both glycoproteins[73] with molecular weight of 52,000 and 725,000[79]. The major difference between them is that whilst the α-1-antitrypsin.PPE complex is inactive against elastin[80] and synthetic substrates/inhibitors[81], the α_2-macroglobulin.PPE complex[82] still shows 80-100% of its activity with small synthetic substrates/inhibitors although it is totally inactive against elastin. It can be assumed that α-1-antitrypsin binds at the active site while α_2-macroglobulin binds somewhere else. The esterolytic[81] and elastolytic[80,82] activities of PPE are inhibited by α-1-antitrypsin compared to α_2-macroglobulin which inhibits only the esterolytic activity[81]. Meyer et al[83] have studied the properties of α-1-antitrypsin and α_2-macroglobulin and have shown that PPE binds preferentially to α_2-macroglobulin. Also the dissociation constant between α-1-antitrypsin.PPE complex is much less than that of α_2-macroglobulin.PPE complex.

Elastinal, N[(S)-1-carboxy-isopentyl]-carbamoyl-(2-imino-hexahydro-4(S)-pyrimidyl]-(S)-glycyl-(S)-glutaminyl-(S)-alanal[75,76] was found to be, a specific inhibitor of PPE. It exists in various species of actinomycete and is found in the culture filtrate of Streptomyces griseoruber.

Elastinal shows biological activity only towards PPE and not with the other serine proteases. Its inhibition is competitive with the substrate. K_i is 0.24 μM with Ac-Ala-Ala-Ala-NH-Ph-NO_2 and 0.2 μM with Ac-Ala-Ala-Ala-OMe.

Synthetically produced substrate analogues have already been applied with great success by James et al[39], Stroud et al[84] and at lower resolution by Henderson et al[85] and Davies et al[86]. This approach was pursued in this work, since there was controversy as to the exact mode of binding of substrate/inhibitor in PPE as mentioned in section 1.9.

Another possible way to determine the binding sites for substrates (both to N-terminal or C-terminal of the substrate's scissile bond), is to modify the enzyme active site thus allowing binding of a complete peptide. The active site serine-195 is first chemically inhibited followed by an elimination reaction to convert $-CH_2OH$ into $=CH_2$. This totally inactivates PPE. Low resolution studies (3.5 Å) on the modified enzyme[87], anhydroelastase, have shown some changes in the active site. Similar to those with anhydrotrypsin and anhydrochymotrypsin, chemically and also by X-ray studies[88-91].

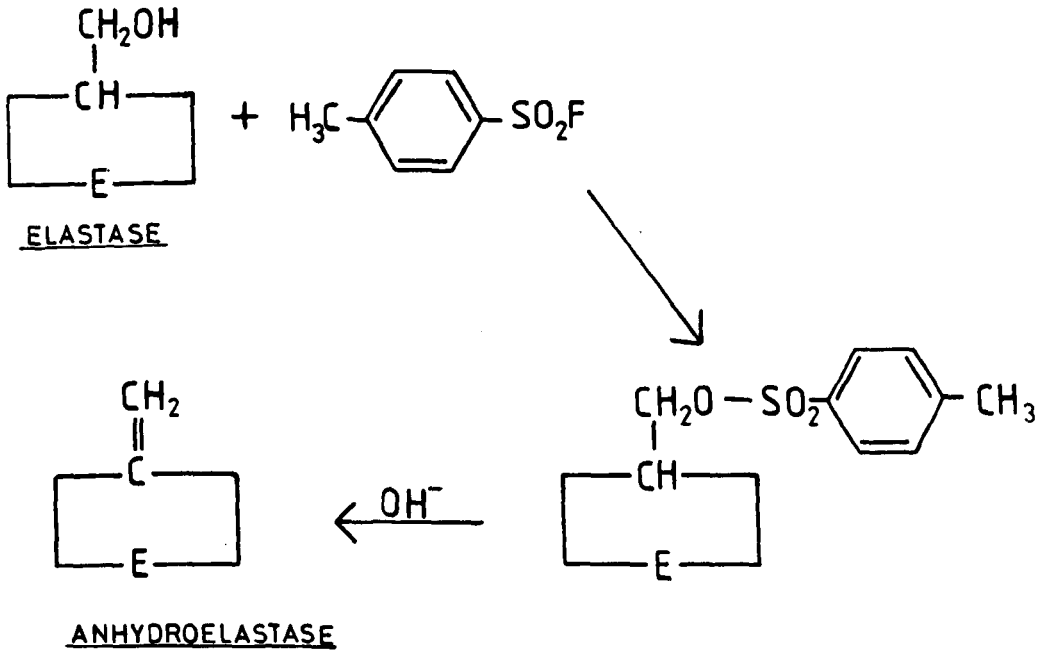
The refined structure of anhydroelastase and that of PPE modified by diffusion of substrate/inhibitors into the crystals would distinguish the binding sites on both sides of the scissile bond, thus establishing the required orientation for peptide cleavage. Using the well established techniques of protein crystallography[92] it should be able to resolve the controversy between various workers about the binding sites of PPE.

1.11 STUDIES TO BE CARRIED OUT TO ESTABLISH THE BINDING MODE IN PPE USING SYNTHETIC SUBSTRATE/INHIBITORS

Various studies to block the active site serine-195 will be carried out firstly, followed by the elimination to convert $-CH_2OH$ into $=CH_2$. This totally inactivates the enzyme and hence the whole substrate should be able to be bound intact to the active site.

Previous studies with the serine pancreatic proteases[87-91] where the enzyme was inhibited using radioactive labelled inhibitor i.e., ^{14}C labelled tosyl fluoride or phenylmethane sulphonyl fluoride (PMSF) (the carbon atom labelled is attached to the sulphonyl group) have been used in the preparation of anhydroenzyme.

The preparation of anhydroenzyme is a two stage process as shown in figure 1.5. The first stage can be performed using non-radioactive label inhibitor, because the per cent inhibition will give an indication of the inhibition measured using a substrate. However, the second stage is crucial since the base elimination of the inhibitor is difficult to detect. Moreover, since the anhydroenzyme formed is inactive there is no certainty of the inhibitor has been removed.



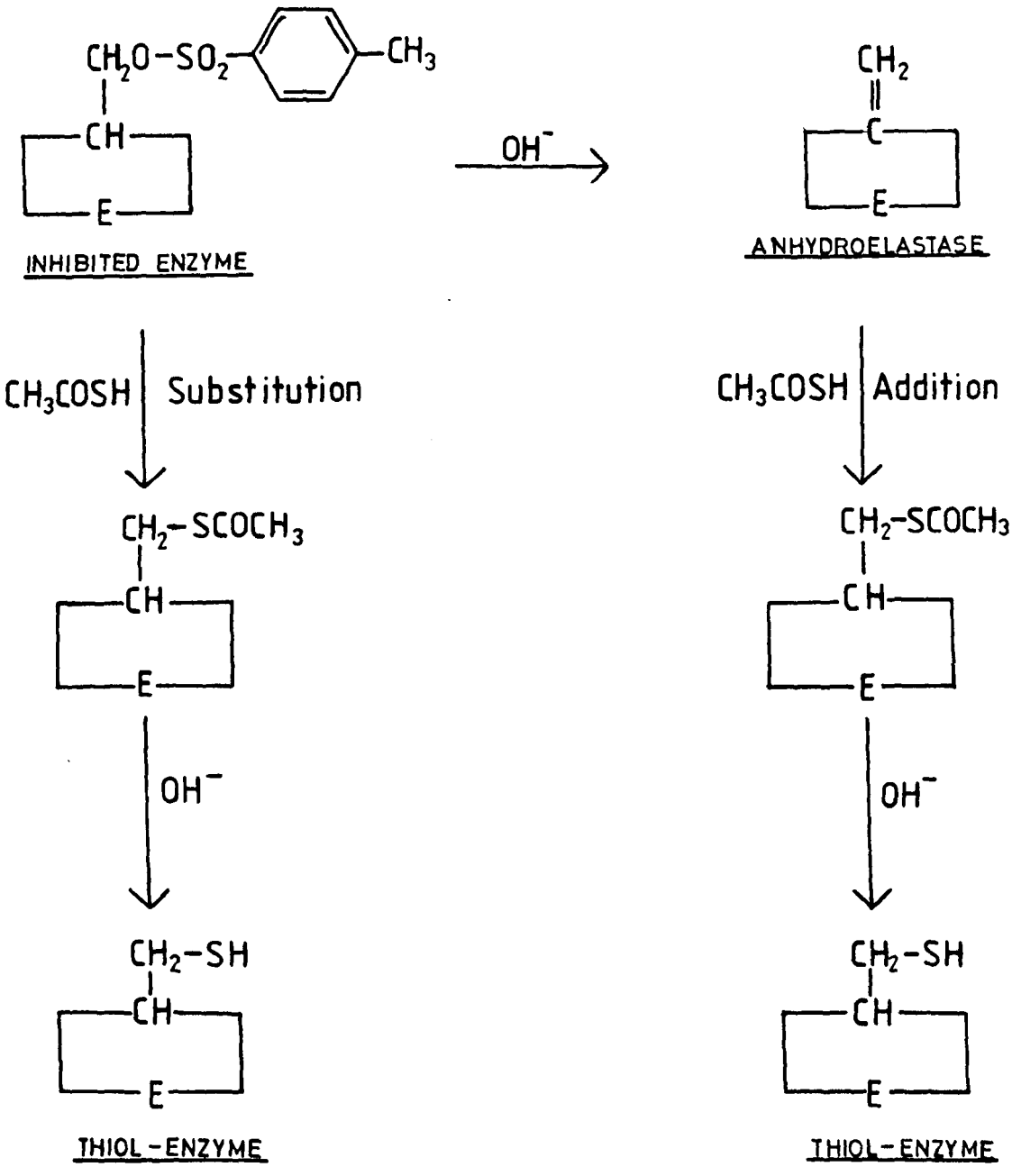
A schematic representation of reaction steps

FIGURE 1·5

However, with the present work, non-radioactive tosyl fluoride and PMSF inhibitor were used for several reasons. Firstly the procedure was more straight forward as the optimum condition for the preparation of anhydroelastase had already been found by Murphy[87]. Secondly, the major problem was ^{14}C labelled inhibitor was not available on the market and therefore required special preparation by the supplier, which meant the inhibitor was prohibitively expensive. Lastly, there were local restrictions on non-qualified persons using labelled compounds.

Thus, during the preliminary studies potential fluorescent probes, 1-dimethylamino-naphthalene-5-sulphonyl halide (Dansyl fluoride or chloride) and p-Nitrophenyl anthranilate (NPA) were also examined, since they had been found to inhibit α-chymotrypsin[93,94]. These could be detected by fluorescent emission at 505 and 422 mμ when bound to the enzyme. On base elimination no emission spectra should be obtained.

The other modification attempted was to convert serine-195 to cysteine, either by substitution of the inhibited PPE or Michael type addition to the double bond in anhydroelastase with thiolacetate (this latter giving another means of detecting the presence of dehydroalanine). The reaction steps are summarised in figure 1.6.



A schematic representation of reaction steps involved in preparation of Thiol-Enzyme

FIGURE 1.6

This is of interest, since thiol proteases are analogous to the serine proteases. The thiol proteases are another widely distributed group of plant enzymes, (papin, actinidin, ficin and bromelain) which contain an active site cysteine. This plays a role analogous to that of serine-195 in the serine proteases. The conversion from OH to SH in PPE may provide completely different insight into the reactivity of PPE.

Once the anhydroelastase is prepared and crystallised, the crystals will be soaked with the hexapeptide substrate (according to section 1.8) i.e., H.Pro-Ala-Pro-Ala-Lys -Phe.OH. The substrate will be examined to reveal the binding sites to the carboxyl side of the scissile bond i.e., P_1', P_2' etc. Also a number of tetrapeptide inhibitors Ac-Pro-Ala-Pro-Ala.OH, Ac-Pro-Ala-Pro-Alaninal (peptide aldehyde) and TFA-Pro-Ala-Pro-Alaninal will be investigated. The product inhibitor Ac-Pro-Ala-Pro-Ala.OH will be examined for the anomalous binding since previous work (Sawyer unpublished) has failed to reproduce the reported results of Shotton et al[48]. The purpose of peptide aldehydes is firstly to distinguish the different mode of binding between acetyl and TFA group. Since, it has been suggested that the TFA group binds to PPE in a unique binding mode[50,52]. Secondly, they bind to PPE more tightly than the substrates (table 1.7). This is probably due to the formation of a transition state analogue, a hemiacetal[95], which is like the tetrahedral intermediate in figure 1.7.

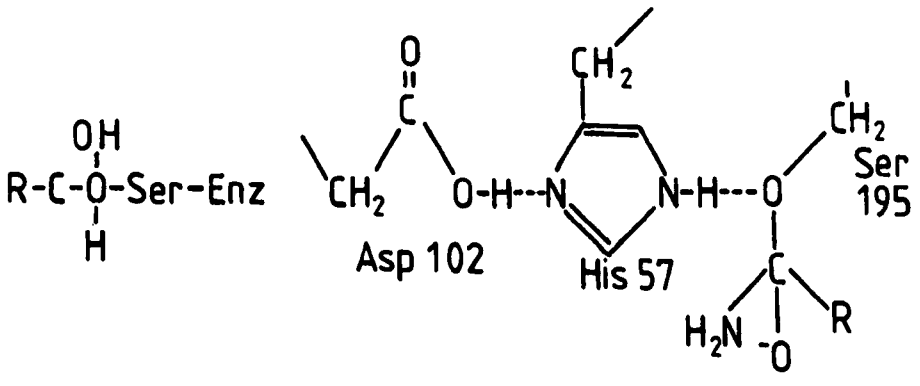
TABLE 1.7

KINETIC CONSTANTS FOR PPE CATALYSED HYDROLYSIS OF PEPTIDE AMIDES AND PPE BINDING TO ALDEHYDES

Peptide	K_i (mM)	K_m (mM)
Ac-Ala-Pro-Ala-NH_2 [64,f]		4.2
Ac-Ala-Pro-Alaninal [95,g]	0.062	
Ac-Pro-Ala-Pro-Ala-NH_2 [64,f]		3.9
Ac-Pro-Ala-Pro-Alaninal [95,h]	0.002	

[f] pH 9.0 and 37° C
[g] pH 7.0
[h] pH 4.0

The inhibitor TFA-Ala_3-CMK prepared by Bieth et al[59] will be examined for anomalous binding at 2.5 Å resolution. (Data collected by Dr Sawyer). Beith et al[59] have investigated this inhibitor using ^{19}F and ^{1}H N.M.R. spectroscopy. They have shown the TFA-peptide binds directly to the active centre of PPE which is irreversibly inhibited.



Hemiacetal compared to enzyme catalysis intermediate

FIGURE 1·7

REFERENCES

1. Stroud, R.M.
Scientific American, 231, No.1, 77 (1974).

2. Kraut, J.
Ann.Rev.of Biochemistry, 96, 331 (1977).

3. Narahashi, Y. and Funkunaga, J.
J.Biochem., 66, 743 (1969).

4. Gertler, A., and Trop, M.
Eur.J.Biochem., 19, 90 (1971).

5. Johnson, P., and Smillie, L.B.
Can.J.Biochem., 49, 548 (1971).

6. Bode, W., Chen, Z., Bartels, C.K., Schmidt-Kastner, G., and Bartunik, H.
J.Molec.Biol., 164, 237 (1983).

7. Chen, Z., and Bode, W.
Ibid., 164, 283 (1983).

8. Kraut, H., Frey, E.K., and Werk, E.
Hoppe-Seyler's Z-physiol.Chem., 189, 97 (1930).

9. Balo, J., and Banga, I.
Nature, 164, 491 (1949).

10. Lewis, U.J., Williams, D.E, and Brink., N.G.
J,Biol.Chem., 222, 705 (1956).

11. Hartley, B.S.
Ann.Rev.of Biochemistry, 29, 45 (1960).

12. Gertler, A., and Birk, Y.
Eur.J.Biochem., 12, 170 (1970).

13. Shotton, D.M., and Hartley, B.S.
Nature, 225, 806 (1970).

14. Hartley, B.S., and Kauffman, D.L.
Biochem.J., 101, 229 (1966).

15. Smillie, L.B., Furka, A., Nagabhushan, N., Stevenson, K.J., and Parkes, C.O.
Nature, 218, 343 (1968).

16. James, M.N.G., Delbaere, L.T.J., and Brayer, G.D.
Can.J.Biochem., 56, 396 (1978).

17. Grant, G.A., Henderson, K.O., Eisen, A.Z., and Bradshaw, R.
Bicohemistry, 19, 4658 (1980).

18. Ariaud, G.J., and Gagnon, J.
Ibid., Wash., 22, 1758 (1983).

19. Carter, P.E., Dunbar, B., and Fothergill, J.E.
Biochem.J., 215, 565 (1983).

20. Ibid., 113, 1 (1969).

21. Partridge, S.M., and Davis, H.F.
Ibid., 61, 21 (1955).

22. Naughton, M.A., and Sanger, F.
Ibid., 78, 156 (1961).

23. Naughton, M.A., Sanger, F., Hartley, B.S., and Shaw, D.C.
Ibid., 77, 149 (1960).

24. Hartley, B.S., Naughton, M.A., and Sanger, F.
Biochim.et.Biophys.Acta., 34, 243 (1959).

25. Smillie, L.M., and Hartley, B.S.
J.Molec.Biol., 10, 183 (1964).

26. Shotton, D.M.
Methods in Enzymology., 19, 113 (1970).

27. Smillie, L.M., and Hartley, B.S.
Biochem.J., 101, 232 (1966).

28. Shotton, D.M.
Ph.D. Thesis., University of Cambridge (1969).

29. Gertler, A., and Hofmann, T.
J.Biol.Chem., 242, 2522 (1967).

30. Wasi, S., and Hofmann, T.
Biochem.J., 106, 926 (1968).

31. Hartley, B.S.
Phil.Trans.Roy.Soc.,Lond.B., 257, 77 (1970).

32. Matthew, B.W., Sigler, P.M., Henderson, R., and Blow, D.M.
Nature, 214, 652 (1967).

33. Shotton, D.M., and Watson, H.C.
Phil.Tran.Roy.Soc.,Lond.B., 257, 111 (1970).

34. Watson, H.C., Shotton, D.M., Cox, J.C., and Muirhead, H.
Nature, 225, 806 (1970).

35. Shotton, D.M., and Watson, H.C.
Ibid., 225, 811 (1970).

36. Ghelis, C., Gaillourdet, T.M., and Yon, J.M.
Biochim.Biophys.Res.Commu., 84, 31 (1978)

37. Shotton, D.M., Sawyer, L.S., and Watson, H.C.
Ibid., 53, 944 (1973)

38. Sawyer, L.S., Shotton, D.M., Campbell, J.W., Wendell, P.L., Muirhead, H., and Watson, H.C.
J.Molec.Biol., 118, 137 (1978)

39. James, M.N.G., Sielecki, G.D., Brayer, L.T., Delbaere, L.T.J., and Bauer, C.A.
Ibid., 144, 43 (1980).

40. Fersht, A.
Enzyme Structure and Mechanism, Freeman (1977)

41. Bieth, J.
Front.Matrix.Biol., 6, 1 (1978).

42. Atlas, D., and Berger, A.
Biochemistry, 11, 4719 (1972)

43. Atlas, D., and Berger, A.
Ibid., 12, 2573 (1973)

44. Schecter, I., and Berger, A.
Biochim.Biophys.Res.Commu., 27, 157 (1967)

45. Atlas, D., Levit, S., Schecter, I., and Berger, A.
Febs.Lett., 11, 281 (1970).

46. Thompson, R.C., and Blout, E.R.
Proc.Natn.Acad.Sci., U.S.A., 67, 1734 (1970).

47. Thompson, R.C., and Blout, E.R.
Biochemistry, 12, 44 (1973).

48. Shotton, D.M., White, N.J., and Watson, H.C.
Cold.Spring.Harbor.Symp.Quant.Biol., 36, 91 (1971).

49. Thompson, R.C., and Blout, E.R.
Biochemistry, 12, 66 (1973).

50. Hughes, D.L., Sieker, L.C., Bieth, J., and Dimicoli, J.L.
J.Molec.Biol., 162, 645 (1982).

51. Renaud, A., Hughes, D.L., Liestienne, P., Bieth, J., and Dimicoli, J.L.
J.Biol.Chem., 258, 8312 (1983).

52. Dimicoli, J.L., Bieth, J., and Lhoste, J.M.
Biochemistry, 15, 2230 (1976).

53. Sampath-Narayanan, A., and Anwar, R.A.
Ibid., 114, 11 (1969).

54. Geneste, P., and Bender, M.L.
Proc.Natn.Acad.Sci., U.S.A., 64, 683 (1969).

55. Marshall, T., Whitaker, J., and Bender, M.L.
Biochemistry, 8, 4671 (1969).

56. Gold, R., and Shaltin, Y.
Biochim.et.Biophys.Acta., 410, 421 (1975).

57. Kaplan, H., Symonds, V.B., Dugas, H., Whitaker, D.R.
Can.J.Biochem., 48, 649 (1970).

58. Dimicoli, J.L., and Bieth, J.
Biochemistry, 16, 5532 (1977).

59. Dimicoli, J.L., Renaud, A., Lestienne, P., and Bieth, J.
J.Biol.Chem., 254, 5208 (1979).

60. Renaud, A., Dimicoli, J.L., Lestienne, P., and Bieth, J.
J.Amer.Chem.Soc., 100, 1005 (1977).

61. Dimicoli, J.L., Renaud, A., and Bieth, J.
Eur.J.Biochem., 107, 423 (1980).

62. Thompson, R.C., and Blout, E.R.
Biochemistry, 12, 51 (1973).

63. Thompson, R.C., and Blout, E.R.
Ibid., 12, 57 (1973).

64. Powers, J.C., and Tuhy, P.M.
Ibid., 12, 4767 (1973).

65. Steitz, T.A., Henderson, R., and Blow, D.M.
J.Molec.Biol., 46, 337 (1969).

66. Segal, D.M., Powers, J.C., Cohen, G.H., Davies, D.R., and Wilcox, P.E.
Biochemistry, 10, 3728 (1972).

67. Johnson, P., and Smillie, L.B.
Febs.Lett., 47, 1 (1974).

68. Hassall, C.H., Johnson, W.H., and Roberts, N.
Bio-Organic Chemistry, 8, 229 (1979).

69. Petsko, G.A., Alber, T., and Tsernogluo, D.
Nature, 263, 297 (1976).

70. Xuong, N.H., Cork, C., Hamlin, R., Howard, A., Katz, B., Kutter, P., and Nielsen, C.
Abstract XII International Congress of Crystallography, 02. 7.03 (1981).

71. Meyer, G., Jr. Presta, L., and Cole, G.
Ibid., 02.1.28 (1981).

72. Gertler, A., and Feinstein, G.
Eur.J.Biochem., 20, 547 (1971).

73. Cohen, A.B.
Biochim.et.Biophys.Acta., 391, 193 (1975).

74. Starkey, P.M., and Barrett, A.J.
Research Monographys in cell and tissue Physiology., 2, 661 (1977).

75. Umezawa, H., Aoyagi, T., Okura, A., Morishima, H., Takeuchi, T., and Okami, Y.
J.Antibiot., 26, 787 (1973).

76. Okura, A., Morishima, H., Takita, Y., Aoyagi, T., Takeuchi, T., and Umezawa, H.
Ibid., 28, 337 (1975).

77. Stevens, F.C., and Feeney, R.E.
Biochemistry, 2, 1346 (1963).

78. Tomimatsu, Y., Clary, J.J., and Bartulovich, J.J.
Arch.Biochem.Biophys., 115, 536 (1966).

79. Jones, J.M., Creeth, J.M., and Kekwick, R.A.
Biochem.J., 127, 187 (1972).

80. Baumstark, J.S.
Arch.Biochem.Biophys., 118, 619 (1967).

81. Bieth, J., Pichoir, M., and Metais, P.
Febs.Lett., 8, 319 (1970).

82. Bagdy, D., Faik, M., and Toinaly, P.
Acta.Physiol.Acad.Sci., Hung., 21, 123 (1962).

83. Meyer, J.F., Bieth, J., and Metais, P.
Clinica.Chimica.Acta., 62, 43 (1975).

84. Stroud, R.M., Kay, L.M., and Dickerson, R.E.
J.Molec.Biol., 83, 185 (1974).

85. Henderson, R.
Ibid., 54, 341 (1970).

86. Segal, D.M., Cohen, G.H., Davies, D.R., Powers, J.C., and Wilcox, P.E.
Cold Spring.Harbor.Symp.Quant.Biol., 36, 85 (1972).

87. Murphy, S.
Ph.D. Thesis, University of Sheffield (1978).

88. Huber, R., Bode, W., Kukla, D., Kohe, U., and Ryan, C.A.
Biophys.Struct.Mech., 1, 189 (1975).

89. Sayers, C.A., and Barrett., A.J.
Biochem.J., 189, 225 (1980).

90. Weiner, H., White, W.N., Hoare, D.G., and Koshland, D.E., Jr.
J.Amer.Chem.Soc., 88, 3851 (1966).

91. Strumeyer, D.H., White, W.N., and Koshland, D.E., Jr.
Biochemistry, 50, 931 (1963).

92. Blundell, T.L., Johnson, L.N.
Protein Crystallography, Academic Press, New York, (1976).

93. Horton, H.R., and Koshland, D.E., Jr.
Methods in Enzymology., Vol.XI., 856 (1967).

94. Haugland, R.P., and Stryer, L., Ramachandran, G.N.
"Conformation of Biopolymers", 321, (1967).

95. Thompson, R.C., and Blout, E.R.
Biochemistry, 12, 47 (1973).

CHAPTER TWO

PREPARATION OF PPE AND ANHYDROELASTASE

Anhydroelastase is prepared by converting the active site serine-195 of PPE into dehydroalanine by an elimination reaction. This is performed by inhibiting PPE with various inhibitors; tosyl fluoride, PMSF, NPA, and Dansyl fluoride or chloride, followed by a base elimination reaction involving treatment of the inhibited PPE with alkali. Also, modification of serine-195 to cysteine either by substitution of the inhibited PPE or addition to the double bond in anhydroelastase was attempted.

2.1 MATERIALS AND METHODS

PPE (EC 3.4.21.11) was purified from trypsin 1-300 (United States Biochemical Corporation, Cleveland, Ohio, U.S.A.) by the method of Shotton[1]. Thrice crystallised chymotrypsin (EC 3.4.21.1 activity 8895 A.T.E.E. units/mg) was purchased from Koch Light Ltd and trypsin (EC 3.4.21.4 activity 15,000 B.A.E.E. units/mg) from Sigma Chemical Ltd.

N-t-Butyloxycarbonyl-L-alaninate-p-nitrophenyl ester (NBA), N-α-Benzoyl-L-arginine ethyl ester (B.A.E.E.), N-Acetyl-L-tyrosine ethyl ester (A.T.E.E.), NPA, PMSF and Dansyl fluoride or chloride were all obtained from Sigma Chemicals Ltd. Tosyl fluoride was purchased from Aldrich Chemicals and potassium thiolacetate was obtained from Fluoro Chemicals Ltd. 5,5' -dithio-bis-(2-nitrobenzoic acid) (Ellman's reagent) was purchased from Koch Light Ltd. DEAE-Sephadex A-50 was purchased from Pharmacia Fine Chemicals. Buffers and reagents were prepared according to the tables published in Data of Biochemical Research[2]. A radio-meter, pH M 7010 with glass electrode was used for all pH measurements.

2.2 PROTEIN CONCENTRATION AND ENZYME ASSAY

Concentrations were determined by measuring absorbances at 280 nm using a Beckman Acta MIV Spectrophotometer. The

following standard absorbances were used : PPE, $E^{1\%}_{1cm}$ = 20.2, molecular weight 25,900; chymotrypsin, $E^{1\%}_{1cm}$ = 20.0, molecular weight 25,100; trypsin, $E^{1\%}_{1cm}$ = 15.9, molecular weight 23,900.

PPE was assayed using NBA at 347.5 nm by the method of Visser and Blout[3]. The standard assay conditions for PPE were as follows : An exact volume of 30 µl of 10^{-2}M stock solution of NBA in spectroquality methanol was added to 2.95 ml of 0.05M sodium phosphate buffer, pH 6.5 in quartz reference and sample cuvettes (1-cm pathlengths). The absorbances were balanced and after 3 minutes of incubation, 20 µl of PPE solution (1 mg/ml) was added. The increase in absorbance was followed for 2 to 3 minutes.

The assay conditions for trypsin were as follows[4]: To the quartz reference and sample cuvettes (1-cm pathlengths) 2.80 ml of 10^{-4}M stock solution of B.A.E.E. in 0.05M 2-amino-2-hydroxy-methyl-propane-1,3-diol (Tris) buffer, pH 8.0 containing 0.2 mg $CaCl_2$/ml was added to the reference and sample cuvettes and the absorbance balanced at 253 nm. After incubation for 3 minutes, 0.2 ml of 10^{-4} M hydrochloric acid (HCl) was added to the reference cuvette and to the sample cuvette, 0.2 ml of stock solution of trypsin (0.003 mg/ml) in 10^{-4} M HCl was then added and the increase in absorbance followed for 2 to 3 minutes.

To assay for trypsin contamination in PPE, similar assay conditions were used except 0.2 ml of stock solution of PPE (1 mg/ml) was added to the sample cuvette.

Similar assay conditions were used for chymotrypsin[4]. This time the substrate, A.T.E.E., dissolved in 0.05M sodium phosphate buffer pH 7.0 was used. In this case the decrease in absorbance was followed for 2 to 3 minutes at 237 nm.

2.3 PREPARATION AND MODIFICATION OF PPE

2.3.1 PURIFICATION OF PPE

All stages are conducted at +4° C. AnalaR reagents were used throughout.

Two different batches (100g and 200g) of trypsin 1-300 were used in separate purifications of PPE. The yields obtained were considerably lower (115mg for twice-crystallised and 120mg for thrice-crystallised) than those obtained by Shotton[1], who quotes 2.7g of thrice-crystallised material from a 500g batch. It was found at the crystallisation stage that cooling to 4° C was necessary for crystallisation. (In the case of the first batch the DEAE-Sephadex filtrate (freeze-dried) had to be put through stage 2[1] again since the enzyme would not crystallise with 0.1M sodium sulphate). The percentage contamination by trypsin and chymotrypsin was shown to be less than 0.02% for these enzymes.

2.3.2 PREPARATION OF TOSYL ELASTASE AND PHENYLMETHANE SULPHONYL ELASTASE (PMSE)

These were prepared by the method of Murphy[5]. Preliminary studies (once crystallised PPE was used, purchased from Worthington Biochemical Corporation, the activity was 20 times less than the fresh batch) revealed that the PMSF was found to be more reactive and PPE ≃ 98% inhibited within 4 hours. Tosyl fluoride was less reactive and ≃ 98% inhibition was achieved only after 24 hours. The measure of percentage inhibition was by assaying PPE with NBA.

METHOD

Once crystallised PPE (50.4mg, 1.96 μ mol) was dissolved in 19.5 ml of ice cold 0.05M sodium phosphate buffer, pH 7.2,

containing 0.5 ml of dioxane. The sulphonyl fluoride inhibitor (0.9mg, 5.17 μ mol) was dissolved in 0.5 ml dioxane and added over 1 minute. The solution was left stirring in the cold room for 24 hours. During this time the solution turned faintly cloudly. It was dialysed extensively against 1mM acetic acid and finally against 10mM sodium acetate buffer, pH 5.5. The solution was spun for half-an-hour at 21,00 rpm (Superspeed 50) and freeze-dried. The enzyme was ≃ 98% inhibited.

2.3.3 PREPARATION OF ANTHRANILOYL ELASTASE[6].

Anthraniloyl elastase was prepared by addition of 0.75 ml of 4.8mM NPA (4.86 μmol) in acetonitrile to 10 ml of PPE (51.8mg, 2.0 μmol) in 0.1M sodium phosphate buffer, pH 6.8 at 6° C. The NPA was added in five aliquots at one and a half-an-hour intervals. A yellow cream precipitate occurred which was dialysed extensively against distilled water, spun for half-an-hour at 21,000 rpm (Superspeed 50) to remove the precipitate and freeze-dried. The enzyme was ≃ 98% active.

2.3.4 PREPARATION OF DANSYL ELASTASE

The Horton et al[7] method was carried out for the dansylation of PPE at pH 8.4 using 0.03M sodium phosphate buffer. This method was also used at various pH's (5.5, 6.0, 6.5, 7.2 and 7.5) using 0.05M sodium phosphate buffer. A solution of PPE (32.1mg, 1.24μ mol) in 8 ml of 0.05M sodium phosphate buffer was cooled in ice and treated with 2.0 ml of Dansyl halide (5.1 μmol) in dioxane. When the reactions were carried out in alkaline conditions dansyl halide was dissolved in acetone to stop the protein denaturation. The solution was left stirring for 24 hours in the cold room, spun for half-an-hour at 21,000 rpm (Superspeed 50) before dialyising exhaustively for 2 days against 1mM acetic acid and freeze-dried. PPE was ≃ 20% inhibited at pH 8.4, while no inhibition occurred at pH's 5.5, 6.0, 6.5, 7.2 and 7.5.

Dansylation was also carried out on anhydroelastase (prepared in 2.3.5).

2.3.5 PREPARATION OF ANHYDROELASTASE FROM TOSYL ELASTASE OR PMSE

The procedure used to modify the inhibited PPE to anhydroelastase was that of Murphy[5] and Ako et al[8].

The inhibited PPE (15.2mg) was dissolved in 1mM acetic acid (15.2 ml) and cooled in ice. Potassium hydroxide solution (0.1M) was added dropwise over 1 minute until the pH of the solution was 12.4. This gave a final concentration of hydroxide of 0.025M. After 4 hours the pH of the solution was adjusted to 5.0 with HCl (0.5M). 10 ml of 0.01M sodium acetate buffer, pH 5.0 was added and the solution dialysed exhaustively against 1mM acetic acid, spun for half-an-hour at 21,000 rpm (Superspeed 50) and freeze-dried. As a control the native PPE was also treated by the above method with 0.1M potassium hydroxide and found to be ≃ 97% active, while the inhibited PPE remained ≃ 98% inactive.

2.3.6 PREPARATION OF THIOL-ELASTASE BY THE ACTION OF THIOL-ACETATE

2.3.6.1 FROM TOSYL ELASTASE OR PMSE

The reaction was carried out at two different pH's (6.5 and 7.5) and temperature (0° C and 20° C). The inhibited PPE (16mg, ≃ 98% inhibited) was dissolved in 8 ml of 1mM acetic acid and treated with potassium thiolacetate (9 ml, 143 mg/ml) in 0.6M acetate buffer at the appropriate pH and temperature. The solution was left stirring for 2 hours and the solution hydrolysed with dilute ammonium hydroxide to pH 12 and the pH brought to 13 with sodium hydroxide. After four and a half hours, it was dialysed extensively against 1mM acetic acid and freeze-dried. The enzyme was ≃ 98% inhibited.

2.3.6.2 FROM ANHYDROELASTASE

The procedure was as for 2.3.6.1 but with anhydroelastase instead of inhibited PPE.

2.4 DISCUSSION

The percentage inhibition for an inhibited PPE was determined by assaying PPE with NBA. Tosyl elastase and PMSE were $\simeq$ 98% inhibited and PMSF was found to be a better inhibitor than tosyl fluoride since inhibition was completed within 5 hours compared to 24 hours for the latter.

The use of fluorescent chromophore, (NPA), to inhibit PPE was not successful as it had been found by Haugland et al[6] with chymotrypsin, where the active site serine-195 was acylated with this reagent. They found the enzyme was inhibited $\simeq$ 95% within 24 hours.

One of the major problems in the case of PPE, was that activity could not be measured during the reaction. This was because both the inhibitor and the substrate, (NBA), produced p-nitrophenol when cleaved by PPE and also by the buffer. This resulted in a considerable increase in the absorbance. However, the activity of the inhibited PPE after dialysis was found to be the same as the native enzyme. This result confirmed that the anthraniloyl group was not attached to the active site.

Further evidence for the reactions having occurred could be obtained by measuring the fluorescence spectrum after excitation at 290 mμ. Haugland et al[6] observed a change in the tryptophan emission spectrum (all the tryptophans in the enzyme molecule) on excitation at 290 mμ. The indication of the anthraniloyl group being bound to the active site was observed by the tryptophan emission at 333 mμ being decreased, in the case of α-chymotrypsin. This enhanced

the fluorescence maximum at 422 mμ due to the release of p-nitrophenol by the inhibitor. This evidence was inconclusive and the peak at 333 mμ remained the same over several hours for PPE.

The other fluorescent chromophore, Dansyl halide, used by Horton et al[7] and Gold[9] to inhibit α-chymotrypsin by dansylating serine-195 was investigated with PPE. Dansylation was carried out on PPE and also on the anhydroelastase (pH 8.4). The activity of PPE under different pH conditions, including PPE treated under the same conditions but without the inhibitor, was measured throughout the reaction. A considerable increase in the activity of PPE was read at first which gradually dropped over 24 hours. After incubation for a further 48 hours, the activity dropped by 40% for PPE between pH 5.5-7.5 and by 20% at pH 8.4. It was impossible to detect with anhydroelastase, if dansylation had taken place since the enzyme was inactive at the beginning.

Fluorescence spectra were run for all the dansyl elastase preparations at different pH's. Horton et al[7] indicated that the absorption maximum was around 325-360 mμ and the fluorescence maximum around 525 mμ. The dansyl elastase prepared at pH 8.4 was found to be excited at 340 mμ and the fluorescence maximum at 505 mμ. The wavelength at 340 mμ was taken as a reference for the rest of dansyl elastase preparations, but no fluorescence maximum around 525 mμ was observed.

It was concluded that dansylation did not take place in acidic or neutral conditions because no change in the activity and no fluorescence spectrum were obtained. Dansylation did take place in alkaline conditions (pH 8.4) and this was also observed by Horton et al[7]. Since there was no change in the activity at this pH it was concluded the serine-195 was not dansylated as Horton et al[7] and Gold[9] had found in the case of α- chymotrypsin.

There was no indication of a fluorescence maximum for dansylated anhydroelastase probably because the enzyme had the $-CH_2OH$ converted into $=CH_2$ and hence the OH was not available for dansylation. Also there was no indication of dansylation taking place anywhere else in the molecule. This suggested that the amino-acids prone to dansylation were affected under basic conditions in the preparation of anhydroelastase.

TABLE 2.1
SUMMARY OF PERCENTAGE INHIBITION OF PPE BY VARIOUS INHIBITORS

Inhibited PPE	Inhibition
Native PPE	0%
Tosyl elastase	≃ 98%
PMSE	≃ 98%
Anthraniloyl elastase	≃ 2%
Dansyl elastase (only at pH 8.4)	≃ 20%

The use of the fluorescent chromophores gives some indication of the mode of binding since these inhibitors were found to inhibit α-chymotrypsin very quickly. The binding site of PPE (P_1) is different and can only accommodate a small group such as methyl because of the occulsion by valine-216 and threonine-226, whilst in α-chymotrypsin these amino-acids are replaced by glycines-216 and -226.

Attempts were carried out to convert the active center serine to cysteine by thiolation of the inhibited PPE and of the anhydroelastase. The activity of the enzyme was measured against the substrate, NBA, and no change in the activity was observed in either case.

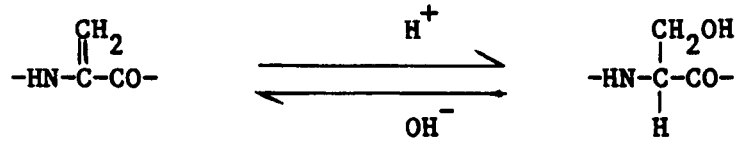
Experimental evidence to detect the presence of the SH group in the active site was carried out using Ellman's reagent[10]. This reagent reacts with the SH group at pH 8.0 and release one p-nitrophenol anion per mole thiol. From this, one can measure the thiol concentration. All of the thiol derivatives prepared were reacted with Ellman's reagent to detect the presence of the SH group. No change in absorbance was noticed. Hence it can be concluded that neither substitution nor addition at serine-195 had taken place. This was also observed by Murphy[5], when preparing from tosyl elastase, Varying the conditions from his, that is, from pH 5.3 to pH 6.5 and 7.5 and also performing the reaction at 0^{o} C and room temperature rather than at 2^{o} C did not result in any change in the reaction.

From these results, two possible explanations are possible about the formation of anhydroelastase. Firstly, it is possible, that the alkaline treatment on the inhibited PPE did not eliminate the tosyl or the PMS groups from the serine-195. Secondly, if the anhydroelastase were formed then conditions used for the addition of the -SH group were not sufficient to cause addition at the double bond.

Hence, to distinguish these two possibilities, a 2.4 Å resolution data set was collected. After processing and comparing the electron density around the active site region with the native PPE no change was observed, (figure 2.0-2.1). Therefore the tosyl and the PMS groups were removed.

However, the anhydroelastase prepared was inactive and moreover it crystallised indicating no gross conformational changes or denaturation. One possible explanation for no difference in the active site could be, the dehydroalanine is formed under basic conditions, but changing the pH 5.0 causes addition to the double bond. If addition takes place

in terms of Anti-Markownikoff rule then it results in the formation of serine. That is,



It was decided to investigate further by soaking the anhydroelastase crystals (prepared from a fresh batch of PPE) in the hexapeptide substrate H.Pro-Ala-Pro-Ala-Lys-Phe.OH (discussed later in chapter eight). Collecting the data to 2.5 Å resolution and processing would reveal the mode of binding and in theory the whole peptide should be intact.

After the work reported was carried out, Tasi et al[11] have reported studies on the inhibition of PPE and anhydroelastase by Boronic acids. They used a similar method for the preparation of anhydroelastase using [^{14}C] PMSF, and proved chemically the formation of dehyroalanine at the active site by the addition of radioactive tritium sodium borohydride (NaB^3H_4). They noticed considerably higher radioactivity in the NaB^3H_4 - reduced anhydroelastase than in the control. Further study using NaB^3H_4 could be used to justify the formation of dehydroalanine.

$\Delta F\alpha_{NAT}$ ANHYDROELASTASE-NATIVE

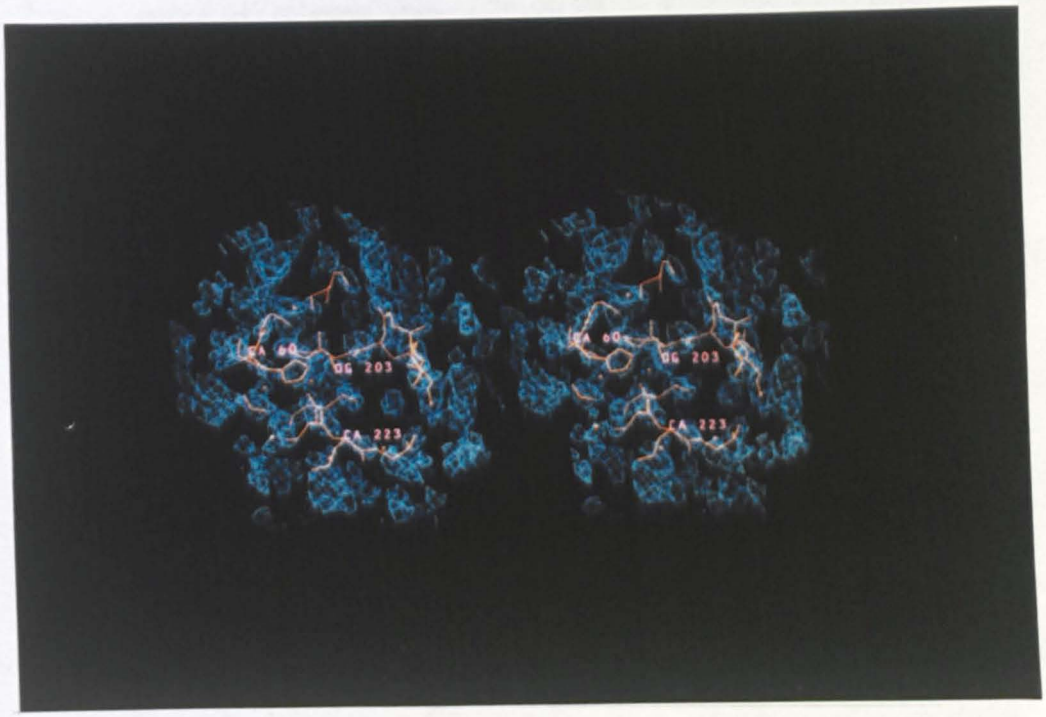


FIGURE 2.0

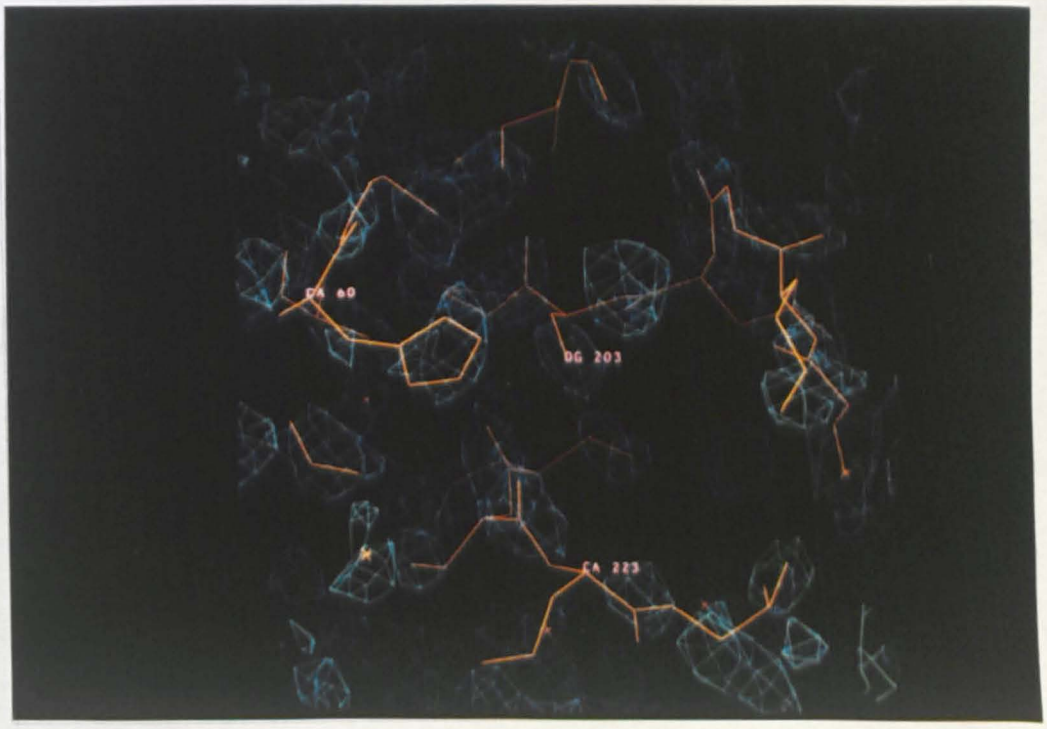


FIGURE 2.1

REFERENCES

1. Shotton, D.M.
Methods in Enzymology., 19, 113 (1970).

2. Dawson, R.M.C., Eillot, D.C., Eillot, W.H., and Jones, K.M.
Data of Biochemical Research, 2nd Edition. Oxford University Press (1969).

3. Visser, L., and Blout, E.R.
Biochim. et. Biophys Acta., 268, 257 (1972).

4. Schwert, G.W., and Takenaka, Y.
Ibid., 16, 570 (1955).

5. Murphy, S.
Ph.D. Thesis, University of Sheffield, (1973).

6. Haugland, R.P., and Stryer, L., Ramachandran, G.N.
"Conformation of Biopolymers", 321 (1967).

7. Horton, H.R., and Koshland, D.E. Jr.
Methods in Enzymology, Vol.XI, 856 (1967).

8. Ako, H., Foster, R.J., and Ryan, C.A.
Biochim. Biophys. Res. Commun., 47, 1402 (1972).

9. Gold, A.M.
Methods in Enzymology, Vol.XI, 706, (1967).

10. Ellmans, G.L.
Arch. Biochem. Biophys., 82, 70 (1959).

11. Tasi, I.H., and Bender, M.L.
Ibid., 228, 555 (1984).

CHAPTER THREE

PEPTIDE SYNTHESIS

Interest in the chemical synthesis of peptides stems principally from (1) use of such synthesis for the unambiguous assignment of structures to naturally occurring peptides and (2) use of such synthesis in the production of chemical analogues of the naturally occurring peptides. Peptides of the latter class are of great interest in terms of aiding the understanding of the relationships between chemical structure and biological function and for their potential medicinal value.

3.1 PROBLEMS OF PEPTIDE SYNTHESIS

Peptides and proteins consist of chains of amino-acids linked to one another by amide bonds. Peptide synthesis is the joining together of chiral amino-acids to form chains of pre-determined length and sequence. Formally peptide synthesis consists simply of elimination of water from two amino-acids or peptides, figure 3.0.

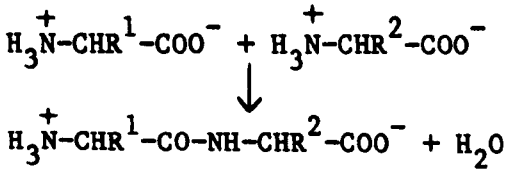


FIGURE 3.0

but due to the dipolar nature of the amino-acids, this reaction is thermodynamically unfavourable and requires unacceptably high temperatures. It would also lead to a mixture of peptides, cyclisation and polycondensation reactions (figure 3.1).

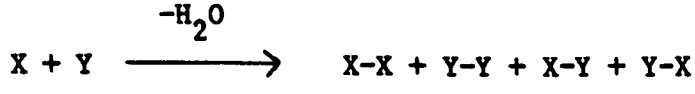


FIGURE 3.1

and each of the dipeptides formed could then react further to give an even greater mixture of peptides.

Natural biosynthesis overcomes these problems, at the ribosomal level where in DNA is translated into a specific peptide but in the laboratory a rational peptide synthesis is required :

(i) destruction of the dipolar nature of reacting amino-acids.

(ii) differentiation of the amino and the carboxyl components. Since amino-acids are di- or polyfunctional, and the coupling of the carboxyl group of one to the amino group of the other is achieved only by protecting some of the functional groups.

(ii) activation of the carboxyl component so that coupling to form the peptide bond can occur with high efficiency under mild conditions.

Steps (i) and (ii) are usually met by use of reversible protecting groups which have been developed for both the amino and carboxyl groups, as well as for the groups occurring in the side-chains.

A suitable protecting group must :

(i) be easily introduced into the molecule.

(ii) protect the functional groups under the conditions of amide formation.

(iii) be easily removable under the conditions that leave the newly formed amide link intact.

A major requirement of peptide synthesis is that all reactions should lead to a single product. However, this is complicated by unprotected functional groups and the danger of racemisation during synthesis, since all the amino-acids with the exception of glycine contain at least one asymmetric carbon atom.

The general scheme involved in peptide synthesis[1] is shown in figure 3.2 and the individual steps will be discussed in the subsequent sections.

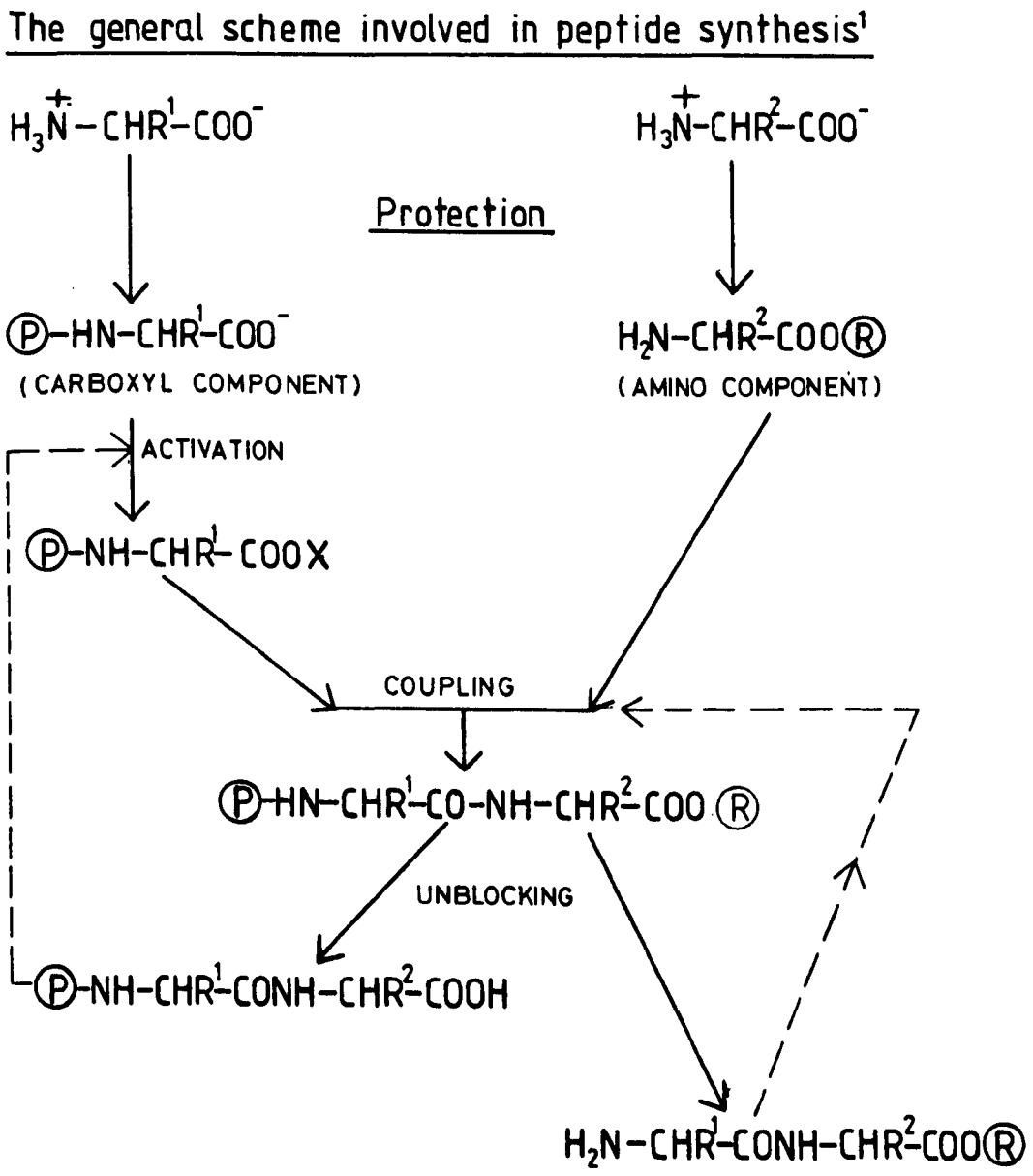


FIGURE 3·2

3.2 PROTECTING GROUPS

There is a whole host of protecting agents for amino and carboxyl groups available and those of importance will be discussed in the following section. Their choice depends on their ease of introduction, the adequacy of the protection provided, their stability during peptide synthesis, the protection of adjacent chiral centre from racemisation and finally, their ease of removal at the completion of the synthesis.

3.2.1 AMINO PROTECTING GROUPS

An amino group is usually acylated. In general the protecting groups function by suppressing the basicity and the nucleophicity of the nitrogen atom.

Emil Fischer[2] used amino-acid chloride-hydrochloride for simultaneous carboxyl activation and amino protection. However, this resulted in impure products in solution, due to the equilibrium which occurs (figure 3.3). An amino group is released free and which is susceptible to further acylation.

$$H_3\overset{+}{N}\text{-CHR-COCl} + H_2N\text{-}R^1 \longrightarrow H_2N\text{-CHR-COCl} + H_3\overset{+}{N}\text{-}R^1$$

FIGURE 3.3

One of the first successful groups, tosyl chloride, that could be exclusively removed from the amino group was suggested by du Vigneaud et al[3]. α-tosyl amino-acids are prepared by the reaction of an amino-acid with tosyl chloride in alkali. Reduction with sodium of a solution of the tosyl peptide in liquid ammonia yields the free amine[3] and this group gained enough popularity for general application in peptide synthesis. However, the tosyl group has its disadvantages. The severe conditions used in the removal of

the N-tosyl group in the intermediates also removes other blocking groups. Hence, these groups require to be reintroduced afterwards. There are also problems in the isolation of pure peptides, since a large amount of inorganic salt must be removed, and the α-tosyl amino-acids cannot be coupled using the mixed anhydride method[4].

3.2.1.1 BENZYLOXYCARBONYL GROUP (Z)

In 1932, Bergmann and Zervas[5] introduced the Z group, which made modern peptide synthesis possible. This group is the first choice for the protection of amino nitrogen. N-Benzyloxycarbonyl derivatives are stable compounds readily obtained in high yields (figure 3.4).

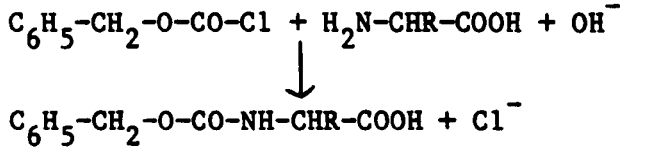


FIGURE 3.4

Several deprotection methods exist:- hydrogenolysis[5], reduction with liquid ammonia[6], HCl solutions[7], HCl[8] or Hydrobromic (HBr)[9] or Hydroiodic (HI) acid[10] in acetic acid and p-toluenesulphonic acid[11]. The Z-group protects amino-acids not only against unwanted acylation but also against racemisation. Of the above methods, solvolysis in HBr and hydrogenolysis are the most commonly used.

However, the method does have a few limitations. Firstly, the Z-group is sensitive to alkali and if Z-peptide esters contain glycine as a second amino-acid from the amino terminal, the formation of hydantoin derivatives[12] results (figure 3.5).

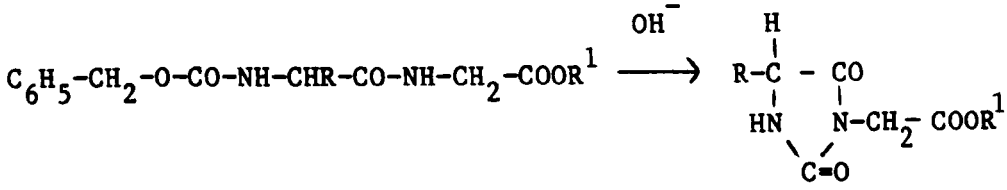


FIGURE 3.5

Excess alkali opens the hydantoin, with the formation of a urea derivative[12] (figure 3.6).

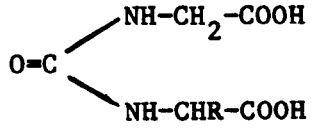


FIGURE 3.6

Secondly, N-carboxyanhydrides (Leuchs anhydrides)[13] are readily formed on heating acid chloride derivatives of Z-amino-acids which are stable in the cold, however.

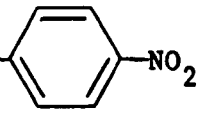
Next, catalytic hydrogenation fails in the presence of sulphur containing amino-acids because of catalyst poisoning but this can be overcome by the use of dry liquid ammonia as solvent[14]. Lastly, if HBr in acetic acid is used for the removal of the Z-group in methionine derivatives then the methyl group is displaced by the benzyl group to form S-benzylhomocysteine[8,15]. Addition of methyl ethyl sulphide can prevent this side reaction[16].

The success of this blocking group, has led to the preparation of a number of modified versions of the Z-group. These include the halogen and nitro substituted forms which are more resistant to acid and are more suitable for the use alongside the t-butyloxycarbonyl group.

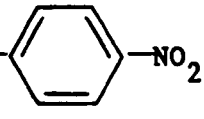
3.2.1.1 TERT-BUTYLOXYCARBONYL GROUP (BOC)

In 1957, Carpino[17] reported the BOC-group as a general amino protecting group and shortly afterwards it was applied to peptide synthesis[15,18]. Introduction of the BOC-group is not as simple as that of the Z-group, since t-butyl-chloroformate (^{t}Bu-O-CO-Cl) is much too unstable except at temperatures below -10° C[19]. Therefore, alternative methods for the synthesis of BOC-amino-acids have been suggested[15,18,20,21] (figure 3.7a, b, c and d).

(a) $(CH_3)_3$-C-OH + O=C=N-CHR-COOR^1 ⟶ $(CH_3)_3$-C-O-CO-NH-CHR-COOR^1

(b) $(CH_3)_3$-C-O-COO-  -NO_2 + H_2N-CHR-COONa

↓

$(CH_3)_3$-C-O-CO-NH-CHR-COONa + HO-  -NO_2

(c) $(CH_3)_3$-C-O-CO-NH-NH_2 + HNO_2 ⟶ $(CH_3)_3$-C-O-CO-N_3 + H_2N-CHR-COOR^1

↓ -5° C

$(CH_3)_3$-C-O-CO-NH-CHR-COOR^1

(d) $(CH_3)_3$-C-O-CO

O + H_2N-CHR-COOR ⟶ (Dioxane, $NaHCO_3$) $(CH_3)_3$-C-O-CO-NH-CHR-COOR^1

$(CH_3)_3$-C-O-CO

FIGURE 3.7

The synthesis of the BOC-protected amino-acids by the azide method (figure 3.7c) proceeds efficiently if the acyl-azide is coupled with the amino component in organic solvents below 5° C, otherwise there is the likely chance of explosion[22]. The method is complicated, but is still one of the most commonly used coupling techniques.

The BOC-group is stable to hydrogenation and also extremely resistant to hydrolysis by base but is cleaved easily by mild acids. Thus, the two groups Z and BOC can be used

together in the same molecule which allows selective protection/deprotection of the amino groups within a molecule. This provides the peptide chemist with more versatility in complex peptide synthesis. Further both groups can be removed simultaneously with HBr in acetic acid. Acid cleavage of the BOC from the derivative is shown in figure 3.8.

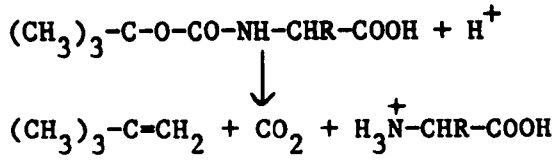


FIGURE 3.8

3.2.1.3 TRIFLUOROACETYL GROUP (TFA)

Another means proposed by Weygand et al[23] for the N-protection of amino-acids was the TFA group. This group is easily cleaved by treatment with alkali. TFA-amino-acids are obtained by the action of trifluoroacetic anhydride on amino-acids. However, there is danger of racemisation occurring if excess trifluoroacetic acid anhydride is used as a solvent. Alternatively, if acylation is performed in ethyl thiotrifluoroacetate in aqueous media at pH 8-9, optically pure derivatives can be obtained[24]. Another possibility for the synthesis of trifluoroacetylated amino-acids is the aminolysis of trifluoroacetic acid esters, for example phenyl and methyl esters. Sodium, barium or ammonium hydroxide[24] or piperdine solutions are most commonly used for the deprotection. The groups mentioned above for N-protection are those most generally used by the peptide chemist although others have been reported.

3.2.2 CARBOXYL PROTECTING GROUPS

Peptide coupling can be carried out, without protection of the carboxyl group of the amino component only if activation and coupling are not used together. Otherwise, the amino component participates as a second carboxyl, and a mixture of products results. The simplest form of carboxyl protection is by salt formation with a strong base, but the most widely used is the methyl or ethyl esterification of the amino-acids.

3.2.2.1 METHYL AND ETHYL ESTERS

The most convenient method used for esterification is mixing methanol or ethanol with thionyl chloride at -10^{o} C and adding the amino-acid[25]. Other catalysts which can be used are sulphuric acid, p-toluenesulphonic acid and phosphorus pentachloride. The deprotection of alkyl esters after the completion of peptide synthesis, is usually done by mild alkaline hydrolysis in organic solvents, (methanol, dioxane or acetone) at or below room temperature. There is a danger, if excess alkali is added, of cleavage of sensitive amide bonds, cyclisation, racemisation and also formation of hydantoin or urea derivatives, (figures 3.5 and 3.6). These esters are also convenient intermediates for the preparation of protected peptide hydrazides.

Amino-acid hydrazides are intermediates of peptide synthesis in the azide coupling method and can also serve as a carboxyl protected form of the acids. The free carboxyl group can be regenerated by oxidation of a diimide, with N-bromosuccinimide (NBS) or I_2, which is then attacked by hydroxyl ions or by an amine to form a peptide (figure 3.9).

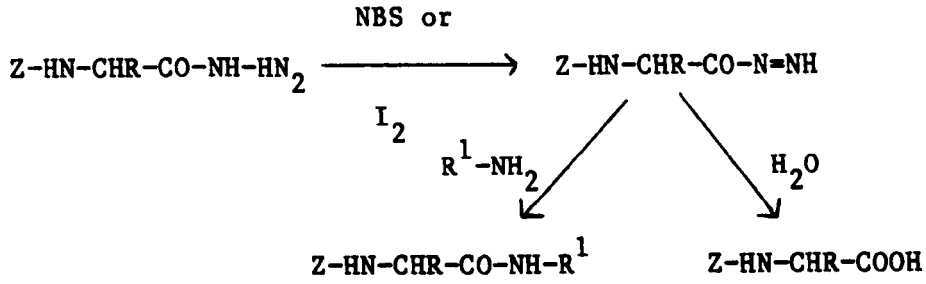


FIGURE 3.9

Besides the methyl and ethyl esters, Bergmann et al[13] introduced benzyl esters.

3.2.2.2 BENZYL ESTERS (-OBz)

This group can be selectively cleaved by catalytic hydrogenolysis in a similar manner to the Z-group. Besides the methyl and ethyl esters, the -OBz group was regarded for a long time as the only carboxyl protecting group. The esterification procedure is simpler and also results in improved yields[25]. However, it is not possible to cleave -OBz esters selectively by catalytic hydrogenation in the presence of the Z-group[27]. Therefore substituted -OBz groups were developed which show increased resistance towards reagents such as HBr in acetic acid and trifluoroacetic acid; for example p-Nitrobenzyl, p-Halobenzyl etc. (figure 3.10).

FIGURE 3.10

The benzyl group can be removed from the C-terminal of a peptide by catalytic hydrogenolysis, HBr in acetic acid[28] or sodium in liquid ammonia[29].

3.2.2.3 TERT-BUTYL ESTERS (O^tBu)

One of the most important advances in protecting the carboxyl group was the use of t-butyl esters (O^tBu)[30]. The advantage was that it could be hydrolysed very readily in acidic reagents, while other esters required more vigorous conditions. The ease of deprotection is due to formation of the "stable" intermediate cation (figure 3.11).

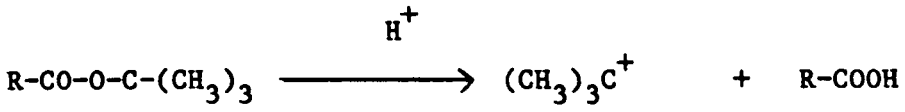
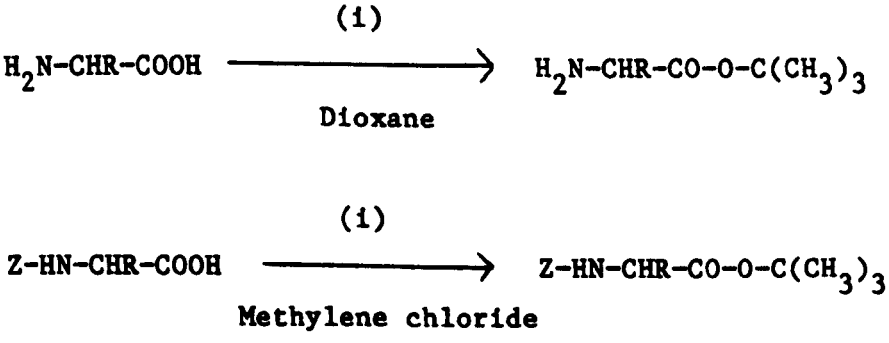


FIGURE 3.11

Other advantages are, the free O^tBU of the amino-acids are stable unlike methyl-, ethyl-, or -OBz esters can be stored and distilled. The O^tBu are also resistant to nucleophilic attack, including attack by amino groups. They are prepared by acid catalysed addition of isobutene to free amino-acids in dioxane[31] or with N-protected amino-acids in methylene chloride[30] (figure 3.12).



(i) Isobutene, H_2SO_4

FIGURE 3.12

The cleavage of the O^tBu esters can be carried out with HBr in acetic acid, p-toluenesulphonic acid in benzene or with trifluoroacetic acid.

3.2.3 SIDE-CHAIN PROTECTION

Consideration must also be given to side-chain protection, if amino-acids with active functional groups on their side-chains are to be used. In the work undertaken here, only the side-chain of lysine needed to be protected.

In the case of lysine, all of the protecting groups discussed in section 3.2.1 for the protection of terminal amino groups, can also be used for the protection of this side-chain. Combinations of groups are particularly useful for peptide synthesis. For example, the Z-group for the terminal amino and the BOC group for the side-chain amino allows the selective deprotection of the terminal amino group whilst the N-protection is left intact.

Selective reaction of the ε-amino group during the protection can be achieved by using the copper complex of the amino-acid and subsequently removing the copper with EDTA (figure 3.13). This method is of general applicability for blocking α-amino and α-carboxyl groups during the introduction of protecting groups to side-chains.

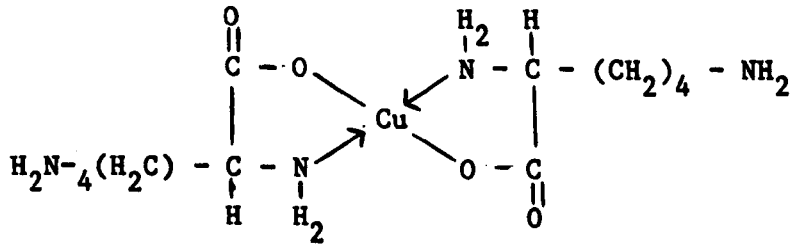


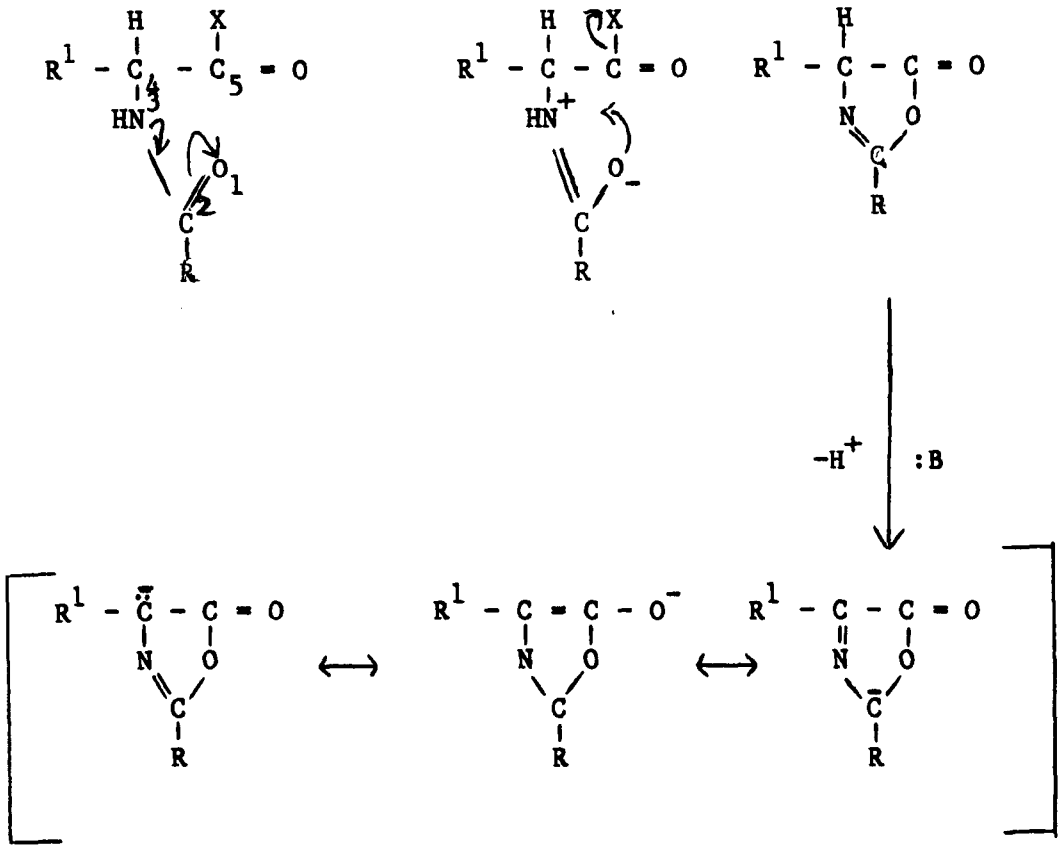
FIGURE 3.13

Other protecting groups have been reviewed[32] and the use of protecting groups in polyfunctional amino-acids has been discussed by Rudinger[33].

3.3 RACEMISATION

High optical purity of the peptide is essential in the synthesis, but there is no guarantee that the peptide synthesis would proceed without racemisation occurring at almost every stage of the preparation. Hence, to prevent or suppress racemisation, careful selection of protecting groups and coupling methods is essential. Two mechanisms have been suggested, which facilitate the loss of optical purity in reactions concerned with the carboxyl group of acylamino-acids or peptides : the withdrawal of proton from the asymmetric α-carbon atom and the formation of an intermediate 5-oxazolone (azlactone) which are known to undergo internal nucleophilic attack and racemise.

The mechanism for the formation of 5-oxazolone (figure 3.14) occurs when any acyl derivative of an N-acyl-α-amino-acid is sufficiently electronegative to undergo internal nucleophilic attack. It is reasonable to suppose that the enhancement in the electrophilic character of the carbonyl carbon, C-5 will facilitate the elimination of XH. There is a strong tendency for the loss of the proton at C-4 in the presence of a base, thus leading to the formation of the resonance stabilised oxazolinone intermediate anions as shown[4].



Oxazlinone intermediate

FIGURE 3.14

Amino protecting groups which are monoacylating such as Z-, BOC-, and other urethane-type protecting groups were not thought to form oxazolinone anions. However, although recent work[35] shows its formation, racemisation does not take place unless coupling is performed in the presence of a tertiary amine base. Since it has already been suggested that oxazolinone formation is base catalysed[34], the major difference between an amide and urethane is the lower acidity of the N-H in the latter. This is because the urethane oxygen can partially donate its electrons to the adjacent carbonyl group (figure 3.15).

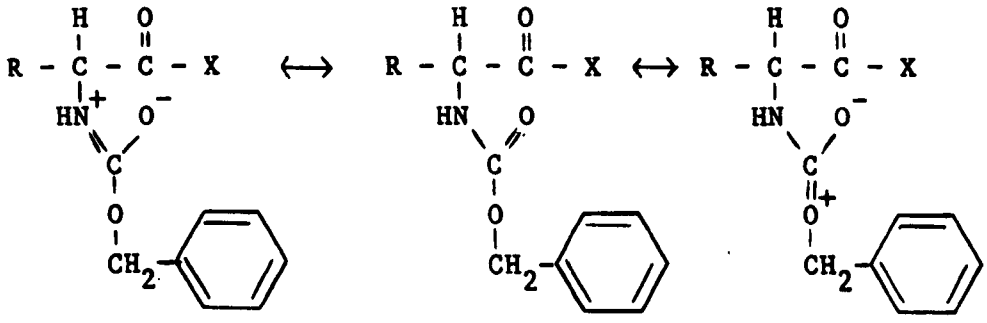
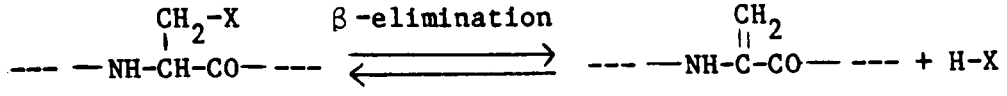


FIGURE 3.15

Extensive experimental evidence for an alternative mechanism has been proposed, which suggests the withdrawal of a proton from the asymmetric α-carbon atom. For example, amino-acids or Z-amino-acids having electronegative substituents in the β-position such as cysteine, serine, threonine, phenylalanine and tyrosine are capable of undergoing base-catalysed racemisation[36,37], and also BOC-benzyl histidine racemised when activated by dicyclohexyl-carbodiimide (DCCI)[38], due to the bascity of the imidazole ring. The β-substituent facilitates the proton abstraction from the α-carbon atom, hence erasing the centre of asymmetry[39](figure 3.16).



X = OH, SH etc.

FIGURE 3.16

With protecting groups other than those of the urethane type no defence against racemisation is available.

Thus general features which lead to racemisation are :

(i) reaction conditions where polar solvents, high temperatures, and the presence of salts[40,41] are prevalent, and

(ii) the nature of the amino-acid at the C-terminal of the carboxyl component, since it has been observed that racemisation almost exclusively occurs there[42].

To reduce the problem of racemisation, various tactics can be used :

(i) use the azide method for coupling as discussed in the next section.

(ii) use glycine or proline as the carboxyl component at the C-terminus.

(iii) use a stepwise strategy for peptide synthesis.

(iv) use of DCCI with additives for coupling.

Various workers have shown that addition of N-hydroxysuccinimide (NHS) also markedly diminishes

racemisation[43,44]. Other new additives where are more efficient than NHS have been proposed, 1-hydroxybenzotriazole[45] and Ethyl-2-hydroximino-2-cyanoacetate[46] (figure 3.17).

FIGURE 3.17

3.4 FORMATION OF THE PEPTIDE BOND

A large number of methods are available but generally only those involving the formation and aminolysis of activated carboxyl acid derivatives have found application in peptide synthesis, where emphasis is on high yield, low racemisation and easy purification of the product.

3.4.1 THE AZIDE METHOD

In this method which is highly regarded[47,48], the reaction proceeds via several steps. The azide is formed by converting the acylamino-acid esters into the corresponding hydrazides, which in turn, are then treated with nitrous acid. The azide formed is then treated under Schotten-Baumann conditions with the amino-acid to form the acyldipeptide (figure 3.18).

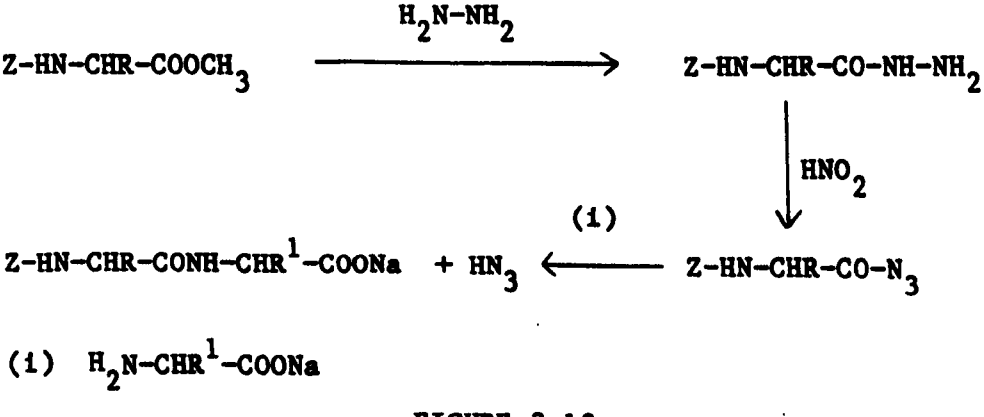


FIGURE 3.18

The formation of acylazide, however, may lead to side reactions because following proton abstraction from the nitrogens, the azide assumes the structure in figure 3.19. The azide is unstable and undergoes a Curtius rearrangement to produce the corresponding isocyanate. The rate of formation of isocyanate depends on the reaction conditions such as solvent and temperature and is a major route in the formation of by-products. Hence, this reduces the peptide yield and the peptide desired is impure[49-51] (figure 3.20). Also there is a possibility of the azides acylating unreacted hydrazide to form symmetrical diacylhydrazines.

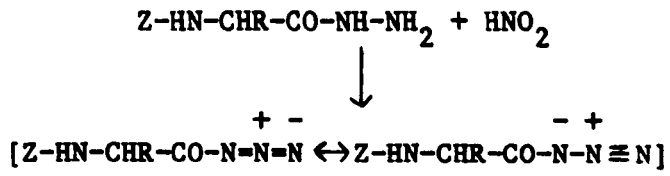


FIGURE 3.19

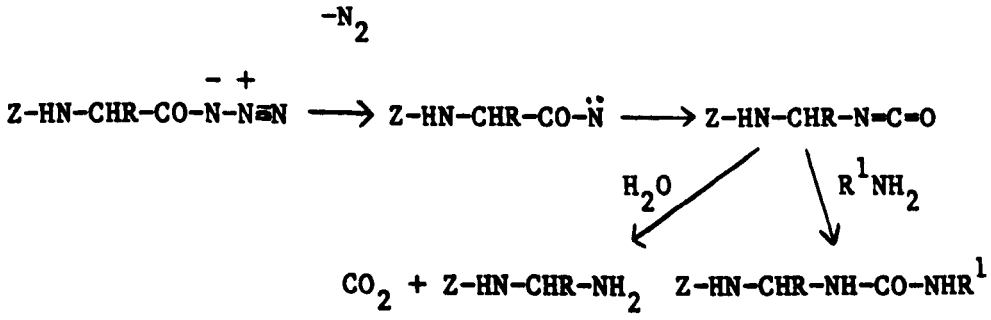


FIGURE 3.20

The azide is usually generated in situ[52] in acidic solution by the addition of sodium nitrite or an organic acid and reacted immediately at -5^{o} C, because of its instability. This method can be applicable for coupling of peptides and for nonaqueous systems. Although, problems exist, it still remains popular, giving products of high hydrazide intermediates and is entirely free from racemisation.

There are a number of methods available for the preparation[53,54] but care must be taken to avoid possible side reactions, especially primary amide formation (figure 3.21).

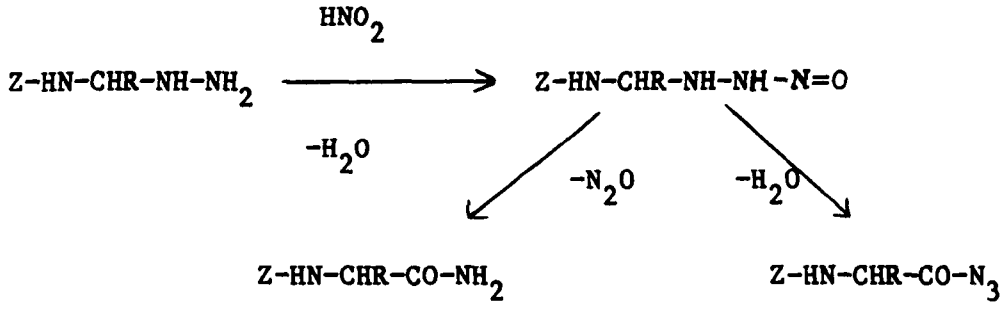


FIGURE 3.21

Optimum conditions for the reactions are low temperature, homogenous solution, high acidity and presence of organic nitrite (e.g. t-butyl nitrite or nitrosyl chloride).

3.4.2 MIXED ANHYDRIDE METHOD

One of the most widely used methods, is the mixed anhydride method[55], which is compatible with many of the protecting groups available and is used when a fast coupling method is required (figure 3.22).

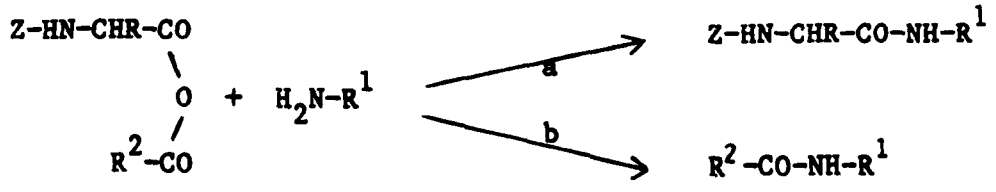


FIGURE 3.22

To gain satisfactory yields with this method, it is necessary to use electron-releasing groups in the activating acid, to compete with the protected aminoacyl part of the

molecule in the acylation of an amine. Suitable groups are long chain aliphatic acids, especially with branched chains. For example, mixed anhydrides containing isovaleroyl[56] or trimethylacetyl residues result in high yields and little side product formation (figure 3.23).

FIGURE 3.23

The mixed anhydrides are prepared by the addition of 1 mole of the acid chloride to a cooled solution of 1 mole of protected amino-acid in the presence of 1 mole of tertiary base (triethylamine or N-ethylmorpholine (NEM)). Toluene, chloroform, dioxane and tetrahydrofuran (THF) are the most frequently used solvents. The formation of mixed anhydrides requires anhydrous conditions and is usually rapid. It can be used directly in the acylation reaction, which can take place under aqueous conditions. A breakthrough in the development of the mixed anhydride method came with the use of monoesters of carbonic acid for coupling[57-59] (figure 3.24).

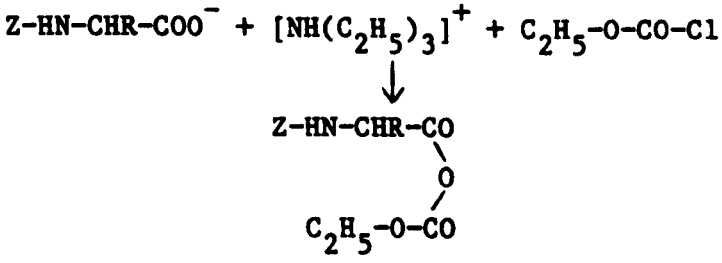


FIGURE 3.24

In this case, the electron-releasing part of the molecule is the alkoxy group, such as ethylchloroformate[57,58] or

isobutylchloroformate[59]. This method has the advantage that no undesired by-products are formed, since the aminolytic cleavage of the anhydride yields the peptide and the corresponding alcohol and CO_2 as shown in figure 3.25.

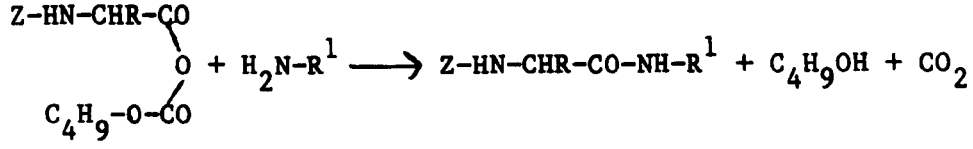


FIGURE 3.25

3.4.3 DICYCLOHEXYLCARBODIIMIDE (DCCI)

Coupling reagents are also required for the formation of peptide bond and the most used reagent is DCCI. It was first used for amide formation by Sheehan and Hess[54,60]. It is highly reactive and gives high yields in a relatively short time. The great advantage of DCCI is that there is very little tendency for racemisation to occur and it is usually used for the preparation of N-protected amino-acid active esters[61,62]. The mechanism was first suggested by Khorana[63-65] and has since been reinvestigated by De-Tar[66]. It involves the addition of N-protected amino-acid to the reagent to form a reactive O-acylisourea intermediate (figure 3.26).

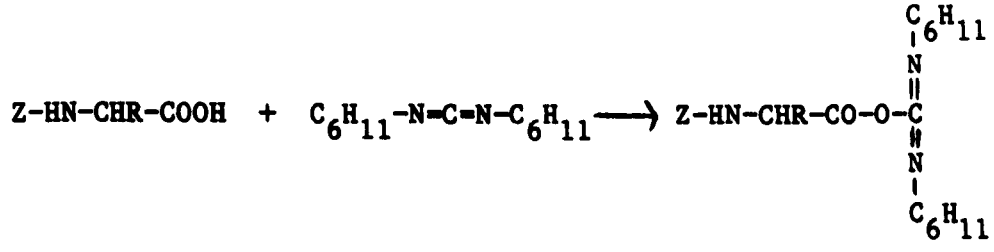


FIGURE 3.26

The resulting intermediate formed can undergo several reactions :

(i) the formation of the peptide bond by the direct attack of the amino component on the reactive intermediate.

(ii) if there is a presence of further acid, then a second addition of a proton takes place on the intermediate resulting in a formation of a symmetrical anhydride and disubstituted urea. The symmetrical anhydride then acylates the amines, and

(iii) lastly O → N acyl migration, with the formation of N-acylurea, which is not an active compound.

The formation of N-acylurea means separation of an undesired by-product is necessary. However, N-acylurea can be suppressed if reactions are carried out in dichloromethane (DCM) or acetonitrile[60]. The N,N^1-dicyclohexylurea formed is insoluble in most of the solvents, but not in the solvents used for coupling. Therefore the peptide formed has a minor amount of urea incorporated. To overcome the problems of separating the impurities, water soluble diimides can be substituted : For example

N-cyclohexyl-N^1-(p-diethylaminocyclohexyl)-carbodiimide.

To reduce racemisation, especially during the activation of acyl peptides with DCCI and to avoid formation of N-acylurea, NHS can be added to the reaction mixture[43,67]. The O-acylurea is converted rapidly into the active ester which then reacts smoothly with the amine. However, NHS coupling is not totally free of racemisation and an alternative was proposed, namely 1-hydroxybenzotriazole. The procedure[45] involving this additive has proved useful in stepwise[68] synthesis as well as fragment condensation[69].

3.4.4 ACTIVE ESTERS

In 1950, Wieland et al[70] described the synthesis of peptides via the active esters, phenyl thioesters which were regarded as asymmetrical anhydrides. Bodanszky[71] observed later that this type of activation was not restricted to the SH group, but aryl and alkyl esters[72] could be used. Active esters are less reactive and are more selective and also more resistant to hydrolysis than are the mixed anhydrides. The difference between the active esters and mixed anhydrides is that only a single product is formed during acylation.

An important development by Nefken et al[73] in the chemistry of active esters was the use of substituted hydroxylamines as the "alcohol" components. The use of N-hydroxy-phthalimide ester and NHS ester become widely accepted[73,74]. These active intermediates resembled the mixed anhydrides than most substituted aryl esters but are sufficiently justified to classify them as active esters. They owe their reactivity towards nucleophiles both to electron-withdrawing effects and to intramolecular hydrogen-bonding as shown in figure 3.27.

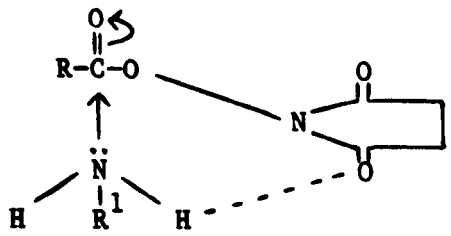


FIGURE 3.27

The rate of acylation with active esters is usually satisfactory but in the case of hindered amino-acids or sparingly soluble peptides, catalytic enhancement is needed. The catalytic effects of tertiary base and carboxylic acid are known[75]. Various reagents are available for the catalysis of the aminolysis reaction such as imidazole[76],

1,2,4-triazole and 2-hydroxypyridine[77] and more recently 1-hydroxybenzotriazole[78]. This was found to be a good catalyst, especially in polar solvents and for solid phase synthesis.

3.5 STRATEGY OF SYNTHESIS

In considering the problems of protection, coupling method and racemisation, it is also essential to take into account the method of assembling of amino-acids into a peptide. Three strategies are available :

(i) Stepwise elongation starting from the N-terminal amino-acid.

(ii) Stepwise elongation starting from the C-terminal amino-acid.

(iii) Fragment condensation (Joining together smaller peptides).

(i) STEPWISE ELONGATION FROM THE N-TERMINUS

The N-terminal, N-protected amino-acid of the peptide chain to be synthesised is activated at its carboxyl group and coupled to the second amino-acid. The partially protected dipeptide so produced is activated and coupled to a third amino-acid to yield an amino protected tripeptide. Elongation in this manner, can progress until the desired peptide is synthesised. The deprotection of the amino group and any side-chain protecting groups, completes the formation of peptide (figure 3.28).

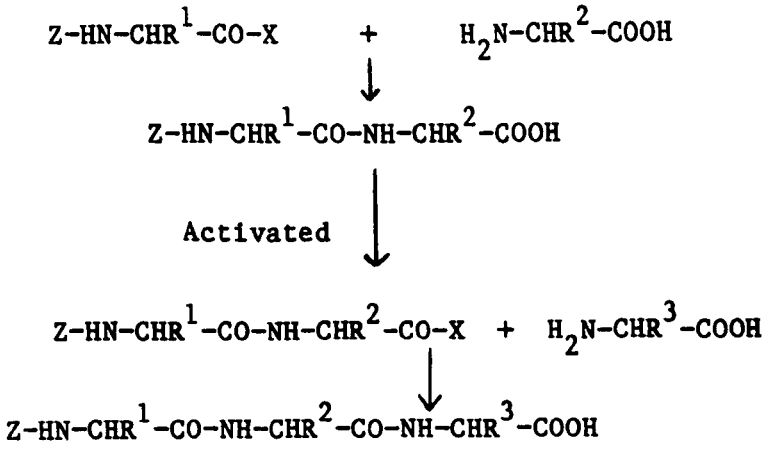


FIGURE 3.28

The scheme assumes that the carboxyl group of the amino component is protected only by the addition of a base, which converts the zwitterion form of the amino-acid to a salt. This leaves the amino group free to act as nucleophile in the acylation. The number of steps involving protection and deprotection is reduced to a minimum and makes this an economical and attractive approach to peptide synthesis.

This type of strategy is regarded as similar to that taken by nature in the synthesis of proteins. However, this procedure has no wide spread application in chemical synthesis, since racemisation can occur at the C-terminal residue of the intermediates during activation and coupling. The first N-protected amino-acid will not be prone to racemisation but all the subsequent peptides formed are, unless glycine or proline are at the C-terminus. Hence, it is very unlikely that this type of stepwise strategy is of any practical use.

(ii) STEPWISE ELONGATION FROM THE C-TERMINUS

One of the most generally applied strategies in which a racemate-free peptide is prepared is to start with the C-terminal residue of the peptide. In this case the carboxyl group is usually protected because of the

zwitterions in the intermediates. One advantage of this procedure is when Z- and BOC-protecting groups are employed, the tendency for racemisation to occur is reduced. Moreover, if active esters are used for the coupling steps, it is reasonable to expect the products to have appreciably different solubilities from the starting material and hence simplify the isolation and purification of the desired peptide. The procedure is shown in figure 3.29.

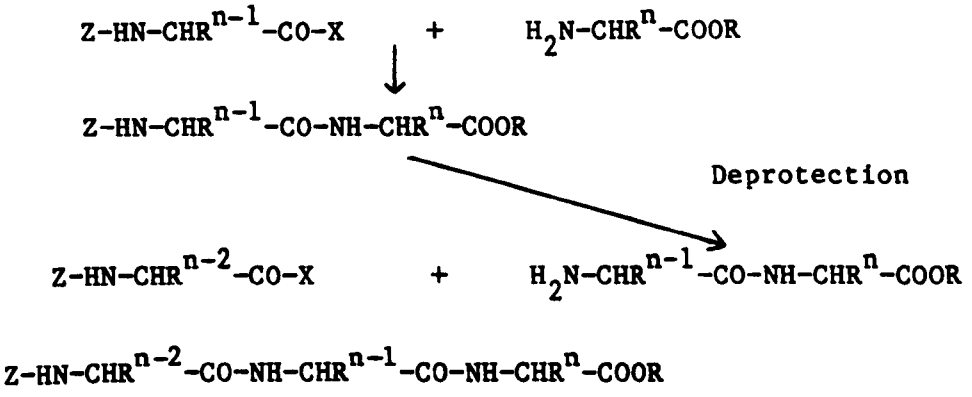


FIGURE 3.29

(iii) FRAGMENT CONDENSATION

In this method fragments of the desired peptide are coupled together, but several points have to be taken into consideration to achieve any success. There is a high risk of racemisation occurring, since it involves the activation of the carboxyl group of the protected peptide rather than a protected amino-acid. To overcome this problem, it would be better to have glycine or proline as the C-terminus residues for reasons mentioned earlier. The major consideration is to avoid aliphatic or aromatic amino-acids having electronegative substitutents in the β-position as the C-terminal amino-acid of the carboxyl component in fragment condensation. If however, they are chosen, the problem of racemisation could be reduced by the introduction of NHS in the coupling stage as discussed earlier.

3.6 SCHEME OF PRESENT SYNTHESIS

The amino-acids required to prepare the hexapeptide and tetrapeptides (mentioned in chapter one), for the study of the binding sites of PPE were alanine, proline, lysine and phenylalanine. The N-terminus was protected with the Z-group and the C-terminus as the active ester in the hexapeptide and one of the tetrapeptides. The other two tetrapeptides had their C-terminus converted to the corresponding alcohol, so that in the final stage of the peptide synthesis it was easier to convert into the aldehyde. The strategy applied to the synthesis of the peptides involved the stepwise addition to the C-terminus except for the hexapeptide, where the last coupling step involved fragment condensation of the tetrapeptide to a dipeptide. The Z-group was used to remove the danger of racemisation via oxazolinone intermediates (figure 3.14). The coupling method employed throughout the peptide synthesis of the hexapeptide and tetrapeptides is illustrated in figure 3.30-3.32.

3.7 EXPERIMENTAL

In all the following synthetic peptide work, all the amino-acids used were of the L-form and obtained from various commercial suppliers.

MATERIALS

L-alanine.
L-proline.
L-lysine.
L-phenylalanine.
N-Benzyloxycarbonyl-L-alanine.

Solvents used were obtained from the various commercial suppliers stored over A4 molecular sieve, if applicable. Melting points were determined with a Tottli-type apparatus.

All evaporations were carried out on a Buchi rotatory evaporator fitted with a cold finger for solid cabondioxide, at temperature of less than 37° C. Hydrogenations were carried out in a round bottom flask fitted with a carbondioxide trap fitted with sodalime. The system used was closed to the atmosphere. All hydrogenations, were flushed with "white spot" N_2 for 5 minutes, before the catalyst, 5% palladium on charcoal (Pd/C) was added. Thin layer chromatography (T.L.C.) was used to determine when the hydrogen uptake was complete. All the products and the intermediates were homogenous on T.L.C. and had the expected ^1H N.M.R. spectra. This spectra were recorded on Perkin-Elmer R34 N.M.R. spectrometer at 220MHz with tetramethylsilane as internal reference.

CHROMATOGRAPHY SOLVENTS

10% Methanol in Chloroform.
5% Methanol in Chloroform.
Butanol : Acetic acid : Water : Ethylacetate (1:1:1:1).
Butanol : Acetic acid : Water (4:1:1).
Chloroform : Methanol : Acetic acid : Water (120:15:3:2).
Chloroform : Methanol : Acetic acid : Water (60:18:2:3)
Chloroform : Methanol : Acetic acid : Water (30:15:3:2).

DETECTION OF AMINO-ACIDS AND PEPTIDES

Detection and identification of the products by T.L.C. is important. Sprays are available for detection of certain groups in the amino-acids and peptides.

(i) NINHYDRIN

A 0.02% solution of ninhydrin in methanol or butan-1-ol is used for detection of a free amino group in a peptide or amino-acid. The spray reagent is developed by heating the

treated plate and ninhydrin positive groups are recognised by the appearance of purple or yellow (for proline) coloured spot.

(ii) CHLORINE/STARCH/POTASSIUM IODIDE

Immersing the treated plate in a chlorine tank for 2-3 seconds and the excess chlorine blown off before being sprayed with a solution containing 1% starch and 1% potassium iodide in water. The appearance of a purple or blue spot indicate the presence of a peptide or amino-acid.

(iii) IODINE

Detection for peptides and amino-acids can also be done by immersing the plate in a tank containing solid iodine for 10 minutes.

(iv) U.V. LIGHT

The method is used in conjunction with others for detection of aromatic ring (e.g. Z, Phe) under the examination of U.V. light, using self-indicating plates.

(v) ALDEHYDE DETECTION

Using a spray made of 2,3,5-triphenyltetrazonium chloride in methanol and 10% sodium hydroxide in methanol. Indication of a aldehyde in the peptide is shown by the formation of a red spot.

3.8 INDEX TO EXPERIMENTAL

(1) N-Benzyloxycarbonyl-L-proline.
(2) N-Benzyloxycarbonyl-L-prolyl-N-hydroxysuccinimide ester.
(3) L-alanine-methyl ester hydrochloride.

(4) N-Benzyloxycarbonyl-L-alanyl-N-hydroxysuccinimide ester.

(5) N-Benzyloxycarbonyl-L-prolyl-L-alanine-methyl ester.

(6) N-Benzyloxycarbonyl-L-alanyl-L-prolyl-L-alanine-methyl ester.

(7) N-Benzyloxycarbonyl-L-prolyl-L-alanyl-L-prolyl-L-alanine-methyl ester.

(8) N-Benzyloxycarbonyl-L-prolyl-L-alanyl-L-prolyl-L-alanine.

(9) N-Benzyloxycarbonyl-L-phenylalanine.

(10) N-Benzyloxycarbonyl-L-phenylalanine-tert-butyl ester.

(11) N-ε-tert-butylcarbonyl-L-lysine.

(12) N-α-Benzyloxycarbonyl-N-ε-tert-butylcarbonyl-L-lysine.

(13) N-α-Benzyloxycarbonyl-N-ε-tert-butylcarbonyl-L-lysyl-N-hydroxysuccinimide ester.

(14) N-α-Benzyloxycarbonyl-N-ε-tert-butylcarbonyl-L-lysyl-L-phenylalanine-tert-butyl ester.

(15) N-Benzyloxycarbonyl-L-prolyl-L-alanyl-L-prolyl-L-alanyl-N-ε-tert-butylcarbonyl-L-lysyl-L-phenylalanine-tert-butyl ester.

(16) L-prolyl-L-alanyl-L-prolyl-L-alanyl-L-lysyl-L-phenylalanine.

(17) N-Acetyl-L-prolyl-L-alanyl-L-prolyl-L-alanine.

(18) N-Benzyloxycarbonyl-L-alanine-methyl ester.

(19) N-Benzyloxycarbonyl-L-alaninol.

(20) N-Benzyloxycarbonyl-L-alanyl-L-proline.

(21) N-Benzyloxycarbonyl-L-alanyl-L-prolyl-N-hydroxy-succinimide ester.

(22) N-Benzyloxycarbonyl-L-alanyl-L-prolyl-L-alaninol.

(23) N-Acetyl-L-proline.

(24) N-Acetyl-L-prolyl-L-alanyl-L-prolyl-L-alaninol.

(25) N-Acetyl-L-prolyl-L-alanyl-L-prolyl-L-alaninal.

(26) N-Trifluoroacetyl-L-proline.

(27) N-Trifluoroacetyl-L-prolyl-L-alanyl-L-prolyl-L-alaninol.

(28) N-Trifluoroacetyl-L-prolyl-L-alanyl-L-prolyl-L-alaninal.

(1) L-proline (51.5g, 0.5 mol) was dissolved in 2M sodium hydroxide solution (250 ml) and cooled to 0° C in an ice/salt bath. The solution was stirred vigorously and benzylchloroformate (85 ml, 0.6 mol) and 2M sodium hydroxide solution (250 ml) added simultaneously over 30 minutes at 0° C. The mixture was stored overnight at 4° C before extracting with diethyl ether (2 x 100 ml) and the aqueous layer acidified with concentrated hydrochloric acid to pH 2.0. The aqueous layer was extracted with ethylacetate (2 x 100 ml), washed with water (100 ml), dried over ($MgSO_4$), filtered, and evaporated to an oil. Addition of petroleum ether 60-80° C gives a white solid (67.97g, 52%), m.p. 76-77° C (Lit[79] 78-80° C).

(2) To a solution of N-Benzyloxycarbonyl-L-proline (62.5g, 0.25 mol) in DME (625 ml), NHS (28.75g, 0.25 ml) was added. The mixture was stirred and cooled to 0° C in an ice/salt bath before DCCI (61.75g, 0.3 mol) was added. The mixture was filtered and DME removed by evaporation. The resulting oil was titurated with ether to give a solid after standing at 0° C for 3 hours. (67.56g, 78%), m.p. 88.5-89.5° C (Lit.[80] 90° C).

(3) L-alanine (62.4g, 0.7 mol) was suspended in methanol (200 ml and cooled to 0° C in an ice/salt bath. Thionylchloride (44.4 ml) was added at 0° C, dropwise over 30 minutes and the mixture then heated to 40° C for 2 hours. The solvent was evaporated and THF (800 ml) added. The mixture was cooled to 0° C to give a white solid. (84.95g, 73.2%), m.p. 109-110° C (Lit.[81] 110.5-111° C).

(4) N-Benzyloxycarbonyl-L-alanine (56.75g, 0.2 mol) was dissolved in DME (625 ml), NHS (28.75g, 0.25 mol) added and the mixture cooled to 0° C in an ice/salt bath. DCCI (61.75g, 0.3 mol) was then added and the mixture stirred at 0° C for 2 hours before allowing it to stand overnight at 0° C. The mixture was filtered and the solution evaporated to an oil,

titurated with ether and left standing at 0° C for 3 hours to give a white solid (65.67g, 83%), m.p. 121.5-123° C (Lit.[80] 123-125° C).

(5) L-alanine-methyl ester hydrochloride (6.98g, 0.05 mol) was dissolved in DMF (80 ml) at 0° C in an ice/salt bath and NEM (6.25 ml) added. The mixture was stirred for 1 minute before N-Benzyloxycarbonyl-L-prolyl-N-hydroxysuccinimide ester (17.3g, 0.05 mol) was added, stirred for a further hour at 0° C and left to stand overnight at room temperature. The solvent was evaporated, the oily residue dissolved in water and the product extracted into ethylacetate (2 x 200 ml). It was washed with 5% citric acid, water, 5% sodium bicarbonate, saturated brine, dried over ($MgSO_4$), filtered and evaporated to give an oil, which was dissolved in minimum of ethylacetate and addition of petroleum ether 40-60° C gave the product. (13.32g, 79%), 75-76° C (Lit.[82] 79-80° C).

(6) N-Benzyloxycarbonyl-L-prolyl-L-alanine-methyl ester (11.7g, 0.035 mol) was dissolved in DMF (95 ml), glacial acetic acid (5 ml) was added and the mixture hydrogenated over Pd/C for 5 hours. After hydrogen uptake was complete the solution was filtered through Hyflo and evaporated under reduced pressure. The oily residue was washed with toluene (3 x 50 ml) and the oily residue redissolved in DMF, N-Benzyloxycarbonyl-L-alanyl-N-hydroxysuccinimide ester, (11.2g, 0.035 mol) was added and stirred for 1 hour at 0° C in an ice/salt bath before allowing to stand overnight at room temperature. The solvent was evaporated, and the oily residue dissolved in water and the product extracted into ethylacetate (2 x 200 ml). It was washed with 5% citric acid, water, 5% sodium bicarbonate, saturated brine, dried over ($MgSO_4$), filtered and evaporated to give N-Benzyloxycarbonyl-L-alanyl-L-prolyl-L-alanine-methyl ester as an oil. (11.5g, 81%).

Anal for: $C_{20}H_{27}N_3O_6$: C, 59.25; H, 6.71; N, 10.37; found C, 59.01; H, 6.55; N, 10.27%

(7) N-Benzyloxycarbonyl-L-alanyl-L-prolyl-L-alanine-methyl ester (11.5g, 0.028 mol) was dissolved in DMF and hydrogenated over Pd/C for 5 hours. When the uptake of hydrogen was complete the solution was filtered through Hyflo and N-Benzyloxycarbonyl-L-prolyl-N-hydroxysuccinimide ester (8.23g, 0.024 mol) added, stirred at 0° C in an ice/salt bath for 1 hour and left to stand overnight at room temperature. The solvent was evaporated, the oily residue dissolved in water and extracted into ethylacetate, dried over ($MgSO_4$), filtered and evaporated to oil. The oil was chromatographed on silica gel (1% methanol/chloroform) and fractions containing the product were evaporated to a white foam (4.2g, 30.5%).

Anal for: $C_{25}H_{34}N_4O_7$: C, 59.75; H, 6.82; N, 11.15; found C, 57.92; H, 6.51; N, 10.86%.

(8) N-Benzyloxycarbonyl-L-prolyl-L-alanyl-L-prolyl-L-alanine-methyl ester (1.5g, 2.98 mmol) was dissolved in 50% methanol/water and sodium hydroxide (0.1g, 3.5 mmol) was added. The reaction stirred for 4 hours at room temperature, the methanol evaporated and the pH of the solution adjusted to 7.0 before extracting with ethylacetate. The aqueous layer was adjusted to pH 2.8 and then extracted with chloroform, the organic layer dried over ($MgSO_4$) and evaporated to a white foam. N.M.R. confirmed that no methyl ester peak was present.

Anal for: $C_{24}H_{32}N_4O_7$: C, 59.00; H, 6.60; N, 11.47; found C, 57.46; H, 6.29; N, 11.10%.

(9) The procedure was as for N-Benzyloxycarbonyl-L-proline (1). Yield : (144.27g, 96.5%), m.p. 83-84° C (Lit.[83] 88-89° C).

(10) N-Benzyloxycarbonyl-L-phenylalanine (71.2g, 0.24 mol) was dissolved in DCM (500 ml) and cooled in a cardice/acetone bath at -10° C. Concentrated sulphuric acid (5 ml) and isobutene (300 ml) distilled at -70° C were added and the mixture left stirring for 4 days at room temperature. The

mixture was cooled and sodium bicarbonate solution (10%) added until the mixture was basic to litmus. An emulsion formed, which was dissolved in DCM and washed with water, brine, dried over ($MgSO_4$), filtered and evaporated to a solid. The crude solid crystallised from DCM-petroleum ether to give a white crystalline product (66.7g, 79%), m.p. 78-79.5° C (Lit.[28] 80.5-81.5° C).

(11) L-lysine hydrochloride (54.8g, 0.3 mol) was dissolved in boiling water (400 ml) and copper(II) carbonate (58g, 0.26 mol) added in portions, then the solution was boiled for 10 minutes, filtered and cooled. Sodium bicarbonate (25.2g, 0.3 mol), Di-tert-butyl-dicarbonate (65.4g, 0.3 mol) and dioxane (200 ml) were added and the mixture stirred for 24 hours. The mixture was filtered and the solid washed with water, ethanol, diethyl ether and dried at 60° C in Vacuo. The solid was suspended in boiling water (600 ml) containing EDTA (di-sodium salt, 75g, 0.20 mol), cooled and filtered, washed with water, ether and the product dried in Vacuo. (42.0g, 57%), m.p. 237-255° C (Lit.[21] 237-255° C).

(12) N-ε-tert-butylcarbonyl-L-lysine (36g, 0.15 mol) was dissolved in 2M sodium hydroxide (73 ml) at 0° C. Benzyl-chloroformate (25 ml, 0.17 mol) and 2M sodium hydroxide (73 ml) were added simultaneously over half-an-hour while stirring at 0° C in an ice/salt bath. The mixture was stirred for a further hour before being allowed to stand for 2 hours at room temperature. The mixture was extracted with diethyl ether (2 x 200 ml) and the aqueous layer acidified (litmus) with 6M hydrochloric acid (100 ml). The product was extracted into ethylacetate, washed with water, brine, dried over ($MgSO_4$) and evaporated to an oil (35.8g, 64.6%)[84].

(13) To N-α-Benzyloxycarbonyl-N-ε-tert-butylcarbonyl-L-lysine (35.8g, 0.09 mol) dissolved in DME (200 ml), NHS (10.85g, 0.09 mol) was added and the solution cooled to 0° C. To the

stirred solution DCCI (24.2g, 0.12 mol) was added and stirred for 1 hour at 0^{o} C in an ice/salt bath before being left overnight at 0^{o} C in the cold room. The mixture was filtered, the DME removed by evaporation and the addition of diethyl ether gave a solid product. (36.6g, 81%), m.p. 94.5-95.5^{o} C (Lit.[85] 94.5-95.5^{o} C).

(14) N-Benzyloxycarbonyl-L-phenylalanine-tert-butylester (9g, 0.025 mol) was dissolved in DMF (150 ml) and hydrogenated over Pd/C for 4 hours. After the intake of hydrogen was complete, the solution was filtered through HyFlo, N-α-Benzyloxycarbonyl-N-ε-tert-butylcarbonyl-L-lysyl-N-hydroxysuccinimide ester (11.9g, 0.025 mol) added and stirred at 0^{o} C in an ice/salt bath for 1 hour. The solution was left to stand overnight at room temperature. On evaporation of DMF, an oily residue resulted which was dissolved in water and extracted with ethylacetate (2 x 200 ml). The organic layer was washed with 5% citric acid, water, 5% sodium carbonate, brine, dried over ($MgSO_4$) and evaporated to give an oil. The oil was dissolved in petroleum ether and the addition of hexane to this solution gave a gel overnight. The solvents were evaporated and the product dried in Vacuo. (13.9g, 95%), m.p. 57-60^{o} C.
Anal. for: $C_{32}H_{45}N_3O_7$: C, 65.85; H, 7.77; N, 7.20 found:
C, 65.56; H, 7.75; N, 7.22%.

(15) N-Benzyloxycarbonyl-N-ε-tert-butylcarbonyl-L-lysyl-L-phenylalanine-tert-butyl ester (1.86g, 3.19 mmol) was dissolved in THF (50 ml) and hydrogenated over Pd/C for 4 hours. When the uptake of hydrogen was complete the solution was filtered through Hyflo, and the solvent evaporated under reduced pressure. The oily residue was redissolved in THF and N-Benzyloxycarbonyl-L-prolyl-L-alanyl-L-prolyl-L-alanine (1.3g, 2.66 mmol) added with stirring in an ice/salt bath at 0^{o} C before being left to stand overnight at room temperature. The mixture was filtered, the THF evaporated and the oily residue dissolved in water and extracted with

ethylacetate (2 x 200 ml). The organic layer was washed with 5% citric acid, water, 5% sodium bicarbonate, brine, dried over ($MgSO_4$) and evaporated to an oil. The oil was dissolved in ethylacetate and hexane added. After standing for 1 hour at 0° C in the fridge the product was obtained. The crude product was chromatographed on silica gel (5% hexane-ethylacetate) and the fractions containing the product were evaporated to a foam.

Anal. for: $C_{48}H_{69}N_7O_{11}$: C, 62.66; H, 7.56; N, 10.66; found: C, 62.89; H, 7.55; N, 10.64%.

(16) N-Benzyloxycarbonyl-L-prolyl-L-alanyl-L-prolyl-L-lysyl-L-phenylalanine-tert-butyl ester (0.5g, 0.54 mmol) was dissolved in a solution of 45% HBr in glacial acetic acid (25 ml) in a flask fitted with moisture trap. The mixture was stirred at room temperature for 2 hours. Anhydrous ether was added to the solution, to precipitate the product and remove excess HBr and acetic acid. The ether layer was decanted and the procedure repeated until all yellow colouration had been removed. The solid was filtered and dried in vacuo overnight, then dried for 3 hours at 100° C at 0.5 vacuum.

Anal. for: $C_{31}H_{49}N_7O_7Br_2$:
C, 45.19; H, 6.03; N, 11.90; found:
C, 44.98; H, 6.01; N, 12.03% + 2.4 mol HBr
Therefore analysis show 0.4 mol of HBr extra.

(17) N-Benzyloxycarbonyl-L-prolyl-L-alanyl-L-prolyl-L-alanine (0.5g, 1 mmol) prepared in (8) was dissolved in acetic acid and hydrogenated overnight over Pd/C. After the intake of hydrogen was complete, the solution was filtered through Hyflo and evaporated to an oil. The oily residue was washed with toluene (4 x 100 ml) and then redissolved in minimum acetic acid, acetic anhydride (0.2 ml) and dry pyridine (1.25 ml) were added. The reaction mixture was stirred for 3 hours at room temperature and on evaporation, an oily residue resulted. This was dissolved in water and

freeze-dried. The crude product was put through a Sephadex G-10 column using 1% acetic acid. The fraction containing the product was identified by T.L.C. using an iodine system (70 mg, 17%).

Anal. for: $C_{18}H_{28}N_4O_6$: C, 54.53; H, 7.12; N, 14.13; found: C, 54.89; H, 6.97; N, 13.98%.

(18) L-alanine-methyl ester hydrochloride (20g, 0.14 mol) was dissolved in water (50 ml) and ethylacetate (300 ml). Potassium bicarbonate (50g) dissolved in minimum water was added to the stirred solution, followed by a dropwise addition of benzylchloroformate (38 ml) over half-an-hour. After stirring for 2 hours, pyridine (14 ml) was added and the organic layer separated and washed with 1M hydrochloric acid, water, 5% sodium bicarbonate, brine, dried over ($MgSO_4$) and on evaporation gave an oil. The oily residue was dissolved in hexane and left to stand overnight at 4° C to give oily crystals (20.49g, 97%).

(19) N-Benzyloxycarbonyl-L-alanine-methyl ester (9.48g, 0.04 mol) was dissolved in dry THF (50 ml), lithium borohydride (1.76g, 2eq) was added and the mixture stirred under nitrogen. After 1 hour ethylacetate (200 ml) and water (100 ml) were added and the mixture acidified with hydrochloric acid. The organic layer was separated and washed with water, 5% sodium bicarbonate, brine, dried over ($MgSO_4$) and evaporated to an oil which crystallised on addition of hexane (6.7g, 80%), m.p. 56-58° C.

Anal. for: $C_{11}H_{15}N_1O_3$: C, 63.14; H, 7.23; N, 6.70; found: C, 63.12; H, 7.21; N, 6.73%.

(20) A solution of L-proline (104g, 0.903 mol) in water (1350 ml) and triethylamine (267 ml, 1.8M) were added to N-Benzyloxycarbonyl-L-alanyl-N-hydroxysuccinimide ester (298g, 0.903 mol, prepared in (4)) in DME (1800 ml). The mixture was stirred for 6 hours and the DME removed by evaporation. The aqueous layer after extraction with

ethylacetate (2 x 300 ml) was acidified to pH 1-2 with 2M hydrochloric acid. The oil was dissolved in ethylacetate (2 x 900 ml), washed with water (2 x 400 ml), dried over (Na_2SO_4) and evaporated to an oil. The product crystallised on addition of diethylether (221.4g, 77%), m.p. 124-125° C (Lit.[86] 120-122° C).

Anal. for: $C_{16}H_{20}N_2O_5$: C, 60.0; H, 6.30; N, 8.70; found: C, 59.90; H, 6.40; N, 8.86%.

(21) The procedure was as for (2) using N-Benzyloxycarbonyl-L-alanyl-L-proline.

(22) N-Benzyloxycarbonyl-L-alaninol (6.27g, 0.03 mol) was dissolved in DMF (50 ml) and hydrogenated over Pd/C for 4 hours and after the uptake of hydrogen was complete, the solution was filtered through Hyflo. N-Benzyloxycarbonyl-L-alanyl-L-prolyl-N-hydroxysuccinimide ester prepared in (21) (12.52g, 0.03 mol) was added and stirred for 1 hour at 0° C in an ice/salt bath before leaving to stand overnight at room temperature. Evaporation of the solvent resulted in an oily residue which dissolved in brine and extracted with chloroform (3 x 150 ml), dried over ($MgSO_4$) and evaporated to a foam. This was dissolved in DCM and the product crystallised to give a white crystalline solid (5.37g, 47.4%), m.p. 157.5-158° C.

Anal. for: $C_{19}H_{27}N_3O_5$: C, 60.46; H, 7.21; N, 11.13; found: C, 59.99; H, 6.98; N, 11.02% + 0.5 mol H_2O.

(23) A solution of L-proline treated in 10-fold amount of boiling acetic acid with exactly 1 mole of acetic anhydride. After evaporation of the solvent the product crystallised from acetone. m.p. 114-116° C (Lit.[87] 118° C).

(24) N-Benzyloxycarbonyl-L-alanyl-L-prolyl-L-alaninol (2.43g, 6.46 mmol), was dissolved in DMF (100 ml) and hydrogenated

over Pd/C for 4 hours. After the intake of hydrogen was complete, the solution was filtered through Hyflo and N-Acetyl-L-prolyl-N-hydroxysuccinimide ester (prepared as (2) from (22)), (1.01g, 6.43 mmol) in situ was filtered into the solution. The reaction was stirred for 1 hour and left to stand overnight at room temperature. The solvent was evaporated and the oily residue was dissolved in water and extracted with chloroform (3 x 100 ml). The aqueous layer was separated and evaporated to an oil. The crude product was chromatographed on silica gel (5-55% methanol-ethylacetate) and the fractions containing the product were evaporated to a foam (1.63g, 66%).

Anal. for: $C_{18}H_{30}N_4O_5$: C, 56.53; H, 7.91; N, 14.65; found: C, 51.40; H, 7.70; N, 12.99% + Ash.

(25) N-Acetyl-L-prolyl-L-alanyl-L-prolyl-L-alaninol (0.5g, 1.31 mmol) was dissolved in DCM and pyridinium chlorochromate (0.74g, 3.43 mmol) was added and stirred at room temperature for 4 hours. The crude product was chromatographed on silica gel by making the column in ethylacetate and eluting with 10% methanol-chloroform. The fraction containing the product were evaporated and redissolved in chloroform and was recolumned in the same system to remove all the oxidising agent. The product after re-evaporation was dissolved in water and freeze-dried (0.12g, 24%).

Anal. for: $C_{18}H_{28}N_4O_5$: C, 56.83; H, 7.42; N, 14.73; found: C, 53.42; H, 7.47; N, 13.60 + 1.35 mol of H_2O.

(26) L-proline (43.6g, 0.04 mol) was suspended in methyl trifluoroacetate (24.26g, 0.21 mol) and N,N,N',N',-tetramethylguanidinium (6.44g, 7 ml) was added. The reaction mixture refluxed for 20 minutes and excess methyl trifluoroacetate removed by evaporation. The mixture was dissolved in water and made acidic (litmus) with concentrated sulphuric acid and extracted with ethylacetate and evaporated to an oil. (7.97g, 96%, (Lit.[88] 46-48° C).

(27) The procedure was as for (24), N-Trifluoroacetyl-L-prolyl-N-hydroxysuccinimide ester substituted for N-Acetyl-L-prolyl-N-hydroxysuccinimide ester.

Yield : (0.94g, 33%).

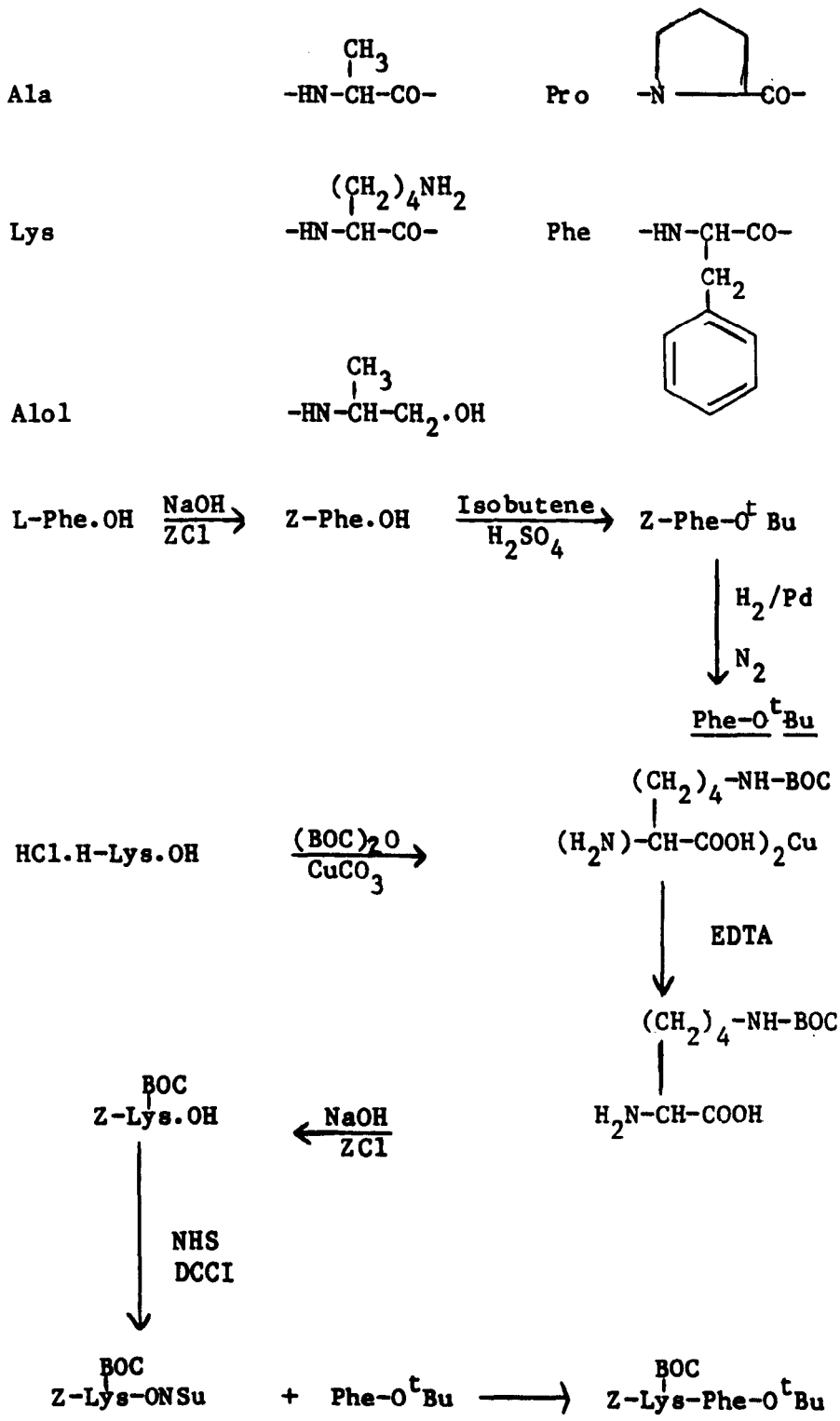
Anal. for: $C_{18}H_{27}N_4O_5F_3$: C, 49.54; H, 6.24; N, 12.84; found: C, 47.87; H, 6.14; N, 12.24% + Ash.

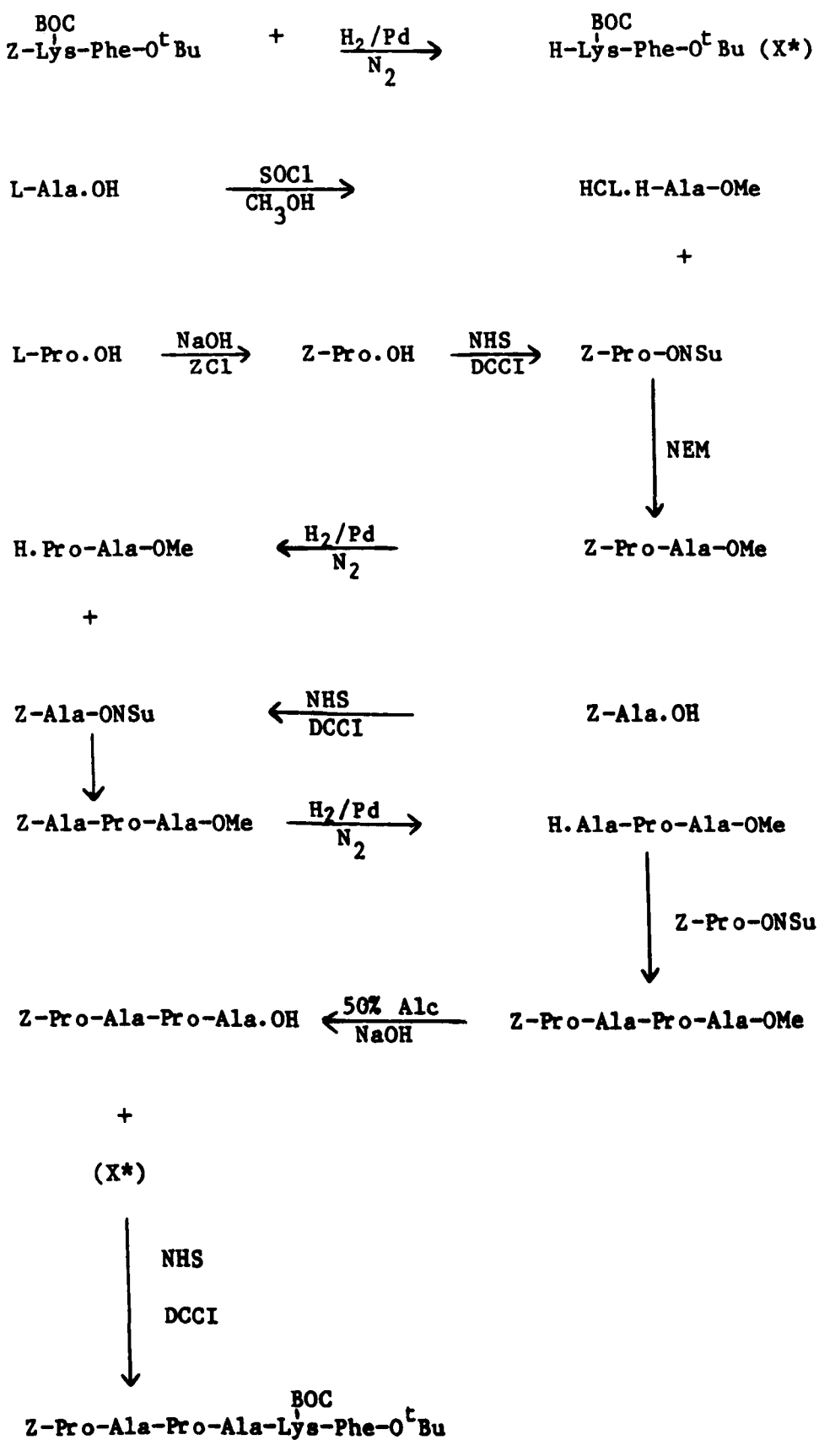
(28) The procedure was as for (25).

Yield : (0.26g, 56%).

Anal. for: $C_{18}H_{25}N_4O_5F_3$: C, 49.47; H, 5.80; N, 12.90; found: C, 47.72; H, 6.20; N, 12.25% + 1 mol of H_2O.

The complete reaction scheme for the hexapeptide (16) is summarised in Figure 3.30.




BOC
Z-Lys-Phe-OtBu
+
H2/Pd
N2
BOC
H-Lys-Phe-OtBu (X*)
L-Ala.OH
SOCl
CH3OH
HCL.H-Ala-OMe
+
L-Pro.OH
NaOH
ZCl
Z-Pro.OH
NHS
DCCI
Z-Pro-ONSu
NEM
H.Pro-Ala-OMe
H2/Pd
N2
Z-Pro-Ala-OMe
+
Z-Ala-ONSu
NHS
DCCI
Z-Ala.OH
Z-Ala-Pro-Ala-OMe
H2/Pd
N2
H.Ala-Pro-Ala-OMe
Z-Pro-ONSu
Z-Pro-Ala-Pro-Ala.OH
50% Alc
NaOH
Z-Pro-Ala-Pro-Ala-OMe
+
(X*)
NHS
DCCI
BOC
Z-Pro-Ala-Pro-Ala-Lys-Phe-OtBu

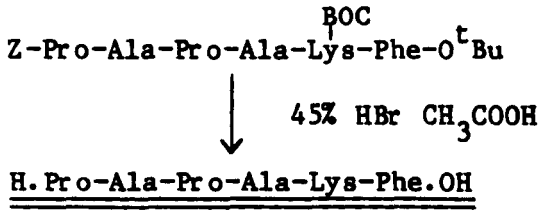


FIGURE 3.30

The complete reaction scheme for the tetrapeptide (17) is summarised in Figure 3.31.

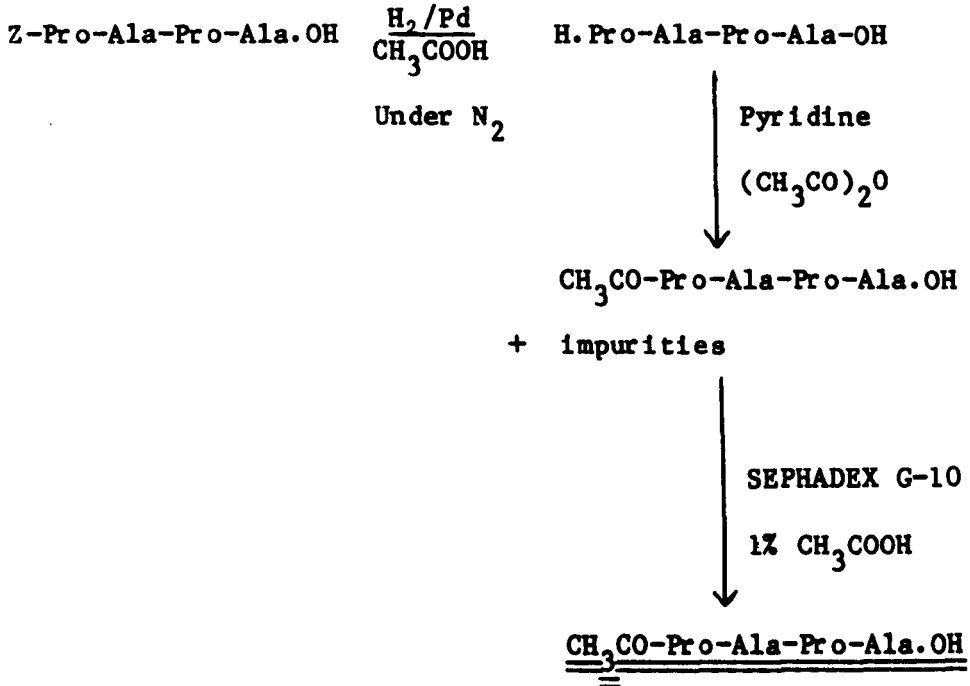
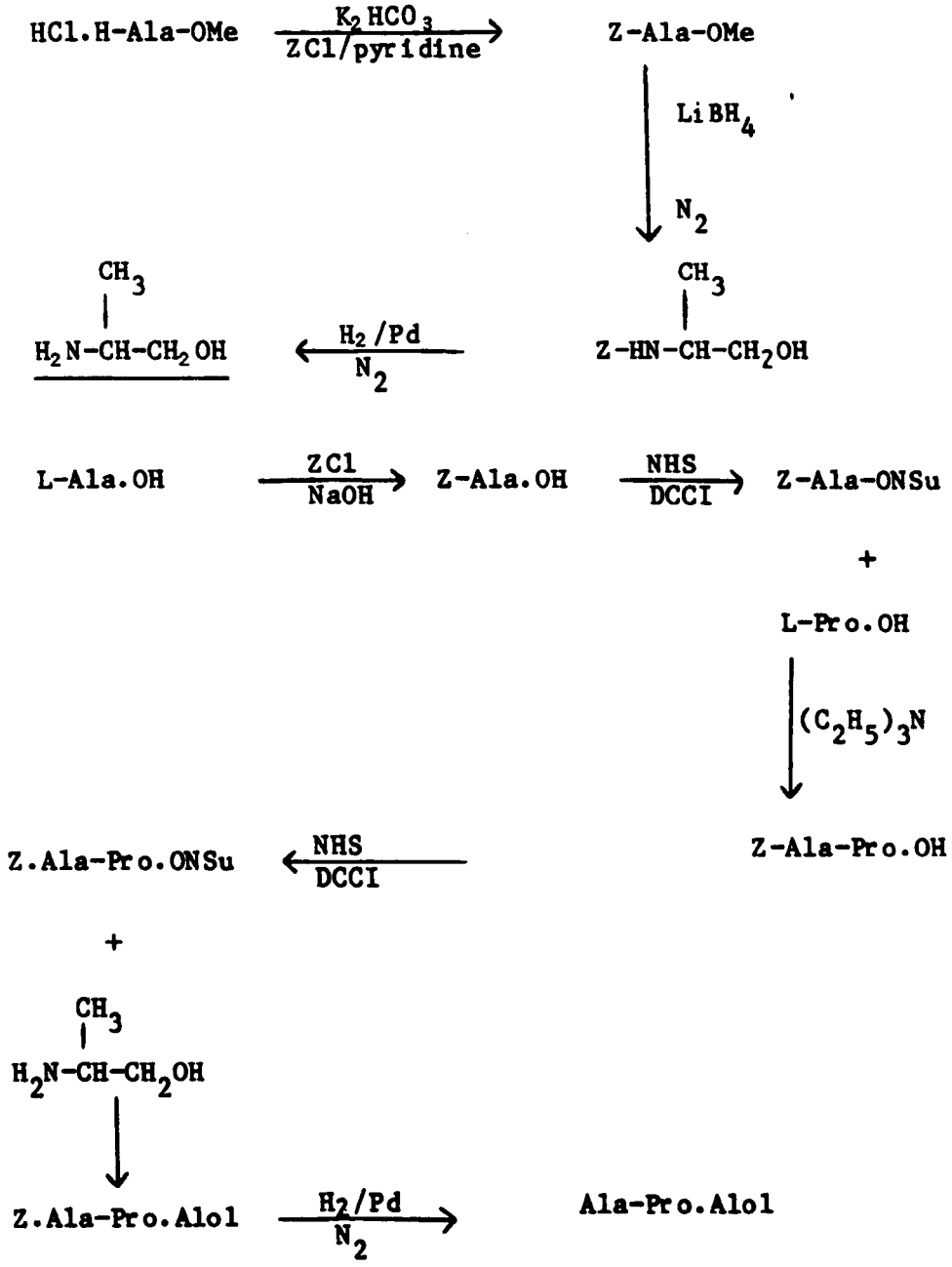
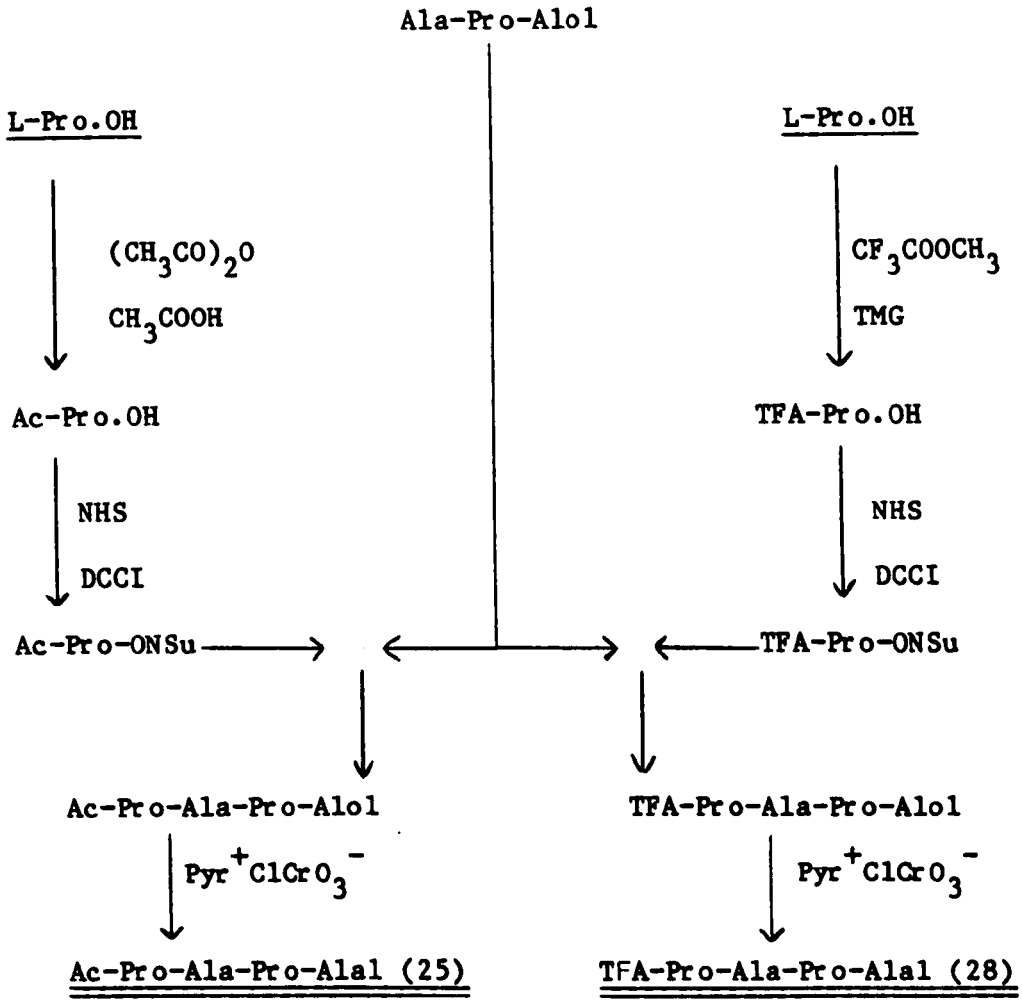


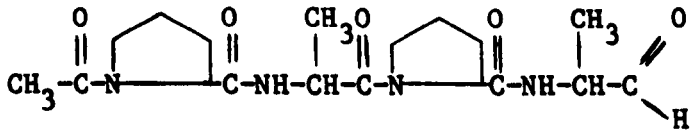
FIGURE 3.31

The complete reaction scheme for the tetrapeptides (25 and 28) are summarised in Figure 3.32.





(25)



(28)

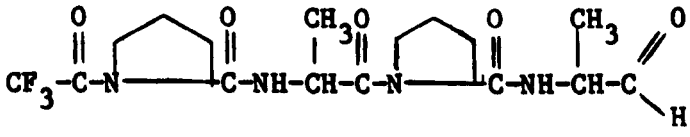


FIGURE 3.32

REFERENCES

1. Koppie, K.D.
Peptides and amino-acids. W.J. Benjamin, Inc. New York (1966).

2. Fischer, E.
Ber.Deut.Chem.Ges., 38, 605 (1905).

3. du Vigneaud, V., and Behrens, O.K.
J.Biol.Chem., 117, 27 (1937).

4. Barrass, B.C., and Elmore, D.T.
J.Chem.Soc., 3134 (1957).

5. Bergmann, M., and Zervas, L.
Ber.Deut.Chem.Ges., 65, 1192 (1932).

6. Sifferd, R.H., and du Vigneaud, V.
J.Biol.Chem., 108, 753 (1935).

7. White, J.
Ibid., 106, 141 (1934).

8. Albertson, N.F., and McKay, F.C.
J.Amer.Chem.Soc., 75, 5323 (1953).

9. Ben-Ishai, D., and Berger, A.
J.Org.Chem., 17, 1564 (1952).

10. Waldschmidtt-Leitz, E., and Kühn, K.
Chem.Ber., 84, 381 (1951).

11. Taschmer, E., and Liberek, B.
Abstr.Int.Cong.Biochem., Vienna (1958).

12. MacLaren, J.A.
Aust.J.Chem., 11, 360 (1958).

13. Bergmann, M., Zervas, L., and Ross, W.F.
J.Biol.Chem., 111, 245 (1935).

14. Kuromizu, K., and Meienhofer, J.
J.Amer.Chem.Soc., 96, 4978 (1974).

15. McKay, F.C., and Albertson, N.F.
Ibid., 79, 4686 (1957).

16. Guttmann, S., and Boissonnas, R.A.
Helv.Chim.Acta., 41, 1852 (1958).

17. Carpino, L.A.
J.Amer.Chem.Soc., 79, 98 (1957).

18. Anderson, G.W., and McGregor, A.C.
Ibid., 79, 6180 (1957).

19. Woodward, R.B., Heusler, K., Gosteli, H., Naegeli, P., Oppolzer, W., Ramage, R., Ranganathan, S., and Vobruggen, H.
Ibid., 88, 852 (1966).

20. Schwyzer, R., Sieber, P., and Kappeler, H.
Helv.Chim.Acta., 42, 2622 (1959).

21. Schwyzer, R., and Rittel, W.
Ibid., 44, 159 (1961).

22. Schnabel, E.
Ann., 702, 188 (1967).

23. Weygand, F., and Csendes, E.
Agnew.Chem., 64, 136 (1952).

24. Schallenberg, E.E., and Calvin, M.
J.Amer.Chem.Soc., 77, 277 (1955).

25. Brenner, M., and Huber, W.
Helv.Chim.Acta., 36, 1109 (1953).

26. Brand, E., Erlanger, B.F., Sachs, H., and Polatnick, J.
J.Amer.Chem.Soc., 73, 3510 (1951).

27. Häussler, R.
Chimica., (Aarau), 14, 369 (1960).

28. Ben-Ishai, D., and Berger, A.
J.Org.Chem., 19, 62 (1954).

29. Roberts, E.W.,
J.Amer.Chem.Soc., 76, 6203 (1954).

30. Anderson, G.W., and Callahan, F.M.
Ibid., 82, 3359 (1960).

31. Roeske, R.W.
Chem.Ind., (London), 1121 (1959).

32. McOmie, J.F.W.
Protective groups in Organic Chemistry, Phenum Press, 183 (1973).

33. Rudinger, J.
Record.Chem.Progr., (Kresge-Hooker.Sci.Lib.), 23, 7 (1962).

34. Young, G.T.
In Peptides. Proceeding of the Eigth European Peptide Symposium, Noordwijk, 1966, Beyermann, H.C., Van de Linde, A., and Massen, Van de Brink, W., Edition, North Holland, Amsterdam. 55 (1967).

35. Bentoin, N.L.
Pept., Struct.Biol.Funct., Proc.Amer.Pept.Symp., 6th edition, Gross, E., and Meienhofer, J., Pierce.Chem.Co., Rochford III., 261 (1979).

36. Bohak, Z., and Katchalski, E.
Biochemistry, 2, 228 (1963).

37. Liberek, B.
Tetrahedron. Lett., 14, 925 (1963).

38. Windridge, G.C., and Jorgensen, E.C.
J.Amer.Chem.Soc., 93, 6318 (1971).

39. MacLaren, J.A., Savige, W.E., and Swan, J.H.
Aust.J.Chem., 11, 345 (1958).

40. Farrington, J.A., Hextall, P.J., Kenner, G.W., and Turner, J.M.
J.Chem.Soc., 1407 (1957).

41. Vaughan, J.R., Jr.
J.Amer.Chem.Soc., 74, 6137 (1952).

42. Iselin, B., Feurer, M., and Schwyer, R.
Helv.Chim.Acta., 38, 1508 (1955).

43. Wünsch, E., and Drees, F.
Chem.Ber., 99, 110 (1966).

44. Zimmermann, J.E., and Anderson, G.W.
J.Amer.Chem.Soc., 89, 7151 (1967).

45. König, W., and Geiger, R.
Berchite., 103, 788 (1970).

46. Itoh, M.
Bull.Chem.Soc.Jap., 46, 2219 (1973).

47. Hofmann, K., Jöhi, A., Furlenmeier, A.E., and Kappeler, H.
J.Amer.Chem.Soc., 79, 1636 (1957).

48. Hofmann, K., Thompson, T.A., Yajima, H., Schwartz, E.T., and Inouye, H.
Ibid., 82, 3715 (1960).

49. Curtius, T.
J.Prakt.Chem., 70, 57 (1904).

50. Curtius, T., and Curtius, H.
Ibid., 70, 158 (1904).

51. Schnabel, E.
Ann.Chem.Liebigs., 659, 168 (1962).

52. Boissannas, R.A., Guttmann, S., Jaquenoud, P.-A., and Waller, J.-P.
Helv.Chim.Acta., 38, 1491 (1955).

53. Hofmann, K., Magee, M.Z., Lindenmann, A.
J.Amer.Chem.Soc., 72, 2814 (1950).

54. Sheehan, J.C., and Hess, G.P.
Ibid., 77, 1067 (1955).

55. Wieland, Th., Kern, W., and Sehring, R.
Ann.Chem.Wiebigs., 569, 117 (1950).

56. Vaughan, J.R., Jr. and Osato, R.L.
J.Amerc.Chem.Soc., 73, 5553 (1951).

57. Wieland, Th., and Bernhard, H.
Ann.Chem.Liebigs., 572, 190 (1951).

58. Boissonnas, R.A.
Helv.Chim.Acta., 34, 874 (1951).

59. Vaughan, J.R., Jr. and Osato, R.L.
J.Amer.Chem.Soc., 74, 676 (1952).

60. Sheehan, J.C., Goodman, M., and Hess, G.P.
Ibid., 78, 1367 (1956).

61. Elliot, D.F., and Russell, D.W.
Biochem.J., 66, 49p (1957).

62. Bodansky, M., and du Vigneaud, V.
J.Amer.Chem.Soc., 81, 5688 (1959).

63. Khorana, H.G.
Chem.Rev., 53, 145 (1953).

64. Khorana, H.G.
Chem.Ind., (London), 1087 (1955).

65. Smith, M., Moffat, J.G., and Khorana, H.G.
J.Amer.Chem.Soc., 80, 6204 (1958).

66. De-Tar, D.F., and Silverstein, R.
Ibid., 88, 1013, 1020 (1966).

67. Weygand, F., Hoffmann, D., and Wünsch, E.
Z.Naturfosch, 21b, 426 (1966).

68. Fletcher, G.A., and Young, G.T.
J.Chem.Soc., Perkin.Trans., 1, 1867 (1972).

69. Wünsch, E., Jaeger, E., and Scharf, R.
Chem.Ber., 101, 3664 (1968).

70. Wieland, T., Schäfer, W., and Bockelmann, E.
Ann.Chem.Leibigs., 573, 99 (1951).

71. Bodanszky, M.
Acta.Chim.Acad.Sci.Hung., 10, 335 (1956).

72. Schwyzer, R.
Helv.Chim.Acta., 36, 414 (1953).

73. Nefkens, G.H.L., and Tesser, G.I.
J.Amer.Chem.Soc., 83, 1263 (1961).

74. Anderson, G.W., Zimmerman, J.E., and Callahan, F.M.
Ibid., 85, 3039 (1963).

75. Weygand, F., Steglich, W., and Chytil, N.
Z.Naturforsch., 236, 1391 (1968).

76. Wieland, T., and Vogler, K.
Agnew.Chem., 74, 904 (1962).

77. Nakamizo, N.
Bull.Chem.Soc.Jap., 44, 2006 (1971).

78. König, W., and Geiger, R.
Chem.Ber., 106, 3626 (1973).

79. Berger, A., Kurtz, J., and Katchalski, E.
J.Amer.Chem.Soc., 76, 5552 (1954).

80. Anderson, G.W., Zimmerman, J.E., and Callahan, F.M.
Ibid., 86, 1839 (1964).

81. Goldschmidt, S., and Krishan, G.K.
Chem.Ber., 98, 2831 (1965).

82. Schröder, E.
Ann.Chem.Liebigs., 679, 207 (1964).

83. Grassmann, W., and Wünsch, E.
Chem.Ber., 91, 462 (1958).

84. Wünsch, E., Jaeger, E.
Hoppe-Seyler's, Z.Physiol.Chem., 345, 193 (1966).

85. Otsuka, H., Inouye, K., Kanayama, M., et al
Bull.Chem.Soc.Jap., 39, 882 (1966).

86. Hassall, C.H., Johnston, W.H., and Roberts, N.
Bio-Organic Chemistry., 8, 229 (1979).

87. Hammer, D., Greenstein, J.P.
J.Biol.Chem., 193, 81 (1951).

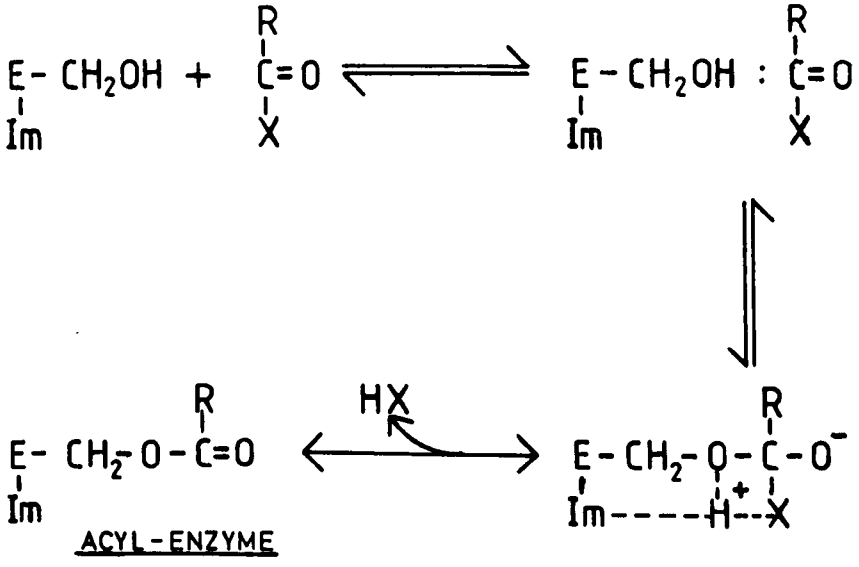
88. Weygand, F., Klinke, P., and Eigen, I.
Chem.Ber., 90, 1896 (1957).

CHAPTER FOUR

KINETIC STUDIES

4.1 INTRODUCTION

The hydrolysis of ester or amide substrates catalysed by serine proteases involves an acyl-enzyme intermediate in which the substrate acylates the hydroxyl group of serine-195[1]. The attack of serine-195 on the carbonyl group of the substrate involves the formation of a high energy tetrahedral intermediate[2].

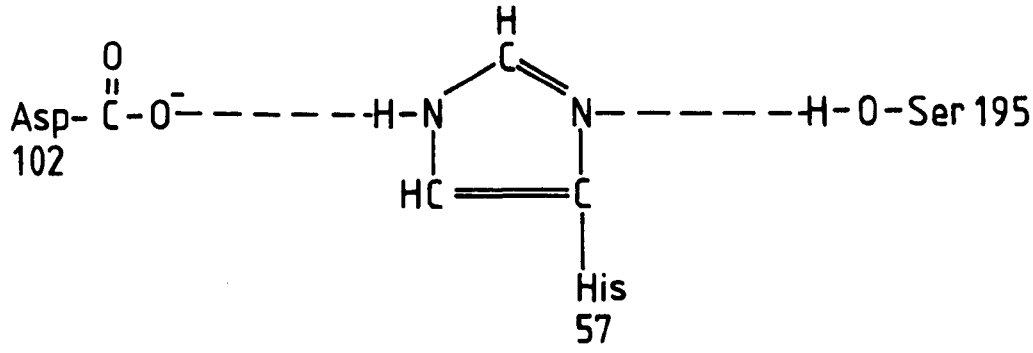


Reaction pathway for serine protease catalysis

FIGURE 4·0

The enzyme represented by E has two important functional groups the hydroxyl of serine-195 ($-CH_2OH$) and the imidazole of histidine-57 (Im) for catalysis.

From X-ray data, Blow et al[3] proposed that the catalytic centre of serine proteases (chymotrypsin) involved in "charge-relay system". This involved the three residues, the buried asparate-102 hydrogen bonded to histidine-57 which in turn hydrogen bonded to serine-195.



Charge relay system in elastase without the substrate.

In this case transfer of electrons from the buried carboxyl group to the oxygen of serine-195 occurs so that the hydroxyl becomes a powerful nucleophile and is able to attack the carbon atom of the substrate (figure 4.1).

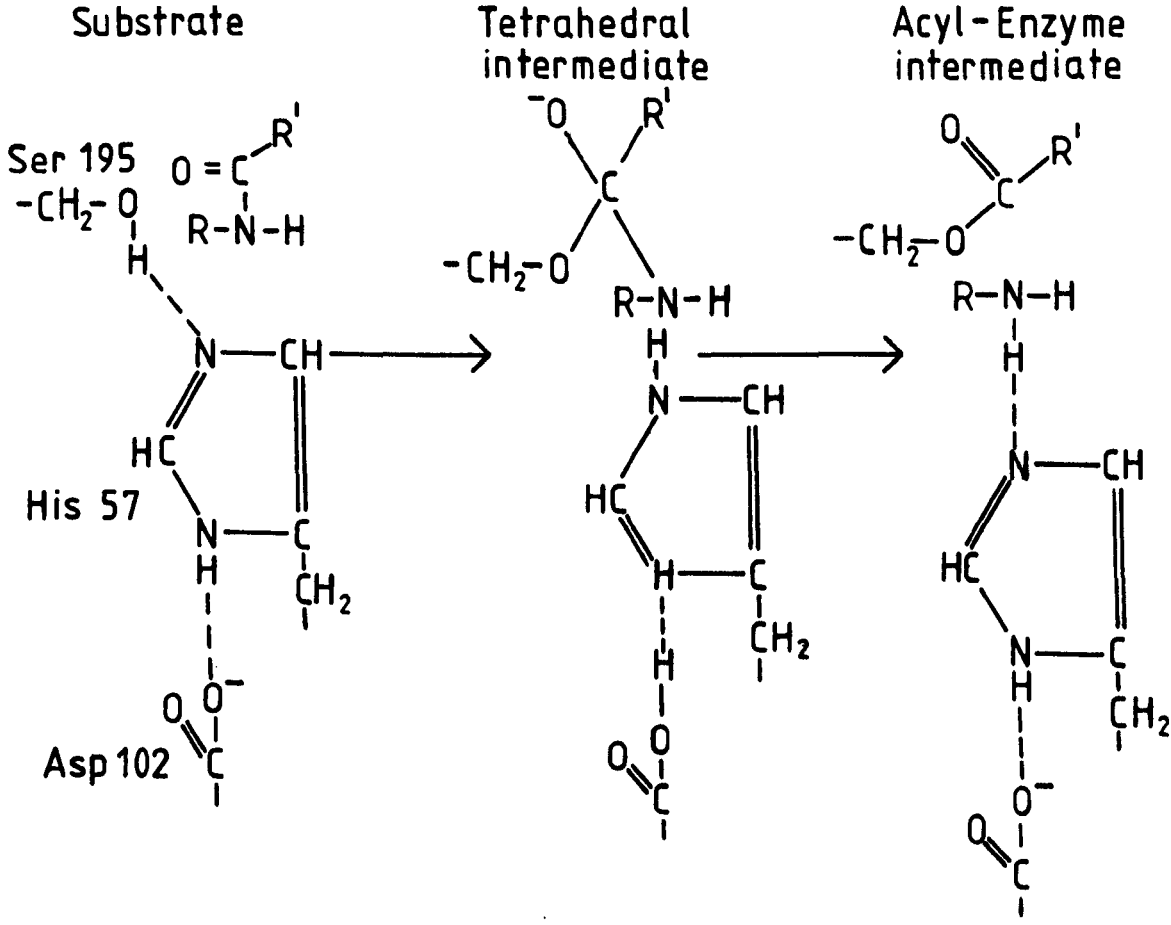


FIGURE 4·1

During this acylation step, which proceeds through a tetrahedral intermediate, a proton is transferred from the imidazole ring to the nitrogen atom of the substrate. In turn the carbonyl carbon atom binds covalently to the oxygen atom of the serine.

However, in 1980 James et al[2] proposed that there was no "charge-relay system" as such, which contributed to the catalytic rate enhancement of hydrolysis by the serine proteases. Their deductions were based upon SGPA which has a similar structure in the active centre to the pancreatic serine proteases. James et al[2] proposed that the intermediate came from the potentially strong electrostatic interaction provided by the two dipole moments of the oxyanion binding site. This enhances the polarization of the carbonyl group of the substrate. The carbonyl oxygen approach towards the oxyanion binding site is coupled with an enhancement of polarization of the peptide bond with the carbonyl carbon atom developing a partial positive charge. It is this positively charged carbon atom of the substrate which induces the departure of the hydroxyl proton of serine-195. The pathway that the proton takes is as mentioned previously, that is, the asparate-histidine couplet (figure 4.1). This is then followed by the collapse of the tetrahedral intermediate to the acyl-enzyme (figure 4.1).

The next stage in the catalysis is the deacylation step shown in figure 4.2.[2]

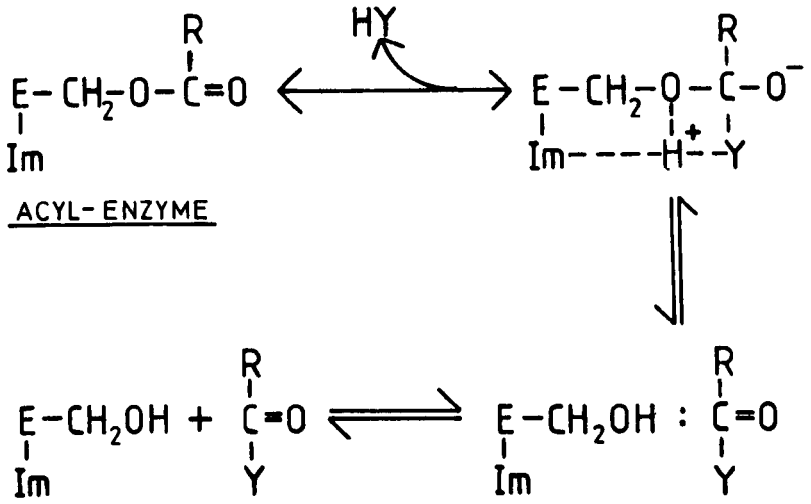


FIGURE 4·2

James et al[2] proposed that the acyl-enzyme formed has a pyramidal carbonyl carbon atom and the carbonyl oxygen atom remains in the strongly polarizing electrostatic environment of the oxyanion binding site. As in the acylation step, formation of an enzyme-substrate tetrahedral intermediate occurs by partial positive charge on the carbonyl carbon atom allowing the second substrate (water molecule) to dissociate a proton. This creates a strong nucleophilic hydroxide ion and finally the break down of the intermediate to the substrate and enzyme. Figure 4.3[4].

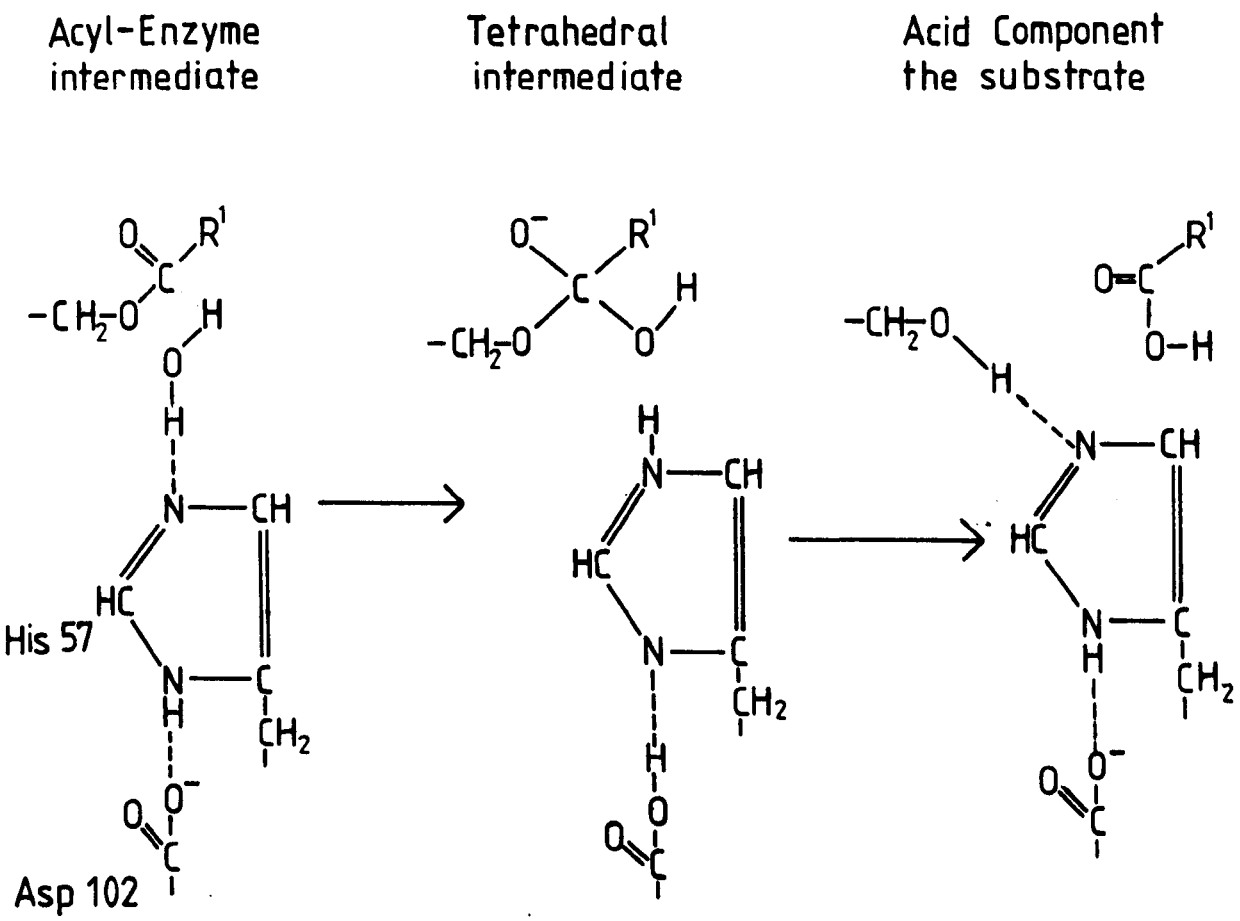

Acyl-Enzyme intermediate
Tetrahedral intermediate
Acid Component the substrate
His 57
Asp 102

FIGURE 4.3

4.2 THE EFFECTS OF ENZYME INHIBITORS ON SERINE PROTEASES

Enzyme inhibitors can be classified into two general categories : irreversible and reversible.

The irreversible inhibitors are so called because on binding to the protein they form covalent bonds, thus obstructing entry of the substrate either by steric hindrance at a site near the active site or within the active site itself. Excess substrate added to such a system will not remove the inhibitor which has now become irreversibly bound to the protein e.g. active site-directed inhibitors and succide inhibitors are usually of this type.

Reversible inhibitors, on the other hand, will bind non-covalently to the protein and addition of excess substrate in this case will remove the inhibitor. Generally, reversible inhibitors can be sub-divided into three groups as follows :

a) COMPETITIVE :

These compete with the substrate for binding to the active site but, once bound, cannot be transformed by the enzyme and vice-versa. Competitive inhibition can be reversed or relieved simply by increasing the susbtrate concentration.

These inhibitors generally resemble the substrate in three-dimensional structure and "trick" the enzyme into binding it. Michaelis-Menten theory can be used to analyse quantitatively the inhibition. The competitive inhibitor, I, binds exclusively to E and not enzyme-substrate complex [ES], to form enzyme-inhibitor complex [EI]. In this case V_{max} is unchanged whilst K_m is changed.

b) UNCOMPETITIVE :

These are very rare and the inhibitor only binds to the [ES]. The inhibitor therefore binds to a secondary site. Both V_{max} and K_m are changed.

c) NON-COMPETITIVE :

Substrate and inhibitor bind simultaneously to the enzyme rather than competing for the active site. V_{max} is decreased but K_m does not change.

In PPE, a number of synthetic irreversible inhibitors can form a stable bond with a catalytically important amino-residue, i.e. those which block serine-195 and those that alkylate histidine-57.

The former groups include DFP[5], PMSF, tosyl fluoride, p-chloromercuri-benzene-sulphonyl-fluoride[6], and p-fluoro-phenyl sulphonyl-fluoride[7]. The latter groups includes peptide chloromethyl-ketones[8,9].

Reversible inhibitors (peptide aldehydes, TFA-peptides) of PPE have been synthesised which are also thought to generate tetrahedral intermediate analogues to PPE catalysis[10]. In this case, the hemiacetal (figure 1.6) which is formed is relatively stable and breakdown of this complex will give an aldehyde. The aldehydes are unique among carbonyl compounds in preferring to exist as tetrahedral addition complexes.

The inhibition constant (K_i) for the various competitive inhibitors can be calculated from Dixon plots[11] using the substrate (NBA). This allows the use of very low concentrations of enzyme and the maintenance of a high inhibitor : enzyme ratios.

4.3 TO MEASURE $1/V_{max}$ AND K_m USING THE MICAHELIS-MENTEN EQUATIONS IN THE FORM OF A LINEWEAVER-BURK PLOT

The inhibition constants, K_i, for the inhibitors Ac-Pro-Ala-Pro-Alaninal (25) and TFA-Pro-Ala-Pro-Alaninal (28) (refer to Chapter Three for the inhibitor numbering) were evaluated from Dixon Plots[11] at two substrate concentrations using the intercept with $1/V_{max}$ which was determined by Lineweaver-Burk plot[12].

4.4 EXPERIMENTAL PROCEDURE :

Buffer : 50mM sodium phosphate, pH 6.5
Substrate : NBA
Enzyme : PPE (200 µg/ml)
Temperature : 30° C

The standard conditions were those of Visser and Blout[13]. A volume of 10mM stock solution of NBA (10, 15, 25, 50 and 100 µl) in spectroquality methanol was added to quartz reference and sample cuvettes (1-cm pathlength), containing 2.95 ml of buffer. The cuvettes were incubated for 5 minutes at 30° C before 10 µl of the stock enzyme solution was added to the sample cuvette. The rate of release of p-nitrophenol was measured for 3 minutes at 347.5 nm in a unicam SP 1800 ultraviolet spectrophotometer. The experiment was repeated until three consistent results for each substrate concentration were obtained.

4.5 EXPERIMENTAL PROCEDURE FOR INHIBITION CONSTANT, K_i, DETERMINATION

For the inhibitors (25) and (28). This time the measurements were carried out at two different substrate concentrations i.e. 10mM and 5mM.

4.5.1 A similar procedure to the above was used, but this time an exact volume of substrate (10mM, 100 µl) was added to 2.95 ml of buffer and the inhibitor volume was varied. The inhibitor volume used, from the stock solution (10mM) were 3, 10, 20, 30 and 50 µl.

4.5.2 As above, except this time 5mM substrate concentration was used.

4.6 RESULTS

FOR 4.4 :

[S] mM	Rate (V) ΔOD/min	1/[S] mM^{-1}	1/V Min
0.0337	0.0037	29.67	270.27
0.0504	0.0053	19.84	188.68
0.0837	0.0080	11.95	125.00
0.1661	0.0160	6.02	62.50
0.3268	0.0273	3.06	36.63

4.5.1 and 4.5.2 FOR INHIBITOR (25)

SUBSTRATE 10mM and 5mM

	Substrate 10mM		Substrate 5mM	
[I] µM	Rate (V)x 10^{-3} ΔOD/min	1/V Min	Rate(V)x 10^{-3} ΔOD/min	1/V Min
0.9775	14.870	67.25	7.000	142.86
3.2501	7.887	126.79	4.333	230.79
6.4805	5.667	176.46	3.333	300.00
9.6890	4.333	230.79	2.667	374.95
16.0450	3.333	300.00	2.000	500.00

4.5.1 and 4.5.2 FOR INHIBITOR (28)

SUBSTRATE 10mM and 5mM

[I]	Substrate 10mM Rate (V)x 10^{-3}	1/V	Substrate 5mM Rate(V)x 10^{-3}	1/V
μM	Δ OD/min	Min	Δ OD/min	Min
0.9650	2.467	40.54	1.600	62.50
3.2094	1.867	53.56	1.167	85.69
6.3981	1.717	58.24	1.133	88.26
9.5660	1.400	71.43	0.950	105.26
15.8408	1.100	90.90	0.500	200.00

4.6 RESULTS FROM THE GRAPH :

GRAPH 1 PPE	K_m (mM) 1.023 + 0.146	V_{max}(Min^{-1}) 0.111 + 0.013
Graph	Inhibitor	K_i (μM)
2	(25)	6.1 (5.8)*
3	(28)	8.6 (8.3)*

* corrected for impurities

4.7 DISCUSSION

The value K_m obtained (1.023 mM)[14], for the PPE varies significantly from the results of Visser and Blout[13] (K_m = 0.3 mM). This discrepancy is probably the result of a difference in temperature (not specified by Visser and Blout[13]) and for different batch of PPE used (Visser and Blout used PPE from Worthington Biochemical Corporation, whereas that used here was from Sigma Chemicals Ltd and was approximately three years old).

The inhibition constant, K_i, was measured for the two aldehydes are not significantly different. Comparing the results with those of Thompson et al[10] for Ac-Pro-Ala-Pro-Alaninal (25), the K_i value is significantly different, i.e.

Ac-Pro-Ala-Pro-Alaninal $K_i = 0.8 \mu M$, 37° C, pH 7.0,
Ac-Pro-Ala-Pro-Alaninal $K_i = 2.0 \mu M$, 37° C, pH 4.0.

However, it was noted by Thompson et al[10] that the K_i for aldehydes was not greatly affected by pH over the range of 4.00-7.00. The difference in the inhibition constant for (25) is probably the result of reaction conditions. The lower temperature than Thompson et al[10] could affect the binding of the inhibitor. Also they used a different substrate (acetyltrialanine methyl ester)[15] which has a difference K_m (0.44 mM) and $kcat/K_m$ (170,000 M^{-1} sec^{-1}) compared to a K_m (0.3 mM) and $kcat/K_m$ (19,000 M^{-1} sec^{-1}) which in turn would affect the K_i value.

The purity of the aldehyde will also affect the K_i since any impurities in the aldehyde would raise or lower the K_i value. The aldehyde prepared was approximately 94% pure, while Thompson et al[10] have not carried out any analysis to check the purity of their aldehyde.

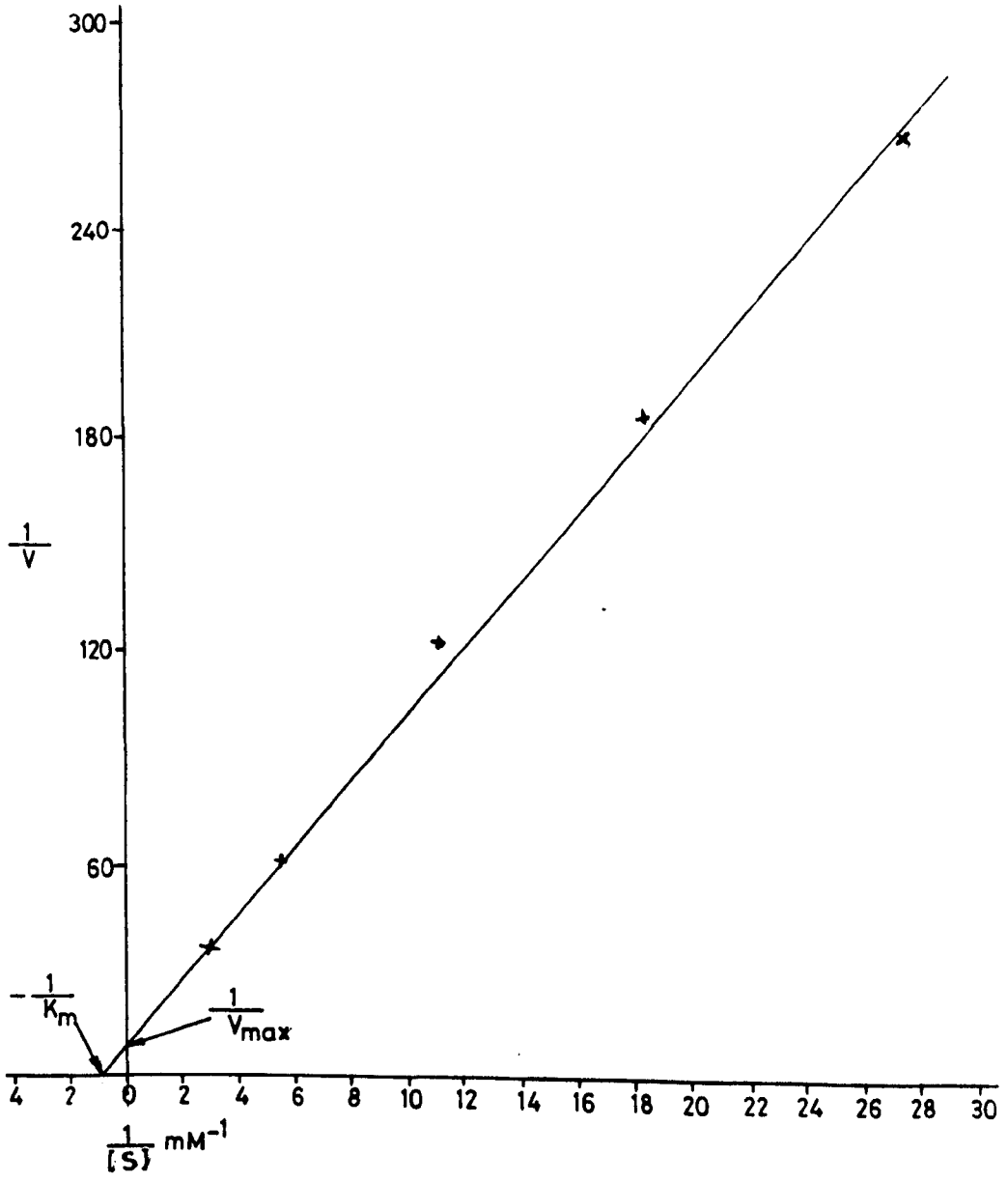
The inhibition constant for TFA-aldehyde (28) has not been reported elsewhere, but in comparison with other reversible TFA-peptides (table 4.0) it agrees quite well.

TABLE 4.0

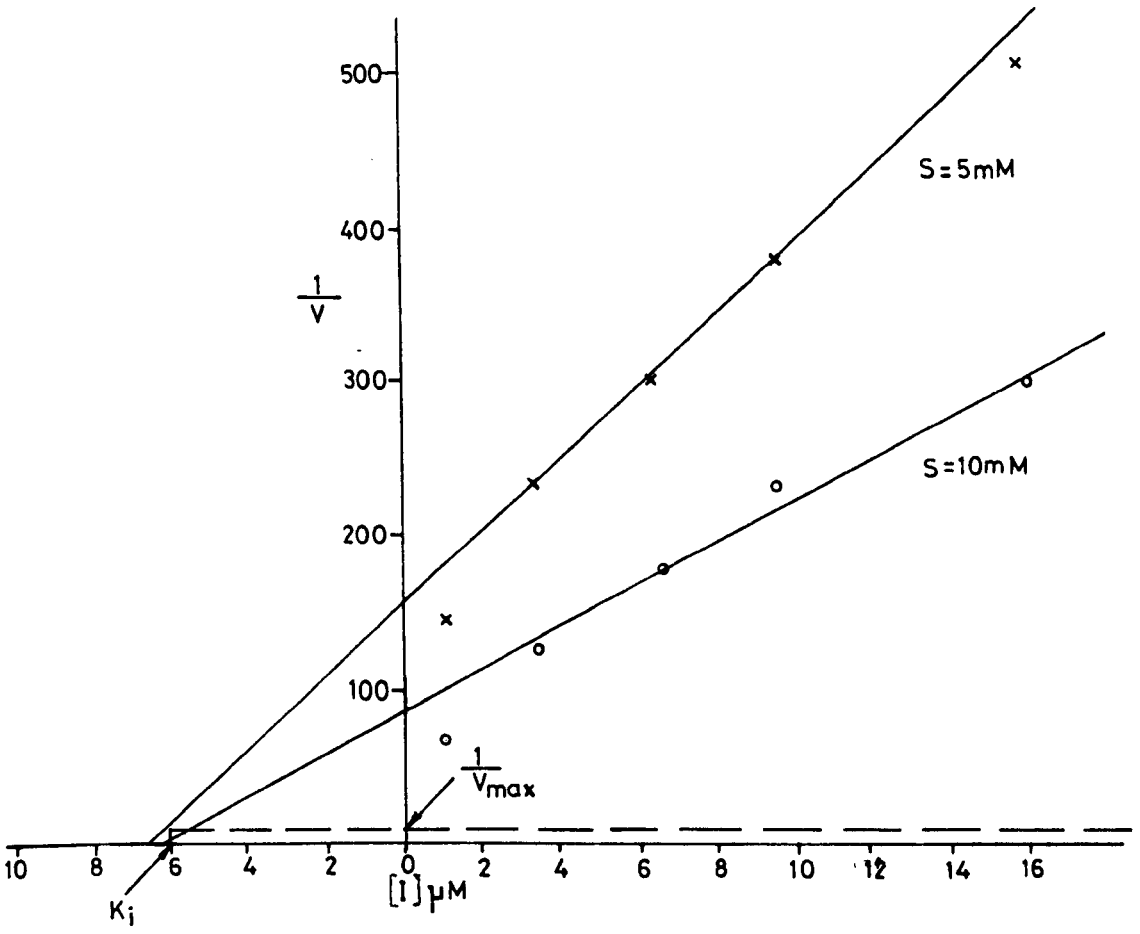
COMPARISON OF THE INHIBITION CONSTANT K_i, OF VARIOUS TFA-PEPTIDES WITH TFA-PEPTIDE ALDEHYDE

Peptide	K_i (μM)
CF_3CO-Ala-Ala-Ala.OH[7]	7.9
CF_3CO-Lys-Ala-Ala-Ala.OH[16]	10.0
CF_3CO-Pro-Ala-Pro-Alaninal	8.3

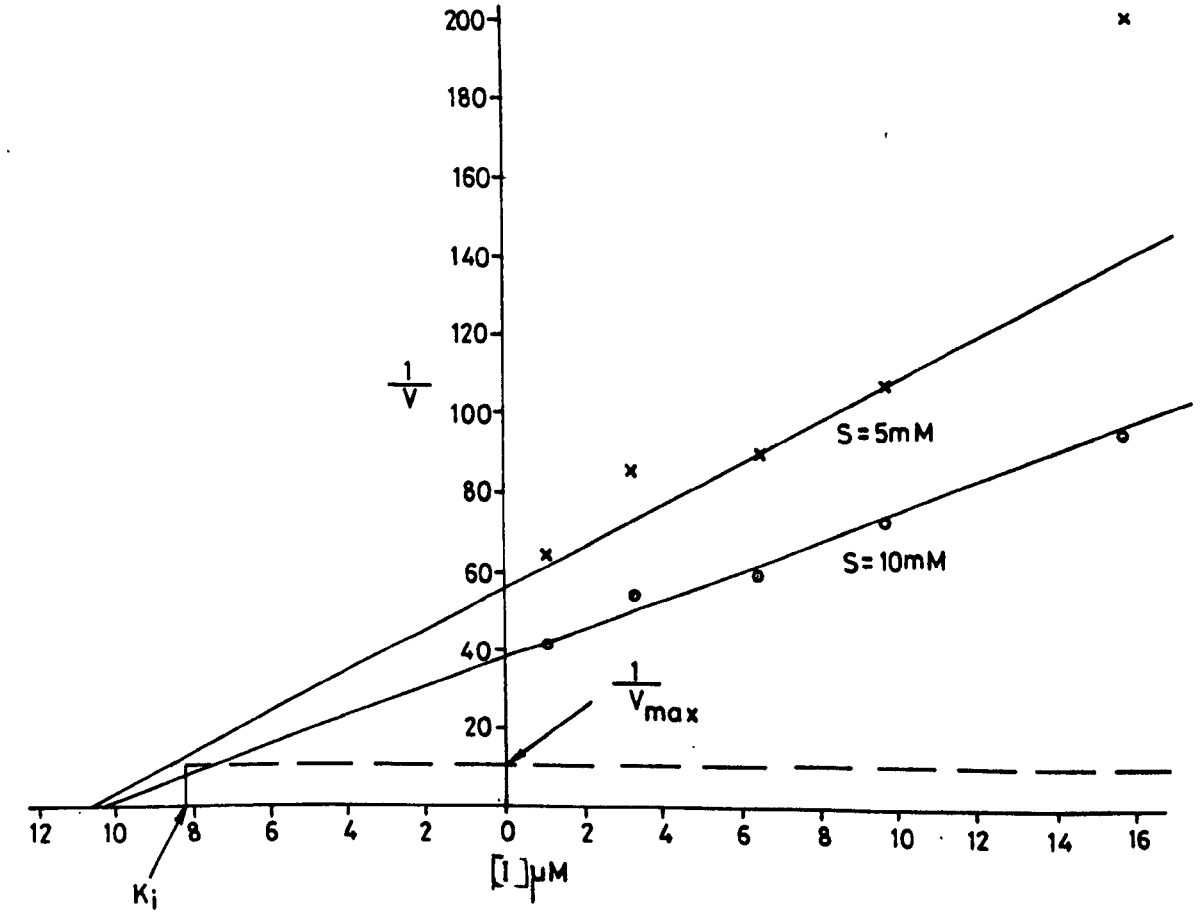
GRAPH 1: A DOUBLE-RECIPROCAL PLOT $\frac{1}{[S]}$ Vs $\frac{1}{V}$



GRAPH 2 : FOR INHIBITOR(25)



GRAPH3 : FOR INHIBITOR (28)



REFERENCES

1. Fersht, A.
Enzyme Structure and Mechanism. (1977).

2. James, M.N.G., Sielecki, A.R., Brayer, G.D., Delbaere, L.T.J. and Bauer, C.A.
J.Molec.Biol., 144, 43 (1980).

3. Blow, D.M., Birktoft, J.J., and Hartley, B.S.
Nature, 221, 337 (1969).

4. Stryer, L.
Biochemistry, 2nd Edition, Freeman. (1975).

5. Naughton, M.A., and Sanger, S.
Biochem.J., 78, 156 (1961).

6. Watson, H.C., Shotton, D.M., Cox, J.C., and Muirhead, H.
Nature, 225, 806 (1970).

7. Dimicoli, J.L., Bieth, J., and Lhoste, J.M.
Biochemistry, 15, 2230 (1976).

8. Powers, J.C., and Tuhy, P.M.
Ibid., 12, 4767 (1973).

9. Thompson, R.C., and Blout, E.R.
Ibid., 12, 47 (1973).

10. Thompson, R.C.
Ibid., 12, 44 (1973).

11. Dixon, M.
Biochem.J., 55, 170 (1953).

12. Lineweaver, H., and Burk, D.
J.Amer.Chem.Soc., 56, 658 (1934).

13. Visser, L., and Blout, E.R.
Biochim. et. Biophys. Acta., 268, 257 (1972).

14. Wilkinson, G.N.
Biochem.J., 80, 324 (1961).

15. Gertler, A., and Hoffmann, T.
Can.J.Biochem., 48, 384 (1970).

16. Dimicoli, J.L., Renaud, A., Bieth, J.
Eur.J.Biochem., 107, 423 (1980).

CHAPTER FIVE

HUMAN LEUKOCYTE ELASTASE

5.1 INTRODUCTION

HLE was first recognised as a serine protease by Janoff and co-workers[1]. This elastase is involved in pulmonary emphysema which causes progressive loss of lung elasticity due to the destruction of lung elastin[2]. This in turn causes respiration problems and eventually leads to death. The main group of sufferers from this disease, is people who are deficient in the serum protease inhibitor α_1-antitrypsin[3-8] which inhibits a variety of proteases including elastase and which is produced in the granules of human leukocytes[9-12]. α_1-antitrypsin, protects the lung from digestion by any of the proteases in normal person, but persons with insufficient inhibitor proteolyse their lung elastin and so develop emphysema. Other factors reported in the literature which aggravate this problem are smoking and air pollution.

Individuals who smoke cigarettes partially inactivate α_1-antitrypsin resulting in lung damage by elastase due to the presence of oxidant present in the smoke[13,14]. The cause for this is, the modification of the methionine-358 residue at the active centre of α_1-antitrypsin resulting in a dramatic decrease in the inhibitor activity towards elastase. This in turn inactivates the protective function[15].

Recently studies were carried out with the possibility of modulating the biological properties of α_1-antitrypsin by substituting the methionine-358 residue at the active centre to valine[16,17]. α_1-antitrypsin (Met$\xrightarrow{358}$Val) proved to be resistant to oxidation as while as an efficient inhibitor. Moreover, α_1-antitrypsin (Met$\xrightarrow{358}$Val) could be a particularly effective therapeutic agent in the treatment of α_1-antitrypsin deficiency or in protecting the lung in situation of increased oxygen burden[17].

HLE consists of four isoenzymes, which are all glycoproteins, with a carbohydrate content mainly composed of neutral sugar[18]. They have been shown to cleave structural proteins such as elastin[19], azo-casein[20], collagen and proteoglycan[21]. HLE has been purified from leukocytes[22-26], obtained by leukaphoresis from patients with chronic leukemia or from normal donors[1]. The "buffy-coat residues" containing the leukocytes can also be obtained from outdated blood. Other useful sources of this form of HLE are purulent sputum[27] and also the human spleen[28]. The elastases obtained from these sources are immunologically identical to that of leukocytes.

The HLE used in this work, a gift from Dr Roberts,[25] was found to be homogeneous by SDS electrophoresis and had a molecular weight of 22,000 by gel-filtration. This was in agreement with other workers (20,000 Starkey and Barrett[28], 22,000 Taylor and Crawford[24]). However, it was inconsistent with other data obtained by SDS electrophoresis (30,000 Baugh and Travis[18]). The reasons for these discrepancies in the molecular weight reported for HLE, (for which one expects a molecular weight of approximately 25,000 for the 210 amino-acids plus 20% for the carbohydrate content giving a molecular weight of approximately 30,000[23,29]) is probably due to gel-filtration. There is also a disagreement in the percentage of neutral sugar in the isoenzyme, E_4. A 22.8% sugar content has been reported by Baugh and Travis[18], and Taylor and Crawford[24] of which 21.2% was due to hexose. Twumasi and Liener[27] have obtained half this value and also the figure obtained by Roberts et al[25] is considerably lower. The isoenzyme, E_4 has an isoelectric point in the range of 8.77 - 9.14[30].

5.2 COMPARISON WITH PPE

The amino-acid compositions of E_4, obtained by various workers have been compared with that of PPE. Table 5.1 shows the amino-acid residues per molecule and also as a percentage of the total amino-acids.

TABLE 5.1

AMINO-ACID COMPOSITION OF PPE AND HLE REPORTED IN LITERATURE

AMINO-ACID	PPE(31)	REF(27)	REF(18)	PPE(31)	REF(18)	REF(27)	REF (23)
	RESIDUE/MOL			% MOL AMINO-ACID			
Aspartic acid	24	22	24	10.0	10.5	10.5	9.6
Threonine	19	6	7	7.9	3.1	2.9	2.8
Serine	22	12	13	9.2	5.7	5.7	5.2
Glutamic acid	19	16	18	7.9	7.9	7.6	7.0
Proline	7	9	10	2.9	4.4	4.3	7.4
Glycine	25	25	28	10.4	12.2	11.9	11.2
Alanine	17	22	24	7.1	10.5	10.5	11.1
Half-Cystine	8	4	6	3.3	2.6	1.9	-
Valine	27	23	25	11.3	10.9	11.0	11.4
Methionine	2	2	2	0.8	0.9	0.9	0.7
Isoleucine	10	9	11	4.2	4.8	4.3	4.2
Leucine	18	21	20	7.5	8.7	10.0	9.7
Tyrosine	11	3	3	4.6	1.3	1.4	1.1
Phenylalanine	3	9[32]	9[33]	1.3	3.9	4.3	3.9
Tryptophan[a]	7	2	2	2.9	0.9	0.9	-
Histidine	6	5	4	2.5	1.7	2.4	2.8
Lysine	3	1	1	1.3	0.4	0.5	0.5
Arginine	12	19	22	5.0	9.6	9.0	11.7
	240	210	229	100	100	100	100

[a] Determined spectorphotometrically.

The amino-acid compositions for HLE when compared with PPE, show quite a difference in the total amino-acids per molecule, but agree most closely when expressed as the number of residues/mol of protein. Compared to PPE there are major differences in the individual amino-acids i.e., threonine, serine, half-cystine, tyrosine, phenylalanine, tryptophan and arginine. The majority of these amino-acids are bulky groups and may affect the specificity around the active centre.

Another major difference is the disulphide bond content which plays an important role in stabilising the three-dimensional structure. HLE shows two or three disulphide bridges as indicated from the amino-acid composition compared to four in PPE, where the whole molecule is compact and globular. However, HLE shows the same number of disulphide bridges as SGPA and SGPB[34,35]. These bacterial enzymes have the same features in the active site region as PPE[34], but one or two important deletions. Hence, with the differences in amino-acid and the carbohydrate content, the specificity and the catalytic activity for substrate and inhibitors are likely to be affected.

Various studies comparing the specificity of HLE with PPE have been made using peptide chloromethyl ketone inhibitors[19,36,37], synthetic substrates[38,39] and the insulin B chain[40]. Peptide chloromethyl ketones have proved to be very effective inhibitors of PPE[41,42]. Hence, several tri and tetra peptide chloromethyl ketones have been prepared by research groups[36,37,41] and tested on HLE.

On the basis of the result obtained by Tuhy et al[36] that Ac-Ala-Ala-Pro-Ala-CMK was the most effective inhibitor for PPE and HLE. Powers et al[37] synthesised various inhibitors with the sequence Ac-Ala-Ala-Pro-X-CMK (X = Ile, Val and Thr) and they compared the inhibition of both elastases.

The inhibition rates obtained for PPE and HLE are listed in table 5.2 and 5.3, for HLE rates (at pH 6.5) were also considered relative to kobsd/[I] for Ac-Ala-Ala-Pro-Ala -CMK[36]. The P_5 residue in the sequence was also changed from acetyl to a succinyl (Suc) and methyl succinyl (MeO-Suc).

TABLE 5.2

INHIBITION OF PPE WITH PEPTIDE CHLOROMETHYL KETONES

Inhibitor P_5 P_4 P_3 P_2 P_1	[I] (mM)	kobsd/[I] (M^{-1} Sec^{-1})	kobsd/[I] (rel)
Ac-Ala-Ala-Pro-Ala-CMK	0.05	40*	1.0
Ac-Ala-Ala-Pro-Val-CMK	0.05	35*	0.9
Suc-Ala-Ala-Pro-Val-CMK	0.05	73*,**	1.8
MeO-Suc-Ala-Ala-Pro-Val-CMK	0.05	55*,**	1.4
Ac-Ala-Ala-Pro-Ile-CMK	0.05	48*	1.2
Ac-Ala-Ala-Pro-Thr-CMK	5.00	0.08*	0.002

PPE concentration 5 μM, 0.05M acetate buffer, pH 5.0, 5% (v/v) methanol, 30° C

* Average of three runs

** No methanol was used in the buffer

TABLE 5.3

INHIBITION OF HLE WITH PEPTIDE CHLOROMETHYL KETONES

Inhibitor P_5 P_4 P_3 P_2 P_1	pH	[I] (mM)	kobsd/[I] (M^{-1} Sec^{-1})	kobsd/[I] (rel)
Ac-Ala-Ala-Pro-Ala-CMK	6.5	0.20	3.4	
	6.5	0.40	3.3	1.0
	7.5	0.11	9.7	
Ac-Ala-Ala-Pro-Val-CMK	6.5	0.05	160*	49
	7.5	0.10	219	
Suc-Ala-Ala-Pro-Val-CMK**	6.5	0.025	320	97
	7.5	0.025	1400	
MeO-Suc-Ala-Ala-Pro-Val-CMK**	6.5	0.025	922	280
	7.5	0.025	1560	
Ac-Ala-Ala-Pro-Ile-CMK	6.5	0.05	133***	40
Ac-Ala-Ala-Pro-Thr-CMK	6.5	10.0	0.15	0.05
Z-Gly-Leu-Phe-CMK	7.5	0.10	0[a]	

HLE concentration 10 μM, 0.1M phosphate, 0.06M NaCl pH 6.5, 5% (v/v) methanol, 30° C

HLE concentration 21 μM, 0.05M phosphate, 0.035M NaCl, pH 7.5, 5% (v/v) methanol, 30° C

* Average of three runs

** 0.1M phosphate, 0.035M NaCl, elastase 0.94 μM, 30° C

*** Average of two runs

[a] No measurable activity after 4 days period.

From tables 5.2 and 5.3 it was concluded that threonine compounds were unreactive towards both elastases. For HLE, the valine and isoleucine chloromethyl ketones were respectively 49 and 40 times more reactive at pH 6.5 than the alanine chloromethyl ketones. However, at pH 7.5 the difference between valine and alanine was less marked. These results compare well with the substrates Ac-Ala-Ala-Pro-X-p-nitroanilide hydrolysis in table 5.4[38].

TABLE 5.4

HYDROLYSIS OF SUBSTRATES BY BOTH ELASTASES

Substrates : Ac-Ala-Ala-Pro-X-p-nitroanilide elastases, rates relative to Ala is 100.

Enzyme	X					
	Ala	Gly	Val	Leu	Ile	Phe
PPE	100	1	5	25	0	0
HLE	100	2-20	800	20	150	0

Also in this case, the substrates with valine and isoleucine were hydrolysed 8 and 1.5 times faster respectively than Ala-p-nitroanilide by HLE. Hence the overall conclusions which emerge from these comparisons between the HLE and PPE subsite specificities are : firstly, HLE prefers valine to alanine at P_1. This indication has also been observed in the digestion of the oxidised insulin B chain with HLE[40]. It was directed towards valine compared to alanine with PPE[43]. Hence the S_1 subsite prefers larger side chains than PPE. Secondly, subsite S_5 for HLE is observed to be structurally different, since it interacts more favourably with Suc or Meo-Suc than acetyl. From, table 5.2 the inhibition rate of PPE is not affected by this change. Powers et al[37] found that the most effective inhibitor for HLE was MeO-Suc- Ala-Ala-Pro-Val-CMK (table 5.3).

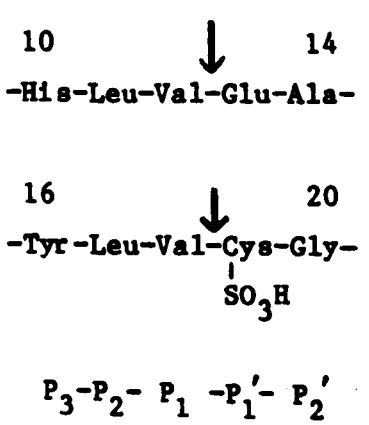
Other subsites have also been compared by Tuhy et al[36] with tri and tetra peptide chloromethyl ketones inhibitors, table 5.5.

TABLE 5.5

INHIBITION OF HLE AND PPE WITH PEPTIDE CHLOROMETHYL KETONES

Inhibitor	[I]	kobsd/[I] (M^{-1} Sec^{-1})	
P_4 P_3 P_2 P_1	(mM)	HLE	PPE[41,44]
Ac-Ala-Ala-Ala-CMK	1.0	1.0	1.0
Ac-Ala-Pro-Ala-CMK	1.0	4.0	1.6
Ac-Pro-Ala-Ala-CMK	1.0	0	0
Z-Gly-Leu-Ala-CMK	1.0	7.3	1.9
Z-Gly-Leu-Phe-CMK	1.0	0	0
Ac-Ala-Ala-Ala-Ala-CMK	0.2	3.5	18
Ac-Ala-Ala-Pro-Ala-CMK	0.2	15.8	71
Ac-Ala-Pro-Ala-Ala-CMK	0.2	8.1	0
Ac-Ala-Ala-Phe-Ala-CMK	0.2	4.8	2.2
Ac-Phe-Gly-Ala-Leu-CMK	0.2	2.2	-

Tetrapeptides were found to be better inhibitors than the tripeptides with 5-fold lower inhibitor concentrations. From the inhibition studies, it is evident that an extended binding site in HLE exists. Position P_2 is favoured by proline rather than alanine, from table 5.5, Ac-Ala-Ala-Pro-Ala-CMK has 4 to 5 time larger kobsd/[I] than Ac-Ala-Ala-Ala-Ala-CMK. A similar indication is obtained in the case of tripeptides. Leucine is also accommodated in position P_2 as was also observed in the hydrolysis of insulin B chain[40].



The subsite S_3 for HLE was found to differ from PPE i.e. the subsite was able to accommodate proline at this position if the substrate or inhibitor was a tetrapeptide but not a tripeptide. However, in the case of PPE, proline at this subsite caused no inhibition[41,44].

When the overall inhibitions are compared on the basis of the value of kobsd/[I][41,44], HLE is inhibited less rapidly with peptide chloromethyl ketones than is PPE. In going from tri- to tetrapeptides there is a considerable increase in the inhibition for PPE but not for HLE. Hence, this indicates differences in the detailed geometry of the extended binding sites and hence on the catalytic activity of the two elastases.

Starkey and Barrett[19] have also carried out inhibition studies using various protein inhibitors from plant and animal tissues, table 5.6.

TABLE 5.6

INHIBITION OF HLE AND PPE WITH VARIOUS PROTEIN INHIBITOR

Enzyme activity is expressed as a percentage of activity in the absence of inhibitor.

Inhibitor	Final concentration (mM)	Activity (%) HLE	PPE
Soya-bean trypsin inhibitor (Kunitz)	1 mg/ml	3	37
	0.1 mg/ml	27	67
Lima-bean trypsin inhibitor	0.1 mg/ml	63	78
Bovine pancreatic trypsin inhibitor (Kunitz)	0.5 mg/ml	35	92
Turkey ovomucoid	0.1 mg/ml	8	4
Chicken ovoinhibitor	0.1 mg/ml	14	0

From table 5.6, soya-bean trypsin inhibitor (Kunitz) and Lima-bean trypsin inhibitor are only moderate inhibitors of PPE but powerful inhibitors of HLE as are α_2-macroglobulin and α_1-antitrypsin from human plasma[23]. HLE is also inhibited by turkey ovomucoid and chicken ovoinhibitor, which have also shown to inhibit PPE[45].

Hence, it is very important to explain the specificity of HLE by structural studies. This allows one to distinguish the major differences and the similarities known from solution studies and also to compare with PPE.

5.3 CRYSTALLISATION OF HLE

The crystallisation technique that has so far been employed, vapor diffusion[46], is ideal for scanning a large number of possible conditions with a minimum of material (typically 200 trials on 10mg enzyme). The presence of carbohydrate was expected almost certainly to give a different form of crystal. This would lead to an independent structure determination and interpretation of the electron density would show the full sequence of HLE which is not yet available.

5.4 EXPERIMENTAL

Glass-distilled water and "AnalaR" grade reagents were used throughout, except for polyethylene-glycol 6000 and 2-methyl -2,4-pentane diol. Buffer solutions were prepared according to the tables published in "Data of Biochemical Research"[11,47]. A radio-meter pH M 7010 with glass electrode was used for all pH measurements. All the inhibitors and substrates used were a gift from D M Shotton.

The following experimental conditions have been tried.

A. PRECIPITATING AGENTS	B. TEMPERATURE	C. DIVALENT ION
Ammonium sulphate	0^{o} C	Mg^{2+}
Ammonium citrate	4^{o} C	Ca^{2+}
Sodium chloride	9^{o} C	Zn^{2+}
Magnesium sulphate		None
Sodium sulphate		
Polyethylene glycol 6000		
2-methyl-2,4-pentane diol		

D. BUFFERS AND pH's

0.05M Sodium phosphate buffer	pH 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5 and 9.0
0.05M Tris/HCl buffer	pH 6.5, 7.5, 8.0, 8.5 and 9.0
0.10M Sodium acetate buffer	pH 5.0, 6.0 and 6.7

E. INHIBITORS AND SUBSTRATES

L-Alanine
$(Ala)_4$
$Z\text{-}Ala_3\text{.}OH$
$H.Pro\text{-}Leu\text{-}Gly\text{-}NH_2$
Ac-Val-Val.OH
H.Ala-Pro-Ala.OH

The concentration of inhibitor or substrate used was 0.1mM in all cases.

The enzyme concentration was kept constant at 8.57 mg/ml. All buffers and solutions were filtered through Millipore cellulose acetate membrane filters (size : 13mm, grade 0.22 µm).

5.5 RESULTS

The precipitating agent that seemed most hopeful from the above list was ammonium sulphate at 4^{o} C. The divalent ions Mg^{2+} tended to produce crystals of salt in phosphate and Tris buffers but ca^{2+} appeared to enhance crystal growth. None of the inhibitors or substrates used produced very hopeful results.

The optimum conditions found where small sized crystals grew were in the pH range 8.0-9.0 with 0.05M sodium phosphate or 0.05M Tris/HCl buffer and between 1.5-2.35M ammonium sulphate concentration. Crystal growth appeared to take about five days. Similar conditions were used where the enzyme was substituted by the appropriate buffer in which the enzyme was dissolved. These blank determinations indicated no crystal growth. These experimental conditions have been repeated several times and also with two different batches of enzyme with similar results.

The experimental conditions were varied to try to increase the size of crystal to that suitable for data collection, i.e., varying the pH between 8.0 and 9.0 with both buffers in steps of 0.2 and the ammonium sulphate concentration in steps of 0.1M. However, this did not appear to enhance crystal growth.

Sealed tube X-ray experiments were carried out on the small crystals but did not reveal a diffraction pattern. The synchroton at Daresbury was also used but no diffraction pattern was obtained.

Finally, studies were carried out to remove some or all of the sugar content using various glycosidases at different pH values, in an attempt to enhance crystallisation.

5.6 PROCEDURE CARRIED OUT TO REMOVE THE SUGAR CONTENT

All the glycosidases : α-Glucosidase (EC 3.2.1.20), β-Glucosidase (EC 3.2.1.21), α-Galactosidase (EC 3.2.1.22), β-Galactosidase (EC 3.2.1.23), α-Mannosidase (EC 3.2.1.24), β-Glucoronidase (EC 3.2.1.31), Almond emulsin and all the materials for gel electrophoresis were all supplied by Roche Products Ltd., Welwyn Garden City.

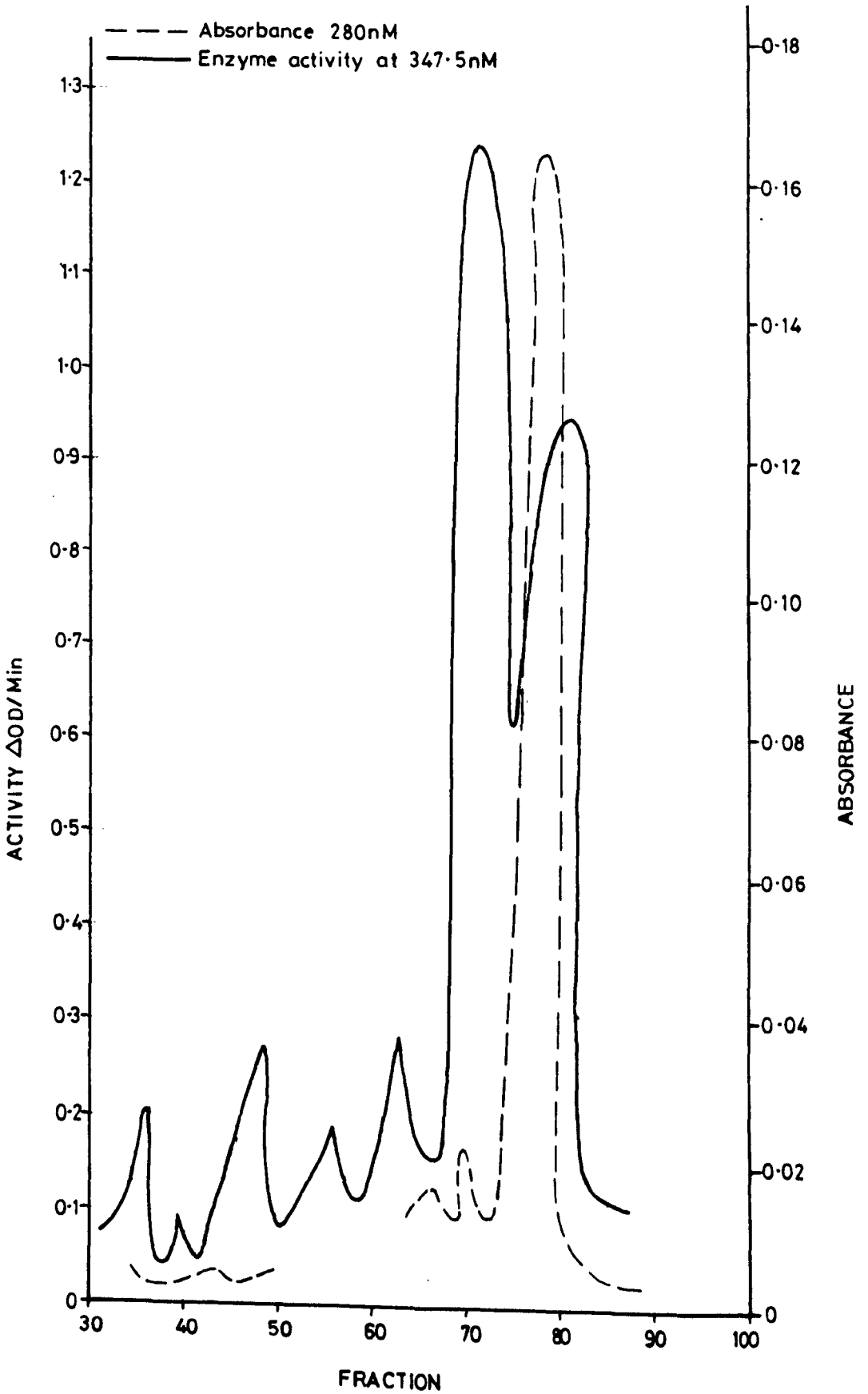
HLE was purified from the crude product supplied by Dr Roberts, which was obtained from purulent sputum[27]. The dried material (30.5 mg) was dissolved in a small volume of 0.05M Tris/HCl pH 7.5 buffer containing 0.2M NaCl and applied to a column (25 x 2 cm) of carboxymethyl cellulose (Whatman CM52) which had been equilibrated and washed with the same buffer at 4^{o} C. After sample application the column was washed with 50 ml of buffer containing 0.2M Nacl followed by a linear salt gradient 0.2 to 1M NaCl in 0.05M Tris/HCl pH 7.5 (2 x 250 ml). The eluate was monitored for extinction at 280 nm, collected in 4 ml fractions at 4^{o} C and screened for elastase activity using NBA as a substrate[48].

The four isoenzymes were pooled separately from the elution profile shown on the graph 5.0. Fraction 75-85, isoenzyme E_4; Fraction 68-73, isoenzyme E_3; Fraction 62-67, isoenzyme E_2; and Fraction 45-51, isoenzyme E_1. The fractions up to 34 showed no elastase activity.

The isoenzymes were dialysed separately against distilled water before freeze-drying. Percentage recovery for isoenzyme E_4, was $\simeq$ 43%.

GRAPH 5.0

SEPARATION OF HLE ISOENZYMES BY CM52 ION-EXCHANGE CHROMATOGRAPHY



5.7 POLYACRYLAMIDE GEL ELECTROPHORESIS ON THE ISOENZYMES

Slab gel electrophoresis was carried out in 7.5% cross-linked polyacrylamide gels at pH 4.3. 100g of protein per track was loaded and electrophoresis performed at 5mA per track for 130 minutes, migration was towards the cathode. Standard HLE prepared by Dr Roberts[25] was also loaded. Gels were stained with 0.05% coomassie blue in 10% methanol/acetic acid for 30 minutes before destaining in 10% acetic acid/ methanol overnight.

The isoenzyme E_4 showed a single band identical to that of the standard HLE, pools containing E_2 and E_3 showed several bands, while there was no indication of protein band for isoenzyme E_1.

5.8 PROCEDURE FOR TREATING ISOENZYME E_4 WITH GLYCOSIDASES

The glycosidases were divided into three groups :

(1) β-Glucosidase, α-Galactosidase and α-Mannosidase at pH 4.5 in 0.1M citric acid/0.2M sodium dihydrogen phosphate buffer.

(2) α-Glucosidase, β-Galactosidase and β-Glucoronidase at pH 6.8 in 0.075M potassium phosphate buffer.

(3) Almond meal in 20 mM acetate buffer pH 5.0. This was only purified through stage one by the procedure of Taga et al[49].

Isoenzyme, E_4 was only treated with the glycosidases. The concentration of isoenzyme E_4 used was 100 µg/ml and the reaction carried out in duplicate. The isoenzyme, E_4 was treated with the inhibitor R03-9782 (Roche inhibitor) so that it did not denature when incubated at 37° C for 24 hours for group 1 and 2, and 4 hours for group 3.

To 500 µl of isoenzyme, E_4, 5 µl of 1 mg/ml of glycosidase (except for almond meal 50 µl of 1 mg/ml) were added and incubated. Control isoenzyme, E_4 and control, glycosidase mixture were also incubated. After incubation they were dialysed at 4^{o} C for 4 hours against distilled water to remove the inhibitor. Slab gel electrophoresis was then carried out in 7.5% cross-linked polyacrylamide at pH 4.3.

The isoenzyme E_4, treated with three groups of glycosidases showed no change in the protein migration bands. Hence, it was decided to determine the sugar content left in the isoenzyme E_4, by the method of Dubois et al[50] using the standard curve of D-galactosidase. The procedure already described was carried out in duplicate and the protein content determined before and after dialysis. The only conclusive result from this, was the isoenzyme E_4, treated with almond meal showed some loss in sugar content. However, it was decided that the almond emulsin would need to be purified through all the stages of Taga et al[33] before treating isoenzyme E_4.

Since this work was carried out during a short stay at the collaborating institute it was not possible to pursue this. It however is, hoped that in the future this experimental conditions can be pursued further.

The optimum conditions found previously at the crystallisation stage were not tried at crystallising the treated isoenzyme E_4, but could be tried once successful in removing the sugar content. This would eventually open up a whole host of studies relating to the treatment of disease, and comparison of binding studies with PPE.

REFERENCES

1. Janoff, A., and Scherer, J.
J.Exp.Med., 128, 1137 (1968).

2. Powers, J.C.
Trends in Biochem.Sci., 211 (1976).

3. Eriksson, S.
Acta.Med.Scand., 175, 197 (1964).

4. Kuhn, C., and Senior, R.M.
Lung., 155, 185 (1978).

5. Gadek, J.E., Hunninghake, G.W., Fells, G.A., Zimmerman, R.L.,
Keogh, B.A., Crystal, R.G.,
Bull.Europ.Physiopath.Resp., 16, (suppl), 27 (1980).

6. Kueppers, F.
Genetic Determinants of Pulmonary Disease., edition.
Litwin, S.D., Dekker, New York, 23 (1978).

7. Morse, J.O.
New.Eng.J.Med., 299, 1045 (1978).

8. Larsson, C.
Acta.Med.Scand., 204, 345 (1978).

9. Carrell, R.W., Jeppsson, J.O., Laurell, C.B., Brennan, S.O., Owen, M.C., Vaughan, L., and Bosewell, D.R.
Nature, 298, 329, (1982).

10. Beatty, K.G.
Ph.D. Thesis, University of Georgia (1982).

11. Gadek, J.E, and Crystal, R.G.
The Metabolic Basis of Inherited Diseases,
5th Edition, Stanbury, J.B., Wyngaarden, J.B.,
Fredrickson, D.S., Goldstein, J.L.,
and Brown, M.S., McGraw-Hill, New York, 1450 (1981).

12. Gadek, J.E., Klein, H.G., Holland, P.V., and Crystal, R.G.
J.Clin.Invest., 68, 1158 (1981).

13. Gadek, J.E., Fells, G.A., and Crystal, R.G.
Science, 206, 1315 (1979).

14. Carp, H., Miller, F., Hoidal, J.R., and Janoff, A.
Proc.Nath.Acad.Sci., U.S.A., 79, 2041 (1982).

15. Beatty, K., Beith, J. and Travis, J.
J.Biol.Chem., 255, 3931 (1980).

16. Rosenberg, S., Barr, P.J., Najarian, R.C., and Hallwell, R.A.
Nature, 312, 77 (1984).

17. Courtney, M., Jallat, S., Tessier, L.H., Benavente, A., and Crystal, R.G.
Ibid., 313, 149 (1985).

18. Baugh, R.J., and Travis, J.
Biochemistry, 15, 836 (1976).

19. Starkey, P.M., and Barrett, A.J.
Biochem.J., 155, 265 (1976).

20. Starkey, P.M., Barrett, A.J., and Burleigh, M.C.
Biochim.et.Biophys.Acta. 483, 386 (1977).

21. Raughley, P.J., and Barrett, A.J.
Biochem.J., 167, 629 (1977).

22. Janoff, A.
Lab.Invest., 29, 458 (1973).

23. Ohlsson, K., Olsson, I.
Eur.J.Biochem., 42, 519 (1974).

24. Taylor, J.C., and Crawford, I.P.
Arch.Biochem.Biophys., 169, 91 (1975).

25. Roberts, N.A., Budd, J.M., and Robinson, P.A.
Unpublished Work (1977).

26. Schmidt, W., and Havemann, K.
Hoppe-Seyler's Z.Physiol.Chem., 355, 1077 (1974).

27. Twumasi, D.Y., and Liener, I.E.
J.Biol.Chem., 252, 1917 (1977).

28. Starkey, P.M., and Barrett, A.J.
Biochem.J., 155, 255 (1976a).

29. Campbell, G.J., Senior, R.M., McDonald, J.A., and Cox, D.J.
J.Clin.Invest., 70, 845 (1982).

30. Taylor, J.C., and Tlougan, J.
Anal.Biochem., 90, 481 (1978).

31. Shotton, D.M.
Methods in Enzymology, 19, 113 (1970).

32. Goodwin, T.W., and Morton, R.A.
Biochem.J., 40, 628 (1946).

33. Benee, W.L., and Schmid, K.
Anal.Biochem., 29, 1193 (1957).

34. Johnson, P., and Smillie, L.B.
Febs.Lett., 47, 1 (1974).

35. James, M.N.G., Delbaere, L.T., and Brayer, G.D.
Can.J.Biochem., 56, 396 (1978).

36. Tuhy, P.M., and Powers, J.C.
Febs.Lett., 50, 359 (1977).

37. Powers, J.C., Gupton, B.F., Hartley, A.D., Nishino, N., and Whitley, R.J.
Biochim.et.Biophys.Acta., 485, 156 (1977).

38. Zimmerman, M., and Ashe, D.M.
Ibid., 480, 241 (1977).

39. Nakajima, K., and Powers, J.C.
J.Biol.Chem., 254, 4027 (1979).

40. Blow, A.M.J.
Biochem.J., 161, 13 (1977).

41. Powers, J.C., and Tuhy, P.M.
Biochemistry., 12, 4767 (1973).

42. Thompson, R.C., and Blout, E.R.
Ibid., 12, 44 (1973).

43. Sampath-Narayanan, A., and Anwar, R.A.
Biochem.J., 114, 11 (1969).

44. Powers, J.C., and Tuhy, P.M.
J.Amer.Chem.Soc., 94, 6544 (1972).

45. Gertler, A., and Feinstein, G.
Eur.J.Biochem., 20, 547 (1971).

46. McPherson, A., Jr.
Methods of Biochemical Analysis., 23, 249.

47. Dawson, R.M.C., Elliot, D.C., Elliot, W.H., and Jones, K.M.
Data of Biochemical Research, 2nd Edition, Oxford University Press (1969).

48. Visser, L., and Blout, E.R.
Biochim.et.Biophys.Acta., 268, 257 (1972).

49. Taga, E.M., Waheed, A., and Van Etten, R.L.
Biochemistry, 23, 815 (1984).

50. Dubois, M., Gilles, K.A., Hamilton, J.K., Rebers, P.A., and Smith, F.
Anal.Chem., 28, 350 (1956).

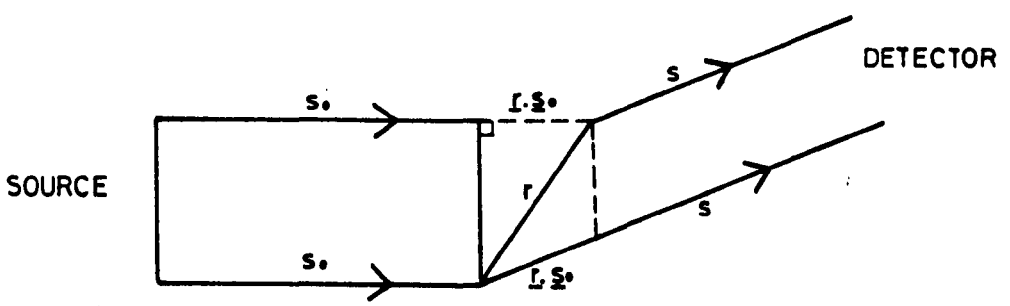
CHAPTER SIX

X-RAY CRYSTALLOGRAPHY

6.1 INTRODUCTION

The most powerful technique available to study the structures of small and large molecules is X-ray diffraction, since it provides a three-dimensional model of the structure. The normal procedure involved in the determination of a structure involves growing single crystals, measuring the structure factor amplitudes from the diffraction pattern, followed by evaluating their phases, obtaining an electron density map by Fourier transformation and finally fitting an atomic model to the density distribution. Comparison of the calculated versus the observed amplitudes is used to indicate the validity of the model structure.

X-rays are photons with an approximate range of wavelength of 0.1 to 100 Å and are generated by decelerating rapidly moving electrons, converting their energy of motion into radiation. When a beam of X-rays interacts with the electrons in the atom, which occupy a finite volume and are bound in well defined energy states, a fraction of incident X-rays is scattered. This leads to radiation propagating away from the sample in all directions, Figure 6.0[1]. The total wave scattered in a given direction depends on the phase differences between the various rays scattered by individual electrons.



Scattering shown by the path difference between radiation scattered by an electron at r and that scattered by an electron at the origin.

Figure 6.0

From figure 6.0, the path difference between the two rays is $\underline{r}.\underline{s} - \underline{r}.\underline{s}_o$. It is convenient to define the scattering vector $\underline{S}$ as $(s/\lambda - s_o/\lambda)$.

Therefore the path difference is equal to $\lambda\, \underline{r}.\underline{S}$. Hence, the phase difference =

$$\frac{2\pi}{\lambda} \times \text{path difference}$$

$$= \quad 2\pi\, \underline{r}.\underline{S} \qquad (6.0)$$

Since the electrons are not localised, it is better to describe an electron density $\rho(\underline{r})$ in a small unit of volume dv at a position r relative to the origin, as proportional to $\rho(\underline{r})dr$. Therefore the total wave scattered in terms of amplitude and phase will be $\rho(\underline{r})\, \exp(2\pi i \underline{r}.\underline{S})dv$. For continuous electron density at position r, the structure factor is given by

$$F(\underline{S}) = \rho(\underline{r})\, \exp(2\pi i \underline{r}.\underline{S})dv \qquad (6.1)$$

Hence, the expression for the atomic scattering factor[2] which takes into account the individual contributions over the volume of the atom, is

$$F(\underline{S}) = \int_{\text{vol. of atom}} \rho(\underline{r})\, \exp(2\pi i \underline{r}.\underline{S})dv \qquad (6.2)$$

The expression for the atomic scattering factor takes no account of the thermal vibrations of the atom. In 1914, Debye[3] noticed that a decrease in intensity by a factor $\exp[-B(\sin^2\theta)/\lambda^2]$ where $B = 8\pi^2\bar{U}^2$ was due to the thermal motion of atoms, $\bar{U}$ represents the mean displacement of atoms along the normal to the reflecting planes. Hence, the atomic scattering may take into account this term.

However, in general case the isotropic temperature factor B is usually replaced by an anisotropic temperature factor, since the atoms are not free to vibrate equally in all directions. The anisotropic vibration is represented by an ellipsoid of the vibration in reciprocal space with six parameters. That is, the reduction in intensity is given by

$$\text{Exp}-[b_{11}h^2 + b_{12}hk + b_{13}hl + b_{22}k^2 + b_{23}kl + b_{33}l^2] \qquad (6.3)$$

The expression for scattering by a molecule can be obtained for an assembly of atoms placed at defined positions in the unit cell. Individual contributions by the atoms are considered to arise at distances r_1, r_2, r_3 etc. from the origin r. The shift in origin means that the distance $\underline{r}$ in equation (6.2) has to be substituted by $\underline{r} + \underline{r}_1$ for atom 1 and so on. Therefore the scattering by atom 1 relative to the new origin is

$$F_1 = \int_{\text{vol. of atom}} \rho(\underline{r}) \exp[(2\pi i\underline{r} + \underline{r}_1).\underline{S}]\, dv \qquad (6.4)$$

$$= f_1 \exp(2\pi i\underline{r}_1.S)$$

where f_1 is from equation (6.2).

Hence similar expressions can be obtained for all the other atoms in the unit cell and the total wave scattered will be the vector sum of the individual contribution of all the atoms, that is

$$G(\underline{S}) = F_1 + F_2 + F_3 + \ldots\ldots\ldots + F_n$$

Hence $G(\underline{S}) = \sum_{j=1}^{N} f_j \exp(2\pi i\underline{r}_j.\underline{S})$ (6.5) is the scattering of the unit cell.

Scattering of X-rays by a crystal, can be obtained from a one-dimensional crystal, which is composed of a linear array of unit cells with a repeat distance $\underline{a}$. The sum of the wave scattered by each unit cell will equal the total wave scattered. $G(\underline{S})$ will be the wave scattered by the first unit cell relative to the origin, but the second unit relative to the same origin will be $G(\underline{S})\exp(2\pi i\underline{a}.\underline{S})$, since there is a shift of $\underline{a}$. Therefore the total wave scattered by n cells will be

$$F(\underline{S}) = \sum_{n=1}^{T} G(\underline{S})\exp 2\pi i(n-1)\underline{a}.\underline{S} \qquad (6.6)$$

T = Total number of unit cells.

However, scattering will only be observed if the phase difference between the waves scattered by successive unit cells is equal to an integral multiple of 2π.

$$\text{i.e.} \quad 2\pi\,\underline{a}.\underline{S} = 2\pi h$$
$$\text{i.e.} \quad \underline{a}.\underline{S} = h \qquad h = \text{an integer.}$$

Hence, for a crystal in three-dimensions, with unit cell dimensions defined by vectors $\underline{a}$, $\underline{b}$ and $\underline{c}$, the diffraction conditions are

$\underline{a}.\underline{S} = h$; $\underline{b}.\underline{S} = k$; $\underline{c}.\underline{S} = 1$ where h,k,l are all integers. These equations are known as the Laue equations.

So the total wave scattered in direction $\underline{S}$ will be

$$F(\underline{S}) = \sum_{j=1}^{N} f_j \exp(2\pi i\underline{r}_j.\underline{S}) \qquad (6.7)$$

Allowing the fractional co-ordinates for the jth atom to be x_j, y_j and z_j

i.e., $$r_j = \underline{a}x_j + \underline{b}y_j + \underline{c}z_j$$

Hence $$r_j \cdot \underline{S} = \underline{a}x_j \cdot \underline{S} + \underline{b}y_j \cdot \underline{S} + \underline{c}z_j \cdot \underline{S}$$

$$= hx_j + ky_j + lz_j \text{ from the Laue equation}$$

Therefore $$F(hkl) = \sum_{j=1}^{N} f_j \exp[2\pi i(hx_j + ky_j + lz_j)] \quad (6.8)$$

This equation represents the structure factor, or the molecular transform sampled at the reciprocal lattice points hkl. The diffractions pattern for the crystal can be calculated if the positions of all the atoms in the unit cell are known.

Unfortunately to measure $F(\underline{S})$ directly is impossible since F is a complex quantity representing the product of two terms : that is,

$$F(hkl) = |F(hkl)| \exp i\alpha(hkl)$$

where $F(hkl)$ is the amplitude and $\alpha(hkl)$ is the phase.

Alternatively, it can be written as the sum of real and imaginary parts, that is

$$F(hkl) = A + iB$$

Where $$A = |F(hkl)| \cos\alpha(hkl) \text{ and}$$
$$B = |F(hkl)| \sin\alpha(hkl)$$

The Argand diagram, figure 6.1 shows the relationship between these two representations of $F(\underline{S})$.

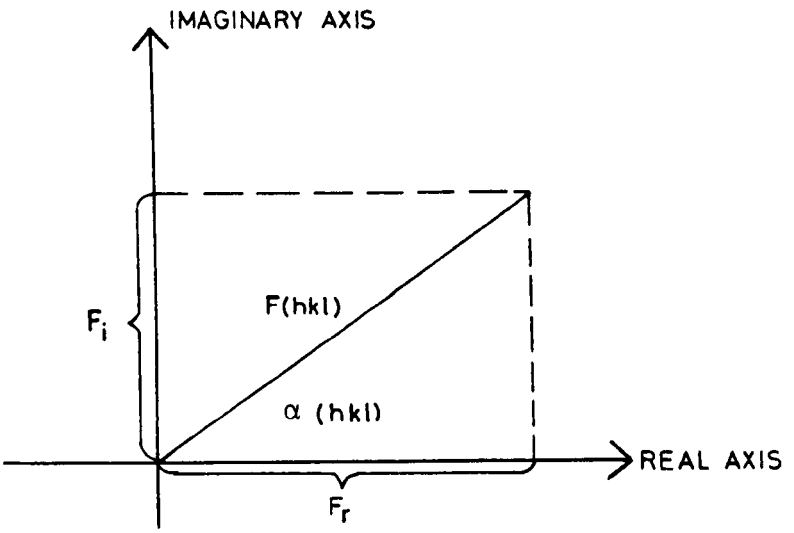


Figure 6.1

Experimentally, all information on the phase is lost and only the intensity of radiation scattered at an angle 2θ can be observed. The intensity is expressed by

$$I(hkl) = F(hkl).F^*(hkl)$$
$$= F(hkl)^2$$

where * indicates the complex conjugate.

The intensity is observable and must be real, however, the phase term of α (hkl) of F(hkl) is not directly measurable and this is a major obstacle in X-ray diffraction studies. To solve a structure from its diffraction pattern one must use the theory of Fourier transforms. To show this, the expression, equation (6.7) can be rewritten as a continuous summation over the volume of the unit cell, that is,

$$F(\underline{S}) = \int_{\text{vol. of unit cell}} \rho(\underline{r}) \exp(2\pi i \underline{r}.\underline{S}) dv$$

By multiplying both sides of equation (6.7) by $\exp(-2\pi i\underline{r}'.\underline{S})$ and integrating over the volume of diffraction space, it may be shown that

$$\rho(\underline{r}) = \int_{\substack{\text{vol. of}\\ \text{diffraction space}}} F(\underline{S}) \exp(-2\pi i\underline{r}'.\underline{S}) dv_s$$

where dv_s is a small unit of volume in diffraction space.

For a crystal the integration is replaced by a summation since $F(\underline{S})$ is not continuous and is non-zero only at the reciprocal lattice points. Hence,

$$\rho(xyz) = (1/V) \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \underline{F}(hkl) \exp[-2\pi i (hx + ky + lz)] \qquad (6.9)$$

V = volume of unit cell, V = a.bxc

This equation represents the electron density and to calculate it, $F(hkl)$ must be known, that is, the amplitude $|F(hkl)|$ and the phase $\alpha(hkl)$ of the structure factor. This is emphasised by rewriting equation (6.9) as

$$\rho(xyz) = (1/V) \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} |F(hkl)| \exp i\alpha(hkl) \exp[-2\pi i(hx + ky + lz)] \qquad (6.10)$$

The main problem, then, is to determine $\alpha(hkl)$.

6.2 THE PHASE PROBLEM

Several methods exist for determining the phases.

(1) Direct methods rely on mathematical relationships between the reflections to provide phase information. In this method usually there are no heavy atoms in the structure, and the solution by direct methods simply involves assigning a plus or minus sign to each observed structure amplitude.

(2) The method for small molecules, which contain two or more heavy atoms per unit cell is to use the Patterson or F^2 function[4]. The Patterson summation is a Fourier summation based on the experimentally observable $F(hkl)^2$, which essentially gives a vector map[5,6]. The function used is

$$P(UVW) = (1/V) \sum_{h=\infty}^{\infty} \sum_{k=\infty}^{\infty} \sum_{l=\infty}^{\infty} |F(hkl)|^2 \exp[2\pi i (hU + kV + lW)] \qquad (6.11)$$

An $F(hkl)^2$ calculation shows peaks corresponding to all interatomic vectors. The atomic scattering factor is proportional to the number of z electrons and the resulting Patterson peaks from atoms i and j are proportional to $z_i z_j$. Therefore, the Patterson maps shows N^2 peaks in the vector map for N atoms in the unit cell. N are self vectors, which superimpose at the origin and the remaining N(N-1) vectors are distributed throughout the volume of the unit cell. Hence, the dominant features are vectors between pairs of heavy atoms. It is usually possible to obtain sufficient information about the location of the heavy atoms to proceed further with the structural analysis.

Once the heavy atoms are located, the simplest approach to structure completion is to use the second Fourier function of importance, that is, the difference synthesis or difference Fourier. In this case, the difference in the electron density between

the observed and calculated structures is expressed in the form :

$$\Delta\rho = (1/V) \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \Delta F \exp i\alpha_c [-2\pi i (hx + ky + lz)] \qquad (6.12)$$

where α_c is the phase of F_c and $\Delta F = |F_o| - |F_c|$

The calculation of a difference Fourier, like that of Fourier synthesis, uses the phases of the calculated structure factors. The difference Fourier map shows the atoms not yet located. The two major facts about this Fourier are firstly, it reflects the correctness of the model used and the true structure implied by the $|F_o|$'s. This is highly useful in refinement (discussed later). Secondly if the phases are in error, the ΔF synthesis can provide information not easily obtained from the Fourier synthesis. It may also be possible from the final difference Fourier to locate the positions of the hydrogen atoms.

Using the conventional Patterson and difference Fourier technique, two small molecules containing heavy atoms ruthenium (Ru) and platinium (Pt) have been solved (see Appendix I and II). Patterson map revealed the Ru-Ru vector for the first molecule and Pt-Pt vector for the other molecule. The remaining atoms in both molecules were obtained from difference Fourier synthesis. In fact, the Pt compound was solved in a lower symmetry space group than necessary as pointed out by Marsh[7]. The new coordinates are in Appendix II.

The methods described earlier for small molecules are not successful with proteins. These macromolecules generally do not have heavy atoms in the molecule,

therefore to solve a crystal structure, heavy atoms have to be incorporated, that is, isomorphous replacement[8]. A perfectly isomorphous derivative should only have a change in the electron density between it and the native crystal at the position of the heavy atom. The heavy atom provides observable change in the diffraction pattern and the intensities of both the native and the derivative enables one to overcome the phase problem once the heavy atom is located.

(3) A single isomorphous replacement will result in the sum of electron densities of the parent crystal and of the heavy atom substitution. Thus the structure factor F_{PH} for the heavy atom isomorphous derivative is related to the structure factor F_P of the native protein and the structure factor F_H of the heavy atom, that is,

$$F_{PH}(hkl) = F_P(hkl) + F_H(hkl) \qquad (6.13)$$

and illustrated by a vector diagram, figure 6.2.

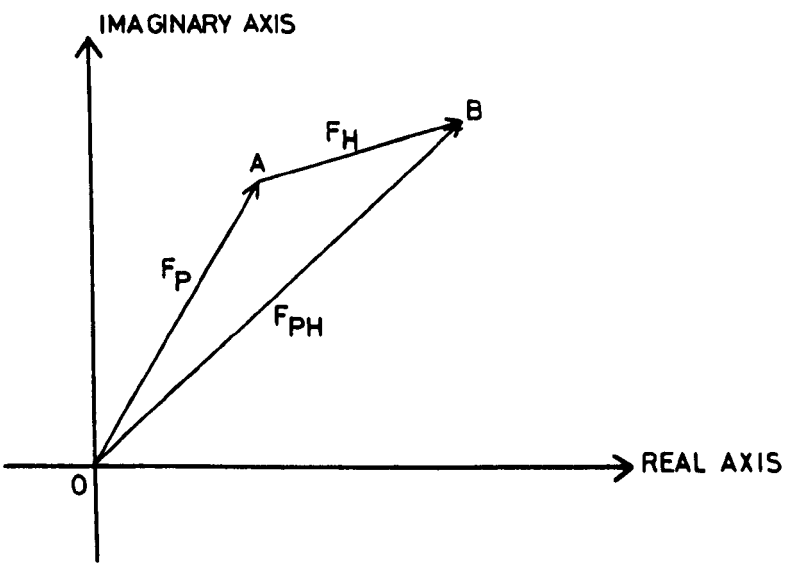


Figure 6.2

Structure factors F_{PH} and F_P are measured and the arrangement of the heavy atom in the crystal unit cell is known, the vector F_H can then be calculated. From figure 6.2, applying the cosine law, the phase α_P can be calculated :

$$\alpha_P = \alpha_H + \cos^{-1}\left(\frac{F_{PH}^2 - F_P^2 - F_H^2}{2F_P F_H}\right) = \alpha_H \pm \alpha' \qquad (6.14)$$

This equation, however, results in two possible values for α_P. However, the ambiguity can be overcome by using the multiple isomorphous replacement (M.I.R.) technique which involves using at least two heavy atom derivatives[9-11], identical with the native protein, except for the presence of heavy atoms.

(4) One additional technique which can be used for the calculation of phases for proteins in the presence of heavy atom is "anomalous scattering". It occurs when the frequency of incident X-rays lies near the frequency to the absorption edge of an atom. The atomic scattering factor previously treated as a real number, has in this case two additional factors : a real part $\Delta f^1(\underline{S})$ and an imaginary part $\Delta f^{11}(\underline{S})$ such that

$$f(\underline{S}) = f(\underline{S}) + \Delta f^1(\underline{S}) + \Delta f^{11}(\underline{S}) \qquad (6.15)$$

6.3 LOCATION OF HEAVY ATOMS BY DIFFERENCE PATTERSON

When both an isomorphous heavy atom derivative and a parent crystal are available with measured structure factor amplitudes $|F_{PH}(hkl)|$ and $|F_P(hkl)|$ it is possible to evaluate an isomorphous difference Patterson map. Since

$|F_H(hkl)|$ is not experimentally measurable, it is estimated fairly well by equation (6.13) from which an isomorphous difference Patterson function ΔP can be calculated.

$$\Delta P = (1/V) \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \left| |F_{PH}| - |F_P| \right|^2 \exp[-2\pi i(hx + ky + lz)] \qquad (6.16)$$

Anomalous scattering difference can also be used to estimate F_H using the expression :

$$|F_{PH}^+| - |F_{PH}^-| = \Delta F_{ano} \simeq |F_H \sin(\alpha_{PH} - \alpha_H)| \qquad (6.17)$$

However, neither of these methods gives a particularly good estimate of F_H, but combining the two expressions for F_H will improve it[12].

$$\Delta F_{iso} + \Delta F_{ano} \simeq F_H \cos^2(\alpha_{PH} - \alpha_H) + F_H \sin^2(\alpha_{PH} - \alpha_H)$$

$$\text{Thus } F_H \simeq (\Delta F_{iso}^2 + \Delta F_{ano}^2)^{\frac{1}{2}} \qquad (6.18)$$

Once the structure factor amplitude $|F_H(hkl)|$ is calculated, it is possible to obtain the phase α_P of the native crystal.

6.4 PHASE CALCULATION

Once F_H is known, then F_P, F_H and F_{PH} should form a closed triangle in which the magnitudes of three sides and direction of one of them are known (Figure 6.2). This implies perfect isomorphism and no experimental error. In practice, however ideal conditions do not exist and errors occur from inaccuracy in the measurement of intensity, lack of isomorphism and incorrect heavy atom positions. Blow et al[13] have considered the errors occurring in isomorphous replacement and provide a proper treatment of errors and best choice of weights in the calculation of electron

density map. They considered errors arose from two distinct sources :

(1) an error, e, for the calculated heavy atom contribution, $F_H(calc)$, to the derivative structure factor.

(2) inaccuracies in the determination of the amplitudes F_P and F_{PH}.

The "lack of closure" error is defined by

$$e_j = (F_{PH}(obs) - F_{PH}(calc)) \qquad (6.19)$$

Blow et al[13] showed that the probability of the phase angle being α for a single isomorphous replacement is

$$P_j(\alpha) = \exp[-e_j(\alpha)^2/2E_j^2] \qquad (6.20)$$

or, where a number of heavy atom derivatives is used simultaneously, the total probability of the phase angle is given by the product of the individual probabilities :

$$P(\alpha) = \prod_j P_j(\alpha) = \exp[-\sum_j e_j(\alpha)^2/2E_j^2] \qquad (6.21)$$

Blow et al[13] have shown that the phase corresponding to the centroid of this distribution, results in the least mean square error in electron density over the unit cell. This can be shown by considering the errors in one reflection for the electron density. If the value of the coefficient used in the synthesis is F_s and its true value is F_t, then the mean square error over the unit cell from the reflection is

$$\langle\Delta\rho\rangle^2 = (1/V^2)(F_s - F_t)^2$$

where $F_t = F\exp i\alpha$.

Hence, the mean square error is

$$\langle \Delta\rho \rangle^2 = (1/V)^2 \int_{\alpha=0}^{2\pi} \left(F_s - (F\exp i\alpha) \right)^2 P(\alpha) d(\alpha) \Big/ \int_{\alpha=0}^{2\pi} P(\alpha) d\alpha \qquad (6.22)$$

The numerator in equation (6.22) is equal to the moment of inertia of ring mass $P(\alpha)$ and radius $F\exp i\alpha$. By the parallel axis theorem, when F_s is at the centre of the ring, the integral has a minimum value, that is

$$F_s(\text{best}) = F \int_{\alpha=0}^{2\pi} \exp i\alpha \, P(\alpha) d\alpha \Big/ \int_{\alpha=0}^{2\pi} P(\alpha) d\alpha$$

$$= mF\exp(i\alpha_{best}) \qquad (6.23)$$

This equation represents for the centre of gravity of the probability distribution with polar co-ordinates (mF, α_{best}). Blow el at[13] pointed out that, m acts as a weight, dependent not on the absolute probability of the phase determination but on its sharpness or unambiguity. They further showed that m, called the "Figure of merit" and α_{best} could be obtained from the expression,

$$m\cos\alpha_{best} = \sum_j P_j(\alpha_j)\cos\alpha_j \Big/ \sum_j P(\alpha_j)$$

$$m\sin\alpha_{best} = \sum_j P_j(\alpha_j)\sin\alpha_j \Big/ \sum_j P(\alpha_j)$$

The most commonly used measure of the goodness of match between the observed and the calculated structure amplitudes is the disagreement ratio, R

$$R = \frac{\sum_{hkl} |F_{hkl}(Obs)| - |F_{hkl}(calc)|}{\sum_{hkl} |F_{hkl}(Obs)|} \qquad (6.24)$$

R = Residual index.

$F_{hkl}(obs)$ represents the square root of diffraction intensity after various corrections are applied for the experimental factors. Thus R gives a guide to how well the observed data $|F_{hkl}(obs)|$ compares with the data calculated from the structure, $|F_{hkl}(calc)|$.

R also gives indication about the molecules being centric, R would equal to 0.83 and to 0.59 would mean the molecule is without centre of symmetry[14,15]. For small molecules it is possible to refine to the R value of < 0.05 but for proteins the R value are usually high in the early stages because solvent and thermal motion are not taken into consideration in the structure determination. A good protein structure determination at high resolution should result in an $R \simeq 0.15$.

6.5 INTERPRETATION OF THE ELECTRON DENSITY MAP

Once the best phases and the figure of merit, m have been calculated the electron density map can be computed. The expression used is

$$\rho(\underline{r}) = \Sigma\ mF_p \exp(i\alpha_{best}) \exp[-2\pi i h.\underline{r}] \qquad (6.26)$$

The map is calculated to a certain resolution using the amplitudes and phases. The resolution (that is, the number of measureable reflections) is a function of the actual crystal and also on the isomorphous derivatives. To obtain a high resolution structure one usually requires excellent phases to at least 2.5 Å spacings.

The interpretation of the map is usually made with the aid of plotting the electron density (on a scale 2Å/cm) on perspex sheets. These sheets are stacked together and viewed on a light box. The interpretation of the amino-acid residues in the protein will depend on the knowledge of the sequence obtained from chemical studies.

More, recently computer graphics are used to view the electron density map. This technique being best in which an existing interpretation of the molecule is available.

6.6 DIFFERENCE FOURIER SYNTHESIS IN STUDYING LIGAND-MACROMOLECULE INTERACTION

Once the structure of the protein is known, then studies relating to the binding of small molecules, (substrates or inhibitors), can be accomplished by means of a difference Fourier synthesis. Stryer et al[15] were first to apply this technique in protein crystallography by binding azide ion to myoglobin. Since then, it has been applied to a wide number of proteins and has provided much information about the proteins active sites. Difference Fourier maps also provide detailed information regarding any changes that may occur in the conformation of the native protein structure and atomic shifts of less than 1 Å can be detected.

The diffraction intensities are measured for the liganded protein, then the difference Fourier is calculated by using these intensities with the phases calculated for the native protein. This results, in showing the electron density of the bound ligand and also any difference in the native enzyme caused by the binding of the small molecule. The disadvantage is that it is only useful where there is a small change.

The technique works well for locating a bound ligand since the presence of the ligand changes the phase of most structure factors relatively little and, by analogy to the equation for a heavy atom,

$$F_{PL} = F_P + F_L \qquad (6.26)$$

where F_{PL} represents the parent-ligand complex and the structure factor for the ligand is F_L (Figure 6.3). From the figure 6.3, F_P is known, and $|F_{PL}|$ is known. Equation (6.26) will only work if $F_L < F_P$ and $F_L < F_{PL}$, hence the possible value of F_{PL} will lie in a small range of phase angles. Thus to a first approximation, α_P, will be good estimate of α_{PL}. Since the native phases which are determined from the heavy atom isomorphous replacement method are associated with a figure of merit, m, it is conventional to weight the difference Fourier map also by m. Thus the difference Fourier $\Delta\rho$ is usually defined by

$$\Delta\rho = (1/V) \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} m[|F_{PL}(hkl)| - |F_P(hkl)|]$$

$$\exp i\alpha(hkl)\exp[-2\pi i(hx + ky + lz)]$$

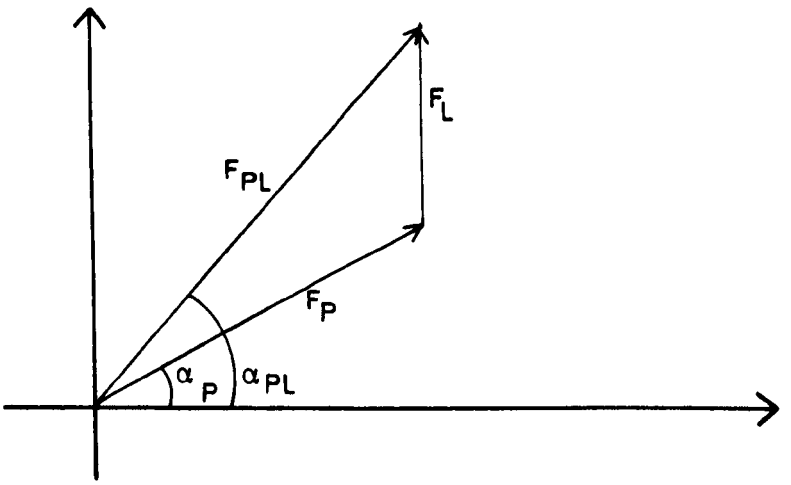


Figure 6.3

6.7 REFINEMENT

(1) SMALL MOLECULES

The refinement technique most often used to give the best fit between a set of observed and a set of

calculated data on a trial set of parameter is a least-square procedure[17]. This procedure refines the atomic co-ordinates, the atomic vibration parameters and scale factor. The observed data are the structure factors obtained from the intensity data whose phases have been determined by direct or indirect methods.

In least-squares procedure the quantity, that is, minimised is given by :

$$D = \sum_{r=1}^{m} W_r (f_{o,r} - f_{c,r})^2 \qquad (6.27)$$

where W_r represents the weight assigned to the observation, $f_{o,r}$, the value of the function and also to the corresponding calculated value, $f_{c,r}$. To gain a satisfactory fit, the parameters, P, are regarded as variables, so that they can be adjusted to minimize D.

The function minimised in the equation (6.27), takes into account a weighting factor[18] for each observation which gives a measure of the reliability in the observation.

(2) PROTEINS

The least-square refinement of a protein is usually only effectively applied to high resolution data. There are several sources of error in any crystallographic analysis. (1) The phases determined from M.I.R. and anomalous scattering contain both random and non-random errors (inaccuracies in measurement and lack of isomorphism between the native derivative crystals). (2) The phases for $d < 2$ Å may not have been calculated by M.I.R. even though the observable native data may extend to higher resolution. (3) The model obtained may not represent the best fit to the calculated electron density.

Least-squares refinement provides the opportunity to refine the phases, to extend them to higher resolution and to obtain a model which is the best fit to the observed data.

For a protein anywhere between three and nine parameters, P_j, are needed for each atom. This means, that a large number is required for the entire asymmetric unit.

As before, the function minimised is

$$D = \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} W_{hkl}[\,|F_o(hkl)| - |KF_c(hkl)|\,]^2 \quad (6.28)$$

Where W_{hkl} is a weighting factor giving the reliability between the observed and calculated points, K is a scaling parameter.

One of the most useful approaches in the initial cycles of refinement, is to treat both the protein phases, and structure factor amplitudes, $\alpha p(calc)$ and F_p as observations. This makes it possible to adjust the model to the density and then, the protein phases can be recalculated from the atomic position and the fitting process repeated.

However, recently, two approaches to reciprocal space refinement methods have been considered.

(1) To reduce the number of parameters by treating groups of atoms as rigid bodies, where each group is described by only six (three translational and three rotational) positional parameter[19].

(2) Increase the number of observations by introducing subsidiary distance restraints to maintain proper sterochemistry[20].

(1) CONSTRAINED LEAST-SQUARES

For each group of atoms, three translational and three rotational parameters are used to describe the geometry. For a set of co-ordinates X given in Å, the fractional co-ordinates x are obtained from

$$x_i = t_g + |u|\,|R_g|\,X_i$$

Where t_g is the translational vector in fractional co-ordinates of the origin of the group relative to the unit cell origin. $|u|$ is the matrix to transform cartesian co-ordinates to fractional crystallographic co-ordinates[21]. $|R_g|$ is the rotation matrix[22].

The derivatives of F_c with respect to the group parameters are obtained from the individual atom derivatives by application of the chain rule[21]. The normal matrix is constructed and the equations solved in the usual way.

This procedure reduces the number of parameters dramatically, since only six positional parameters for each rigid group. Hence the method is applicable to low resolution.

(2) RESTRAINED LEAST-SQUARES

The function to be minimised is a sum of structure factor and restraint terms.

$$D = \sum_{hkl} W_{hkl}(|F_c(hkl)|-|F_o(hkl)^2|)+$$

$$\Sigma_r W_r(d_{o,r}-d_{c,r})^2$$

The first term is the normal structure factor component of least-squares and the second term is a sum over the subsidiary distance.

It is the restraints of the weighted square of the difference between an "ideal" distance, $d_{o,r}$ between pairs of atoms and the calculated distance, $d_{c,r}$.

The distance restraint may include terms for bonded atoms, next nearest neighbour atoms to define bond angles, C_α .....O distances to restrain the ω torsion angle, planar groups of atoms and repulsive restraints for non-bonded contact distances.

6.8 OTHER METHODS USED IN REFINEMENT OF PROTEIN MOLECULES

Other techniques are available which could be used for refinement, these are classified into two broad categories : real space least-squares methods, difference Fouriers procedures and the reciprocal space least-squares methods[23,24] already described.

(1) Real space least-squares refinement was developed by Diamond[23,25]. This procedure uses a constrained geometry based on conformational angles. In this case, the electron density is considered to be an objective expression of experimentally quantity minimised, that is, $(\rho_o - \rho_m)$. ρ_o is the observed electron density and ρ_m is the calculated electron density from the model.

(2) Difference Fourier methods are mainly used for locating side-chains which are not positioned properly and are also useful in locating solvents atoms in the structure.

REFERENCES

1. Cantor, C.R., and Schimmel, P.R.
Biophysical Chemistry, Part II, Freeman, W.H., and Company (1980.

2. Blundell, T.L. and Johnson, L.N.
Protein Crystallography, Academic Press, New York, (1976).

3. Debye, P.
Annl.Phys., 43, 49 (1914).

4. Stout, G.H., and Jensen, L.H.
X-ray Structure Determination, The Macmillan Company, (1968).

5. Patterson, A.L.
Phys.Rev., 46, 372 (1934).

6. Patterson, A.L.
Z. Kristallogr., 90, 517 (1935).

7. Marsh, R.E.
Acta.Cryst., C41, 1383 (1985).

8. Green, D.W., Ingram, V.M., and Perutz, M.F.
Proc.Roy.Soc., A225, 287 (1954).

9. Harker, D.
Acta.Cryst., 9, 1 (1956).

10. Blow, D.M.
Proc.Roy.Soc., A247, 302 (1958).

11. Bodo, G., Dintzis, H.M., Kendrew, J.C., and Wyckoff, H.W.
Ibid, A253, 70 (1959).

12. Phillips, D.C.
Adv.Res.Diff.Meth., 2, 75 (1966).

13. Blow D.M., and Crick, F.H.C.
Acta.Cryst., 12, 794 (1959).

14. Wilson, A.J.C.
Ibid., 2, 318 (1949).

15. Wilson, A.J.C.
Ibid., 3, 397 (1950).

16. Stryer, L., Kendrew, J.C., and Watson, H.C.
J.Molec.Biol., 8, 96 (1964).

17. Whittaker, E.T., and Robinson, G.
The Calculus of Observations, 4th Edition, Blackie & Son, Glasgow, 209 (1944).

18. Mellor, J.W.
Higher Mathematics for Students of Chemistry and Physics, Dover, New York, 548 (1955).

19. Scheringer, C.
Acta.Cryst., 16, 546 (1963).

20, Waser, J.
Ibid., 16, 1091 (1963).

21. Rossman, M.G., and Blow, D.M.
Ibid., 15, 24 (1962).

22. Doedens, R.
"Crystallographic Computing", Copenhagen, Munksgaard, 198 (1970).

23. Diamond, R.
Acta.Cryst., A27, 436 (1971).

24. Watenpaugh, K.D., Sieker, L.C., Herriott, J.R., and Jensen, L.H.
Ibid., B29, 943 (1973).

25. Diamond, R.
J.Molec.Biol., 82, 371 (1974).

CHAPTER SEVEN

INTENSITY-DATA COLLECTION AND DATA REDUCTION

7.1 INTENSITY-DATA COLLECTION

The diffracted intensities can be measured using either photographic or diffractometer (counter) methods. The photographic method involves measuring the degree of the blackening of spots on the diffraction photographs which is taken as proportional to the intensity of beam. However, the diffractometer method has advantages. It provides a more precise intensity measurements and the data output is in a form ready to be processed by a computer, which relieves the tedium involved with photography. Moreover, with protein molecules crystallising with large unit cells, the number of reflection to be measured at any given resolution is corresponding large and therefore the diffractometer is usually ideal for it.

The data are collected using a computer controlled four-circle diffractometer[1] operating with "Normal-beam equatorial" geometry. The crystal orientation is determined by three Eulerian angles which need to be calculated for each reflection. The detector is constrained to move in the equatorial horizontal plane which contains the incident beam and is set to the 2θ value appropriate for each reflection. A small on-line computer calculates the crystal and detector setting and also controls the digital diffractometer shaft setting.

The intensities for the two small molecules were measured on a Nonius CAD4 diffractometer with monochromatised Molybdenum (Mo-K_α) radiation (see details in the Appendix I and II). Data collection for the PPE crystals soaked in substrate/inhibitor was also carried out on an Erap-Nonius CAD4 diffractometer but with filtered copper (Cu-K_α) radiation (details in Chapter Eight).

7.2 DATA REDUCTION

The raw data which are collected on the diffractometer are not in the form used in the determination of the crystal structure. Several corrections have to be applied which are, correction for Lorentz-polarisation, absorption and radiation damage effects. After correction, if different sets of data are collected they need to be scaled together and for protein molecules the scaling of heavy atom derivative or substrate bound data need to be scaled to the native data. After all the corrections and scaling have been done, the data set obtained can be used for the methods discussed in Chapter Six for solving the crystal structure.

7.2.1 LORENTZ-POLARIZATION CORRECTION[2-4]

The Lorentz factor, L, arises because the time required by a reciprocal lattice point to pass through the reflection sphere is not constant but varies with its position in reciprocal space and the direction in which it approaches the sphere.

The L, depends on the geometry which is used to measure the X-ray intensity. The expression for a four circle symmetrical A setting diffractometer where the χ-circle is arranged to bisect the angle between the incident and diffracted beams. Hence, the L is given by $1/\sin 2\theta$[5].

The polarization factor, p, is a function of 2θ and arises because of the nature of the X-ray beam. Since the usual X-ray beam is unpolarized, which means that the azimuth of the electric vector assumes all directions with time. The effective amplitude of the radiation after it has reflected by the crystal at the angle 2θ consists only of the components of these azimuths after reflection. This in turn will affect the reflections for the crystal under study and the factor, p, must be accounted for the correction in the structure amplitudes equation (7.1)[2].

$$|F_{hkl}| = \sqrt{I_{hkl}/Lp}$$

where $$p = \frac{1 + \cos^2 2\theta}{2}$$

7.2.2 ABSORPTION

The corrections for absorption are the most difficult to determine especially when a three-dimensional data set is collected by the diffractometer[6]. In this case the crystal presents many different facets to the X-ray beam. In applying the correction for small molecules, computer methods[7,8] are available which calculates from a knowledge of crystal shape, the absorption for the actual path length travelled within the crystal by the beam reflecting from each infinitesimal portion of the crystal. However, for protein molecules this is not suitable since various factors need to be considered which are difficult to estimate, these are, crystal shape, the surrounding mother liquor and the glass capillary tube. Also, due to large numbers of reflections measured, the time required for calculation of the path length for each reflection becomes prohibitive.

Hence, the most used method in protein crystallography is that described by North et al[9] where an empirical absorption curve is obtained from the variation in the intensity of strong Bragg reflection as the crystal is rocked about the normal to the corresponding planes.

7.2.3 RADIATION DAMAGE[5]

Radiation damage corrections have to be applied since it is observed that the intensity decreases as the crystal is exposed. This correction is simple to apply and is constructed from a set of reference reflections measured at intervals throughout the data collection. Radiation damage

usually affects high angle reflections more than the lower ones. Hence for high resolution data it is essential to measure intensity controls throughout data collection for h00, 0k0 and 00l rows of reflections at various angles and various direction.

7.2.4 SCATTERING FACTORS[10]

In the case of small molecules it is essential for the data reduction program to know the necessary X-ray scattering power of each kind of atom in the cell, since the output serves as input to the structure factor program. This program calculates the structure factors on the basis of some assumed arrangement of atoms and compares these calculated values with those actually observed.

The scattering curves have been calculated for all elements using theoretical electron distributions of varying degrees of elegance. The results are available in tabular form in the International Tables and also in the literature[11,12]. It is economical to have tabulated scattering factor data for the atom types in the crystal included in the input to the data reduction program. This is because the structure factors are usually calculated several times during the course of a structure analysis. Hence the output associates with each reflection an interpolated scattering factor which corresponds to the value of $(\sin\theta/\lambda)$

7.3 SCALING DIFFERENT SETS OF DATA

Intensity data collection for small molecules usually requires only one crystal and therefore there is no scaling problem. However, for protein molecules, an intensity data collection at high resolution will certainly involve several crystals, and so it is necessary to scale these different sets of data together on the basis of the common reflections between the sets. A method is required which accounts for

the fact that some sets may contain reflections which may overlap with other sets but not with every set. Hamilton et al[13] have developed a method which has proved useful and reliable.

7.4 ABSOLUTE SCALING

In 1949, Wilson[14] derived a plot which gave an approximate value of the absolute scale of data and also including an estimate of the effects of the thermal motions of the atoms.

It is essential to take into consideration the effect of temperature on the X-ray intensities[15]. The normal scattering-factor curves are calculated on the assumption that the atoms are stationary, but in fact they are vibrating about their rest points in the crystal. The degree of vibration is dependent on factors, such as temperature, the mass of the atom and the firmness with which it is held in place by covalent bonds or other forces. It has been shown both theoretically and practically that thermal vibration may be represented by an average isotropic temperature factor B equation (7.2).

$$\mathrm{Exp}[-2B(\sin^2\theta)/\lambda^2] \qquad (7.2)$$

It is convenient to have an estimate of the average value of B for the whole structure before beginning the actual analysis and a value is provided by the Wilson plot.

$$\ln \frac{I_{rel}}{\sum_{i=1}^{n} f_{oi}^2} = \ln C - 2B(\sin^2\theta)/\lambda^2$$

where I_{rel} and $\sum_{i=1}^{n} f_{oi}^2$

are the average values of the observed intensities and the sum of the squares of the scattering factors of the atoms in the molecule. C is a scale factor. A plot of

$$\ln(I_{rel}/\sum_{i=1}^{n} f_{oi}^2) \quad \text{versus} \quad (\sin^2\theta)/\lambda^2$$

should give a straight line with slope of -2B and intercept of lnC.

Hence to convert $|F_{rel}|$ to $|F_{abs}|$, C is related to the scale constant, K, by equation (7.3).

$$K = \sqrt{\frac{1}{C}}$$

where $$|F_{abs}| = K|F_{rel}| \qquad (7.3)$$

REFERENCES

1. Wooster, W.A.
J.Sci.Instrum., 42, 685 (1965).

2. Buerger, M.J.
Crystal Structure Analysis, Wiley, New York, 156 (1960).

3. Arndt, U.W., and Willis, B.T.M.
Single Crystal Diffractometry, Cambridge University Press, 278 (1966).

4. International Tables of X-ray Crystallography, Vol.II, 265.

5. Blundell, T.L., and Johnson, L.N.
Protein Crystallography, Academic Press, New York, 323 (1976).

6. Buerger, M.J.
Crystal Structure Analysis, Wiley, New York, 204 (1960).

7. Busing, W.R., and Levy, H.A.
Acta.Cryst., 10, 180 (1957).

8. Coppens, P., Menlenaer, J. de., and Tompa, H.
Ibid., 22, 601 (1967) and references cited therein.

9. North, A.C.T., and Phillips, D.C., and Matthews, F.S.
Ibid., A24, 351 (1968).

10. Stout, G.H., and Jensen, L.H.
X-ray Structure Determination, Collier-MacMillan Ltd., London (1968).

11. International Tables for X-ray Crystallography, Vol.III, 202.

12. Cromer, D.T.
Acta.Cryst., 19, 224 (1965).

13. Hamilton, W.C., Rollett, J.S., and Sparks, R.A.
Ibid., 18, 129 (1965).

14. Wilson, A.J.C.
Nature, 150, 151 (1942).

15. Buerger, M.J.
Crystal Structure Analysis, Wiley, New York, 231 (1960).

16. Green, D.W., Ingram, V.M., and Perutz, M.F.
Pro.Roy.Soc., A225, 287 (1954).

17. Blow, D.M.
Ibid., A247, 302 (1958).

CHAPTER EIGHT

CRYSTALLOGRAPHIC STUDIES OF SUBSTRATE BINDING

8.1 INTRODUCTION

Several studies were undertaken to complex PPE with various substrates (Ac-Pro-Ala-Pro-Ala.OH, Ac-Pro-Ala-Pro-Alaninal and TFA-Pro-Ala-Pro-Alaninal prepared in Chapter Three) and the inhibitor (TFA-Ala_3-CMK, a gift from Bieth et al[1]). Concurrently studies were carried out between the anhydroelastase (prepared in Chapter Two) and the hexapeptide inhibitor H.Pro-Ala-Pro-Ala-Lys-Phe.OH.

However, data were only able to be collected on two crystal complexes to 2.5 Å resolution: Ac-Pro-Ala-Pro-Ala.OH and TFA-Ala_3-CMK. In the case of the others, crystals soaked in the substrates and inhibitors are mounted and ready for data collection.

8.2 EXPERIMENTAL

Crystals of native PPE and of the anhydroelastase preparation were grown from dilute sodium sulphate solutions as described by Shotton et al[2]. They were then transferred to 1.2M sodium sulphate-0.01M sodium acetate pH 5.0, in which they were stable indefinitely at room temperature. The conditions used for preparing each of the derivatives was as follows :

(1) Ac-Pro-Ala-Pro-Ala.OH

The native PPE crystals were transferred into 2M potassium phosphate, pH 8.5 for 24 hours before being soaked in 10mM substrate (in the same buffer) at room temperature.

(2) Ac-Pro-Ala-Pro-Alaninal and TFA-Pro-Ala-Pro-Alaninal

The inhibitor/PPE complex were crystallised using the technique of co-crystallisation, by incubating for

half-an-hour 0.25 ml of 10 mg/ml PPE solution (made up in 0.01M sodium acetate buffer, pH 5.0), with 12.5 λ of 10mM inhibitor in the same buffer. Followed by adding 5 λ of 1M aqueous solution of sodium sulphate, pH 5.0. Crystals grew in 3 days at 4^{o} C. Inhibitor.PPE complex crystals were then transferred to a 1M sodium sulphate solution at pH 5.0 containing 500 μM of the inhibitor at room temperature.

(3) H.Pro-Ala-Pro-Ala-Lys-Phe.OH

Similar conditions and concentrations to that of Ac-Pro-Ala-Pro-Ala.OH were used to crystallise H.Pro-Ala-Pro-Ala-Lys-Phe.OH.Complex. In this case, anhydroelastase crystals were used which were obtained in similar conditions to those of native PPE.

(4) TFA-Ala_3-CMK

The native PPE crystals were soaked for 3 days in 20 mM inhibitor in 1.2M sodium sulphate-0.01M sodium acetate pH 5.0 buffer.

The optimum conditions described above were selected for each of the peptides in the soaking experiments. The maximum peptide concentrations were used which did not crack the crystals.

Selected crystals of the native PPE (0.4 x 0.2 x 0.3 mm^3), Ac-Pro-Ala-Pro-Ala.OH (0.4 x 0.2 x 0.3mm^3) and TFA-Pro-Ala-Pro-Alaninal (0.5 x 0.2 x 0.3mm^3) were mounted in their mother liquors, in thin-walled capillary tubes for analysis by X-ray diffraction methods. Precession photographs of the native PPE confirmed the space group ($P2_12_12_1$) and the cell dimensions to be those of Shotton et al[2] (table 8.1). Similar X-ray photographs of the crystals of Ac-Pro-Ala-Pro-

Ala.OH and TFA-Pro-Ala-Pro-Ala-Alaninal derivatives showed that they were both isomorphous. A comparison of the cell dimensions are shown in table 8.1.

TABLE 8.1

CELL PARAMETERS

Compound	Method	a (Å)	b (Å)	c (Å)	Reference No.
Native PPE	P	51.50	58.00	75.50	(2) and (3)
	P	50.80	58.15	75.20	(4)
	D	51.48(43)	57.99(20)	75.19(16)	(5)
	P	50.76	57.80	75.19	
Ac-Pro-Ala-Pro-Ala.OH/ PPE Complex	P	50.66	57.34	75.59	this work
TFA-Pro-Ala-Pro-Alaninal/ PPE Complex	P	50.72	57.25	75.83	this work
TFA-Ala_3-CMK/PPE Complex	P	51.12(1)	58.00(1)	75.08(1)	this work
TFA-Lys-Ala-NH-C_6H_4-p-CF_3/ PPE Complex	D	52.53(42)	57.47(10)	75.26(5)	(5)

N.B. 1) The values in parentheses indicate the root mean square error.
2) P = Precession measurements.
D = Diffractometer measurements.
3) The dimensions measured by diffractometer will be a mean values of several measurements.

8.3 DATA COLLECTION AND PROCESSING

Data for the TFA-Ala_3-CMK crystal complex was collected on an Enraf-Nonius CAD4 diffractometer (computer controlled) to 2.5 Å resolution by Dr. L. Sawyer. The data collection was performed at 15° C with nickel filtered Cu-K_α radiation and were measured on one crystal. A total of 8162 unique reflections ($2\theta \leq 36^\circ$) were recorded by the 2θ-ω scan technique.

A computer controlled Enraf-Nonius CAD4 diffractometer was used for the data collection of Ac-Pro-Ala-Pro-Ala.OH to 2.5 Å resolution. The data collection was performed at 18° C with nickel filtered Cu-K_α radiation and were measured on two crystal. A total of 7249 unique reflections ($2\theta \leq 36^\circ$) were recorded by the ω scan.

The data set for TFA-Ala_3-CMK and Ac-Pro-Ala-Pro-Ala.OH were corrected for Lorentz polarization and absorption correction. Data collection parameters are shown in Table 8.2.

TABLE 8.2

DATA COLLECTION PARAMETERS

	TFA-Ala_3-CMK	Ac-Pro-Ala-Pro-Ala.OH
Resolution Limits	2.5 Å	2.5 Å
No. of recorded reflections (including controls).	11529	9424
No. of radiation damaged control reflections	6	3
No. of absorption correction curves	146	1

8.4 RESULTS AND DISCUSSION

(i) TFA-Ala$_3$-CMK

The phase set for the native PPE structure[4] was used with the reflections of TFA-Ala$_3$-CMK in order to calculate the difference Fourier map [F_{DER}-F_{NAT}]. The Evans and Sutherland graphics system (at the University of Leeds) was used for positioning the inhibitor molecule, in the active site region.

The program used, FRODO[6], allowed the atomic parameters to be adjusted to fit best both to the X-ray data and for a given set of geometrical dimensions expected in the amino-acid residues. The inhibitor molecule will be described in the text in two orientations: one with the N-trifluoroacetyl group in S_1' and the other with the CMK group in S_1.

The atomic labels for the inhibitor molecule are in figure 8.1 and the shorter inhibitor-PPE interactions are in table 8.3. Table 8.4 shows the hydrogen-bonding between the hydrogen atom attached to the nitrogen atom of the inhibitor and PPE molecules. Atomic parameters for the inhibitor molecule and selected neighbouring PPE residues are shown in table 8.5. The inhibitor molecule co-ordinates are in diamond format.

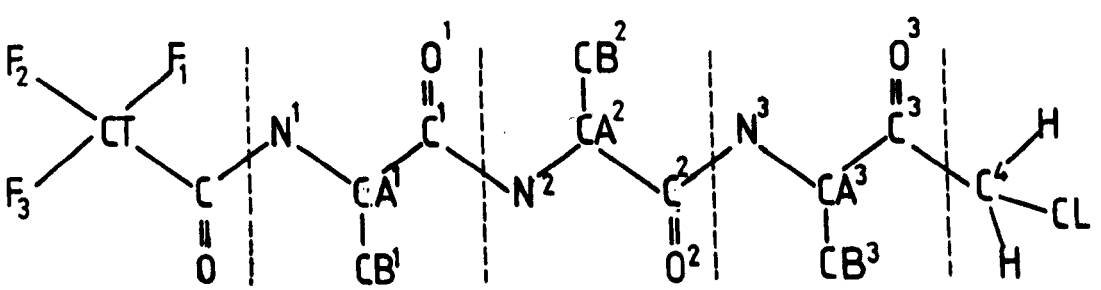


Figure 8.1

TABLE 8.3

SHORTER INTERACTIONS BETWEEN THE INHIBITOR AND PPE MOLECULES

Atom (A)	Atom (B)	Subsite[7]	Distance A-B Å
CT-TFA	OG -Ser 195		3.03
	O-Ser 214		2.94
F^1-TFA	OG -Ser 195		2.12
	O-Ser 214		3.17
F^3-TFA	O-Ser 214		2.44
	C-Ser 214		2.71
	N-Phe 215	S'_1	2.75
	CA-Phe 215		2.44
	C-Phe 215		3.08
C	OG -Ser 195		3.13
	O-Ser 214		2.88
O	OG -Ser 195		3.04
	CE1-HiS 57		3.18
N^1	CA-Phe 215		3.13
	CB-Phe 215		3.22
	O-Ser 214		2.48
CB^1	CD2-HiS 57	S_1	3.30
C^1	CB-Phe 215		3.29
O^1	CA-Phe 215		3.06
	CB-Phe 215		2.53
CA^2	O-Val 216		2.93
CB^2	O-Val 216	S_2	2.79
N^3	CD2-Phe 215		3.22
	O-Val 216	S_3	3.14
C^5	CE2-Phe 215		3.27
	CG1-Val 99		3.11
C1	CG1-Val 99	S_4	2.02

TABLE 8.4

HYDROGEN-BONDING BETWEEN THE HYDROGEN ATOM ATTACHED TO THE NITROGEN ATOM OF THE INHIBITOR AND PPE MOLECULES

Atom (H)	Atom (B)	Distance (H).....(B) Å
N^1-H	N-Val 216	3.91
	N-Phe 215	2.82
	O-Ser 214	1.86
	NE2-His 57	3.64
	CG-Phe 215	3.83
N^3-H	O-Val 216	2.34
	CG-Phe 215	3.51
	CZ3-Trp 172	3.97

TABLE 8.5

ATOMIC CO-ORDINATES (A) FOR (i) THE TFA-Ala_3-CMK and (ii) RESIDUES OF THE PPE MOLECULES WHICH SURROUND THE INHIBITOR MOLECULE

(i) THE INHIBITOR MOLECULE, TFA-Ala_3-CMK

	X	Y	Z
CT-CF_3	-10.698	23.625	38.726
F_1-CF_3	-11.353	24.152	37.942
F_2-CF_3	-11.162	23.686	39.751
F_3-CF_3	-9.695	24.189	38.759
C-CF_3	-10.444	22.187	38.269
O-CF_3	-9.799	21.961	37.247
N^1-Ala_1	-10.964	21.220	39.013
CA^1-Ala_1	-10.950	19.820	38.504
CB^1-Ala_1	-10.177	18.934	39.582
C^1-Ala_1	-12.387	19.520	38.572
O^1-Ala_1	-12.757	20.586	38.915

	X	Y	Z
N^2-Ala_2	-13.630	18.882	38.327
CA^2-Ala_2	-14.923	19.444	38.677
CB^2-Ala_2	-15.653	19.882	37.406
C^2-Ala_2	-15.720	18.414	39.482
O^2-Ala_2	-15.662	17.220	39.196
N^3-Ala_3	-16.445	18.916	40.470
CA^3-Ala_3	-17.120	18.091	41.458
CB^3-Ala_3	-18.373	17.471	40.837
C^3-Ala_3	-16.147	17.035	41.984
O^3-Ala_3	-16.129	15.906	41.497
C^4-CMK	65.505	-100.000	5.843
H-CMK	90.048	-100.000	36.717
H-CMK	66.729	-100.000	100.00
Cl-CMK	-15.391	17.433	42.935
(ii) RESIDUES OF PPE MOLECULE			
N-His 57	-3.817	18.052	41.621
CA-His 57	-4.612	17.733	40.431
C-His 57	-3.922	18.195	39.152
O-His 57	-3.791	17.397	38.213
CB-His 57	-6.121	17.688	40.203
CG-His 57	-6.666	19.092	39.910
ND1-His 57	-6.625	19.629	38.652
CD2-His 57	-7.259	19.973	40.733
CE1-His 57	-7.914	20.873	38.712
NE2-His 57	-7.584	21.084	40.005
N-Val 99	-11.543	13.717	43.722
CA-Val 99	-11.178	15.101	44.049
C-Val 99	-12.121	15.687	45.098
O-Val 99	-11.664	16.435	45.973
CB-Val 99	-11.153	15.984	42.797

	X	Y	Z
CG1-Val 99	-12.568	16.295	42.315
CG2-Val 99	-10.351	17.272	42.975
N-Thr 175	-20.362	16.958	48.611
CA-Thr 175	-19.186	16.976	47.726
C-Thr 175	-18.035	17.767	48.338
O-Thr 175	-16.885	17.308	48.330
CB-Thr 175	-19.561	17.575	46.372
OG1-Thr 175	-20.599	16.791	47.795
CG2-Thr 175	-18.377	17.634	45.407
N-Gln 192	-11.130	26.644	34.322
CA-Gln 192	-10.012	25.856	33.784
C-Gln 192	-8.750	26.675	33.523
O-Gln 192	-8.804	27.692	32.822
CB-Gln 192	-10.447	24.934	32.645
CG-Gln 192	-11.273	23.734	33.112
CD-Gln 192	-12.340	24.170	34.110
OE1-Gln 192	-12.943	25.236	33.929
NE2-Gln 192	-12.817	23.183	34.854
N-Ser 195	-6.651	26.604	38.533
CA-Ser 195	-6.289	25.523	39.474
C-Ser 195	-4.490	25.680	40.163
O-Ser 195	-3.922	25.926	39.513
CB-Ser 195	-6.335	24.137	38.837
OG-Ser 195	-7.669	23.561	38.722
N-Ser 214	-9.288	24.330	43.566
CA-Ser 214	-9.451	22.904	43.287
C-Ser 214	-10.379	22.606	42.112
O-Ser 214	-9.925	22.064	41.097
CB-Ser 214	-9.926	22.170	44.535
OG-Ser 214	-9.912	20.776	44.267
N-Phe 215	-11.672	22.837	42.314
CA-Phe 215	-12.677	22.533	41.278

	X	Y	Z
C-Phe 215	-13.874	23.481	41.194
O-Phe 215	-13.972	24.445	41.966
CB-Phe 215	-13.158	21.077	41.361
CG-Phe 215	-13.870	20.728	42.682
CD1-Phe 215	-13.140	20.258	43.761
CD2-Phe 215	-15.250	20.818	42.775
CE1-Phe 215	-13.781	19.931	44.961
CE2-Phe 215	-15.900	20.493	43.971
CZ-Phe 215	-15.164	20.054	45.070
N-Val 216	-14.553	23.353	40.059
CA Val 216	-15.833	24.009	39.746
C-Val 216	-16.885	22.989	39.316
O-Val 216	-16.586	21.792	39.217
CB-Val 216	-15.668	25.055	38.642
CG1-Val 216	-14.934	26.305	39.121
CG2-Val 216	-15.067	24.480	37.361

The discussion will be mainly directed towards the binding of the inhibitor molecule. The inhibitor is found to bind at sites adjacent to the primary specificity pocket. The N-terminus shows it to be bound at S_1' subsite close to the active centre serine-195 and three amino-acid residues are associated in a parallel pleated-fashion with the protein chain at subsites S_1, S_2 and S_3 from residues 214 to 216. The CMK group was found close to the subsite S_4 (Figure 8.2-8.4). This arrangement is similar to that found by Hughes et al[51] regarding the inhibitor TFA-Lys-Ala-NH-C_6H_4-p-CF_3.

Various groups[1,8-10] have previously predicted the mode of binding of TFA-peptides using ^{19}F and ^{1}H N.M.R. spectroscopy and also by kinetic measurements. Dimicoli et al[1] predicted the position of TFA group in TFA-peptide-CMK to be close to the serine-195 in the S_1' subsites. They also showed that the reaction of the CMK with histidine-57 was hindered, which is also evident from X-ray crystallographic study.

Hence, it seems that the TFA-peptides bind in the opposite direction to that of the acetylated peptides[11]. Further PPE shows the direction of binding, as defined from amino to carboxyl termini of the ligand to be reversed compared with other serine protease inhibitor complexes[12-14]. However, recently Meyer et al[15] have complexed Ac-Ala-Pro-Ala-p-nitroanilide with PPE at 1.65 Å resolution and found the inhibitor to complex in a similar manner to that of Hughes et al[5] and TFA-Ala_3-CMK.

An alternative mode of binding for the TFA-Ala_3-CMK was fitted compared to that of Hughes et al[5]. The whole of the peptide was positioned in the similar manner except the TFA group was directed away from the

active site serine-195. There was no interaction between the two groups, (Figure 8.5-8.7).

(ii) Ac-Pro-Ala-Pro-Ala.OH

A difference Fourier map [F_{DER}-F_{NAT}] and a Fourier map were calculated using the phases of the native PPE[4]. Both maps were viewed on the Evans and Sutherland graphics system. Neither revealed continuous density around the serine-195, that is, neither of the S_1 or S'_1 subsites was occupied and it seems that perhaps the substrate did not bind by soaking the native crystals. To obtain adequate binding the technique of co-crystallisation should be tried.

(iii) CONCLUSION

It is evident from the study carried out, the TFA-peptides binds in a unique way compared to the acetyl-peptides. Further evidence to clarify the affinity of the TFA group is to compare the complexed structure of the TFA-Pro-Ala-Pro-Alaninal with PPE. This would also show the formation of transition state analogue as mentioned in Chapter One.

Since recently Meyer et al[15] have found the substrate Ac-Ala-Pro-Ala-p-nitroanilide binding in a similar manner to the TFA-peptides, further confusion as to the binding sites is raised.

However, more recently the inhibitor, Ac-Pro-Ala-Pro-Alaninal was found to bind in a similar manner as to the other serine proteases. That is at subsites S_1-S_4 along the protein chain 214-216, figure 8.8-8.9 (unpublished results of Dr Sawyer). This therefore furthers the issue about the whole binding sites in PPE and hence various studies need to be carried out before reaching a definite conclusion.

It is essential to pursue firstly, with the substrate Ac-Pro-Ala-Pro-Ala.OH and the inhibitor TFA-Pro-Ala-Pro-Alaninal to distinguish the S subsites. Secondly, with the complex between the anhydroelastase and H.Pro-Ala-Pro-Ala-Lys-Phe.OH to also establish the S and S′ subsites. The S′ subsites can also be found from complexes between the protease and protein inhibitor and then by analogy, possibly through model building studies.

PPE RESIDUES (ORANGE) + TFA-ALA$_3$-CMK INHIBITOR

(GREEN) BOUND AT SERINE-195

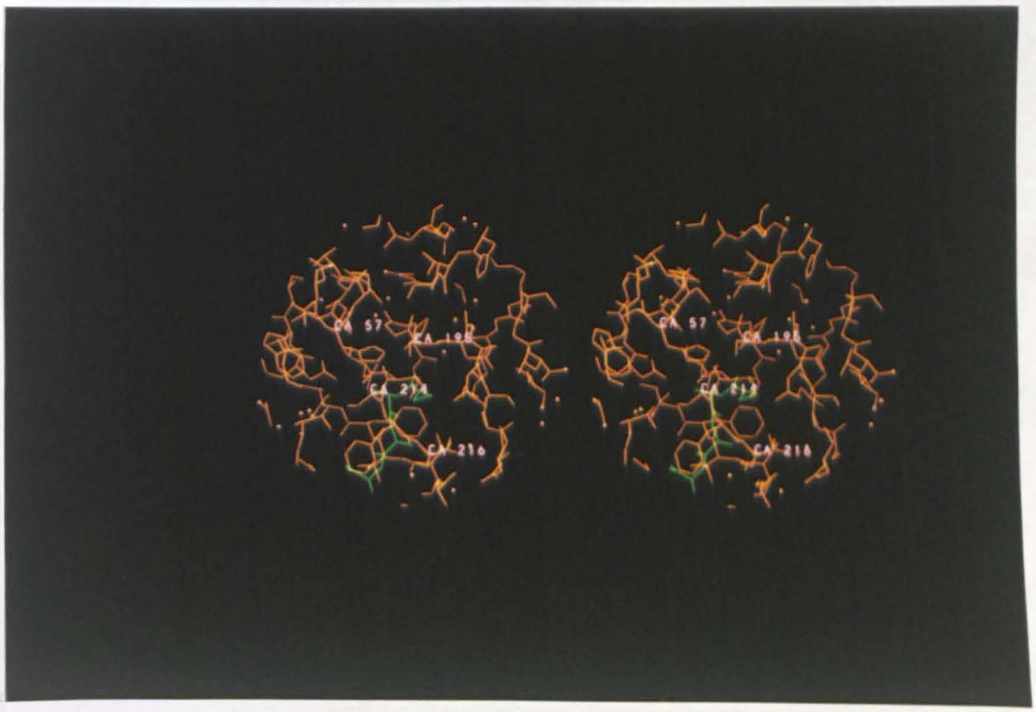


FIGURE 8.2

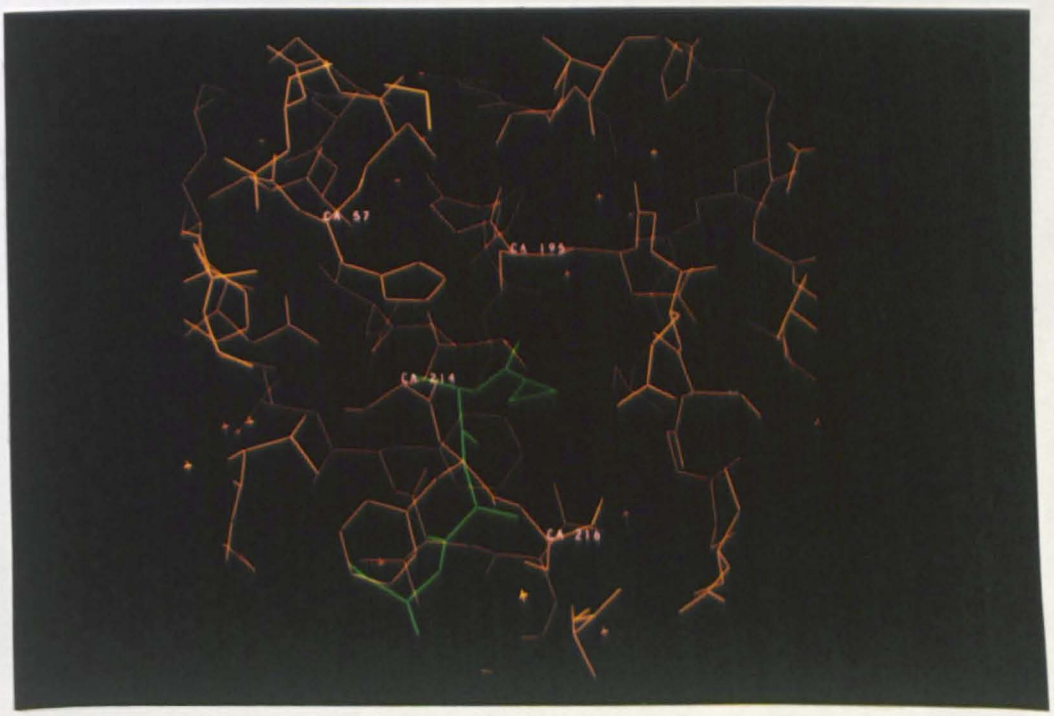


FIGURE 8.3

THE DENSITY AROUND THE TFA-ALA$_3$-CMK INHIBITOR BOUND AT SERINE-195

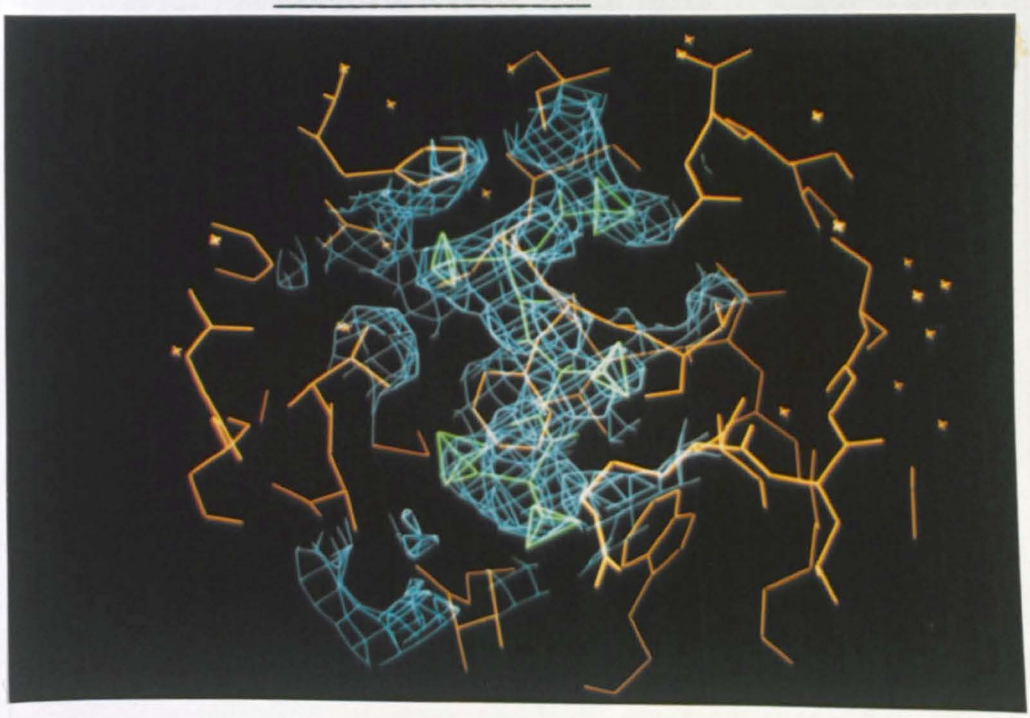


FIGURE 8.4

PPE MOLECULE (ORANGE) + TFA-ALA$_3$-CMK INHIBITOR (GREEN) BOUND AWAY FROM SERINE-195

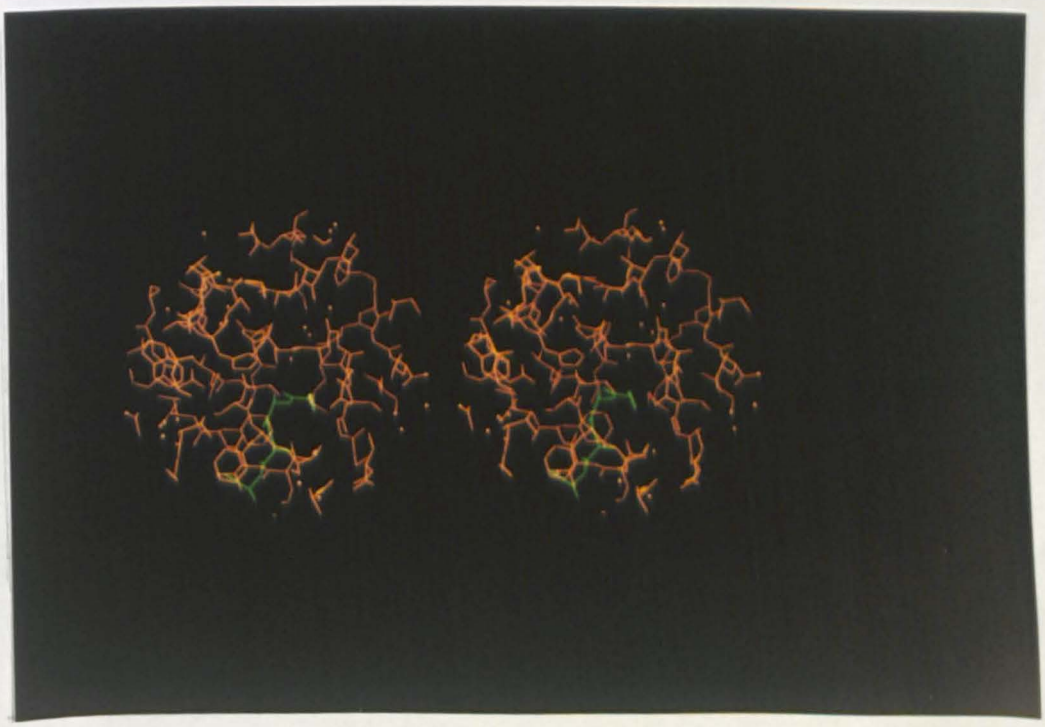


FIGURE 8.5

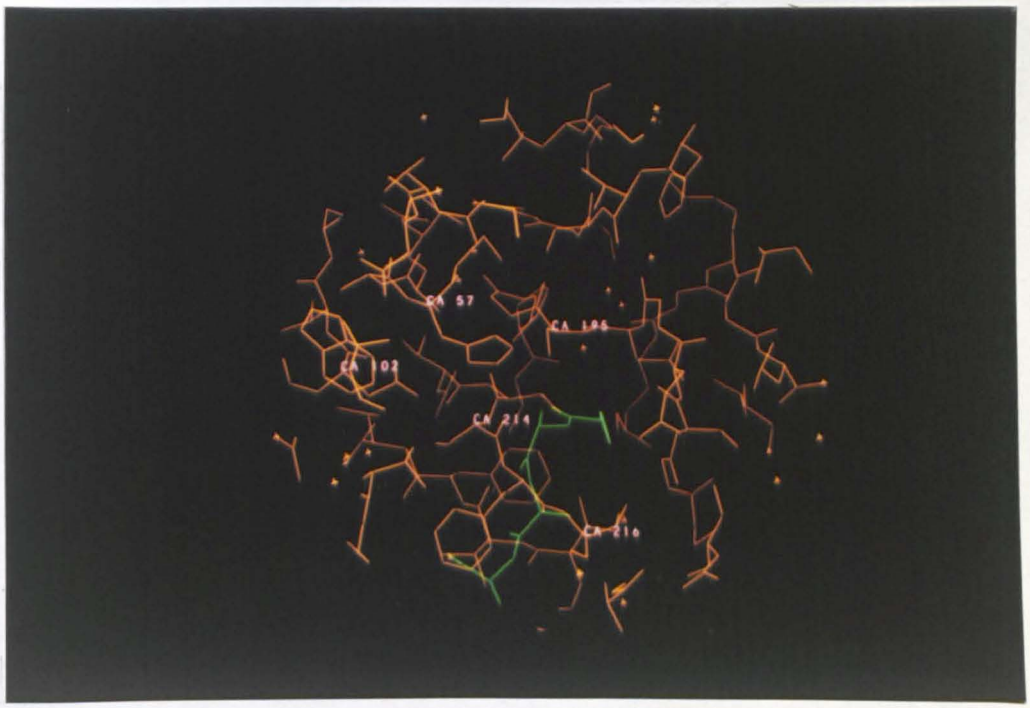


FIGURE 8.6

THE DENSITY AROUND THE TFA-ALA$_3$-CMK INHIBITOR BOUND AWAY FROM SERINE-195

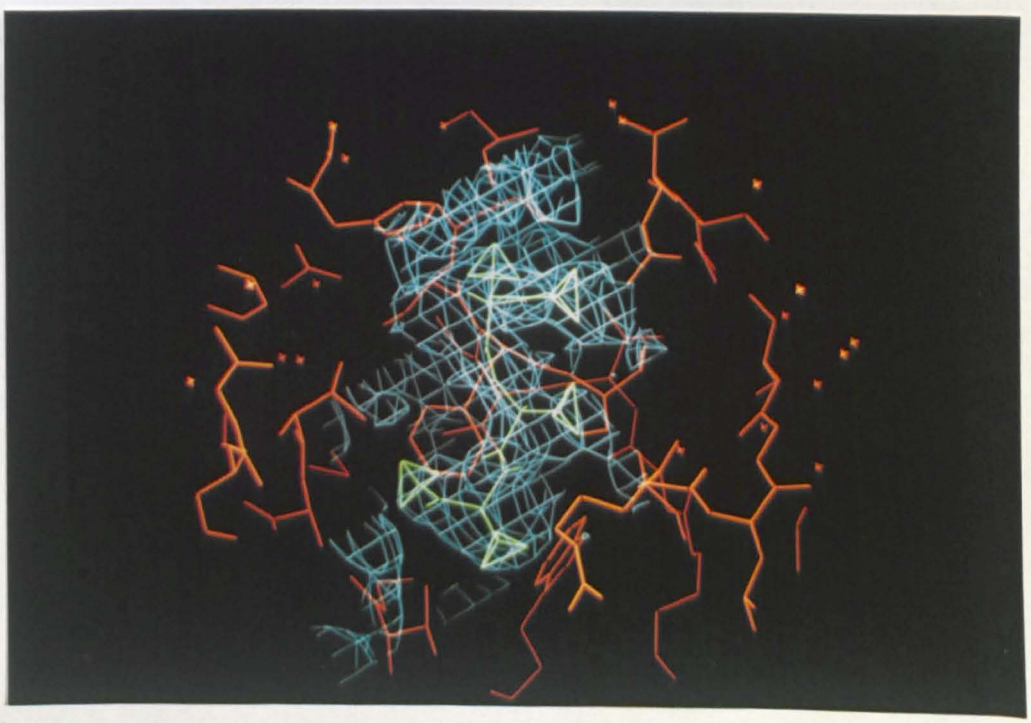


FIGURE 8.7

PPE RESIDUES (ORANGE) + AC-PRO-ALA-PRO-ALANINAL

INHIBITOR (GREEN) BOUND AT SERINE-195

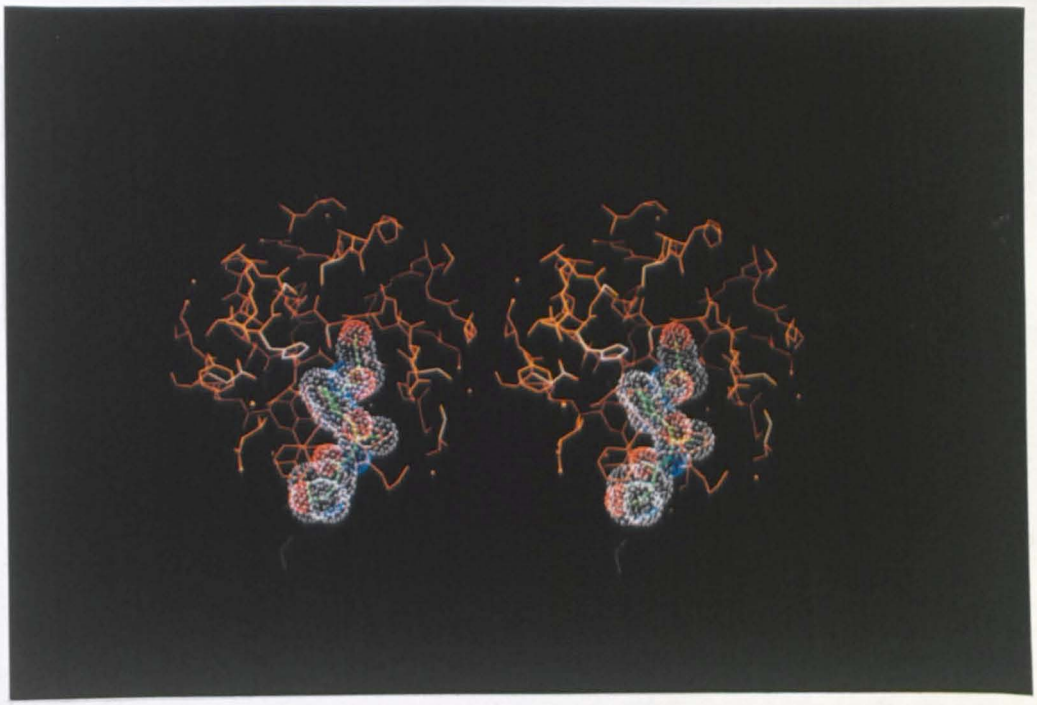


FIGURE 8.8

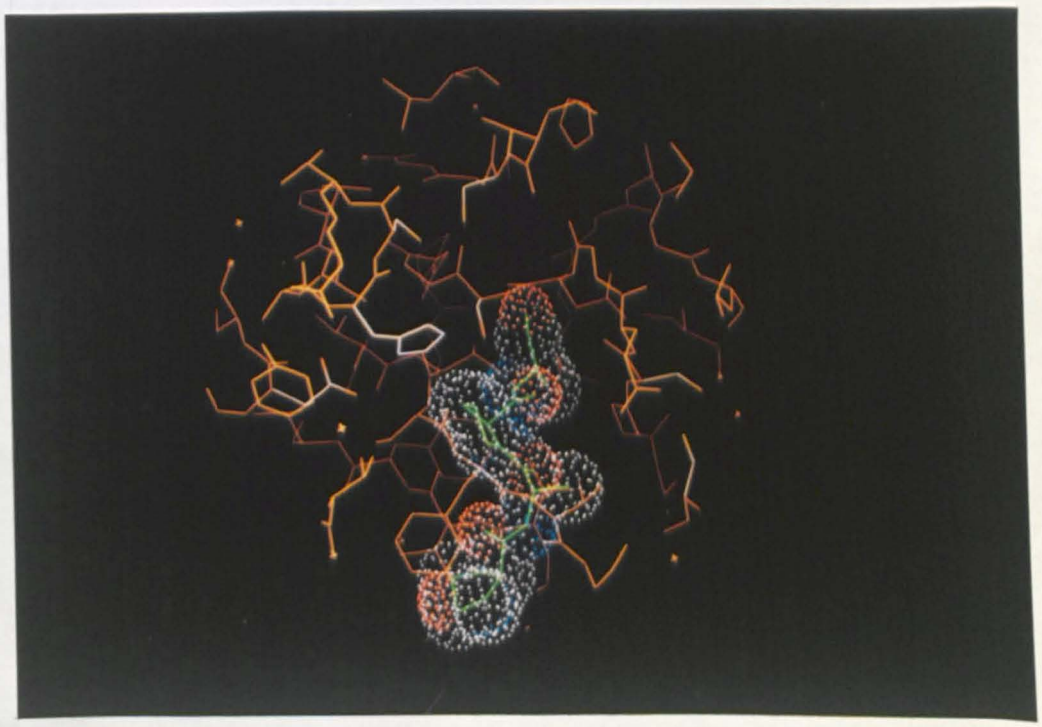


FIGURE 8.9

REFERENCES

1. Dimicoli, J.L., Renaud, A., Lestienne, P., and Bieth, J.G.
J.Biol.Chem., 254, 5208 (1979).

2. Shotton, D.M., Hartley, B.S., Camerman, N., Hofmann, T., Nyburg, S.C., and Rao, L.
J.Molec.Biol., 32, 155 (1968).

3. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Jr., Brice, M.D., Rodger, J.R., Kennard, O., Shimanouchi, T., and Tasumi, M.
Ibid., 112, 535 (1977).

4. Sawyer, L., Shotton, D.M., Campbell, J.W., Wendell, P.L., Muirhead, H., Watson, H.C., Diamond, R., and Ladner, R.C.
Ibid., 118, 137 (1978).

5. Hughes, D.L., Sieker, L.C., Bieth, J.G., Dimicoli, J.L.
Ibid., 162, 645 (1982).

6. Jones, A.
J.Appl.Cryst., 11, 268 (1978).

7. Schecter, I., and Berger, A.
Biochim.Biophys.Res.Commu., 27, 157 (1967).

8. Dimicoli, J.L., Bieth, J.G., and Lhoste, J.H.
Biochemistry, 15, 2230 (1976).

9. Dimicoli, J.L., and Bieth, J.G.
Ibid., 16, 5532 (1977).

10. Renaud, A., Dimicoli, J.L., Lestienne, P., and Bieth, J.G.
J.Amer.Chem.Soc., 100, 1005 (1977).

11. Shotton, D.M., White, N.J., and Watson, H.C.
Cold Spring Harbor Symp.Quant.Biol., 36, 91 (1971).

12. Sielecki, A.R., Hendrickson, W.A., Broughton, C.G., Delbaere, L.T.J., Brayer, G.D., and James, M.N.G.
J.Molec.Biol., 134, 781 (1971).

13. Steitz, T.A., Henderson, R., and Blow, D.M.
Ibid., 46, 337 (1969).

14. Segal, D.M., Powers, J.C., Cohen, G.H., Davies, D.R., and Wilcox, P.E.
Biochemistry, 10, 3728 (1972).

15. Meyer, E.F., Jr., Radhakrishnan, R., Cole, G.M., Presta, L.G.
In publication.

CHAPTER NINE

CONCLUSIONS

9.1 INTRODUCTION

PPE might be expected to exhibit similar substrate binding sites to those of the other serine proteases. However, low resolution studies (3.5 Å) by Shotton et al[1] have shown that the binding sites for substrate are apparently different to those of chymotrypsin and trypsin[2,3].

Therefore, the objective was to map out conclusively the substrate binding sites both to the N-terminal and C-terminal sides of the scissile bond. This could be investigated by; firstly binding synthetically produced substrate and inhibitor analogues with native PPE, secondly co-crystallisation of native PPE crystals with naturally occurring inhibitors and lastly, the modification of the enzyme's active site, thus allowing the binding of a complete peptide without it being hydrolysed.

Further justification for these studies was the confusion which had arisen[4] involving the binding of a trifluoroacetyl dipeptide inhibitor to PPE. (Details in section 1.9). Hence, the binding of two tetrapeptide inhibitors* TFA-Pro-Ala-Pro-Alaninal and Ac-Pro-Ala-Pro-Alaninal were investigated. The rationale behind these peptide aldehydes was firstly to distinguish the different modes of binding of acetyl and TFA groups. Since, it had been suggested that the TFA group binds to PPE in a unique binding mode[4,5]. Secondly, they bind to PPE more tightly than the substrates (refer to table 1.7). This is probably due to the formation of a transition state analogue, a hemiacetal[6], which is analogous to the tetrahedral intermediate (figure 1.7).

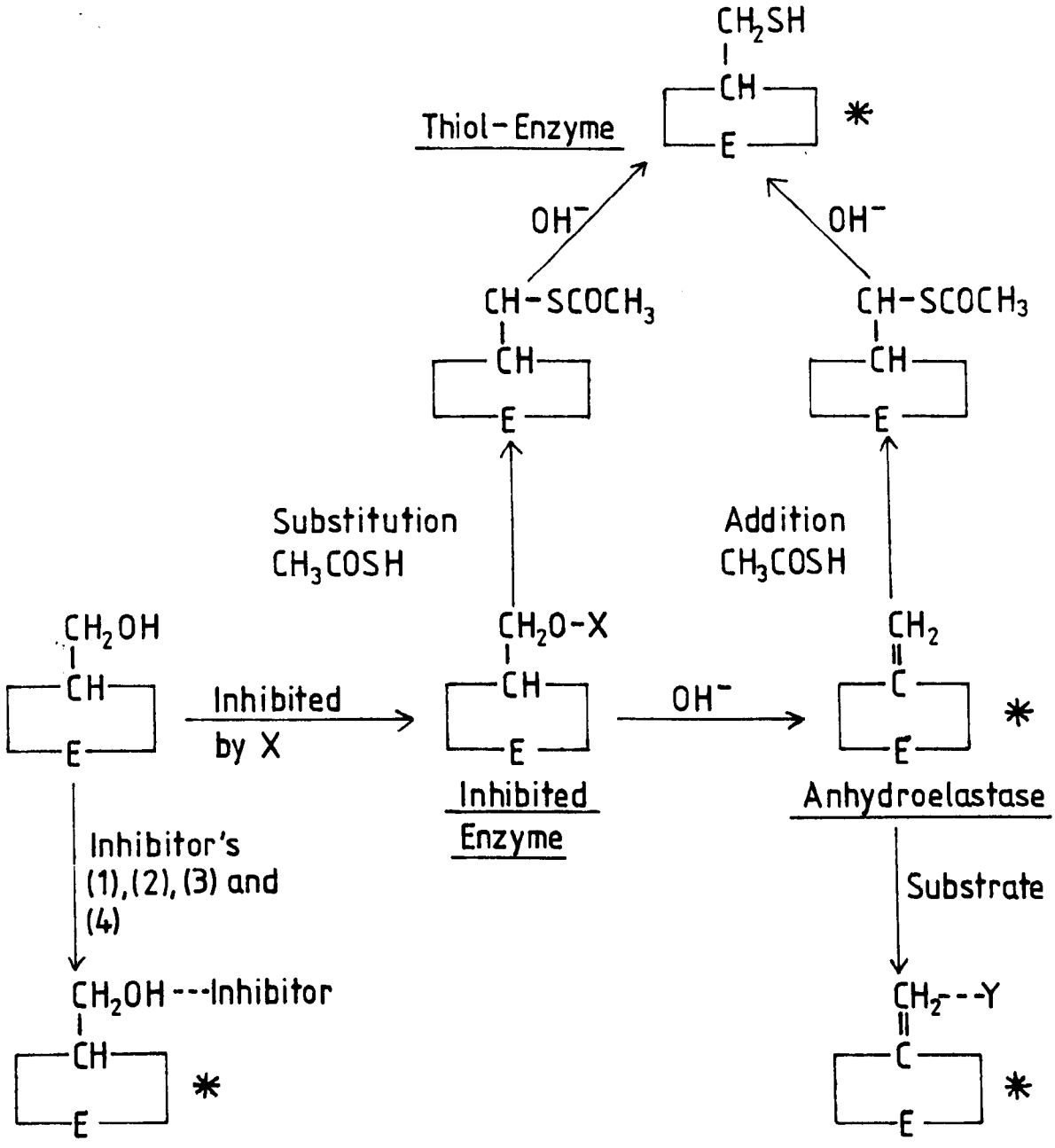
A further two peptide inhibitors; Ac-Pro-Ala-Pro-Ala.OH* and TFA-Ala_3-CMK (prepared by Bieth et al[7]) were examined for anomalous binding at 2.5 Å resolution. Investigation of Ac-Pro-Ala-Pro-Ala.OH was carried out since previous

work by Sawyer[8] (unpublished reported by Johnson et al) had failed to reproduce the results of Shotton et al[1].

The modification of the active site serine-195 entails inhibition by a sulphonyl fluoride inhibitor, followed by an elimination reaction to convert the $-CH_2OH$ into $=CH_2$. This totally inactivates the PPE. Low resolution studies (3.5 Å) on the modified enzyme[9], anhydroelastase have shown some changes in the active site. The refined structure of anhydroelastase together with that of the enzyme modified by diffusion of a substrate, e.g., H.Pro-Ala-Pro-Ala-Lys-Phe.OH* into the crystals should therefore reveal the binding sites on both sides of the scissile bond.

A further modification attempted was to convert serine-195 to cysteine with thiolacetate, either by substitution of sulphonyl derivative of PPE or by Michael-type addition to the double bond in anhydroelastase. This is of interest, since thiol proteases are analogous to the serine proteases and the conversion may provide completely different insight into the reactivity of PPE. Figure 9.0 summarises the overall scheme of the investigation.

* Refer to section 1.8 for the preference of these peptides and to section 3.8 for their synthesis.



X = Tosyl fluoride
= PMSF
= Dansyl halide
= NPA

Y = H·Pro-Ala-Pro-Ala-Lys-Phe·OH
(1) = Ac-Pro-Ala-Pro-Ala·OH
(2) = TFA-Pro-Ala-Pro-Alaninal
(3) = Ac-Pro-Ala-Pro-Alaninal
(4) = TFA-Ala_3-CMK

* Crystal structures to be investigated

FIGURE 9·0

9.2 PREPARATION AND MODIFICATION OF PPE

PPE was purified from trypsin 1-300 by the method of Shotton[10] (details of the preparation in section 2.3.1).

The preparation of anhydroelastase, is a two stage process as shown in figure 9.0. The first stage can be performed using non-radioactive label inhibitor, because the per cent inhibition will give an indication of the inhibition measured using a substrate, NBA. However, the second stage is crucial since the base elimination of the inhibitor is difficult to detect. Moreover, since the anhydroelastase formed is effectively inactive, there is no certainty that the inhibitor has been removed.

Earlier studies on serine pancreatic protease[9,11-14] have already shown, the enzyme can be fully inactivated with radioactive labelled i.e., ^{14}C labelled tosyl fluoride or PMSF and also used in the preparation of anhydroenzyme, so the elimination of inhibitor can be followed. Therefore, it might seem most feasibile to use these inhibitors, but for reasons of expense, availability and local restrictions it was considered desirable to attempt modification by non-radioactive reported groups as optimum conditions had already been found by Murphy[9].

Thus, the possibility of detecting the formation of anhydroelastase by substituting the latter inhibitors with fluorescent chromophores was investigated. The following were tried: Dansyl fluoride or chloride and NPA, since they had been found to inhibit α-chymotrypsin[15-17]. These could easily be detected on binding to the enzyme by fluorescent emission spectra at 505 mμ for dansyl elastase and 422 mμ for anthraniloly elastase. On base elimination to the formation of anhydroelastase no fluorescent emission spectra should be obtained.

Details of the preparation of the inhibited elastase, anhydroelastase and thiol-elastase are given in section 2.3.

9.3 DISCUSSION

PPE was fully inhibited (≃ 98%) by tosyl fluoride and PMSF within a time of 5 hours compared to 24 hours for the latter. However, with the fluorescent inhibitor's Dansyl halide and NPA no inhibition was achieved at the active site serine-195, whereas the serine-195, in the case of α-chymotrypsin was inhibited[15-17].

Although, no dansylation was achieved at the active site at pH 8.4[15], the dansyl elastase showed a fluorescence spectra on excitation at 340 mμ. A fluorescence maximum was observed at 505 mμ (figure 9.1). Thus, indicating that dansylation had probably taken place somewhere else on the molecule. A possible explanation is that under mild alkaline conditions the amino-acids most prone to dansylation were the ε - amino groups of lysine. PPE has exactly three lysine residues in the external positions available for dansylation. These are lysine-87, lysine-177 and lysine-224 which are 23.4 Å , 16.5 Å and 18.0 Å respectively from serine-195. From these, the lysine residues most prone to dansylation would be the second one. The N-terminal NH_2 is not available since it is buried forming a ion pair with asparate-194.

A FLUORESCENCE SPECTRUM OF DANSYL ELASTASE AT pH 8.4 ON EXCITATION AT 340 mμ.

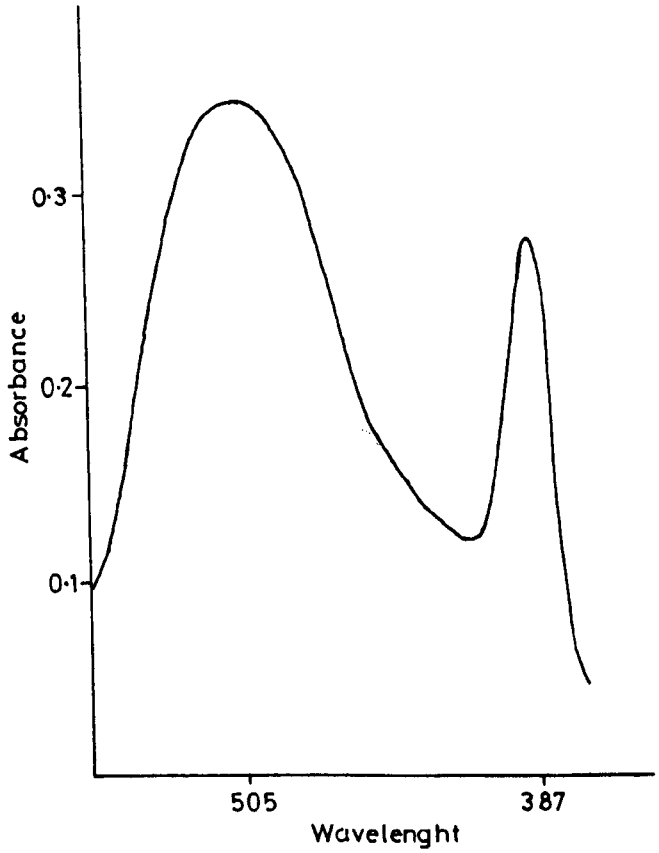


Figure 9.1

Further, attempts were carried out to dansylate PPE by varying the conditions from those of Horton et al[15] and Gold[16]. That is, performing the inhibition between pH 5.5-7.5, but no change in the activity and no fluorescence spectrum were obtained.

Hence, the only successful means that remained for preparing anhydroelastase were from tosyl elastase or PMSE. However, doubt remained in the formation of the double bond, since both the inhibited PPE and anhydroelastase were inactive. Hence, other methods for the detection of the double bond needed to be investigated; using Dansyl halide and thiolacetate.

Attempts to dansylate anhydroelastase, indicated no fluorescence maximum, demonstrating that the serine-195 was converted to dehydroalanine, and hence the OH was not available to dansylation. Also, there was no indication of dansylation taking place anywhere else in the molecule (as happens in the native elastase). This suggested that the amino-acids prone to dansylation were affected under the strong alkaline conditions in the preparation of anhydroelastase.

Secondly, addition to the double bond to convert serine-195 into cysteine by thiolation and substitution of inhibited PPE proved experimentally difficult. This was also observed by Murphy[9] with tosyl elastase. Varying the conditions form his, that is, pH 5.3 to pH 6.5 and 7.5 and performing the reaction at 0^{o} C and room temperature rather than at 2^{o} C did not result in any change in the reaction.

TABLE 9.1

SUMMARY OF THE INHIBITION OF PPE BY VARIOUS INHIBITORS

Inhibited PPE	pH	Time (hours)	Temperature (^{o} C)	Inhibition (%)	Activity After OH^- treatment (%)	Fluorescence (mμ)
Native PPE	7.2	24	4	-	≃97	-
Tosyl elastase	7.2	24	4	≃98	≃ 2	-
PMSE	7.2	4	4	≃98	≃ 2	-
Anthraniloyl elastase	6.8	24	6	≃ 2	*	-
Dansyl elastase	5.5,6.0,6.5 7.2 and 7.5	48	4	≃40	*	-
	8.4	48	4	≃20	*	505
Thiol-elastase from						
Tosyl elastase	6.5	6	0 and 20	≃ 2		
or	and					
PMSE	7.5	6	0 and 20	≃ 2		
Anhydroelastase	6.5 and 7.5	6	0 and 20	≃ 2		

* No alkaline treatment, since no inhibition was observed

Thus, uncertainty still remained as to whether the formation of anhydroelastase had been achieved. Therefore, a 2.4 Å resolution data set was collected, but after processing and comparing the electron density around the active site region showed no change. (Refer to the mono and stereo figure 2.0-2.1 showing a small region around the serine-195 with the electron density around these residues). The serine O_γ is still clearly visible, indicating the tosyl and the PMS groups were removed but the expected elimination had not occurred.

However, the anhydroelastase prepared was inactive and further more it crystallised indicating no gross conformational changes or denaturation. Hence it was decided to investigate further by soaking the crystals in the hexapeptide substrate, H.Pro-Ala-Pro-Ala-Lys-Phe.OH. Collecting the data to 2.5 Å resolution followed by processing should reveal the mode of binding. If residue 195 were dehydroalanine, the whole peptide should be found intact straddling the active site. But if the enzyme was active only the Ac-Pro-Ala-Pro-Ala.OH would be observed.

For further details on the discussion refer to section 2.4

9.4 EXPERIMENTAL CONDITIONS FOR BINDING OF THE SUBSTRATE/ INHIBITOR TO PPE

Details of the experimental conditions and cell parameters are specified in section 8.2 and table 8.1. However, of all the complexes prepared only two data sets to 2.5 Å resolution could be collected (Ac-Pro-Ala-Pro-Ala-OH and TFA-Ala_3-CMK). In the case of the others, crystal complexes are mounted and ready for data collection in the future.

9.5 DATA COLLECTION AND PROCESSING

Intensity data for both derivatives (Ac-Pro-Ala-Pro-Ala.OH and TFA-Ala_3-CMK) were collected on a computer controlled Enraf-Nonius CAD4 diffractometer. All of the data for TFA-Ala_3-CMK were measured on a single crystal; a total of 8162 unique reflections ($2\theta \leqslant 36^{o}$) were recorded by the 2θ-ω scan technique to 2.5 Å resolution by Dr L Sawyer. Two crystals were required for Ac-Pro-Ala-Pro-Ala.OH; with a total of 7249 unique reflections ($2\theta \leqslant 36^{o}$), recorded by the ω scan to 2.5 Å resolution. Both data collections were performed just below room temperature (15^{o} and 18^{o} C) with nickel filtered Cu-K_{α} radiation.

The data sets were TFA-Ala_3-CMK and Ac-Pro-Ala-Pro-Ala.OH were corrected for Lorentz polarization, absorption and time-fall off (Refer to table 8.2 for Data Collection Parameters). After applying all the corrections, the data for the two crystals of Ac-Pro-Ala-Pro-Ala.OH were merged together by scaling the common reflections. The $R_{merge} = \Sigma\ (I(i) - \langle I \rangle)/\ \Sigma\ I(i)$ for this was 0.032, (with I(i) being the intensity value of an individual measurements, $\langle I \rangle$ being the corresponding mean value of the i measurements; the summation is over all reflections with more than one measurement).

The phase set for the native PPE structure[18] was used together with the reflections of TFA-Ala_3-CMK and Ac-Pro-Ala-Pro-Ala.OH in order to calculate difference Fourier maps, with coefficients $|F_{DER}| - |F_{NAT}|$, α_{NAT} and also a Fourier map for the latter derivative. The difference Fourier map for the TFA-Ala_3-CMK revealed only contiguous density around the active site region of the PPE molecule. However, neither map for the Ac-Pro-Ala-Pro-Ala.OH revealed any continuous density around the serine-195 even when the contour level was reduced to the level of the noise.

The Evans and Sutherland graphics system (at the University of Leeds) was used for positioning the TFA-Ala_3-CMK molecule in the active site region. The program, FRODO[19], was used to adjust the atomic parameters to be fitted best both to the X-ray data and for a given set of geometrical dimensions expected in the amino-acid residues. The TFA-Ala_3-CMK molecule will be described in two orientations: One with the N-trifluoroacetyl group as the N-terminal amino-acid and the other with the CMK group as the C-terminus.

Refer to figure 8.1 for labelling of the inhibitor molecule and to tables 8.3-8.5 for shorter inhibitor-PPE interactions, hydrogen-bonding and atomic parameters for the inhibitor molecule and selected neighbouring PPE residues.

9.6 CRYSTALLOGRAPHIC RESULTS FOR THE INHIBITOR TFA-ALA_3-CMK

The inhibitor, TFA-Ala_3-CMK is found to bind at sites adjacent to the primary specificity pocket but in the opposite direction to that of the other serine protease inhibitor complex[20-22]. Figures 8.4 and 8.7 show a continuous chain of electron density running from the region close to histidine-57 and serine-195 and along the main chain of PPE from residues serine-214 up to the vicinity of serine-217. Figure 8.7 shows an alternative mode of binding of the TFA group with the rest of the peptide positioned in the similar manner to figure 8.4.

In figure 9.2, the NE2-histidine-57.....OG-serine-195 contact is 3.3Å in native elastase[18] but the histidine ring is not in the correct orientation to form a good hydrogen bond. This point is still controversial as discussed by Tsukada et al[23].

However in contrast, the binding of TFA-Ala_3-CMK with PPE results in a slight shift of the histidine ring and hence

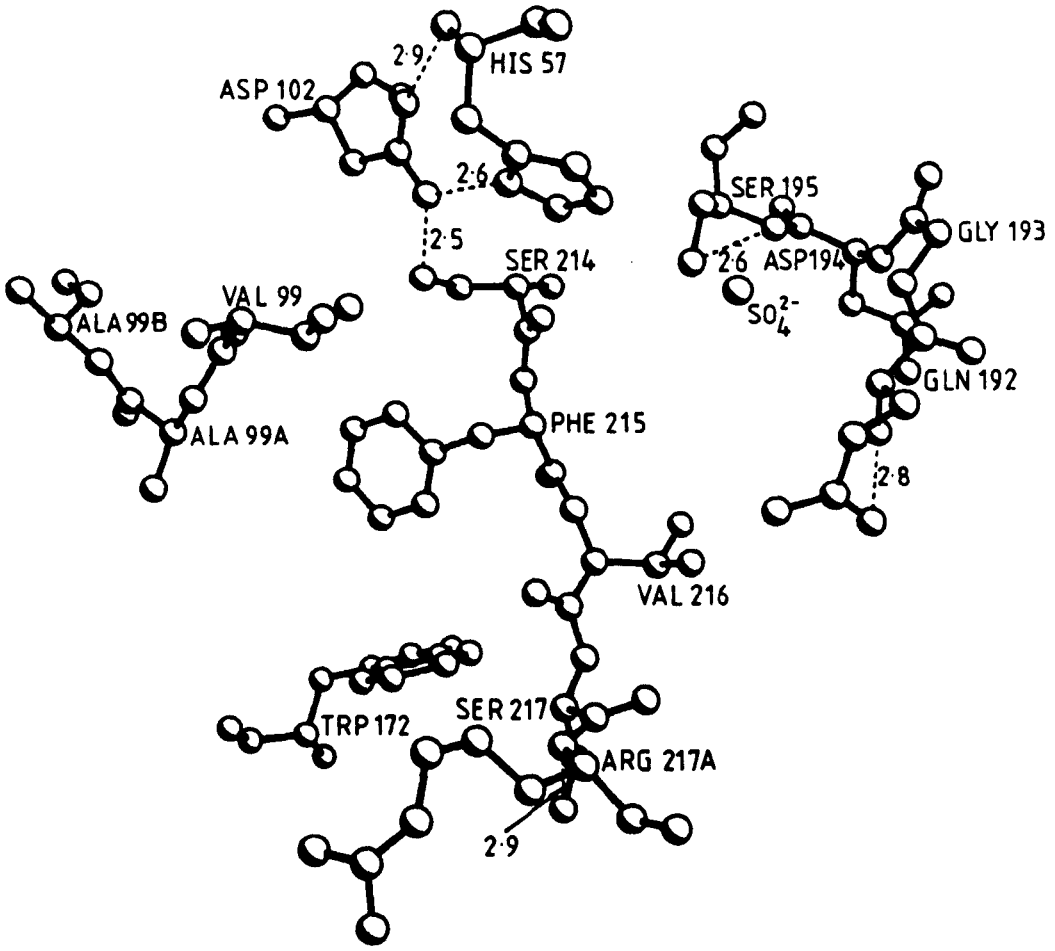
in a good hydrogen bond of 2.6 Å . This was also observed by Hughes et al[4] but not in the complex of Ac-Pro-Ala-Pro-Alaninal (Figure 9.3 (c) and (d)).

However, a weaker hydrogen bond between OD2-asparate-102 .....ND1-histidine-57 of 3.0 Å results but previous crystallographic studies on the other serine proteases have shown to be similar to the native elastase (2.6 Å) [18,25-25]. This was also observed in TFA-Lys-Ala-NH-C_6H_4-p-CF_3 and Ac-Pro-Ala-Pro-Alaninal complexes.

The inhibitor molecule and all its surrounding residues of PPE molecule are shown in stereo figure 9.3 (a) and (b) and figure 9.3 (c) shows the similar binding of TFA-Lys-Ala-NH-C_6H_4-p-CF_3, inhibitor that was observed by Hughes et al[4]. Figure 9.4 shows a schematic representation of the interactions formed between the inhibitor's and PPE at S and S' subsites.

Figure 9.2 shows the active centre region of native PPE (pH 5.0). Hydrogen bonds are shown as dotted and all distances are in Å . Labels are associated with the C_α - atoms. In the figure the SO_4^{2-} ion represents the sulphur atom with the oxygens $\simeq$ 2.0 Å away.

The active centre region showing hydrogen bond contacts (shown as dotted and all distances are in Å) between the inhibitor molecule and PPE residues in Figure 9.3; (a) TFA-Ala_3-CMK (b) TFA-Ala_3-CMK (alternative mode of binding) (c) TFA-Lys-Ala-NH-C_6H_4-p-CF_3 (d) Ac-Pro-Ala-Pro-Alaninal. Labels are associated with the C_α - atoms. Stereoviews of the inhibitor molecules and the surrounding residues are exactly equivalent to the above figures to which reference should be made for labelling.



STEREOVIEW OF NATIVE ELASTASE (pH 5.0)

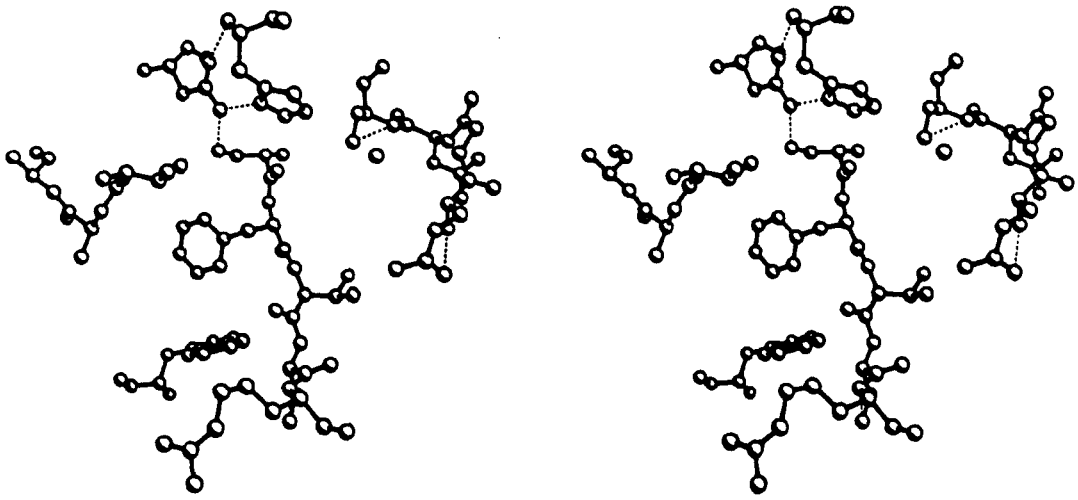
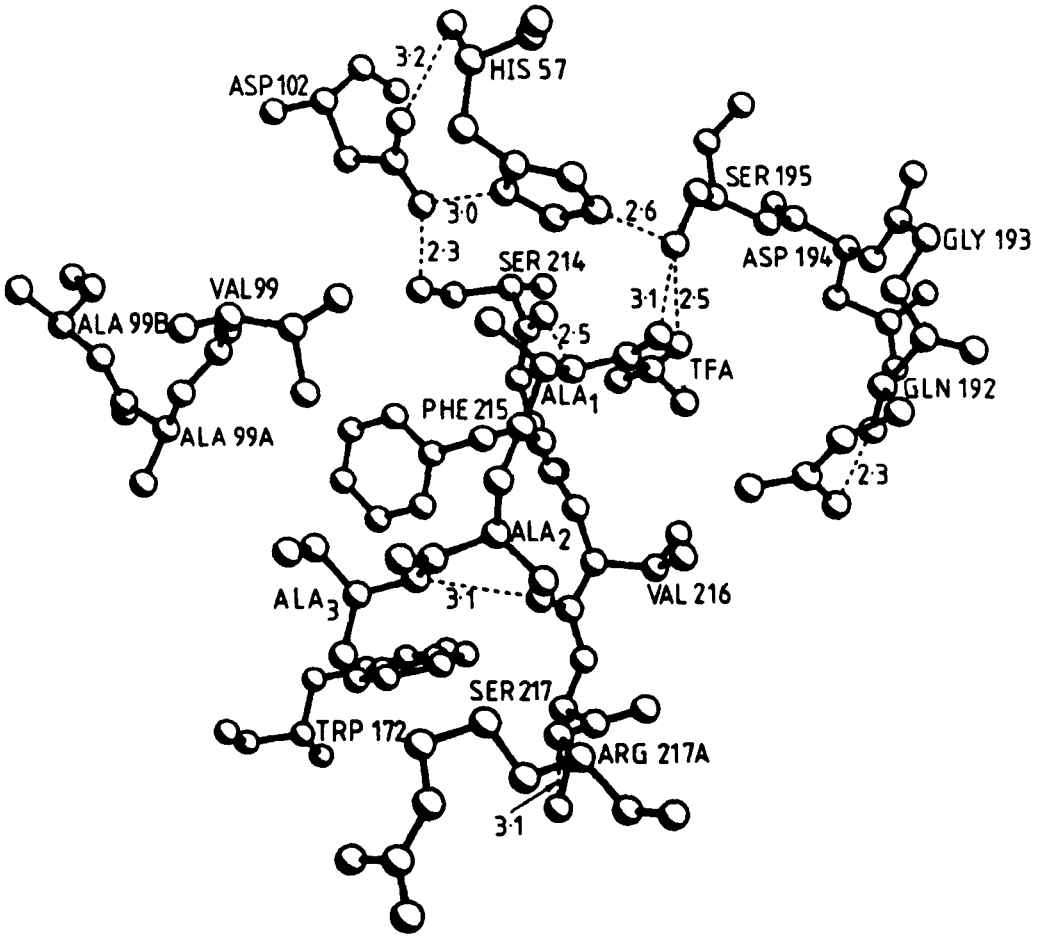
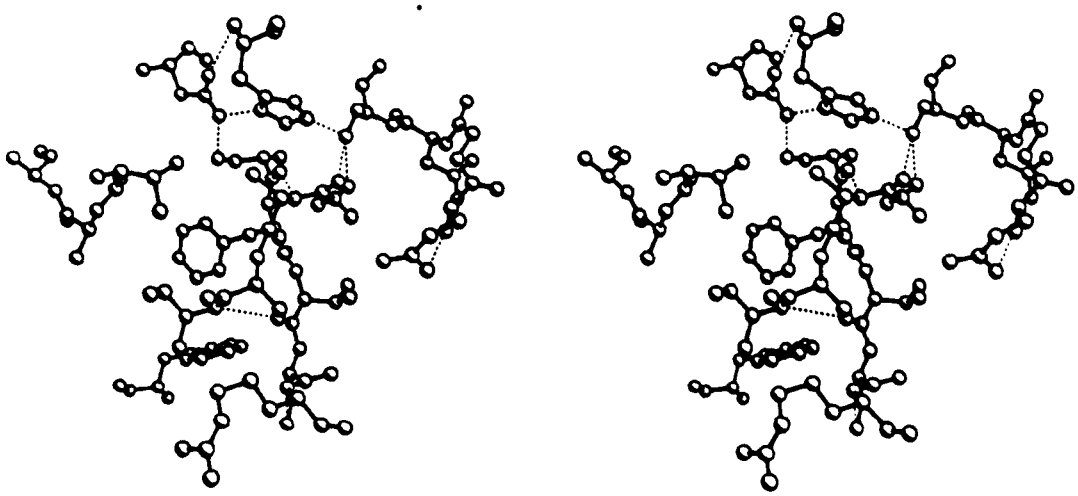


Figure 9.2

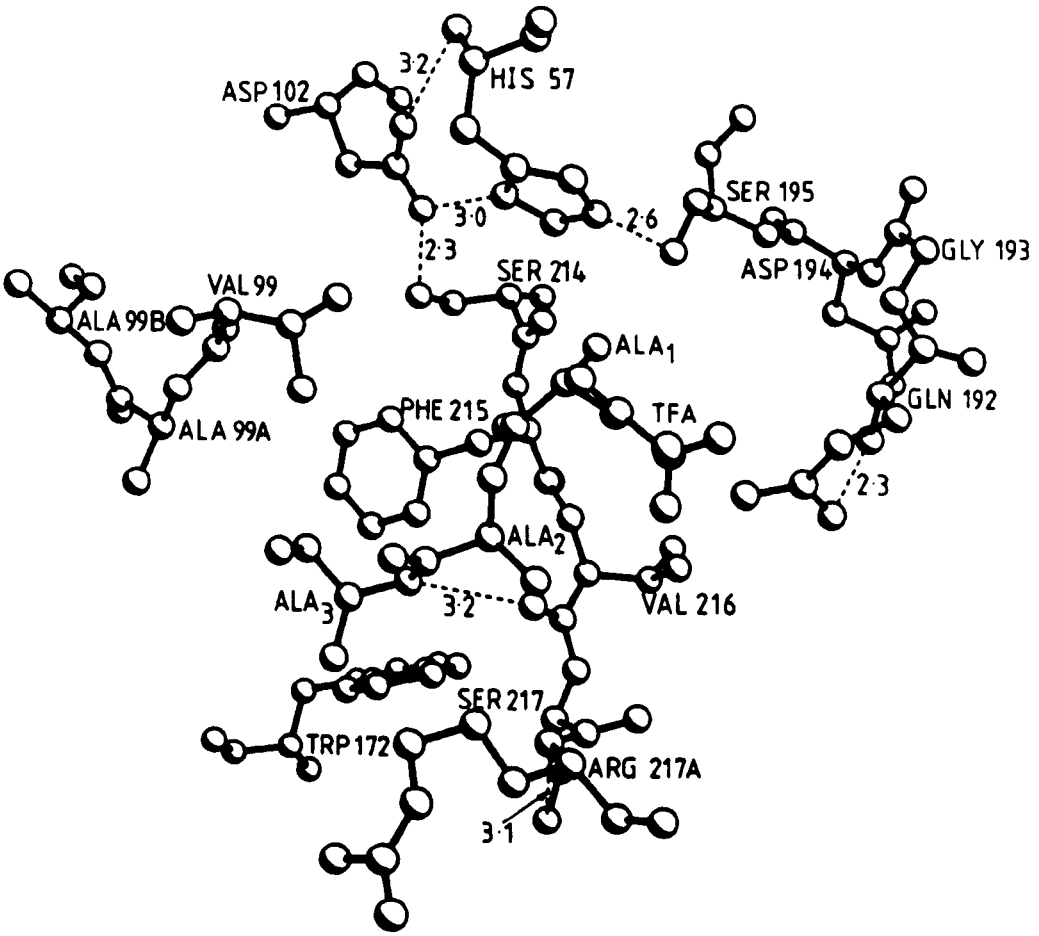
(a)



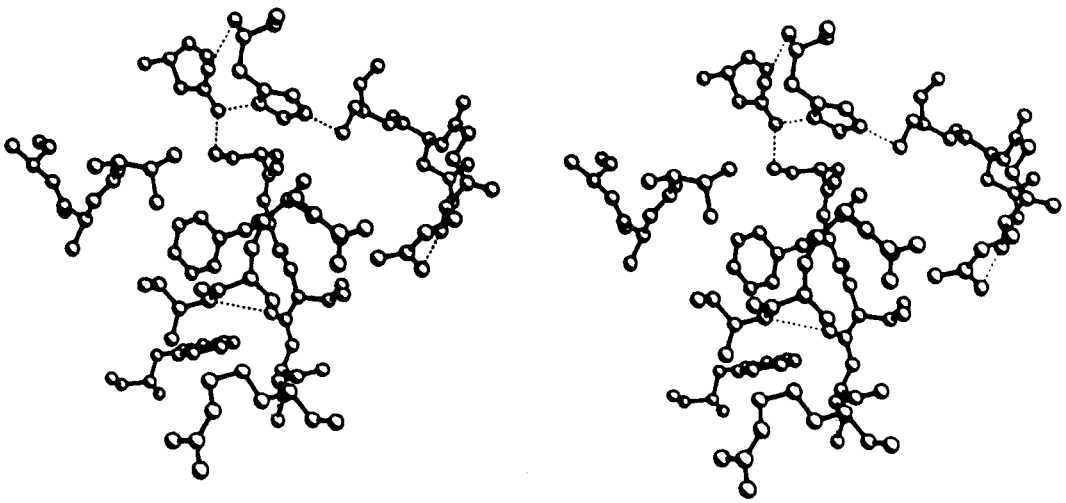
STEREOVIEW OF TFA- ALA_3-CMK



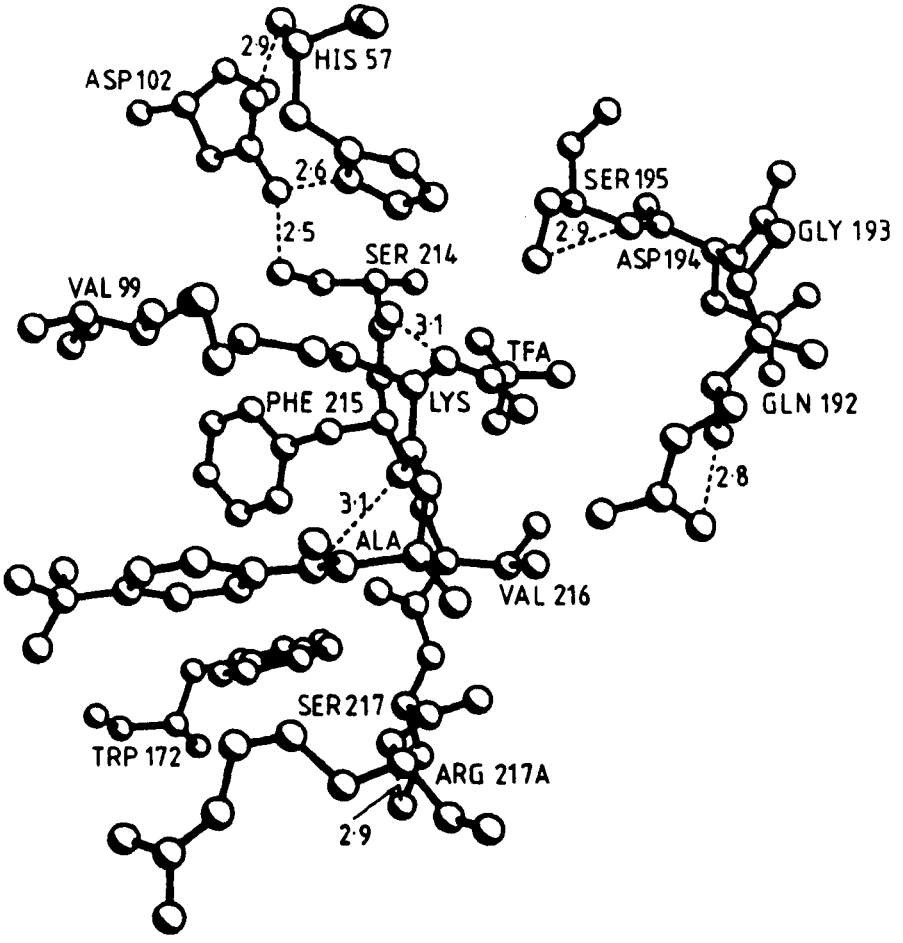
(b)



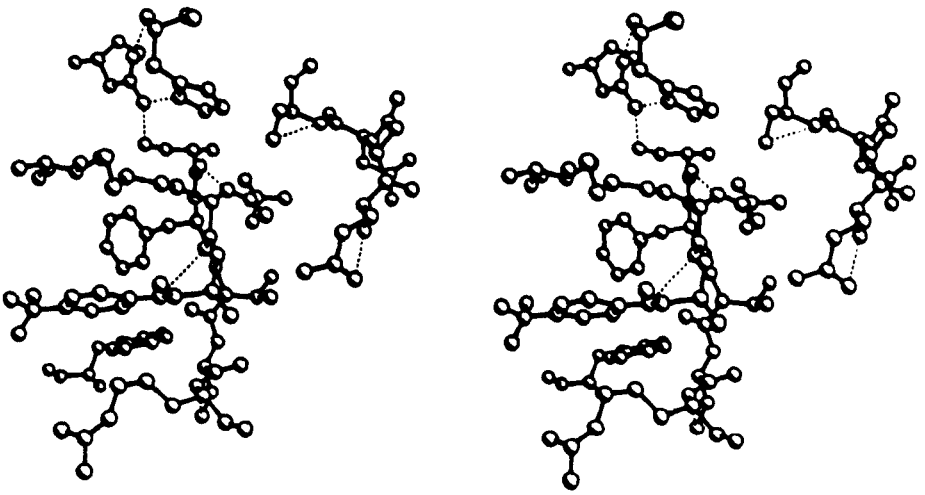
STEREOVIEW OF TFA-ALA$_3$-CMK (ALTERNATIVE MODE)



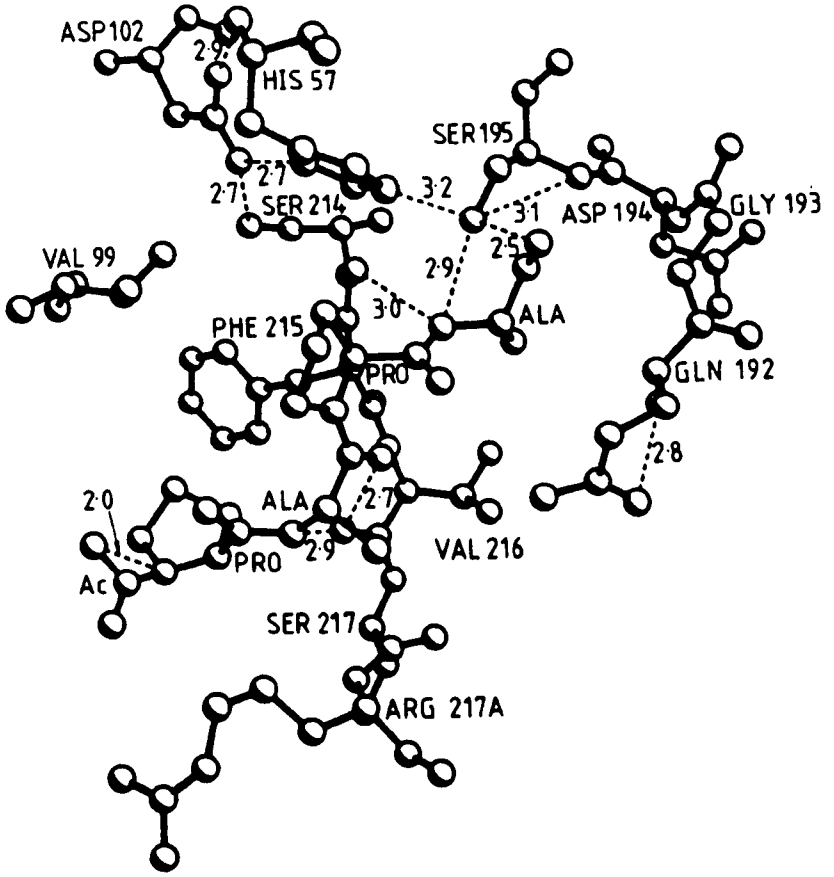
(c)



STEREOVIEW OF TFA-LYS-ALA-NH-C_6H_4p-CF_3



(d)



STEREOVIEW OF Ac-PRO-ALA-PRO-ALANINAL

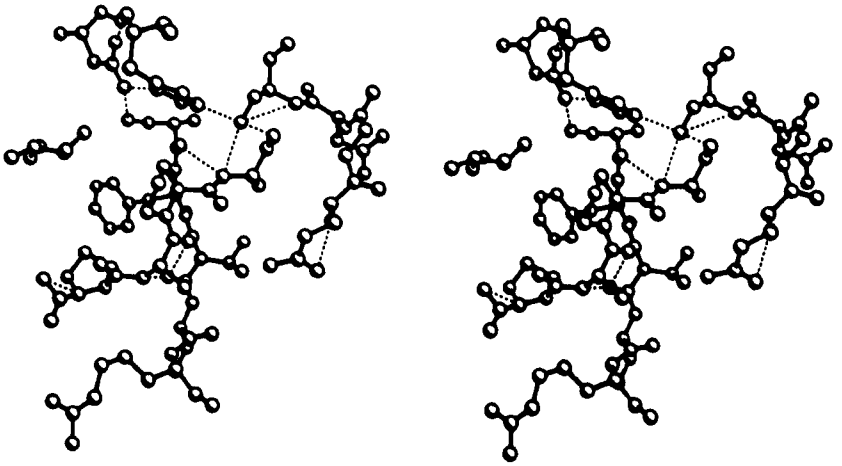
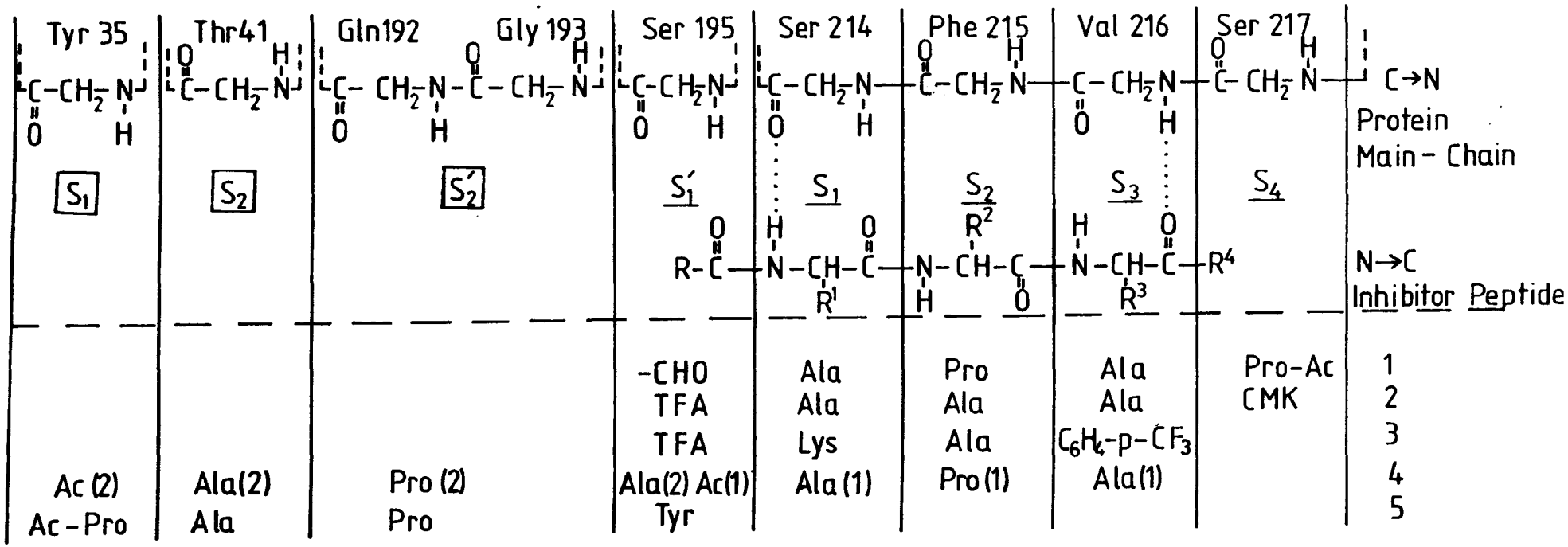


Figure 9.3



Inhibitor Peptide

1 = Ac-Pro-Ala-Pro-Alaninal

2 = TFA-Ala_3-CMK

3 = TFA-Lys-Ala-NH-C_6H_4-p-CF_3

4 = Ac-Ala-Pro-Ala-OH (TWO MODES OF BINDING (1) AND (2))

5 = Ac-Pro-Ala-Pro-Tyr·NH_2

Schematic representation of the interaction between various peptide inhibitors and PPE. Hydrogen bonds are shown as dotted, and all the inhibitors, except (1) take on an approximately parallel-pleated fashion. The subsites, [S], are the one's assigned to peptide inhibitors 4 (2) and 5.

Figure 9.4

The N-terminus shows it to be bound at S'_1 subsite close to the active centre with OG-serine-195 equidistant from each C-, O- and CT-TFA and with the F^2-TFA and F^3-TFA about 3.7 Å away. This contrasts with Hughes et al[4] where the C-, CT- and the two fluorine's of the TFA group were equidistant at 3.3 Å from OG-serine-195. The three alanine residues are associated in a parallel-pleated fashion with the protein chain at subsites S_1, S_2 and S_3 from the main chain residues 214 to 216. This arrangement is stabilised by strong hydrogen bonding of $N\text{-}Ala_1\text{-}H.....O\text{-}serine\text{-}214$ and $N\text{-}Ala_3\text{-}H.....O\text{-}valine\text{-}216$. There is also a twist of about 90^o consistent with that found in β-sheets[26]. This arrangement is that also observed by Hughes et al[4] and very recently on two peptide inhibitors Ac-Pro-Ala-Pro.OH and Ac-Pro-Ala-Pro-Tyr.NH_2 complexed with PPE were found to bind in the similar manner[27-28], that is, parallel to 214-216.

Various groups[5,29-31] have previously predicted the mode of binding of TFA-peptides using ^{19}F and ^1H N.M.R. spectroscopy and also by kinetic measurements. Dimicoli et al[7] predicted the position of TFA group in TFA-peptide-CMK to be close to the serine-195 in the S'_1 subsites. They also showed that the reaction of the CMK with histidine-57 was hindered, which is also evident from X-ray crystallography study (the CMK group is not shown in figure 9.3 (a) and (b) since the resolution was not good and the CMK end is more mobile, being towards the solvent). Hence it seems that the TFA-peptides bind in the opposite direction to that of the acetylated peptides. Further, PPE shows the direction of binding, as defined from the carboxyl termini of the ligand to be reversed compared with other serine proteases inhibitor complexes. This is clearly seen when comparing the TFA and the aldehyde inhibitors.

This result contrasts sharply with those reported for N-acetylated short peptide substrate[1] and inhibitor[32] molecules where the amino-acids residues are associated

at subsites S_1, S_2 etc and the N-acetyl group at subsite S_4, far removed from the active site (figure 9.3 (d)).

9.7 OVERALL CONCLUSION

It is obvious from the studies carried out on the modification of PPE, that the use of the fluorescent chromophores gave some indication of the mode of binding, since these inhibitors were found to inhibit α-chymotrypsin efficiently. It suggested that the binding site of PPE (P_1) is different and can only accommodate a small group such as methyl because of the occlusion by valine-216 and threonine-226, whilst in α-chymotrypsin these amino-acids are replaced by glycines-216 and 226. It has also shown that radioactive labelled tosyl fluoride or PMSF need to be used in the preparation of anhydroelastase to be able to monitor its formation in a fully satisfactory manner.

The X-ray analyses have confirmed the TFA group in the inhibitor, TFA-Ala_3-CMK, binds in a unique mode and in the same orientation as Hughes et al[4]. However, recent cyrstallographic studies on two peptide inhibitors; Ac-Ala-Pro-Ala.OH[27] and Ac-Pro-Ala-Pro-Tyr.NH_2[28] with PPE have contradicted the previous results of acetyl peptides and shown these peptides to bind in similar directions to the TFA-peptides.

However, the results on the inhibitor, Ac-Pro-Ala-Pro-Alaninal, (figure 9.3 (d)) indicates the inhibitor binding in a similar manner to the other serine proteases. That is, at subsites S_1-S_4 along the protein chain 214-216 with acetyl group at subsite S_4 and the inhibitor forming an anti-parallel sheet with the enzyme.

These small peptide inhibitors are apt to be artifacts but in vivo, one excepts a unique binding between the PPE and its natural substrate, elastin[33]. Elastin is large in structure and is elastic fibrous protein of connective tissue.

Thus, the overall picture remains somewhat confused as regards substrate binding to PPE. It seems more obvious now that a great deal of research is required on binding of substrates/inhibitors to PPE and by collating all the results before one can come to some definite conclusion.

This can be pursued, firstly with the substrate Ac-Pro-Ala-Pro-Ala.OH, by ensuring adequate binding, by using the technique of co-crystallisation. Further, data on the inhibitor complex of TFA-Pro-Ala-Pro-Alaninal needs to be collected to clarify the affinity of the TFA group and would also show the formation of transition state analogue as mentioned in chapter one. Secondly, to establish the S and S' subsites the complex between anhydroelastase and the substrate H.Pro-Ala-Pro-Ala-Lys-Phe·OH needs to be investigated. Further, it might help solve the problem of binding by complexing Ac-Pro-Ala-Pro-Ala-Lys-Phe.OH and TFA-Pro-Ala-Pro-Ala-Lys-Phe.OH peptides with anhydroelastase.

Various protein inhibitors[34-36], have already been shown to inhibit PPE, thus enabling one another means of investigating the possible S' subsites. These could be found from the complexes between the protease and protein inhibitor and then by analogy, possibly through model building studies.

REFERENCES

1. Shotton, D.M., White, N.J., and Watson, H.C.
Cold.Spring.Harbor.Symp.Quant.Biol., 36, 91 (1971).

2. Steitz, T.A., Henderson, R., and Blow, D.M.
J.Molec.Biol., 46 337 (1969).

3. Segal, D.M., Powers, J.C., Cohen, G.H., Davies, D.R., and Wilcox, P.E.
Biochemistry, 10, 3728 (1972).

4. Hughes, D.L., Sieker, L.C., Bieth, J., and Dimicoli, J.L.
J.Molec.Biol., 162, 645 (1982).

5. Dimicoli, J.L., Bieth, J., AND Lhoste, J.M.
Biochemistry, 15, 2230 (1976).

6. Thompson, R.C., and Blout, E.R.
Ibid, 12, 47 (1973).

7. Dimicoli, J.L., Renaud, A., Lestienne, P., and Beith, J.
J.Biol.Chem., 254, 5208 (1979).

8. Johnson, P., and Smillie, L.B.
Febs.Lett., 47, 1 (1974).

9. Murphy, S.
Ph.D. Thesis, University of Sheffield (1978).

10. Shotton, D.M.
Methods in Enzymology, 19, 113 (1970).

11. Huber. R., Bode, W., Kukla, D., Kohe, U., and Ryan, C.A.
Biophys.Struct.Mech., 1, 189 (1975).

12. Sayers, C.A., and Barrett, A.J.
Biochem.J., 189, 225 (1980).

13. Weiner, H., White, W.N., Hoare, D.G., and Koshland, D.E., Jr.
J.Amer.Chem.Soc., 88, 3851 (1966).

14. Strumeyer, D.H., White, W.N., and Koshland, D.E., Jr.
Biochemistry, 50, 931 (1963).

15. Horton, D.R., and Koshland, D.E., Jr.
Methods in Enzymology., Vol. XI, 856 (1967).

16. Gold, A.M.
Ibid., Vol. XI, 706 (1967).

17. Haughland, R.P., and Stryer, L., Ramachandran, G.N.
"Conformation of Biopolymers", 321 (1967).

18. Sawyer, L., Shotton, D.M., Campbell, J.W., Wendell, P.L. Muirhead, H., Watson, H.C., Diamond, R. and Ladner, R.C.
J.Molec.Biol., 118, 137 (1978).

19. Jones, A.
J.Appl.Cryst., 11, 268 (1978).

20. Sielecki, A.R., Hendrickson, W.A., Broughton, C.G., Delbaere, L.T.J., Brayer, G.D., and James, M.N.G.
J.Molec.Biol., 134, 781 (1971).

21. Steitz, T.A., Henderson, R., and Blow, D.M.
Ibid, 46, 337 (1969).

22. Segal, D.M., Powers, J.C., Cohen, G.H., Davies, D.R. and Wilcox, D.E.
Biochemistry, 10, 3728 (1972).

23. Tsukada, H., and Blow, D.M.
J.Molec.Biol., 184, 703 (1985).

24. Cohen, G.H., Silverton, D.W., and Davies, D.R.
Ibid., 139, 45 (1980).

25. Kossiakoff, A.A., and Spencer, S.A.
Biochemistry, 20, 6462 (1981).

26. Chothia, C., and Janin, J.
Ibid., 21, 3955 (1982).

27. Meyer, E.F. Jr., Radhakrishnan, R., Cole, G.M. and Presta, L.G.
J.Molec.Biol., 189, 533 (1986).

28. Clore, G.M., Gronenborn, A.M., Carlson, G., and Meyer, E.F. Jr.
Ibid, 190, 259 (1986).

29. Dimicoli, J.L., Renaud, A., Lesteinne, P., and Bieth, J.G.
J.Biol.Chem., 254, 5208 (1979).

30. Dimicoli, J.L., and Bieth, J.G.
Biochemistry, 16, 5532 (1977).

31. Renaud A., Dimicoli, J.L., Lesteinne, P.M. and Bieth, J.G.
J.Amer.Chem.Soc., 100 1005 (1977).

32. Sawyer, L.
Unpublished.

33. Naughton, M.A., and Sanger, F.
Biochem.J., 78, 156 (1961).

34. Gertler, A., and Feinstein, G.
Eur.J.Biochem., 20, 547 (1971).

35. Cohen, A.B.
Biochim.Et.Biophys.Acta., 391, 193 (1975).

36. Starkey, P.M. and Barrett, A.J.
Research. Monographys in cell and tissue physiology,
2, 661 (1977).

PAGE/PAGES
EXCLUDED
UNDER
INSTRUCTION
FROM
UNIVERSITY