A STRUCTURAL STUDY OF HUMAN

COMPLEMENT SUBCOMPONENT C1s

A THESIS PRESENTED IN PARTIAL FULFILMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN BIOCHEMISTRY
AT MASSEY UNIVERSITY

PETER JOHN MORRIS

1981

ABSTRACT

The aim of this project was to determine the carboxyl terminal amino acid sequence of the heavy chain of human complement subcomponent $C\overline{1s}$. The proteolytic cleavage of a peptide bond(s), probably at an Arg-Ile bond, of the single chain C1s yields the active serine protease $C\overline{1s}$ composed of a heavy and light chain. A knowledge of the amino acid sequence preceding the scissile Arg-Ile bond would allow the chemical synthesis of a model peptide substrate for the C1s-activating enzyme $C\overline{1r}$.

Human $C\overline{1s}$ was purified to homogeneity by euglobulin precipitation and repeated ion exchange chromatography. Unactivated C1s, which could be activated by incubation with partially purified $C\overline{1r}$, was isolated by performing all purification steps in the presence of the serine protease inhibitor phenylmethane sulphonylfluoride and at low temperature. The heavy and light chains of activated $C\overline{1s}$ were separated by ion exchange chromatography in the presence of denaturant following thorough disulphide bond reduction.

Isolation of the carboxyl terminal-derived peptide of the $C\overline{1s}$ heavy chain by peptide mapping as well as by chemical modification of protein carboxyl groups was unsuccessful largely due to the high molecular weight of the protein substrate. Digestion of $C\overline{1s}$ by carboxypeptidase B resulted in the very rapid release of arginine in a quantitative yield presumably from the carboxyl terminus of the $C\overline{1s}$ heavy chain.

Affinity chromatography using immobilized anhydrotrypsin was successful in isolating the carboxyl terminal chymotryptic peptide of the $C\overline{1s}$ heavy chain. Anhydro-

trypsin displays a remarkably specific affinity for trypsin product-like peptides possessing a carboxyl-terminal arginine residue. Attempts to determine the entire amino acid sequence of the isolated peptide were prevented by the difficulty in obtaining sufficient material. However, by determining the N-terminal amino acid sequence and amino acid composition of this peptide as well as by performing further peptide fragmentation by trypsin the following partial primary structure is proposed:

Gln-Gln-Lys-Glx-Val-Pro-Glx-Gly-[Thr, Ser, (Leu), Ala]-Lys-Glx-Glx-Asx-Arg.

ACKNOWLEDGEMENTS

I am very grateful to my supervisor Dr. M.J. Hardman for his continued interest and advice throughout this project. The advice of other members of the Department of Chemistry, Biochemistry and Biophysics, notably Drs. J.C. McIntosh and C.H. Moore, is gratefully appreciated. I wish to thank Mr J. R. Reid for performing amino acid analyses, Dr. G.G. Midwinter for initial assistance with high voltage paper electrophoresis and Dr J.W. Tweedie for assistance with immunodiffusion. My thanks to the Palmerston North Hospital for generous supplies of human plasma.

The support of my wife Raewyn and daughters, Fiona and Helene, throughout the course of this project is gratefully acknowledged. I also wish to thank Fay Parsons for her assistance in the preparation of this thesis, in particular her typing of this manuscript which is much appreciated.

CONTENTS

LIST OF FIGURES

LIST OF PLATES

LIST OF TABLES

ABBREVIATIONS

Bapna	α -N-Benzoyl-DL-Arginine-p-nitroanilide. HCl
BAW	Butanol/glacial acetic acid/distilled water
Cbz-Lys-ONp	N-α -carbobenzoyloxyl-L-Lysine-p-nitrophenyl ester
CP	Carboxypeptidase
CT	Chymotryptic digest
Dansyl, DNS	1-dimethylaminonaphthalene-5-sulphonyl
DEAE, DE	Diethylaminoethyl
dip-F	Diisopropylfluorophosphate
dip-$C\overline{1s}$	dipF-treated $C\overline{1s}$
DTT	Dithiothreitol
EDC-HCl	1-Ethyl-3(3-dimethyl-amino-propyl)-carbodiimide HCl
EDTA	Disodium ethylenediaminetetra acetic acid
Em	Electrophoretic mobility
E.U.	Enzyme Unit
M.Wt	Molecular weight
nm	nanometre
PAGE	Polyacrylamide gel electrophoresis
%	All percentages are weight per volume ($^{W}/v$) unless otherwise stated.
PMSF	Phenylmethanesulphonylfluoride
PMS-trypsin	PMSF-treated trypsin
RCM	Reduced and carboxymethylated
RNase	Bovine pancreatic Ribonuclease A
SDS	Sodium dodecyl sulphate
ST-Sepharose 4B	Tryptic digest of Salmine immobilized on Sepharose 4B
S.V8	Staphylococcus Aureus V8
TEMED	N,N,N',N'-tetramethylene ethylene diamine
TFA	Trifluoroacetic acid
Tris	Tris (hydroxymethyl) amino methane

Complement Nomenclature

The nomenclature of complement components and subcomponents is that recommended by the World Health Organisation (1968). The components of classical complement are designated numerically, eg. C1, C2, C3, ...C9. The activated forms of complement are indicated by a bar above the symbol, e.g $C\bar{1}$ is the enzymically active form of C1, $C\overline{5,6}$ is the active complex of C5 plus C6. Fragments are described by adding a lower case letter, e.g. C3 fragments C3a, C3b, C3c, C3d. Subunits of a macromolecular complex are designated by a lower case letter, e.g. C1q, C1r, C1s.

CHAPTER I

INTRODUCTION

1.1 Complement Action

Human complement is a complex system of at least fifteen distinct serum proteins (Reviewed by Fothergill and Anderson, 1978). It is an essential part of an immune person's capacity to recognize foreign substances and to defend himself against infection by invading pathogenic organisms.

The bactericidal activity of fresh serum consists of both a heat-stable specific component, antibody and a heat-labile nonspecific component, complement. Antibody is responsible for the recognition and agglutination of foreign cells as well as the subsequent activation of complement. Complement action results in the lysis of the invading organism as well as the production of potent pharmacological effects. Although nonspecific in immunological terms, complement is highly specific in biochemical terms.

Activation of the complement cascade initiates a sequential series of protein-protein and protein-membrane interactions. Protein-protein interactions consist of either limited zymogen proteolysis to directly form an active protease from complement components or protein binding of activation fragments of several complement components to form a protease or to bind to and modify that protease's substrate specificity. Soluble complement components are able to undergo transition to membrane constituents through the generation of binding regions. The complement components are all high molecular weight glycoproteins (85-200,000 daltons), containing five to ten percent carbohydrate and are present at low concentrations in serum. (Müller-Eberhard, 1975).

Complement action can conveniently be divided into three phases:

(1) The generation of a C3 activating enzyme

(2) C3 activation
(3) Assembly of the multimolecular complex causing cell lysis.

(1) C3 - Activating Enzyme Generation

Activation of the complement cascade may occur by two parallel but independent pathways: (i) the classical pathway and (ii) the alternative pathway, both of which produce a C3 - activating enzyme. Figure 1.

(i) The Classical Pathway

Aggregated antibody or cell-bound antibody of sensitized cells binds to the first component of complement, C1 and results in the activation of the s subcomponent of C1 to yield an active protease, $\overline{C1s}$ (Porter, 1977a).
Activated subcomponent $\overline{C1s}$ then proteolytically converts C2 and C4 into forms capable of interacting and together comprising another protease, the C3 convertase ($\overline{C4,2}$).

C4 is a triple-chain glycoprotein (molecular weight 200,000 daltons) in its inactive precursor form, which is cleaved by $\overline{C1s}$ near the amino terminal of the largest of the three polypeptide chains. This limited proteolysis yields an 8,000 dalton activation peptide, C4a and the remainder of the C4 molecule, C4b (molecular weight 200,000 daltons) binds covalently to the cell membrane or to the Fab part of the antibody (Goers and Porter, 1978) of sensitized cells but possesses no enzymic activity. This membrane binding ability is rapidly lost if C4b is not bound to the membrane within a short time of activation.

The single-chain glycoprotein C2 (molecular weight 115,000 daltons) is cleaved by $\overline{C1s}$ to give $\overline{C2a}$ (molecular weight 80,000 daltons) which associates with cell-bound C4b to form a cell-bound C3 convertase enzyme complex, $\overline{C4b,2a}$. It is the C2a moiety of the serine protease C3 convertase which contains the active site serine residue. The smaller C2b fragment is

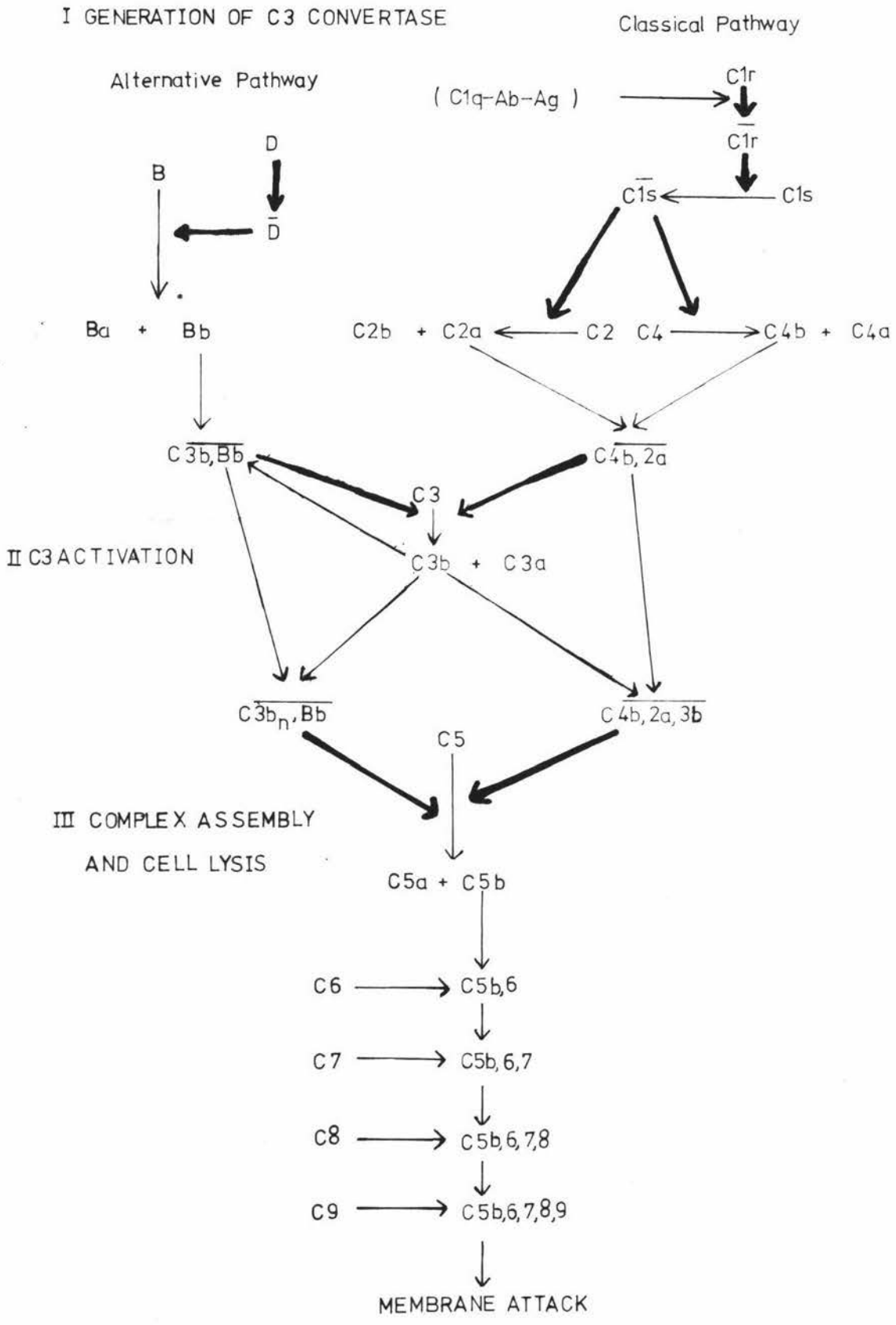


FIGURE 1 The Complement System

↓ Limited Proteolysis

↓ Physicochemical Binding

thought to be released into solution.

(ii) Another distinct C3 convertase can be generated by the alternative pathway of complement. This pathway is activated by complex bacterial, yeast or plant cell wall polysaccharides as well as the IgA immunoglobulin class and results in the activation of factor D (molecular weight 25,000 daltons) to give an active serine protease factor $\overline{D}$. The *in vivo* activator is unknown while *in vitro* trypsin activates factor D. (Volanakis *et al*, 1977). Factor $\overline{D}$ then cleaves the single chain factor B (molecular weight 30,000 daltons). Factor $\overline{Bb}$ is analogous to $C\overline{2a}$ in the alternative pathway C3 convertase complex, $(C\overline{3b,Bb})$. Following production of the C3 convertases, the classical and alternative pathways share common steps.

(2) C3 Activation

C3 (molecular weight 195,000 daltons) is a two chain glycoprotein present in plasma at a ten-fold higher concentration than any of the other complement proteins. It is considered to play a central role in complement action. A single proteolytic cleavage by either the classical pathway C3 convertase $(C\overline{4b,2a})$ or the alternative pathway convertase $(C\overline{3b,Bb})$ liberates a C3a activation peptide (molecular weight 9,000 daltons) from the amino-terminal end of the heavy chain of C3.

C3a is an anaphylatoxin possessing potent pharmacological activity including contraction of smooth muscle, histamine release from mast cells and chemotaxis of leukocytes, thus mediating the inflammatory response. This action is inhibited by a plasma carboxypeptidase B - like anaphylatoxin inactivator which removes the carboxyl-terminal arginine residue essential for activity.

The remaining molecule, C3b, possesses cell surface binding properties. However if C3b is not bound to the cell membrane following activation this membrane-binding ability is lost due to proteolysis by a C3 inactivator (C3 INA) (Harpel and Cooper, 1975). A cell coated with many C3b molecules results in immune adherence and the stimulation of opsonization thereby promoting cell phagocytosis. As C3b is required for C3 activation in the alternative pathway a positive feedback mechanism exists whereby C3 convertase will stimulate its own activation by producing more C3b. As well as binding to cell membranes C3b also binds to its own activating enzyme, C3 convertase through the C4b or factor B moieties. This binding results in a modification of the C3 convertase substrate specificity and the new enzyme, C5 convertase ($\overline{C4b,2a,3b}$ or $\overline{C3b_n,Bb}$) now cleaves an analogous peptide bond in C5.

(3) Complex Assembly and Cell Lysis

C5, (molecular weight 205,000 daltons) is similar in structure to C3. C5 activation also results in the release of an anaphylatoxin from the amino terminal of the C5 heavy chain. C5b (molecular weight 195,000 daltons) is able to bind to cell membranes and also to C6. It is proteolytically inactivated by the C3b INA if not bound soon after activation. C5 convertase cleavage of C5 is the final proteolytic step in complement activation. Subsequent complement action is solely by protein-protein and protein-membrane nonenzymic physicochemical interactions. (Müller-Eberhard, 1975). The C5b generated acts as a nucleus for the self-assembling membrane attack complex of C5b,6,7,8,9.

Both C6 and C7 are single polypeptide chains of molecular weight 125,000 daltons while C8 is a triple-chain protein (molecular weight 155,000 daltons) capable of binding to the membrane-bound C5b,6,7 complex. The large hydrophobic chain located inside native C8 is involved in the initiation of membrane damage. The addition of C9 (molecular weight 75,000 daltons) results in the generation and stabilization of membrane pores. These pores are of 10 nm diameter and are visible by electron microscopy. (Knobel et al, 1975).

Two major complement control proteins are the C3b INA already mentioned and a $C\bar{1}$ inhibitor (molecular weight 100,000 daltons). $C\bar{1}$ inhibitor blocks $C\bar{1}$ activity by stoichiometrically complexing with the serine proteases $C\bar{1}r$ and $C\bar{1}s$. (Ratnoff *et al*, 1969). A further control point is the rapid spontaneous decay of C3b, C4b and C5b membrane binding ability if not membrane bound immediately following activation.

1.2 C1 Subcomponents

The first component of the classical pathway of complement is composed of three subcomponents: C1q, C1r and C1s held in a complex by calcium ions. The C1 subcomponents are so-named because of their order of elution during DEAE anion exchange chromatography in the presence of EDTA using salt gradient elution. (Lepow et al, 1963).

In serum or 5 mM $CaCl_2$ two molecules each of C1r and C1s are bound by calcium in a tetrameric complex of molecular weight 350,000 daltons represented by: $C1r_2$-Ca^{++}-$C1s_2$. An octomeric complex has also been proposed (Nagasawa et al, 1974). Maximal haemolytic activity occurs with a molar ratio of C1q: C1r: C1s of 1:4:4 and as serum molar ratios are approximately 1:2:2 this is suggestive of one $C1r_2$-Ca^{++}-$C1s_2$ tetramer binding to each C1q molecule. (Gigli et al, 1976) (Porter, 1977a) (Ziccardi and Cooper, 1976). Little interaction occurs between the C1r-C1s tetramer and C1q in the unactivated C1s state. (Nagasawa et al, 1974) (Porter, 1977a). However, upon activation of complement this binding is increased with no dissociation of subcomponents occurring and so resultant $\overline{C1s}$ proteolytic activity is associated with the whole C1 complex. $\overline{C1}$ Inhibitor binding does release a $\overline{C1}$-Inhibitor-$\overline{C1s}/\overline{C1r}$ complex leaving C1q-Ag-Ab bound to the cell membrane. (Sim et al, 1979).

Subcomponent C1q (molecular weight 410,000 daltons) is a unique glycoprotein composed of six A,B and C chains each of about 23,000 daltons molecular weight. Each chain is linked by a single disulphide bond through residue cysteine-4 to form six A-B and three C-C subunits. Each C-C subunit is associated noncovalently with two A-B subunits to form a gross C1q structure of six identical A-B-C subunits resembling the so-called 'posy of tulips' in electron micrographs in which each 'tulip' is represented by one subunit. The 'stalk' of each 'tulip' consists of a 78- residue length of collagen - like triple helix formed by the interaction of the A,B and C chain of each subunit. It is this region of C1q which binds the C1r-C1s tetramer through C1r (Porter, 1977). At the carboxyl terminus of each subunit is a globular pod-like

region which resembles the 'petals' or 'flowers' of the 'tulips'. It is this globular region of each C1q subunit which binds to the antibody molecule. Not all antibody is able to activate the classical pathway of complement. IgM and the IgG dimers of subclasses 1, 2 and 3 do while IgG subclass 4, IgA, IgE and IgD do not bind to C1q to activate complement. (Frank, 1979).

The first event leading to complement activation is the binding of the foreign cell-surface antigen to the variable domain of the Fab portion of antibody. This allows the C1q C - terminal globular heads to bind directly to the Fc region of antibody; possibly to a C1q binding site exposed on antibody reaction with antigen. Thus C1q binding and subsequent complement action is independent of antibody specificity which resides only in the antigen-antibody Fab reaction. C1q is unable to bind parent antibody but it is able to bind to heat-aggregated Fc fragments of antibody or heat-aggregated whole antibody in addition to the in vivo antibody-antigen complex. (Taranta and Franklin, 1961), (Ishizaka et al, 1962).

The end result of antibody binding to the C - terminal globular heads of C1q is the activation of C1r to give the active protease, $\overline{C1r}$. The rate of C1r activation in the C1-Ab-Ag complex is the same whether C1s, $\overline{C1s}$ or inactive dipf-treated $\overline{C1s}$ is used. If no C1s is present at all then no C1r activation occurs suggesting that C1s is required in a non-proteolytic role. (Dodds et al, 1978). Dipf added to C1-Ab-Ag complexes prevents C1s activation but not C1r activation indicating that dipf inactivates $\overline{C1r}$ as it is formed. During C1r activation a single-chain reactive intermediate is thought to be formed before chain cleavage occurs. This intermediate is able to activate C1r in the C1 complex but in solution C1r does not self-activate (Dodds et al, 1978). Further evidence has been found consistent with an intramolecular auto-catalytic activation mechanism involving an intermediate C1r prosite of the C1r dimer. (Arlaud et al, 1980). These workers however, found that C1r was able to self-activate in solution. A consequence of C1q binding to

antibody is a firmer binding between C1r and C1q, possibly due to conformational changes transmitted from the antibody-bound C - terminal heads of C1q to the C1r binding site of C1q. As calcium is known to decrease C1r activation C1q may remove calcium from a critical site in C1r thus allowing auto-catalytic activation of C1r by the low catalytic activity often associated with zymogens of proteases. (Kassell and Kay, 1973) (Londsdale-Eccles et al, 1979). A combination of these possible mechanisms for C1r activation may be involved.

Subcomponent C1r in its zymogen form is a single polypeptide chain (molecular weight 85,000 daltons) and is a dimer both in the presence and absence of calcium. (Sim and Porter, 1976) (Ziccardi and Cooper, 1976). Activation is brought about by the cleavage of a peptide bond(s) to yield two disulphide-linked polypeptide chains; the a chain (molecular weight 58,000 daltons) and the b chain (molecular weight 27,000 daltons). Although no detectable loss of mass occurs on activation the possibility of an activation peptide being produced does exist (Sim et al, 1977). N - terminal amino acid sequence determination of the unactivated C1r and the activated $\overline{C1r}$ a and b chains has shown that the heavy a chain is derived from the N - terminus of C1r while the b chain is derived from the C - terminus. (Sim et al, 1977). Figure 2. The b chain contains the active site serine residue which binds dipF and also shows N-terminal and active site amino acid sequence homology with other serine proteases. (Sim et al, 1977). Figure 3. The $\overline{C1r}$ b chain is also of approximately the same molecular weight (24-28,000 daltons) as the active site serine-containing chains of other serine proteases. The $\overline{C1r}$ a chain, which possesses a blocked N-terminus, shows little amino acid sequence homology with other proteins of known sequence. (Sim et al, 1977).

$\overline{C1r}$ possesses an exceptionally restricted esterase and protease substrate specificity spectrum. Some basic amino acid esters are reportedly hydrolysed, albeit at very low rates. Naff and Ratnoff (1968) found the preferred $\overline{C1r}$ substrates to be acetyl arginine methyl ester followed by acetylglycyl lysine methyl ester for partially purified $\overline{C1r}$.

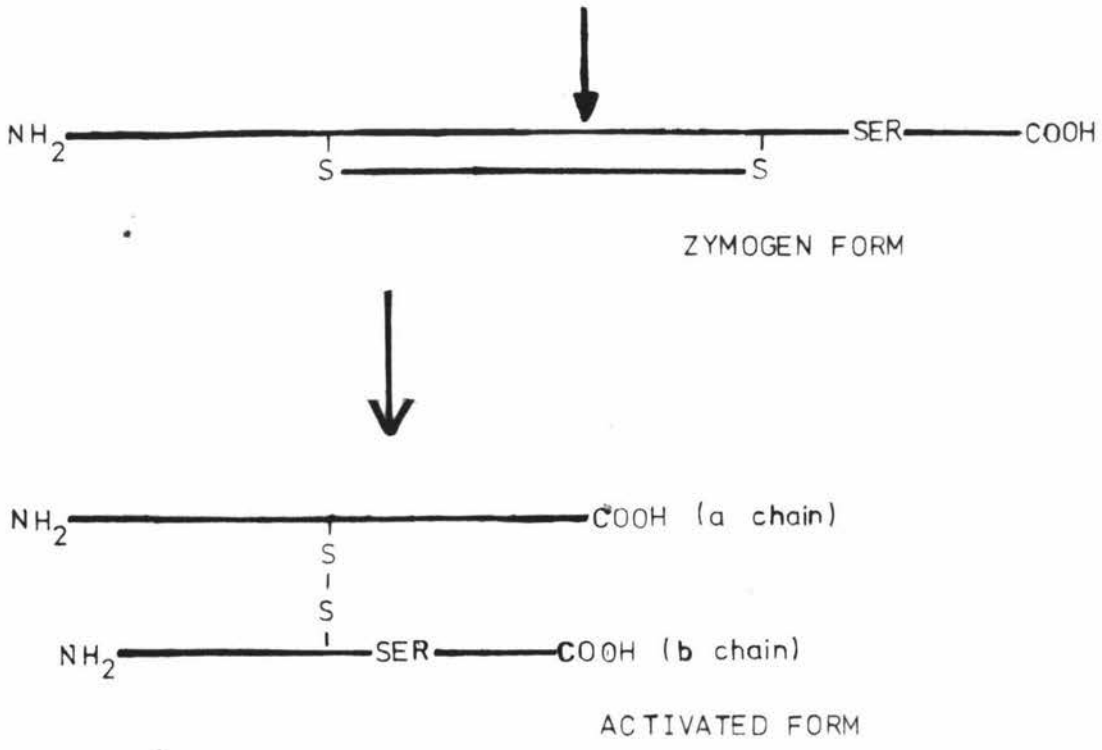


FIGURE 2 **Activation** of C1 Subcomponent Zymogens by Limited Proteolysis

In the zymogen state both Clr and Cls are single chain proteins. Activation by limited proteolytic cleavage (arrow) yields two chains linked by one or more disulphide bonds. The smaller b chain contains the active site serine residue (SER) and possesses the amino terminal sequence Ile - Ile - Gly - Gly common to many serine proteases.

	(1)	(5)	(10)	(15)	(20)
Human $\overline{C1r}$ b chain (a)	Ile-Ile-Gly-Gly-Gln-Lys-Ala-Lys-Met-Gly-Asn-Phe-Pro-Trp-Gln-Val-Phe-Thr-Asn-Glx				
Human $\overline{C1s}$ b chain (a)	Ile-Ile-Gly-Gly-Ser-Asp-Ala-Asp-Ile-Lys-Asn-Phe-Pro-Trp-Gln-Val-Phe-Phe-Asp-Asn				
Bovine Trypsin (a)	Ile-Val-Gly-Gly-Tyr-Thr-Cys-Gly-Ala-Asn-Thr-Val-Pro-Tyr-Gln-Val-Ser-Leu-Asn-Ser				
Bovine Chymotrypsin A(a)	Ile-Val-Asn-Gly-Glu-Glu-Ala-Val-Pro-Gly-Ser-Trp-Pro-Trp-Gln-Val-Ser-Leu-Gln-Asp				
Bovine Thrombin B chain (a)	Ile-Val-Glu-Gly-Gln-Asp-Ala-Glu-Val-Gly-Leu-Ser-Pro-Trp-Gln-Val-Met-Leu-Phe-Arg				
Human Thrombin (b)	Ile-Val-Glu-Gly-Ser-Asn-Ala-Glu-Ile-Gly				
Human Plasmin B chain (a)	Val-Val-Gly-Gly-Cys-Val-Ala-His-Pro-His-Ser-Trp-Pro-Trp-Gln-Val-Val-Leu-Leu-Arg				
Human Factor $\overline{D}$ (b)	Ile-Leu-Gly-Gly-Arg-Glx-Ala*-Glx-Ala*-				
Human Factor XIa (b)	Ile-Val-Gly-Gly-Thr-Val				
Factor Xa (b)	Ile-Val-Gly-Gly-Glu-Glu-Cys-Lys-Asp-Glu				

Figure 3. Amino Terminal Sequence Homology of the "Enzymic Chain" of Serine Proteases.

"Enzymic Chain" is the chain containing the active site serine residue.
An amino acid identical to that in the sequence of $\overline{C1s}$ b chain is indicated by underlining.
(a) from Sim et al (1977)
(b) from Davis et al (1979)
* or serine

Other carbobenzyloxy, benzyl and tosyl basic amino acid esters were not hydrolysed at all under the conditions used. The hydrolysis of the acetyl arginine and acetylglycyl lysine methyl esters by $\overline{C1r}$ was confirmed recently by Andrews and Baillie (1979) who also found that N-cbz-tyrosine and N-α-cbz-lysine p-nitrophenyl esters were also hydrolysed. This latter result is in contrast to work by Sim et al (1977) who found that their preparation of $\overline{C1r}$ would not hydrolyse any of the N-α-cbz p-nitrophenyl esters of tyrosine, phenylalanine, leucine and lysine. In a comparative study of the ester substrate specificities of several serine proteases Volanakis et al (1977) found acetylglycyl lysine methyl ester to be the most specific substrate for $\overline{C1r}$. This ester was hydrolysed by $\overline{C1r}$ at approximately 1% of the $\overline{C1s}$-catalysed rate or 0.3% of the trypsin-catalysed rate. Acetyl arginine methyl ester was hydrolysed by $\overline{C1r}$ at only 2% of the rate of acetylglycyl lysine methyl ester hydrolysis. The only protein substrate known to be proteolytically cleaved by $\overline{C1r}$ is its in vivo substrate C1s.

Subcomponent C1s is similar to C1r with respect to molecular weight, chain structure and amino acid composition. (Sim et al, 1977), (Sim and Porter, 1976). It is a single polypeptide chain having an approximate molecular weight of 83,000 daltons in its zymogen form. (Valet and Cooper, 1974a,b) (Sim et al, 1977) (Ziccardi and Cooper, 1976). In serum or 5 mM calcium C1s is present as a dimer while in the presence of EDTA it is a monomer.

Proteolytic cleavage by $\overline{C1r}$ produces an active protease, $\overline{C1s}$, consisting of two disulphide-linked chains, the a and b chains of molecular weight 56,000 and 27,000 daltons respectively (Sim et al, 1977). Figure 2. N - terminal amino acid sequence analysis of C1s and $\overline{C1s}$ a and b chains shows that the a chain is derived from the N - terminus of the zymogen while the b chain is derived from the C - terminus (Sim et al, 1977) (Takahashi et al, 1975a). As in $\overline{C1r}$, the $\overline{C1s}$ b chain contains the active site serine residue which reacts with dipF. Amino acid sequence homology of the b chain active site region (Barkas et al, 1973) and the N - terminal region

(Sim et al, 1977) with other serine proteases has been noted. Figure 3 and 4. This is particularly apparent when comparing the N - terminal amino acid sequences of the $\overline{C1s}$ and $\overline{C1r}$ b chains where 60% homology is observed for the first twenty residues. Figure 3. The N - terminal amino acid sequence of the $\overline{C1s}$ a chain is also known and as with the $\overline{C1r}$ a chain no sequence homology with other proteins of known amino acid sequence has yet been found.

Although $\overline{C1s}$ and $\overline{C1r}$ are structurally similar the most striking difference between them is their esterase substrate specificities. Whereas $\overline{C1r}$ hydrolyses only slowly acetyl arginine methyl ester and acetylglycyl lysine methyl ester $\overline{C1s}$ hydrolyses the methyl esters of acetyl, benzoyl and tosyl arginine plus benzoyl arginine ethyl ester (Volanakis et al, 1977). $\overline{C1s}$ also hydrolyses the methyl esters of tosyl and cbz lysine and to a much greater degree the acetyl and acetylglycyl lysine methyl esters which are hydrolysed at 30% of the trypsin-catalysed rate (Ratnoff and Lepow, 1957). It is one of the few serum esterases to hydrolyse acetyl tyrosine ethyl ester and is therefore described as exhibiting a plasmin-like esterase substrate specificity (Naff and Ratnoff, 1968) (Volanakis et al, 1977). $\overline{C1s}$ hydrolyses the p-nitrophenyl esters of lysine and tyrosine but has no effect on the glycine, phenylalanine and leucine esters. (Sim et al, 1977). This is clearly distinct from the specificities of trypsin and chymotrypsin. In addition to proteolytically cleaving its in vivo substrates C2 and C4; C4 at an Arg-X bond, $\overline{C1s}$ is able to cleave some denatured protein substrates. (Scott and Fothergill, 1975).

C1s activation is not autocatalytic but requires active $\overline{C1r}$; C1r or inactivated dip-C1r are unable to activate C1s. No detectable loss of mass occurs on proteolytic activation of C1s, indicative of a single proteolytic cleavage by $\overline{C1r}$. (Sim et al, 1977). Lepow et al (1958) have reported no release of trichloro acetic acid - soluble nitrogen during activation of C1. Takahashi et al (1975a) found no evidence for the loss of an activation peptide from the

Human $\overline{C1s}$ (a)	-Ala-Cys-Gly-Lys-Asp-SER-Gly-Glu- X -Arg-
Human Plasmin (a)	-Ser-Cys-Gln-Gly-Asp-SER-Gly-Gly-Pro-Leu-
Bovine Thrombin (b)	-Ala-Cys-Glu-Gly-Asp-SER-Gly-Gly-Pro-Phe
Bovine Factor XIIa (a)	-Ala-Cys-Gln-Gly-Asp-SER-Gly-Gly-Pro-
Bovine Trypsin (b)	-Ser-Cys-Gln-Gly-Asp-SER-Gly-Gly-Pro-Val
Bovine Factor Xa (c)	-Ala-Cys-Gln-Gly-Asp-SER-Gly-Gly-Pro-His
Bovine Chymotrypsin A (b)	-Ser-Cys-Met-Gly-Asp-SER-Gly-Gly-Pro-Leu
Bovine Factor VII (c)	-Ala-Cys-Lys-Gly-Asp-SER-Gly-Gly-Pro-His
Factor D (c)	-Ser-Cys-Lys-Gly-Asp-SER-Gly-Gly-Pro-Leu
Porcine Elastase (b)	-Gly-Cys-Gln-Gly-Asp-SER-Gly-Gly-Pro-Leu

Figure 4. The Active Site Sequences of Some Serine Proteases

A residue identical to that in $\overline{C1s}$ is indicated by underlining.

SER indicates the active site reactive serine residue

(a) from Fothergill and Anderson (1978)

(b) from Dayhoff (1972)

(c) from Davis (1980)

N-terminus of zymogen C1s as is the case in bovine trypsinogen activation where a hexapeptide is released.

It was also pointed out by Takahashi et al (1975a) that the newly appeared α amino group of isoleucine is essential for proteolytic activity of the serine proteases trypsin, chymotrypsin and thrombin. Perhaps the basic mechanism of activation is the same for C1s. In the cases of chymotrypsinogen and trypsinogen the newly appeared α amino group of isoleucine folds into the interior portion of the protein, giving rise to the formation of an ion pair with the carboxyl group of an aspartic acid residue which is adjacent to the active site serine. (Sigler et al, 1968). This leads to the charge relay network which is characteristic of pancreatic serine proteases (Sigler et al, 1968) (Blow et al, 1969). Thus it may be that the newly formed amino terminal Ile residue of the $\overline{C1s}$ b chain also forms an ion pair with aspartic acid adjacent to the active site serine residue in a similar manner giving rise to an active serine protease. Figure 4. An essential step in the formation of active $\overline{C1s}$ is the cleavage of a specific X-Ile peptide bond in zymogen C1s. (Takahashi et al, 1975a). Considering the trypsin-like esterase substrate specificity of $\overline{C1r}$ it is likely that $\overline{C1r}$ acts in a very specific trypsin-like manner to cleave this specific peptide bond in zymogen C1s. The X-Ile bond, highly susceptible to $\overline{C1r}$ cleavage, would therefore be expected to be either an Arg-Ile or Lys-Ile bond. The possibility of more than one cleavage site in C1s with subsequent release of a small internally-derived activation peptide cannot be ignored. Even if this was the case, then one part of the C1s activation process would be cleavage of this Arg-Ile bond. In any event, the cleavage of one specific Arg-Ile or Lys-Ile bond in C1s is an essential step in the activation mechanism of C1s. Assuming the simplest case for C1s activation of one proteolytic cleavage by $\overline{C1r}$, Arginine or Lysine will be the carboxyl terminal amino acid residue of the $\overline{C1s}$ a chain while Ile is the known amino terminal residue of the b chain. Such activation of C1s by $\overline{C1r}$ can be described as in Figure 5.

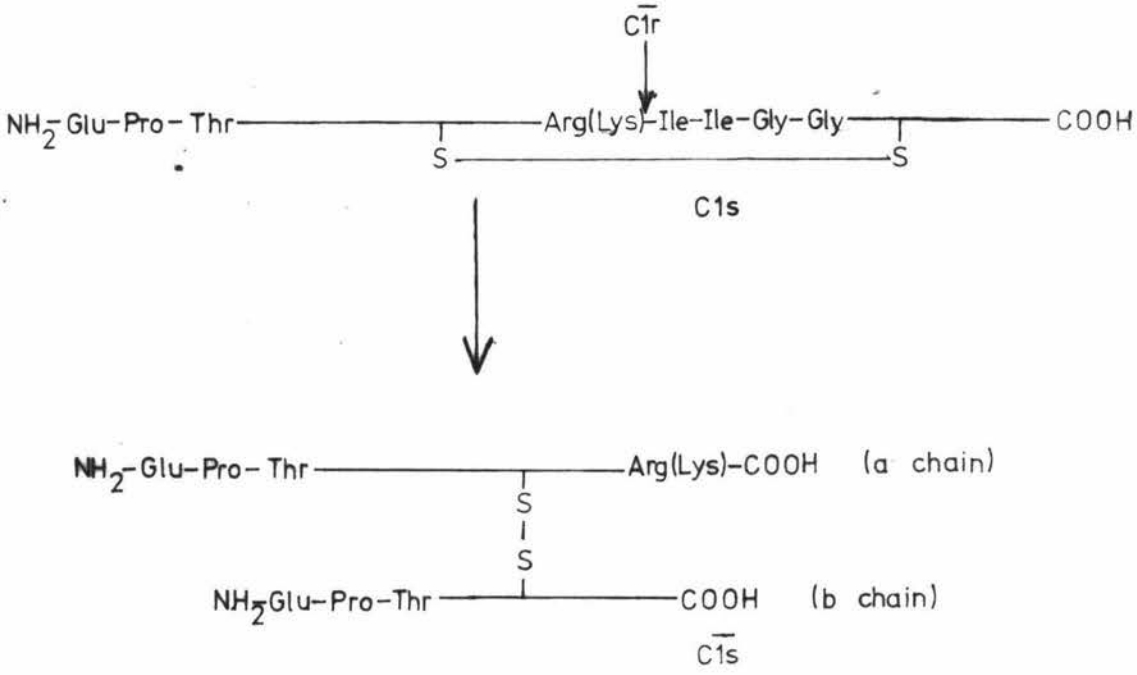


FIGURE 5 Activation of Subcomponent C1s by a Single Proteolytic Cleavage by $\overline{C1r}$

If it is assumed that the simplest case for C1s activation is a single proteolytic cleavage at Arg (Lys) - Ile then the Ile residue will become the N - terminal residue of the b chain while the Arg or Lys residue will become the C - terminal residue of the a chain.

The arrow indicates the site of peptide bond cleavage by $\overline{C1r}$. From Takahashi et al (1975a) with corrections of amino acid sequence from Sim et al (1977)

In the micro-environment of the substrate binding site of the $C\bar{1}r$ active site that portion of the C1s polypeptide which is specifically bound would be expected to have most influence in determining the site of bond cleavage. The amino acids of C1s specifically bound or 'recognized' by the $C\bar{1}r$ substrate binding site would include those residues in the immediate vicinity of the Arg/Lys-Ile bond, that is the amino acid sequence preceding and following the amino acid residues linked by this specific peptide bond. The substrate binding site of $C\bar{1}r$ may be represented by the extended subsite model according to Schechter and Berger (1967). Figure 6. If each subsite of the $C\bar{1}r$ active site accommodates one specific amino acid residue of the C1s substrate and assuming the simplest case of C1s activation involving only one proteolytic event then P_1'-P_2'-P_3'-P_4'-P_5'-etc. will represent Ile-Ile-Gly-Gly-Ser-etc., the N-terminal amino acid sequence of the b chain; while P_1 will represent the postulated C-terminal arginine or lysine residue of the a chain. On the basis of Kinetic studies with model substrates P_2 is predicted to be a small amino acid as the preferred substrates for $C\bar{1}r$ are acetyl arginine and acetylglycyl lysine methyl esters. The more bulky benzoyl, carbobenzyloxy or tosyl derivatives are not hydrolysed by $C\bar{1}r$. Exactly how many subsites are involved in C1s binding is unknown however an enzyme of the size of $C\bar{1}r$ should be able to 'recognize' quite a large portion of a polypeptide substrate.

The reactivity of a specific peptide bond of a protein substrate depends not only on the two residues P_1 and P_1' forming this bond but also on the nature of residues occupying nearby subsites. Although P_1 dominates substrate specificity in some cases, it does not dictate it as additional residues of the substrate act to modulate specificity also. With the exception of thermolysin and pepsin to some extent, all known proteases cleave the peptide bond on the C-terminal side of a specific amino acid residue or class of amino acids (represented by P_1 in Figure 6). This may suggest that amino acid residues on the N-terminal side of the substrate cleavage site (P_1 P_2 P_3 etc.) commonly have a greater influence in determining substrate specificity than those on the C-terminal side (P_1' P_2' P_3' etc).

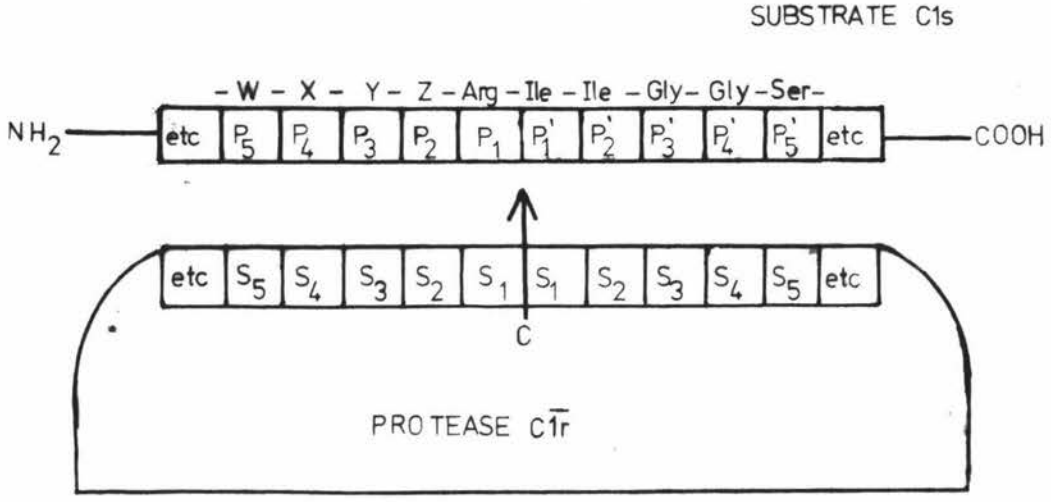


FIGURE 6 Extended Subsite Model Representation of C1s Activation by $\overline{C1r}$

The active site of an enzyme performs both the functions of binding the substrate and catalysing the reaction. The active site of $\overline{C1r}$ is divided into 'subsites', each accommodating one amino acid residue of the polypeptide substrate, C1s. The subsites (S) are located on both sides of the catalytic site (C). The positions (P) on the substrate are counted from the point of cleavage (arrow) and thus have the same numbering as the subsites they occupy. From Schechter and Berger, (1967). The symbol ' indicates positions on the substrate which are on the C - terminal side of the point of cleavage and subsites which specifically bind these residues.

Because the Ile-Ile-Gly-Gly amino acid sequence (P_1' P_2' P_3' P_4') is conserved in many serine proteases, including $\overline{C1r}$ and $\overline{C1s}$ (Figure 3), it is unlikely that this sequence confers on C1s the remarkably specific enzyme - substrate relationship of $\overline{C1r}$ and C1s.

A knowledge of the precise sequence of amino acids at the C-terminus of the $\overline{C1s}$ a chain will give more definite information regarding the substrate specificity of $\overline{C1r}$. The determination of this amino acid sequence will provide information necessary for the design of a highly reactive specific peptide substrate for $\overline{C1r}$.

1.3 Determination of the Carboxyl-Terminal Amino Acid Sequence of Proteins

The various methods available for determining the carboxyl-terminal amino acid sequence of proteins can be conveniently grouped into either of two strategic approaches (Reviewed by Needleman, 1975):

A. Direct amino acid sequencing from the C-terminus of the polypeptide chain.

B. Isolation of the C-terminal peptide which is then sequenced in the usual manner from the N-terminus.

Direct sequencing from the C-terminus of a polypeptide obviates the need for further complex manipulation of the protein once it has been purified to homogeneity. The purified polypeptide chain then directly becomes the substrate for a repetitive amino acid sequencing method or for enzymatic digestion. This procedure is in contrast to C-terminal peptide isolation requiring many steps to cleave the polypeptide chain to give a labelled C-terminal peptide which must then be chromatographically separated from the bulk peptides. Once the single C-terminal peptide of suitable size has been isolated the much superior N-terminal amino acid sequencing methods can be employed to give the complete amino acid sequence of the peptide.

A. Direct Carboxyl Terminal Amino Acid Sequencing Methods

1. Carboxyl Reduction

Limited information can be obtained by carboxyl reduction end group determination which involves reduction of the polypeptide's C-terminal carboxyl group by lithium borohydride. The acid hydrolysate of the modified polypeptide will contain one α-amino primary alcohol corresponding to the C-terminal amino acid residue which can be separated from the remaining free amino acids and identified chromatographically. Figure 7a. Bailey (1955) proposed a procedure for the sequential removal of C-terminal amino acid residues as amino alcohols based on the carboxyl reduction reaction

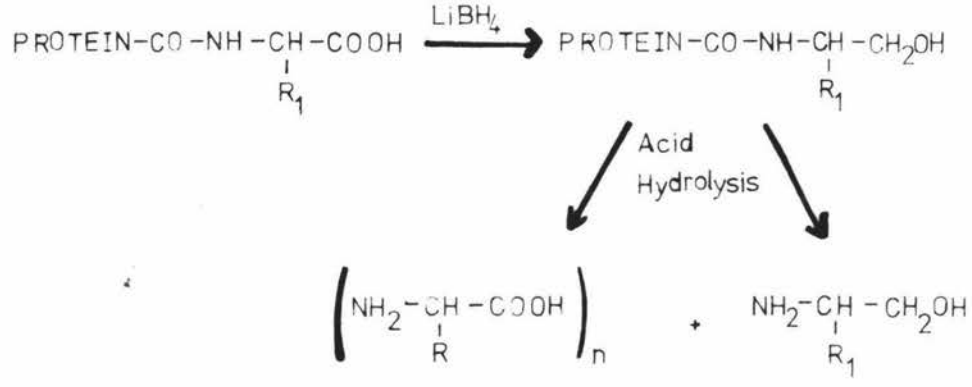


FIGURE 7a Carboxyl Reduction End Group Determination

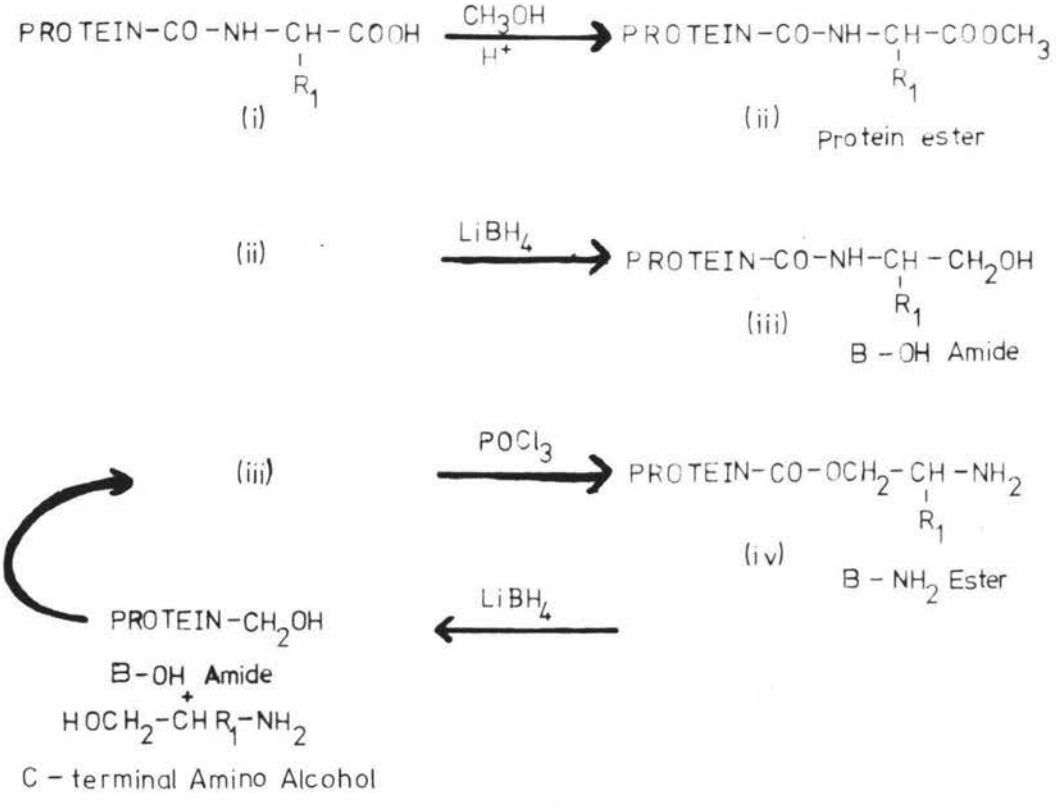


FIGURE 7b The Sequential Carboxyl Reduction Method of Bailey

coupled with an N,O-acyl shift. Figure 7b. The peptide (i) is first converted to an ester (ii) which is split reductively with lithium borohydride before inducing an N→O acyl shift using phosphorous oxychloride to give the β amino ester (iv). The β amino ester is then reductively cleaved to produce a new β hydroxy amide (iii) on which the degradation cycle is repeated. The free amino alcohol is identified chromatographically. The application of this method on proteins or otherthan model peptides has proven unsatisfactory due to the undesirable side reactions of reductive cleavage of peptide bonds (at threonine and serine especially) and also reduction of peptide carbonyl groups to methylene groups.

2. Hydrazinolysis

Hydrazinolysis has found wide application solely as a C-terminal end group determination method and so offers only limited information on amino acid sequences. Treatment of a protein with anhydrous hydrazine at 100°C liberates amino acid hydrazides derived from amino acid residues which possessed carboxyl groups formerly involved in peptide bond formation (Akabori et al, 1952). Thus only the C-terminal amino acid residue is released as the free amino acid. Figure 8. Separation of the free C-terminal derived amino acid from the bulk amino acid hydrazides can be achieved by ion exchange column chromatography or alternatively by dinitrophenylation or treatment with benzaldehyde to leave only the free amino acid in the aqueous phase.

Noncarboxyl-terminal arginine and cysteine are both decomposed to degradative products while asparagine and glutamine are able to form the corresponding β and γ hydrazides even if C-terminal as well as dihydrazides if non C-terminal. Other side reactions are caused by the instability of amino acid hydrazides, especially labile glycine, serine and alanine hydrazides and the dihydrazides of the acidic amino acids which undergo decomposition to the free amino acid. Free amino acids with small sidechain groups are also able to undergo partial conversion to their corresponding

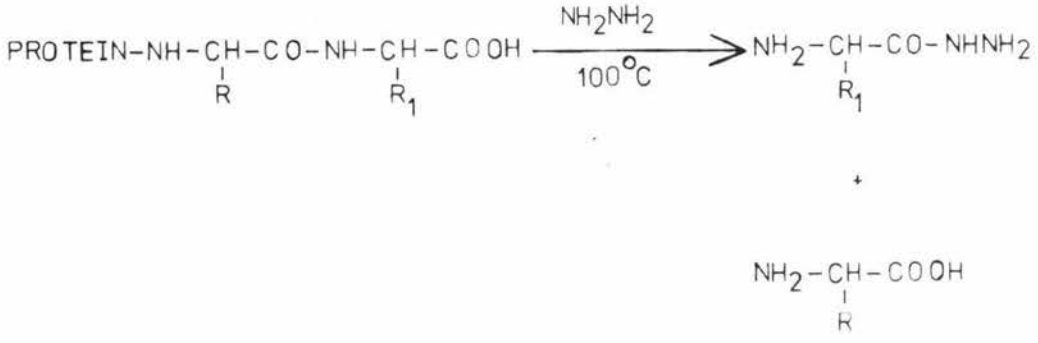


FIGURE 8 Hydrazinolysis

hydrazides. The acidic monohydrazides have also been postulated to convert to their corresponding dihydrazides. Thus conditions must be optimized to minimize conversion of C-terminal amino acids to their hydrazides and the reverse regeneration of the free amino acids as well as maximize hydrazinolytic cleavage of the C-terminal residue.

3. Tritium Labelling

Tritium labelling is a further end group determination method having little scope for development into a repetitive sequence determination method. It is however very useful as a method for specifically radioactively-labelling the C-terminal amino acid or C-terminal peptide of a protein (Matsuo et al, 1966). Selective tritiation is brought about through racemization via the oxazolone intermediate by the action of acetic anhydride in a medium containing $^{3}H_2O$ and pyridine. The tritium-labelled C-terminal amino acid can be identified chromatographically after acid hydrolysis of the tritiated protein. Alternatively carboxypeptidases have been used to enzymically cleave the C-terminal tritiated amino acid from the protein.

The tritium-labelling reaction mechanism is represented by three different reactions which take place in one step. See Figure 9. The overall rate of racemization and thus tritium incorporation is controlled by the substituents R_1 and to some extent R_2. (See Figure 9). Carboxyl terminal serine and threonine incorporate radioactivity poorly, possibly due to unfavourable sidechain dehydration catalysed by acetic anhydride while C-terminal proline, which cannot form the corresponding oxazolone intermediate, does not incorporate tritium at all. Carboxyl terminal penultimate glycine (R_2 in Figure 9) has the effect of decreasing tritium incorporation into the C-terminal residue. A further problem of tritium incorporation is that of "interior labelling" of noncarboxyl terminal aspartic and glutamic acids. Such levels of incorporation vary from

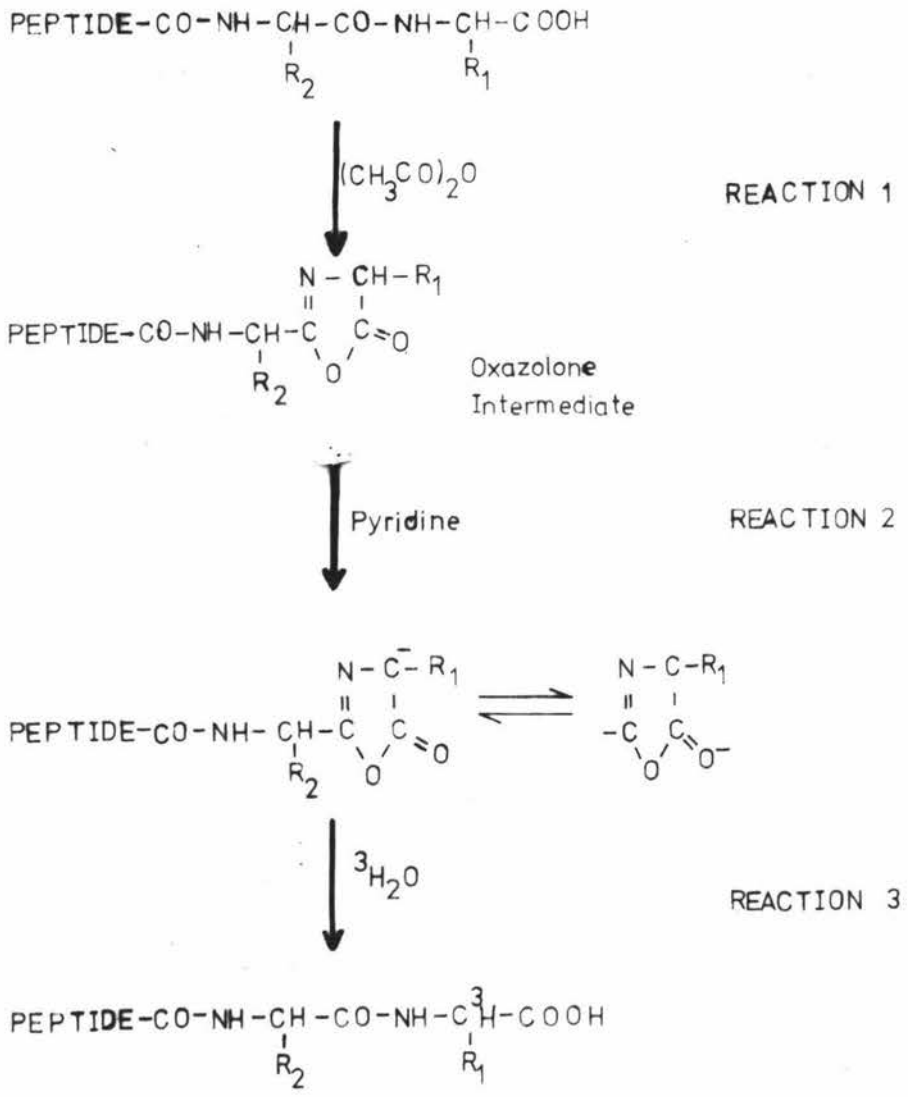


FIGURE 9 Tritium Labelling Method of Matsuo

The tritium-labelling reaction mechanism is represented by three **different reactions** which take place in one step.

Reaction 1. C - terminal oxazolone formation by the action of acetic **anhydride.**

Reaction 2. Base - catalysed racemization at the C4 position of the **oxazolone in the presence** of pyridine.

Reaction 3. Ring opening of oxazolone to regenerate the C - terminal **amino acid which incorporates** tritium at its αC atom. Acid hydrolysis **of the protein gives** the tritiated C - terminal amino acid residue.

less than 1% to more than 25% of C-terminal amino acid labelling levels. Caution must therefore be exercised when aspartic acid or glutamic acid is radioactively labelled.

4. Thiohydantoin Carboxyl Terminal Sequencing

Peptides substituted at the terminal amino group are converted to a peptidyl thiohydantoin derivative by heating with ammonium thiocyanate in acetic anhydride. (Schluck and Kumpf, 1926). Figure 10. The thiohydantoin derivative corresponding to the C-terminal residue is released by mild alkali treatment to hydrolyse the peptide acyl group XCO, on which the procedure is repeated. The released thiohydantoin is extracted with organic solvents and the parent amino acid is regenerated with HBr or alternatively the thiohydantoin derivative of the amino acid can be characterized directly by paper chromatography.

As proline is unable to form a 1-acyl-2-thiohydantoin derivative C-terminal proline cannot be determined. Arginine, aspartic acid, glutamic acid, serine and threonine are also unable to be determined. Success with this method has been achieved with model oligopeptides only. A notable extension by Stark (1968) used milder conditions to form the acyl isothiocyanate derivative and for the deacylation step to sequence all except aspartic acid and proline. Attempts to obtain improved yields and ease of application by coupling the peptide through its amino terminus to an insoluble glass bead support have been made.

5. Carbodiimide-Catalysed Sequential Removal of Carboxyl-Terminal Amino Acids

Khorana (1952) has suggested the use of carbodiimides for the selective removal of C-terminal amino acid residues. The acyl peptide is first reacted with di-(p-tolyl)-carbodiimide to form the corresponding acyl peptidyl urea (i). Figure 11. On treatment with dilute alkali in

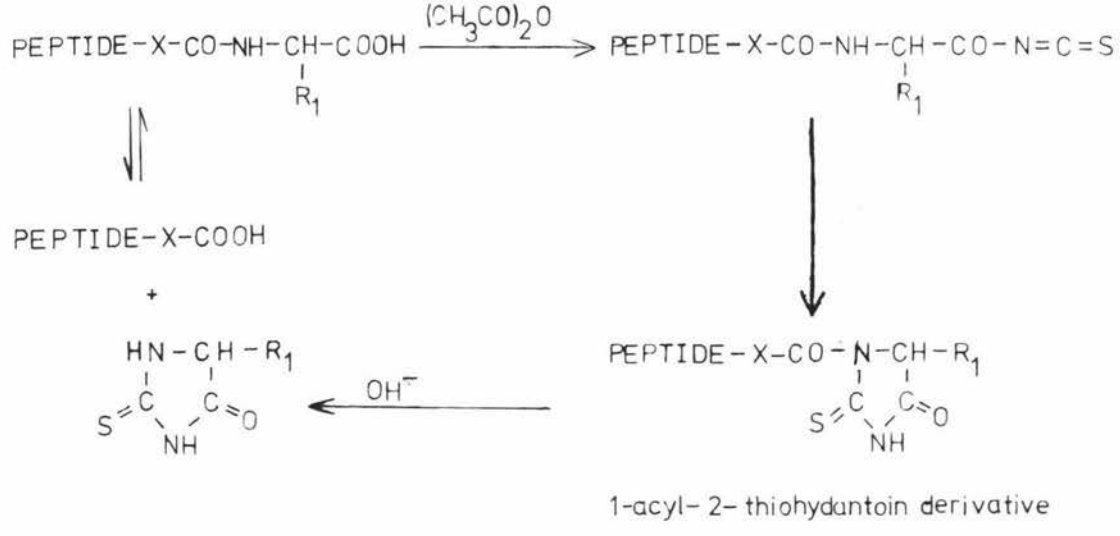


FIGURE 10 Thiohydantoin Carboxyl terminal Sequencing Method

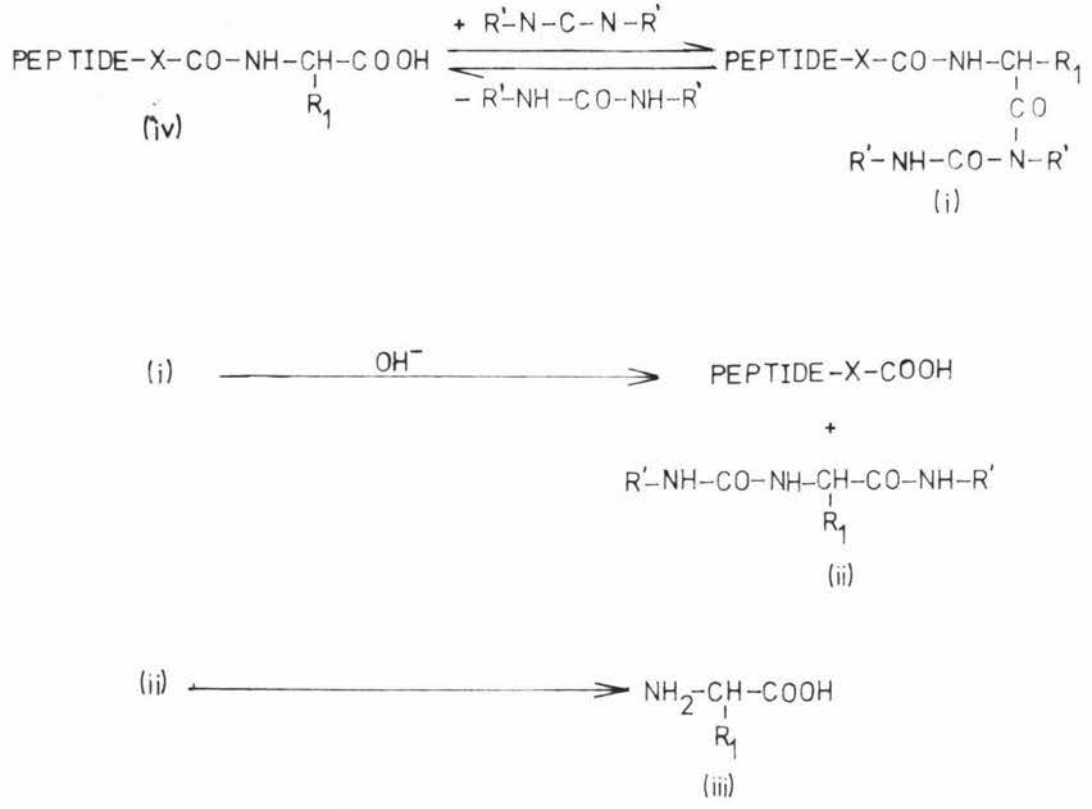


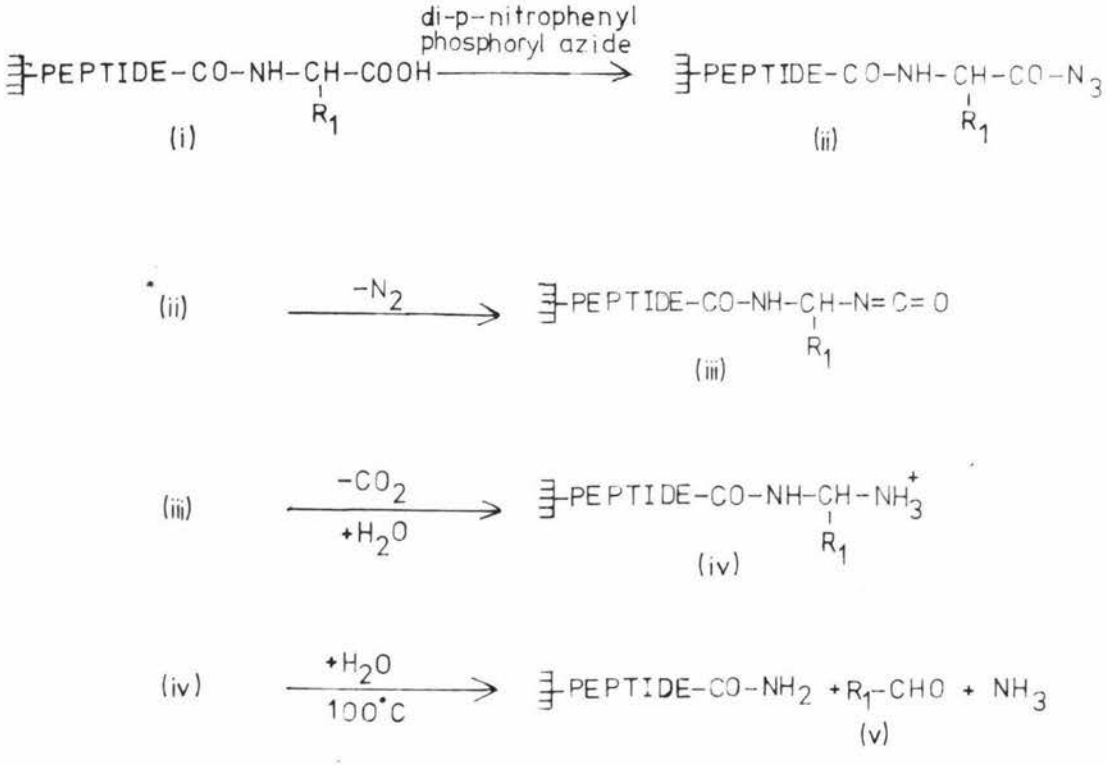
FIGURE 11 Carbodiimide-catalysed Sequential Removal of C-terminal Amino Acids

aqueous ethanol the C-terminal amino acid is split off as an N-p-tolyl-carbamyl toluide derivative (ii). The product derived from the C-terminal residue can be converted into the parent amino acid (iii) by acid or alkali hydrolysis. Unfortunately the acyl urea derivative (i) is hydrolysed in a significantly quantitative side reaction to reform the original peptide (iv). As this will interfere with subsequent carbodiimide coupling cycles this method has found little application in the carboxyl terminal amino acid sequence determination of proteins.

6. Curtius Rearrangement

A sequential degradative procedure beginning at the C-terminus of a polypeptide was proposed by Bergmann and Zervas (1936) using the Curtius rearrangement. The reactions involved were difficult to perform when applied to proteins and occurred in only moderate yields. The reactions were conversion of the benzoyl peptide ester into the corresponding hydrazide, then into the peptide amide using nitrous acid, then into the benzoyl urethane derivative using benzoyl alcohol. Catalytic reduction yields the benzoyl peptide amide with one residue less than the parent peptide plus an aldehyde derivative of the C-terminal residue.

An extension of this method has been developed by Parham and Loùdon (1978) for residue-by-residue sequential degradation from the C-terminus of peptide amides using solid supports and new reagents to produce the azide derivative in high yields. Figure 12. The peptide was coupled to glass beads through their amino terminus (i) and the carboxyl groups of the coupled peptide were converted to the acyl azide derivatives (ii) in a one step reaction using di-p-nitrophenyl phosphoryl azide. Figure 12. I, I-bis (trifluoroacetoxy) iodobenzene brings about the conversion of the peptide



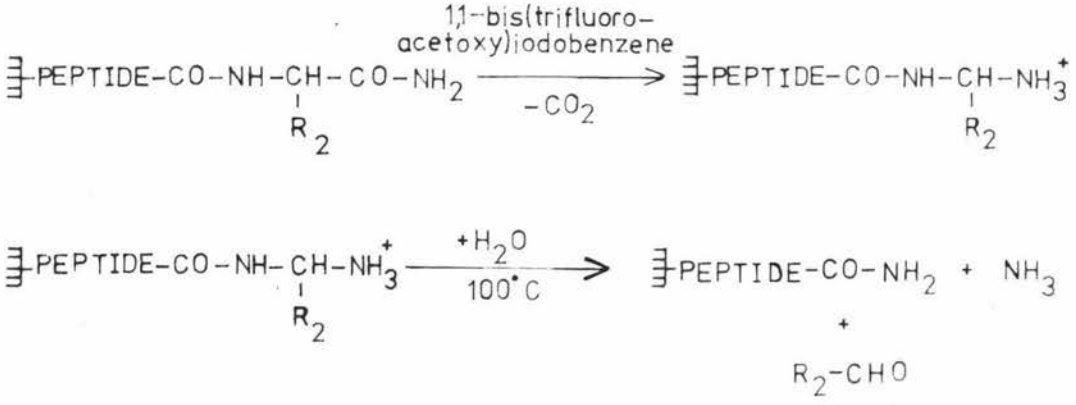


FIGURE 12 The Sequential Method of Parham and Loudon

amide (iv) to the isocyanate derivative (iii), the hydrolysis of which occurs under acidic conditions. The main problem with this method is that no satisfactory means of identification of the aldehyde derivative (v) is available therefore subtractive amino acid analysis of the peptide was performed by Parham and Loudon (1978). Taking suitable precautions for the more labile and reactive sidechain amino acids greater than 80% yield was achieved at each cycle in the sequencing of model peptides.

7. Carboxypeptidase Digestion

It is apparent from methods 1 to 6 that no satisfactory chemical amino acid sequencing method from the carboxyl terminus of polypeptides is available which is free from undesirable side-reactions and gives an easily identifiable end product in sufficient yields. It is therefore currently preferable to use enzymatic methods to sequentially cleave amino acids from the C-terminus of polypeptides. Although attractive because of technical simplicity compared to chemical methods the scope of this approach is limited. In chemical methods all peptide molecules should ideally undergo each degradation cycle before a new cycle is started. This is not the case for enzymatic degradation where each peptide or protein molecule, whether minus one or more C-terminal residues, has an equal opportunity to become a carboxypeptidase substrate. Thus the sequence of amino acid residues is not determined in a stepwise manner but instead from the rate of release of amino acids from the polypeptide. The C-terminal residue will be released at the highest rate with the C-terminal penultimate residue next and so on. Unfortunately the difference in rates of release of amino acid residues becomes smaller as degradation proceeds beyond a few residues thus making the results ambiguous.

The specificity of the exopeptidase can be a further limitation of enzymatic degradation of proteins. Different C-terminal amino acids have varying susceptibilities

to release by a particular carboxypeptidase. Thus while some C-terminal amino acids are removed rapidly others are not removed at all therefore constituting a blockage. If a slowly-released amino acid precedes a rapidly-released amino acid in the sequence from the C-terminus then the two residues will appear in the digest at similar rates. Bovine pancreatic carboxypeptidase A rapidly removes C-terminal aromatic and aliphatic amino acids having long sidechains while glycine and acidic residues are removed only slowly and lysine, arginine, proline and histidine block carboxypeptidase A action completely (Ambler, 1972). Porcine pancreatic carboxy-peptidase B has a restricted specificity and rapidly releases only lysine and arginine. Carboxypeptidase C from citrus fruit removes all C-terminal amino acids including basic and acidic residues and proline. Carboxypeptidase Y from yeast also possesses a broad specificity including proline release (Potts, 1967) (Ambler, 1972) (Folk, 1970).

A serious practical difficulty is contamination of the enzyme preparation with other proteases, especially endopeptidases which produce new carboxyl terminals for the carboxypeptidase. Using suitable inhibitors this contaminant proteolytic activity can be removed. Short peptide substrates are less likely to be affected by endopeptidase contamination compared to larger proteins as there is less likelihood of their containing a susceptible peptide bond. This probability increases greatly with increasing substrate size.

B. Methods of Carboxyl-Terminal Peptide Isolation

1. Complete Amino Acid Sequence Determination of the Protein and Alignment of Peptides

When a protein is being fully sequenced the complex mixtures of peptides obtained by various proteolytic digestions are repeatedly chromatographed until each

peptide is obtained in a homogeneous state. Once the amino acid sequence of these peptides is determined they can be aligned in order using overlapping amino acid sequences from different proteolytic digestions or chemical cleavages to give the complete protein amino acid sequence. A carboxypeptidase digestion or end group determination method on the entire protein will indicate which peptide is derived from the carboxyl terminal end of the protein. Remembering the specificity of proteases used to cleave the intact protein into manageable peptides, the carboxyl terminal peptide can be deduced from the nature of the C-terminal amino acid of each peptide. If the protein is cleaved by chymotrypsin all except the C-terminal peptide will have tyrosine, phenylalanine and isoleucine as the predominant C-terminal residue of each peptide (unless the C-terminal peptide also terminates in one of these residues). If the protein should have an arginine or lysine C-terminal residue; as human $\overline{C1s}$ a chain probably does, then a chymotryptic digestion will produce only one arginine - or lysine - ending peptide, this having been derived from the C-terminus of the polypeptide. Treatment of a protein with cyanogen bromide will specifically fragment the protein on the carboxyl side of methionine residues. Therefore only the C-terminal peptide will not have the cyanogen bromide-derived homoserine lactone as its C-terminal residue (unless the protein contains methionine as its C-terminal residue).

These methods are obviously not applicable in situations where only limited information is known about the protein or where the protein cannot be completely sequenced. In these situations the carboxyl terminal residue of the protein must be chemically altered or labelled in some way to allow differentiation and hence separation from the remaining bulk peptides derived from an enzymatic or chemical cleavage of the protein. This means that only the C-terminal derived peptide needs to be purified for sequence determination.

2. Tritium Labelling

A suitable method for general application to C-terminal

peptide labelling is tritium incorporation into the C-terminal residue of the protein. (See Section 1.3, Part A, Number 3). Subsequent enzymatic cleavage of the protein will give one radioactive peptide which can be conveniently followed as it is separated from the bulk peptides. Once obtained in a homogeneous form the amino acid sequence can be determined by well-established N-terminal sequencing methods. The radioactive C-terminal amino acid residue can also be determined with certainty upon acid hydrolysis of the peptide.

3. Chemical Modification of Protein Carboxyl Groups

A variation of such radioactive labelling methods was used by Carraway and Koshland (1972) (Carraway et al, 1969) who followed the proteolytic activation of chymotrypsinogen by firstly chemically blocking all sidechain and C-terminal carboxyl groups before activating the zymogen to generate a new carboxyl group at the peptide bond cleavage site. The carboxyl group of the newly-found polypeptide chain is radioactively labelled to permit identification and hence easier isolation of the C-terminal-derived peptide. This method is of course only applicable to situations where a new polypeptide chain is generated as a result of zymogen activation by limited proteolysis as in C1s activation.

4. Peptide Mapping

Two dimensional peptide mapping has been employed to isolate the C-terminal peptide of proteins. (Canfield, 1963) (Carlton and Yanofsky, 1963). One-half of the protein is subjected to a short carboxypeptidase digestion while the remaining half is left undigested. After removal of free amino acids both protein samples are proteolytically digested in parallel and subjected to a two-dimensional combination of electrophoresis and paper chromatography. The peptide maps of the sample not previously digested with carboxypeptidase will contain all the peptides constituting the protein. The maps of carboxypeptidase-digested protein will contain the same peptides in the same positions except for one peptide, which due to removal of some C-terminal

amino acid residues will be missing completely or have different electrophoretic and chromatographic mobilities.

This method is only applicable to smaller molecular weight proteins so that after proteolytic digestion a reasonable number of peptides can be identified otherwise the peptide maps or 'fingerprints' tend to be extremely difficult to interpret with any confidence. Very specific proteolysis is required to reduce the number of partial or nonspecific cleavage peptides. Specific cleavage is aided if the protein contains a carboxyl terminal arginine or lysine residue thus permitting the use of carboxypeptidase B to remove only the C-terminal residue. Such a removal of one positively-charged residue would be expected to greatly alter the mobility of this peptide. Once the corresponding peptides having different mobilities have been identified the intact C-terminal derived peptide can be eluted from the noncarboxypeptidase treated map and subjected to amino acid analysis and sequence determination.

5. Affinity Chromatographic Isolation of Arginine-ending Peptides

An interesting application of affinity chromatography has been used by Yokosawa and Ishii (1976, 1979) to purify biologically-active peptides. These workers used anhydrotrypsin, an enzymically inert derivative of trypsin, which is still able to stoichiometrically bind trypsin inhibitors. Anhydrotrypsin is prepared by base elimination of PMS from PMSF-treated trypsin and results in the dehydration of the active site serine-185 residue to dehydroalanine. It exhibits a much higher affinity for the product-type ligand containing an L-arginine residue having a free carboxyl group than for substrate-type ligands containing a substituted carboxyl group.

Immobilization of anhydrotrypsin on agarose gels provides a means of specifically binding and isolating derivatives of L-arginine whose carboxyl group is free. As many biologically-active arginine-ending peptides such as bradykinin, Kallidin,

tuftsin and fibrinopeptide are products of trypsin-like enzymes, chromatography on anhydrotrypsin-Sepharose resulted in retension of these peptides. Carboxypeptidase B treatment of these peptides abolished this binding ability indicating that the C-terminal arginine residue is essential for anhydrotrypsin binding. (Yokosawa and Ishii, 1976). Similarly benzoyl phenylalanine, benzoyl L-arginine-amide, benzoyl D-arginine and free L-arginine are not bound by anhydrotrypsin-Sepharose while benzoylglycyl lysine was slightly bound and benzoyl L-arginine was bound still more. Only benzoylglycyl arginine was bound sufficiently strongly to the column to require elution with 5 mM HCl. Benzoylglycyl arginine could be separated from a 10-fold excess of benzyloxycarbonyl L-phenylalanine by application to anhydrotrypsin-Sepharose and elution with 5 mM HCl. (Yokosawa and Ishii, 1976). This approach to the isolation of arginine-ending peptides has been applied to the fractionation of peptides from a tryptic digestion of erabutoxin (Yokosawa and Ishii, 1979). As was the case using small model substrates, it was found that arginine-ending peptides exhibited stronger affinity for anhydrotrypsin than lysine-ending peptides.

A chymotryptic digestion of a protein possessing a carboxyl terminal arginine residue would yield a single arginine-ending peptide; this being the carboxyl terminal-derived peptide. Anhydrotrypsin affinity chromatography would provide a convenient means for the isolation of such a peptide. This rationale was applied to $\overline{C1s}$ in an attempt to isolate a peptide derived from the carboxyl terminus of the $\overline{C1s}$ a chain which is thought to terminate in an arginine or lysine residue. Other methods which were also applied to $\overline{C1s}$ a chain carboxyl terminal determination were the carboxyl-modification method of Koshland (Number B3), the peptide mapping procedure (Number B4) and carboxypeptidase digestion (Number A7).

CHAPTER 2

MATERIALS AND METHODS

MATERIALS

Sodium acetate, sodium bicarbonate, taurine, sodium dodecyl sulphate, salmine, agarose, ammonia (Analar grade), sodium barbitone, N,N'-methylenebisacrylamide (recrystallized from acetone) and acrylamide (recrystallized from chloroform) were obtained from British Drug Houses Limited. Univar grade trichloroacetic acid and sodium hydroxide and Unilab grade ethanolamine and ethylene diamine tetra-acetic acid-disodium salt were obtained from Ajax Chemicals Limited. Potassium hydroxide, ammonium bicarbonate, cadmium acetate, sodium chloride, disodium hydrogen phosphate and potassium dihydrogen phosphate were obtained from May and Baker. Tris (7-9 Buffer grade), glycine, dithiothreitol, cyanogen bromide, 1-dimethylaminonaphthalene-5-sulphonyl chloride, phenylmethanesulfonylfluoride, bovine serum albumin, Coomassie Brilliant Blue R and G250, glycine methyl ester, L-Norleucine and 1-ethyl-3(3-dimethyl-amino-propyl)-carbodiimide-HCl (EDC-HCl) were purchased from Sigma Chemical Company. Phosphorus pentoxide, trisodium phosphate and calcium chloride were purchased from Riedel de Häen. N,N,N',N'-tetramethyl ethylene diamine (TEMED) and 2-mercaptoethanol were supplied by Fluka. Ninhydrin was supplied by Koch-Light Laboratories Limited.

Urea (Koch-Light Laboratories Limited) was dissolved in water, as a 10 molar solution and deionized before use by passage through a column of Amberlite MB-3 resin. Guanidine HCl (Fisher Scientific Company) was further purified as described by Nozaki (1972). Guanidine HCl (250g) was dissolved in 1 litre hot absolute ethanol, decolourized with activated charcoal and filtered. To the hot ethanol solution 500 ml benzene was slowly added and crystals which

were formed on standing at 4°C overnight were washed with cold absolute ethanol-benzene mixture then with acetone before dessicating. Iodoacetic acid, obtained from Riedel de Häen, was recrystallized from cyclohexane.

Enzymes bovine pancreatic ribonuclease A (Type II-A), bovine pancreatic trypsin (DPCC-treated Type XI for proteolytic digestions; 2 x crystallized Type III for anhydrotrypsin preparation), soybean trypsin inhibitor (Type II-S), thermolysin (Type X) and amino acid substrates α-N-Benzoyl-DL-Arginine p-nitroanilide HCl (Bapna), N-α-cbz-L-Lysine p-nitrophenyl HCl and hippuryl-L-arginine were supplied by Sigma Chemical Company. Staphylococcus aureus V8 protease was purchased from Miles Laboratories Limited, United Kingdom and α-chymotrypsin from British Drug Houses Limited. Anti-$\overline{C1s}$ was obtained from Behringwerke AG, West Germany. Carboxypeptidase Y was kindly donated by C.H. Moore and carboxypeptidase B by D.S. Colls.

Chromatography gels Sephacryl S200, Sepharose 4B, Sephadex G100-120, Sephadex G25, Sephadex G150-40, Sephadex G200-120, Sephadex A-50-120 and Amberlite MB-3 were obtained from Sigma Chemical Company. DE-52 and Whatman No.1 chromatography paper were obtained from Whatman. DE viscose was kindly donated by J. Ayers. Oxygen-free nitrogen was supplied by Industrial Gases, New Zealand Limited. Transfusion packs (200 ml) of human plasma containing anticoagulant were kindly donated by the Immunohaematology Laboratory, Palmerston North Hospital. Phenylisothiocyanate was obtained from Koch-Light Laboratories Limited and was purified by distillation at 55°C at 1 mm Hg pressure. Aliquots of distilled phenylisothiocyanate were sealed in vials and stored at -18°C.

Trifluoroacetic acid was purified by refluxing over chromium oxide for three hours and then distilling. The fraction boiling at 72-73°C was collected, dried over calcium sulphate and then redistilled. Pyridine (Ajax Chemicals Limited) was refluxed over sodium hydroxide (1g

per 100 ml) for 2 hours and then distilled with the fraction boiling at 115-117°C being collected. This fraction was then refluxed over ninhydrin (0.5g per 100 ml) for 2 hours and then redistilled as before. Hydrochloric acid was supplied by British Drug Houses Limited and was purified in an all-glass distillation apparatus by diluting concentrated acid with an equal volume of water and collecting the constant boiling fraction (110°C) by distillation. Acetic acid (British Drug Houses Limited) was distilled and the fraction boiling at 119°C was collected. This fraction was then refluxed over potassium permanganate and redistilled as before. Formic acid, toluene, ethanol, acetone and butanol were all drum grade and distilled before use. Sulphuric acid was supplied by Kempthorne-Prosser, New Zealand and was distilled before use. British Drug Houses Limited supplied methanol, n-ethyl acetate and n-butyl acetate which were also distilled before use.

METHODS

2.1 Protein Purification and Characterisation

(i) Component $\overline{C1s}$ Purification

Activated $\overline{C1s}$ was isolated from human plasma essentially according to the method of Lepow et al (1963) as adapted by Sim et al (1977). Human plasma (4 litres) was made 10 mM in calcium chloride and allowed to clot by standing at 37°C for several hours. After centrifugation the serum was dialysed at 4°C against two or three changes of 5 mM CaCl2 pH 7.4. The precipitate was then dissolved in approximately 400 ml of 0.5M NaCl containing 10 mM EDTA by stirring continuously for at least two hours at 4°C to both solubilize and dissociate C1 into its subcomponents.

Centrifugation at 9,000 rpm, 4°C for 30 mins. yielded an opalescent supernatant, which after filtration through Mira cloth to remove lipid-like material, was used as the starting material for subsequent column chromatography. The filtrate was dialysed against 40 mM Phosphate starting buffer, pH 7.4, containing 1 mM EDTA of ionic strength I=0.1 (conductivity 2.6 mmho/cm). Following centrifugation to remove any precipitate formed during dialysis the resultant supernatant was applied to a DEAE-viscose anion exchange column (3.2cm x 25cm) previously equilibrated with phosphate starting buffer. The column was washed with starting buffer, I=0.1 until the absorbance at 280 nm of the eluate was near zero. Stepwise elution was then employed, first using starting buffer containing 0.1M NaCl (I=0.2) to elute the $\overline{C1r}$ peak then starting buffer containing 0.3M NaCl (I=0.4) to elute the $\overline{C1s}$ peak. DEAE viscose chromatography was performed at room temperature to ensure full activation of C1s. The column eluant flow rate was approximately 50 ml/hour, collecting 10 ml fractions.

Active esterase fractions of the $\overline{C1s}$ peak were pooled, dialysed against phosphate starting buffer at 4°C and

applied to a DE cellulose column (1.6cm x 15cm) previously equilibrated with starting buffer. DE 52 column chromatography was performed at 4°C. After application of dialysed $\overline{C1s}$ the column was washed firstly with a small volume of starting buffer then extensively with phosphate starting buffer containing 0.1M NaCl (I=0.2) until the absorbance at 280 nm of the eluate was zero. $\overline{C1s}$ elution was then performed with a linear salt gradient consisting of 400 ml of I=0.2 buffer and 400 ml of I=0.4 buffer. The column elution flow rate was 30-40 ml/hour, collecting 6-8 ml fractions.

(ii) $\overline{C1s}$ Assays

$\overline{C1s}$ esterase activity was measured by following the rate of hydrolysis of N- α -cbz-L-Lysine p-nitrophenyl ester. HCl (cbz-Lys-ONp) at 348 nm using a dual beam Perkin Elmer 124 spectrophotometer equipped with a pen recorder. Assays were performed at room temperature (19°C). The assay mixture consisted of 0.2 ml enzyme solution, 0.1 ml cbz-Lys-ONp solution and 2.7 ml of 0.08M phosphate buffer, pH 6.0 in a 3 ml quartz cuvette. The cbz-Lys-ONp solution was freshly prepared by dissolving 6mg in 5.0 ml distilled water. The final substrate concentration in the assay mixture was 0.09 mM. Auto-hydrolysis of the substrate was determined using a substrate blank containing 0.2 ml distilled water instead of enzyme solution. An extinction coefficient for p-nitrophenol of 5400 ℓ/mol/cm was used. (Boland and Hardman, 1972.) One unit of enzyme activity (E.U.) is defined as that amount of enzyme able to cause the hydrolysis of 1 μmole substrate/min at 20°C.

(iii) Component C1s Purification

The procedure used was the same as that for $\overline{C1s}$ purification except that proteolytic activation of C1s was prevented by the addition of phenylmethanesulfonylfluoride (PMSF) throughout the purification, which was rigidly performed at 4°C. Appropriate volumes of a 1M PMSF stock

solution in isopropanol were added at each of the initial stages of the C1s purification procedure to give a final PMSF concentration of 5 mM. PMSF was added to serum before and after dialysis against 5 mM $CaCl_2$ as well as the $CaCl_2$ wash step and the NaCl solubilization step. Although PMSF was added for dialysis against starting buffer none was included in column chromatography buffers. All glassware in contact with PMSF was immersed in 1M NaOH solution to ensure complete hydrolysis of PMSF.

(iv) Activation of $\overline{C1s}$ by $\overline{C1r}$

C1s was assayed by incubating appropriate volumes of the C1s fraction (0.5 ml) at 37°C for 30 mins with 0.050-0.100 ml of partially purified $\overline{C1r}$. $\overline{C1r}$ was obtained from the I=0.2-eluted peak of the DEAE-viscose chromatography step of a previous $\overline{C1s}$ preparation. Resultant $\overline{C1s}$ activity was then assayed as described in "Methods" section 2.1 (ii). $\overline{C1r}$ and C1s controls were performed.

(v) Protein Determination

Protein determinations were performed by the Coomassie protein dye-binding method of Bradford (1976). The absorbance of samples at 595 nm was determined using an Hitachi 101 manual spectrophotometer. Bovine serum albumin ($E^{1\%}_{1cm}$= 6.68) (Tanford and Roberts, 1952) was used as a protein standard. The concentration of purified $\overline{C1s}$ was sometimes determined using $E^{1\%}_{1cm}$= 9.4. (Sim et al, 1977). Trypsin concentrations were determined using $E^{1\%}_{1cm}$= 15.4 at 280 nm (Yokosawa and Ishii, 1977) or by the Coomassie method as above using trypsin as the standard protein.

(vi) Polyacrylamide Gel Electrophoresis

Disc gel electrophoresis was performed using the 7.5% polyacrylamide discontinuous gel system of Davis (1964) and Ornstein (1964). The gel solution for 7.5% polyacrylamide gels was prepared as required and consisted of: 7.5 ml of a stock solution of 30% acrylamide and 0.17% N,N'-methylenebis-acrylamide, 3 ml of 1M tris HCl, pH 8.9 and distilled water

to a total volume of 29.4 ml. After de-aerating for several minutes 0.3 ml of 10% sodium dodecyl sulphate (SDS), 0.15 ml of 10% ($^{v}/v$) TEMED in ethanol and 0.15 ml 10% ammonium persulphate was added in the order given. The gel solution was poured into 6mm x 10cm glass tubes and distilled water was layered carefully onto the gel surface. After 45 minutes the tubes containing the solidified gel were placed in an Ortec disc gel electrophoresis apparatus containing electrode buffer of 0.1M tris glycine, pH 8.3 and 0.1% SDS. SDS-polyacrylamide gels and electrode buffer contained 0.1% SDS (Weber and Osborn, 1969). Native gels were prepared in the same manner as for SDS-polyacrylamide gels except that SDS was omitted from both the gel and electrode buffer.

The sample, 10-100 μg protein in approximately 20 μl electrode buffer or distilled water was mixed with 10 μl glycerol containing bromophenol blue dye marker and applied to the gel. Reduction of protein samples (50 μl) was carried out by incubation with 10 μl 2-mercaptoethanol and 20 μl SDS for two minutes at 100°C. Electrophoresis was performed at 1 mA/gel and increased to 5 mA/gel after the sample had entered the gel. Electrophoretic mobilities (Em's) of protein bands were measured relative to the position of the bromophenol blue dye marker which was marked before the commencement of gel staining. Staining and destaining of gels was performed as described by Weber and Osborn (1969). Gels were stained for two hours in solution of 1.25g Coomassie Brilliant Blue R250 dissolved in 454 ml 50% ethanol and 46 ml glacial acetic acid. Native gels were sometimes stained using the standard protein determination stain. Gel destaining was accomplished by repeated changes of a gel destain solution consisting of 75 ml acetic acid, 50 ml methanol and 875 ml distilled water.

(vii) Immunodiffusion

Immunodiffusion was performed essentially by the method of Ouchterlony (1953). Clean microscope slides were coated with 1 ml of hot 0.2% agarose in distilled water before

allowing to dry and layering with 2.5 ml hot 1% agarose in 0.02M Na barbital-HCl buffer, pH 8.2, ionic strength, I=0.025. After allowing the agarose gel to set the slides were stored at 4°C in a water saturated atmosphere for at least four hours prior to well-punching. Immediately before use the well-plugs were removed and the inner well was filled with antibody to $C\bar{1}s$ while the outer wells were filled with suitable aliquots (0.050 - 0.200 ml) of the $C\bar{1}s$ solution. Diffusion was allowed to proceed for one to three days at 4°C. After precipitation lines had formed the plates were washed overnight in phosphate buffered saline (0.01M KH_2PO_4/NaOH, pH 7.4 containing 0.15M NaCl) and stained for 10 minutes with 0.125% Coomassie BB R250 in acetic acid: methanol: water (1:9:10 $^{v/v/v}$) before destaining in 7% acetic acid.

(viii) Reduction and Carboxymethylation of $C\bar{1}s$

An appropriate volume of a 10 molar urea solution, previously deionized by passage through an Amberlite MB-3 column, was added to a volume of 1M tris-HCl, pH 8.0 to give a final buffer concentration of 0.1M tris-HCl, pH 8.0 containing 8M urea. Solid EDTA-disodium salt was added to give a final concentration of 0.2%. Protein was dissolved to a concentration of about 0.5% and the reaction mixture was gassed with oxygen-free nitrogen for at least 15 minutes at room temperature, with constant stirring. Reduction was initiated by the addition of solid DTT to give a final concentration of 50 mM and reduction was allowed to proceed overnight at room temperature. Carboxymethylation was performed by the addition of iodoacetic acid, neutralized in a minimum volume of distilled water, to give a final concentration of 110 mM or an approximate two-fold molar excess over DTT. The reaction was allowed to proceed at room temperature and in the dark for 40 minutes after which time the reaction was terminated by quenching with a minimum ten-fold molar excess of 2-mercaptoethanol over iodoacetic acid added. The reaction mixture was then either desalted by extensive dialysis or gel filtration using Sephadex G25 and freeze dried or alternatively dialysed directly against urea-containing

buffers in preparation for chain separation.

2.2 Chain Separation

Column chromatography in the presence of denaturant was performed at room temperature while in the absence of denaturant 4°C was used. Eluate fractions were collected from the columns using an LKB 7000 Ultrorac fraction collector and were monitored for UV absorbance using an LKB 8300II Uvicord monitor.

(i) Gel Filtration Techniques

Pharmacia jacketed precision bore chromatographic tubes (1.6 cm x 70 cm or 2.5 cm x 100 cm) were used for gel filtration chromatography. Various gel filtration media was used as described by Hirs (1977) for "Chain Separation" by gel filtration. Sephadex G100, G150, and G200 were used with and without denaturants present. Sephacryl S200 was used in the absence of denaturant. Downward elution by gravity flow was carried out at a maximum flow rate of 2 ml/hr/cm^2.

(ii) Ion Exchange Chromatography

Ion exchange chromatography was performed in a glass column (1.5 cm x 10 cm) at a maximum flow rate of 20 ml/hr. The procedure employed was that of Campbell et al (1979) for the separation of bovine reduced and carboxymethylated $\overline{C1s}$ a and b chains except that DEAE Sephadex A-50 was the ion exchange resin used. The RCM-$\overline{C1s}$ sample was dialysed (with changes) against 40 mM phosphate starting buffer, pH 7.4 containing 5 mM NaCl and 8M urea and then applied to the DEAE Sephadex A-50 column previously equilibrated in the same buffer. The column was then washed with a minimum of two column volumes of phosphate starting buffer before chromatographic development of the column was commenced with a linear NaCl gradient consisting of 100 ml of phosphate starting buffer and 100 ml of phosphate starting buffer containing 0.5M NaCl.

2.3 Enzyme Digestions

The protein substrate was dissolved in a suitable buffer at a concentration of 0.5% and digested at 37°C using an enzyme to substrate ratio of 1% (w/w), unless otherwise stated.

(i) Exopeptidase Digestions (Ambler, 1972)

Carboxypeptidase Y was stored as an ammonium sulphate suspension at 4°C and before use 0.1 ml of this suspension was diluted with 0.1 ml of 0.01M phosphate buffer, pH 7.0 and dialysed for 2 hours at 4°C against 0.01M phosphate buffer, pH 7.0. Carboxypeptidase Y digestions were performed in 0.1M pyridine acetate, pH 5.5. Frozen carboxypeptidase B was thawed and added directly to the digest buffer of 50 mM N-ethyl morpholine, pH 8.0 at an enzyme to substrate ratio of 1:60 (w/w). Protein substrate and L-Norleucine internal standard were both dissolved in such a volume of digest buffer that a 0.2 ml aliquot contained 50 nmoles of each. The 0.2 ml aliquot withdrawn at each sampling time was added directly to 1 ml of 15% TCA and centrifuged. The supernatant was removed and extracted three times with diethyl ether and the remaining aqueous phase was freeze dried over phosphorus pentoxide and sodium hydroxide pellets in preparation for subsequent amino acid analysis. The yields of free amino acids obtained were corrected using the yield of L-Norleucine internal standard to give 100% recovery. Parallel enzyme blanks and protein substrate blanks containing the same amount of carboxypeptidase and protein substrate present in the digest sample were also performed.

A carboxypeptidase digestion for peptide mapping consisted of a single time digestion (4 hours) in the appropriate buffer and the digestion was halted by the addition of 15% TCA as for time course experiments. After centrifugation the supernatant was extracted with diethyl ether and the aqueous phase was freeze dried over phosphorus pentoxide and sodium hydroxide pellets for subsequent amino

acid analysis. The precipitate containing undigested protein substrate was resuspended in 3M urea and dialysed overnight against 0.1M NaCl then exhaustively against distilled water before freeze drying over phosphorus pentoxide in preparation for further digestion by an endopeptidase. An enzyme blank and protein substrate blank were also subjected to the above same procedure.

(ii) Endopeptidase Digestions

Digestions using DPCC-treated trypsin and α-chymotrypsin were performed in 0.1M Ammonium bicarbonate, pH 8.0 at 37°C with an enzyme to substrate ratio of 1% ($^{w}/_{w}$) for 16 hours. Soybean trypsin inhibitor was added to chymotryptic digestions (on a 1:1 molar basis with chymotrypsin) to inhibit any contaminant trypsin activity present. S. V8 protease digestions were performed in 50 mM Ammonium bicarbonate, pH 7.8 using an enzyme to substrate ratio of 2% ($^{w}/_{w}$). Digestions were halted by placing in a boiling water bath for 10 minutes and then freeze drying. Digestions of peptide substrates were performed in 1% Ammonium bicarbonate, pH 8.0 for 16 hours at 37°C using an enzyme to substrate ratio of 2.5% ($^{w}/_{w}$) for thermolysin and trypsin and 30% ($^{w}/_{w}$) for S. V8 protease.

2.4 Peptide Mapping

(i) High Voltage Paper Electrophoresis (Shotton and Hartley, 1973)

High voltage paper electrophoresis was performed using a tank electrophoresis apparatus with a coolant. Electrophoresis was carried out at pH 2.1 or pH 6.5. The pH 2.1 buffer consisting of formic acid/glacial acetic acid/distilled water (1:4:45 $^{v/v/v}$) was used for separating peptides which were basic or acidic at pH 6.5 while the pH 6.5 buffer of pyridine/glacial acetic acid/distilled water (25:1:225 $^{v/v/v}$) was used for separating peptides which were basic or acidic at pH 6.5. The electropherograms were subjected to 50 volts per cm of Whatman No.1 chromatography paper (46cm x 57cm) for 40 minutes. To aid the comparison and identification of peptides two 5 mM marker amino acid solutions in 10% isopropanol were used.

R marker: Tyr, Phe, Met, Leu, Val, Arg, Lys

T marker: Asp, Glu, Thr, Ser, Ile, Ala, His, Gly

DNS-Arg and DNS-OH were used as fluorescent markers for electrophoresis and chromatography. For electrophoresis at pH 6.5 mobilities of peptides were measured relative to aspartic acid and at pH 2.1 and BAW chromatography relative to DNS-Arg. Guide strips were cut from electropherograms for staining to locate the positions of peptides where further purification of peptides by electrophoresis was required.

(ii) Paper Chromatography

Descending chromatography on sheets of Whatman No. 1 chromatography paper was carried out using a solvent system of butanol/glacial acetic acid/distilled water (4:1:5 $^{v/v/v}$). The upper phase was used as solvent. (Katz *et al*, 1959). Chromatography papers were equilibrated overnight with the lower aqueous phase.

(iii) Peptide Staining on Paper

After checking for fluorescence under a UV light the electropherograms and two dimensional peptide maps were stained with the ninhydrin-cadmium acetate reagent consisting of 3 parts of solution A (15g cadmium acetate in 300 ml glacial acetic acid and 600 ml distilled water) and 17 parts of solution B (1% ninhydrin in acetone) and then air dried at room temperature. The papers were stored in a chromatography tank containing a small vessel of concentrated sulphuric acid to produce an ammonia-free atmosphere to reduce background colour development.

(iv) Peptide Elution from Paper

Following electrophoresis or chromatography the peptides located by either using the stained guide strips or by direct ninhydrin staining were eluted from the paper with 0.02M ammonia. Peptides to be subjected to acid hydrolysis for subsequent amino acid analysis were eluted using 6M HCl. Paper strips containing peptides to be eluted were supported vertically between glass slides in a perspex peptide elution apparatus. The peptides were eluted in a descending direction into acid-washed glass test tubes (6cm x 1 cm) and 2 ml of eluate was collected. A blank consisting of a similar portion of the same chromatography paper which contained the peptide was eluted with 6M HCl when amino acid analysis of the eluted peptide was to be performed.

(v) Amino Acid Analysis

Determination of the amino acid composition of peptide and protein samples was carried out by firstly lyophilizing the sample in thick walled acid washed pyrex test tubes (10mm x 75mm) followed by addition of 0.2 ml of 6M HCl (containing 0.1% ($^{v}/v$) phenol for protein hydrolysis). The tubes were sealed under vacuum (0.5 mm Hg) and hydrolysed at 110^{o}C overnight then freeze dried over sodium hydroxide pellets and applied to a Beckman 120C amino acid analyser.

(vi) Peptide Mapping

A one-half aliquot of the protein substrate was subjected to a carboxypeptidase digestion while the undigested half was used as a substrate control. (See "Methods", 2.3 (i).) The trichloroacetic acid-precipitated protein of each sample was then resuspended and exhaustively dialysed before being digested by an endopeptidase. (See "Methods", 2.3 (i) and (ii).) The lyophilized enzyme digests were then dissolved in 1% ammonium bicarbonate and applied to Whatman No.1 paper and subjected to pH 6.5 electrophoresis. The neutral band of the electropherogram of each digest sample was cut out and sewn onto a second sheet of Whatman No.1 paper and rerun at pH 2.1. The acidic and basic peptides of the pH 6.5 electropherogram and the neutral peptides rerun at pH 2.1 were cut out and each was sewn onto a second sheet of Whatman No.1 paper for BAW chromatography in the second dimension. The pH 2.1 and pH 6.5 electrophoresis and BAW chromatography of the endopeptidase digests of carboxypeptidase and noncarboxypeptidase-treated protein substrates were each performed at the same time to minimize differences in peptide mobility due to different experimental conditions.

2.5 Chemical Modification

(i) Protein Carboxyl-blocking

Chemical modification was performed as described by Carraway and Koshland (1972). Protein (10 mg/ml) plus glycine methyl ester or taurine (1.0M) were dissolved in distilled water at pH 2 to 3. The pH was adjusted to pH 4.75 prior to the reaction, which is initiated by adding solid EDC-HCl to a concentration of 0.1M . Identical quantities of EDC-HCl were added after 1 and 2 hours. After 3 hours at room temperature the reaction is quenched by the addition of excess 1.0M acetate (pH 4.75) and the pH is lowered to pH 3.0. Excess reagents were removed by exhaustive dialysis at $4^{\circ}C$. An appropriate amount of the modified and unmodified protein was then freeze-dried and subjected to 6M HCl hydrolysis for subsequent amino acid analysis. A comparison of the amino acid compositions gave a measure of the extent of carboxyl-modification.

2.6 Affinity Chromatographic Isolation of a Carboxyl Terminal Peptide

Anhydrotrypsin was prepared according to the method of Yokosawa and Ishii (1977) and purified by affinity chromatography using ST-Sepharose 4B before being immobilized on Sepharose 4B.

(i) Preparation of ST-Sepharose 4B

A tryptic digest of salmine (protamine sulphate) was coupled to Sepharose 4B, previously activated with cyanogen bromide, essentially according to the method of Yokosawa *et al* (1976). Salmine (5g) was digested overnight with 20 mg bovine trypsin (DPCC-treated) in 150 ml of 0.1M phosphate buffer, pH 8.0. The digestion was halted by placing in a boiling water bath. Following centrifugation to remove insoluble material the resulting supernatant was freeze dried and coupled directly to cyanogen bromide-activated Sepharose 4B. One hundred millilitres settled volume of previously washed Sepharose 4B was suspended in 200 ml distilled water and activated in a fumehood with 24g CNBr at pH 11, 10°C for 8 minutes. The pH of the activation mixture was maintained at pH 11 by the addition of 5M NaOH. The activated gel was rapidly washed in an icecold sintered glass funnel with cold distilled water before being added to the freeze dried tryptic digest of salmine which was dissolved in 150 ml of 0.5M sodium bicarbonate, pH 9.0. Coupling was carried out at 4°C for 18 hours by slow end-to-end rotation of the coupling vessel.

The coupled Sepharose gel was then extensively washed with distilled water, 10 mM HCl, 0.5M sodium bicarbonate, pH 9 and distilled water before treatment with 150 ml of 2M ethanolamine, pH 9.0 (adjusted with HCl) for 18 hours at 4°C to block remaining activated groups on the agarose. The gel was then thoroughly washed with distilled water, 0.5M NaCl, 10 mM HCl, 0.5M sodium bicarbonate buffer, pH 9.0, 10 mM HCl and distilled water followed by a final washing with 0.05M sodium acetate buffer, pH 5.0 containing 0.02M

$CaCl_2$. The ST-Sepharose gel was then packed into a glass column (2 cm x 30 cm) and after further equilibration with the sodium acetate buffer was used to purify anhydrotrypsin.

The amount of immobilized ligand was determined by the method of Kasai and Ishii (1975). A known settled volume of ST-Sepharose was washed with distilled water, solutions of increasing acetone content and finally acetone then freeze dried. The dried gel was subjected to acid hydrolysis and subsequent amino acid analysis. (See "Methods", Section 2.4 (v).)

(ii) Preparation and Purification of Anhydrotrypsin

Anhydrotrypsin was prepared by alkaline treatment of PMS-trypsin by the method of Ako et al (1972) as adapted by Yokosawa and Ishii (1977) and was purified by affinity chromatography using ST-Sepharose (Yokosawa et al 1976). Bovine pancreatic trypsin (500 mg) was treated with 200 mg PMSF (dissolved in a minimum volume of isopropanol) in 190 ml of 0.1M tris-HCl buffer, pH 7.0 containing 0.02M $CaCl_2$ at 25°C for 30 minutes. The pH was continually monitored and adjusted with alkali if required. The reaction was halted by the addition of HCl to give pH3. After centrifugation, the supernatant was dialysed overnight against 1 mM HCl at 4°C.

The PMS-trypsin obtained was treated with 0.05M KOH (final concentration) at 0°C for 10 minutes and the reaction was stopped by lowering the pH to pH 5. The anhydrotrypsin solution was made 0.05M in acetate, pH 5.0 and 0.02M in $CaCl_2$ by the addition of appropriate amounts of 1M sodium acetate buffer and 0.2M $CaCl_2$. The solution was then applied to the ST-Sepharose column (2 cm x 30 cm) equilibrated with 0.05M sodium acetate, pH 5.0 containing 0.02M $CaCl_2$ and affinity chromatography was performed at 4°C. After thorough column washing with the above buffer bound anhydrotrypsin was eluted with 5 mM HCl. The flow rate employed was 40 ml/hour and 10 ml fractions were collected.

Trypsin and anhydrotrypsin were assayed using α-N-Benzoyl arginine p-nitroanilide (Bopna) as substrate according to the method of Erlanger et al (1961). The assay mixture consisted of 1 ml of 100 mg Bapna in 100 ml distilled water, 1.8 ml of 0.1M tris-HCl, pH 8.0 and 0.2 ml of trypsin solution. The absorbance at 405 nm was followed for 5 minutes and activity was calculated from 1 E.U. being equivalent to a change in absorbance at 405 nm of 3.32 for a 3 ml assay mixture volume. (Erlanger et al, 1961).

(iii) Preparation of Anhydrotrypsin-Sepharose 4B

Sepharose 4B (25 ml settled volume) was washed thoroughly with 0.1M NaCl followed by distilled water before being suspended in 50 ml distilled water. Activation and coupling was performed as described by Cuatrecasas and coworkers (Parikh et al 1974). All steps involving the use of CNBr were conducted in a fumehood. The gel slurry was added to an equal volume of 2M carbonate buffer and the gel slurry was stirred with a magnetic stirrer. A temperature of 5-10°C was maintained by the addition of ice and the activation process was initiated by the dropwise addition of 5g CNBr dissolved in 2.5 ml acetonitrile. After 2 minutes the gel was rapidly washed with 10-15 volumes of cold distilled water using an ice-cold sintered glass funnel then rapidly with cold coupling buffer, 0.5M sodium bicarbonate, pH 9.5.

The gel was filtered to a moist cake then suspended in coupling buffer to give a total volume of twice the gel settled volume. Solid anhydrotrypsin (120 mg) was added and coupling was carried out by slow end-to-end rotation of the reaction vessel for 24 hours at 4°C. The gel was then washed with cold distilled water and then treated with 100 ml of 2M ethanolamine HCl, pH 9.5 and coupled as above to mask any reactive groups remaining on the agarose gel. The anhydrotrypsin-Sepharose 4B was washed with distilled water followed by alternate washes with 0.1M sodium bicarbondate, pH 9.5 and 10 mM HCl, both containing 0.5M NaCl, then with distilled water and finally with 0.05M

pyridine acetate, pH 5.0 containing 0.02M $CaCl_2$. The gel was then packed into a column (0.8 cm x 20 cm) and further equilibrated with the same pyridine acetate buffer ready for use.

(iv) Peptide Isolation by Affinity Chromatography

Anhydrotrypsin-Sepharose affinity chromatography was performed at 4°C. The sample was applied in a minimum volume (0.1 - 0.5 ml) of 0.05M pyridine acetate, pH 5.0, containing 0.02M $CaCl_2$. The column was washed thoroughly with at least 1 litre of the above buffer before buffer replacement with 5 mM HCl to elute column-bound material. Elution was carried out at a flow rate of 5 ml/hour and 2 ml fractions were collected. The absorbance at 225 nm and 280 nm of each fraction was determined using a Beckman Acta III double beam spectrophotometer to locate those fractions containing peptides. The fractions containing peptides bound to anhydrotrypsin-Sepharose were combined and freeze dried over phosphorus pentoxide and sodium hydroxide pellets. After dissolving the freeze dried peptide in 200 µl distilled water an aliquot (10 - 20 µl) was taken for N-terminal dansylation. If N-terminal end group analysis showed the presence of a single peptide only, the remaining material was used for amino acid sequence determination or amino acid analysis.

(v) Dansyl-Edman Procedure for Peptide Sequencing (Hartley, 1970)

Edman Cycle Procedure for Peptide Sequencing.

The peptide sample (100-200 nmol) was transferred to an acid-washed screw-capped pyrex test tube (1 cm x 4.5 cm) and dried down under vacuum ready for Edman degradation. Distilled water (150 µl) and 150 µl of 5% ($^{v}/v$) phenylisothiocyanate in pyridine were added and the tube was gassed with oxygen-free nitrogen before being sealed. The tube was incubated at 45°C for 90 minutes. The tube contents were then dried over phosphorus pentoxide and sodium hydroxide pellets under high vacuum at 60°C for 20 minutes then at room temperature

for 10 minutes. Trifluoroacetic acid (250 µl) was added and the tube was gassed with oxygen-free nitrogen before being sealed. The tube was incubated at 45°C for 30 minutes to cleave the N-terminal phenylthiocarbamyl amino acid from the peptide. The TFA was removed under vacuum over sodium hydroxide pellets. To remove byproducts the material in the tube was suspended in 0.2 - 0.5 ml distilled water and extracted three times with 2 ml n-butyl acetate. The top n-butyl acetate layer was discarded and the peptide in the aqueous phase was dried down. The peptide sample, now one residue shorter, was dissolved in 100-150 µl distilled water, from which a 10 - 15 µl sample was taken for N-terminal analysis by dansylation. The remaining material was subjected to another cycle of the Edman degradation.

Dansyl Procedure for N-terminal Determination of Peptides (Gray, 1972)

Five to ten nmol (10 µl) of peptide was dried down in a 5 mm x 30 mm glass tube and 10 µl of 0.2M sodium bicarbonate and 10 µl 1-dimethylaminonaphthalene-5-sulphonyl chloride (2.5 mg DNSCl/ml acetone) were added to give final concentrations of about 5 mM DNSCl, 1 mM peptide and 50% acetone at about pH 9.8. The solution in the tube was mixed and incubated for 30 minutes at 45°C. The reaction mixture was evaporated and the peptide hydrolysed under vacuum in 6M HCl for 16 hours at 110°C. The tube was opened and the contents evaporated over sodium hydroxide pellets. The DNS-amino acid was dissolved in 95% ethanol (15 µl) and equal amounts were spotted on each side of a polyamide sheet (7.5 cm x 7.5 cm) for two dimensional chromatography. One-half µl of a DNS-amino acid marker solution (containing 0.1 mg/ml DNS-Gly, DNS-Glu, DNS-Ile, DNS-Phe, DNS-Pro, DNS-Ser and DNS-Arg) was applied to one side of the plate.

Chromatography solvents:

Solvent 1.	1.5% ($^{v}/_{v}$) formic acid in distilled water
Solvent 2.	Toluene/glacial acetic acid (9:1 $^{v}/_{v}$)
Solvent 3.	Ethyl acetate/methanol/glacial acetic acid (20:1:1 $^{v/v/v}$)
Solvent 4.	1M ammonia/Ethanol (1:1 $^{v}/_{v}$)

The plates were run in Solvent 1 in one dimension for 10 minutes, dried 15 minutes, then run in Solvent 2 in the second dimension for 15 minutes, dried and examined under 366 nm wavelength UV light for the separation of DNS-Leu and DNS-Ile. Solvent 3 was then run in the same dimension as Solvent 2 for 10 minutes, dried and examined under UV light for separation of DNS-Thr and DNS-Ser, DNS-Glu and DNS-Asp and DNS-CMCys and DNS-OH. Solvent 4 was run in the same dimension as Solvent 2 for 45 minutes to separate DNS-His, ε-DNS-Lys and DNS-Arg. Prior to Solvent 4 chromatography an ε-DNS-Lys and DNS-Arg marker was applied next to the sample dansyl spot.

CHAPTER 3

EXPERIMENTAL RESULTS AND DISCUSSION

3.1 Protein Purification and Characterisation

(i) Subcomponent $\overline{C1s}$ Purification

Historically the first method used to isolate human $\overline{C1s}$ consisted of euglobulin precipitation followed by C1 solubilization and dissociation using 0.5M sodium chloride and 1 mM EDTA. (Lepow et al, 1963). A single DE-cellulose ion exchange chromatography step using phosphate buffer, pH 7.4 containing 1 mM EDTA and salt gradient elution (0 - 0.5M) was then employed to separate the C1 subcomponents. The subcomponents thus obtained were described as impure and contained up to eight other protein bands by polyacrylamide gel electrophoresis. (Lepow et al, 1963). This original method was adapted by Gigli and coworkers who performed Sepharose 6B gel filtration of the resuspended euglobulin fraction in the presence of calcium to obtain a more pure C1-containing starting material for subsequent DE-cellulose ion exchange chromatography in the presence of EDTA. (Gigli et al, 1976). Elution was performed as described above to give homogeneous $\overline{C1s}$ as determined by polyacrylamide gel electrophoresis.

The method utilized in these studies was essentially that of Lepow et al (1963) and included an additional ion exchange chromatography step (DE-viscose) instead of the gel filtration step used by Gigli et al (1976). The results of a typical $\overline{C1s}$ purification are shown in Table 1. The elution profiles obtained for the DE viscose and DE cellulose ion exchange chromatography steps are shown in Figures 13 and 14 respectively.

Human serum contains a large excess of $\overline{C1}$ inhibitor (100g/litre of serum) (Fothergill and Anderson, 1976) and so prevents the determination of serum $\overline{C1s}$ esterase

Table 1. Purification of Human Subcomponent $\overline{C1s}$

Step	Fraction	Volume (ml)	Protein Conc. (mg/ml)	Total Protein (mg)	Total Activity (units)	Specific Activity (units/mg)	Yield (%)
	serum	4,500	70	315,000	N.D.[a]	N.D.[a]	
1	NaCl/ EDTA	452	8.5	3,842	306.33	0.0901	100
2	dialysed	507	5.8	2,941	215.47	0.0733	70.3
3	DE-viscose	300	0.64	192	156.67	0.8160	50.9
4	DE-52	111	0.36	40	74.00	1.8530	24.1

Note a. Esterase activity of serum could not be determined because of the presence of natural proteinase inhibitors.

b. The activity present at the NaCl/EDTA solubilization step (Step 1) is assigned to be 100%.

A yield of 40 mg $\overline{C1s}$ from 4.5 litres human serum represents an 8% recovery (9 mg per litre serum) using a $\overline{C1s}$ serum concentration of 110 mg per litre (Müller-Eberhard, 1975).

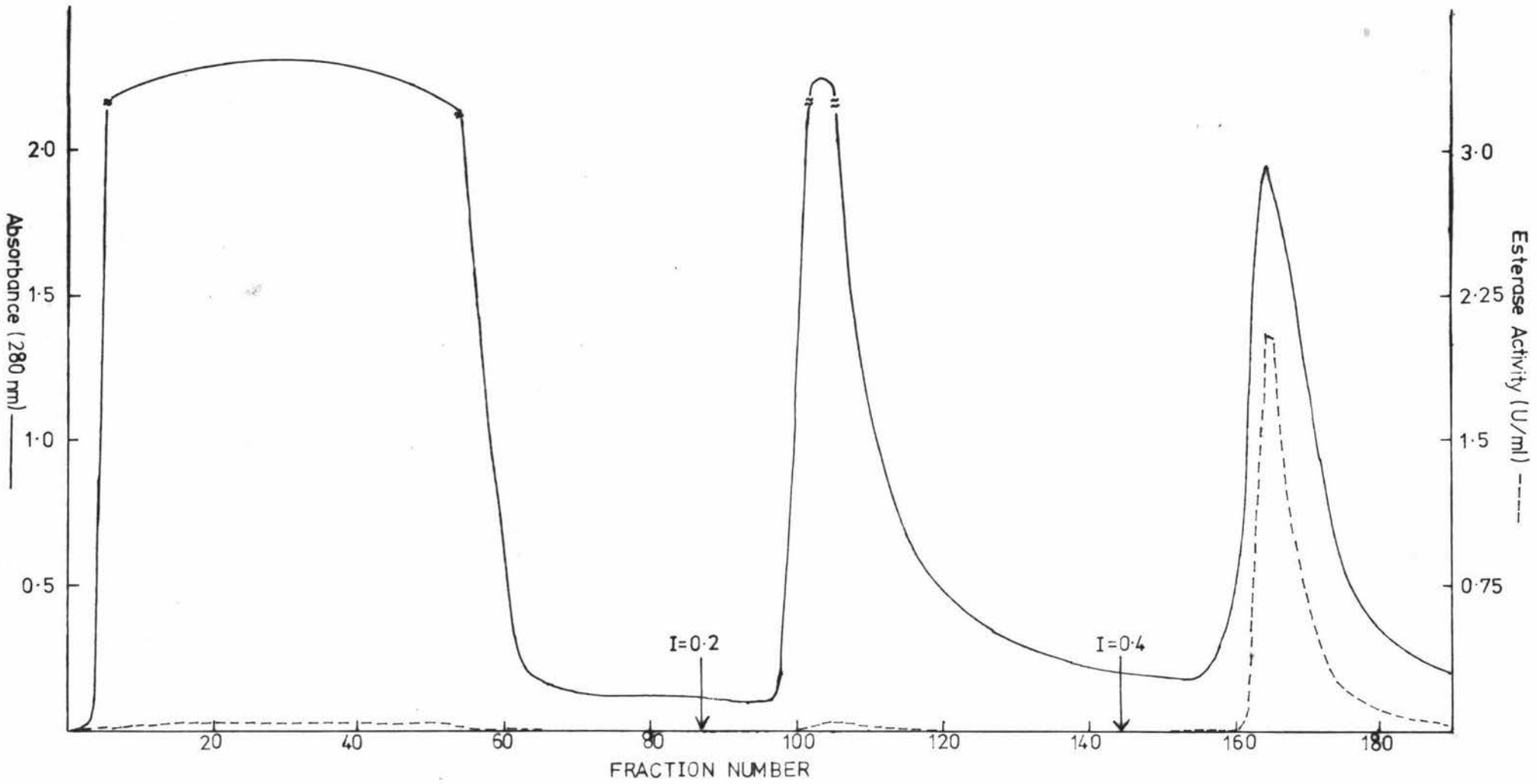


FIGURE 13 DE Viscose Chromatography of $C\bar{1}s$ (flow rate 60 ml/hr, collecting 10 ml fractions)

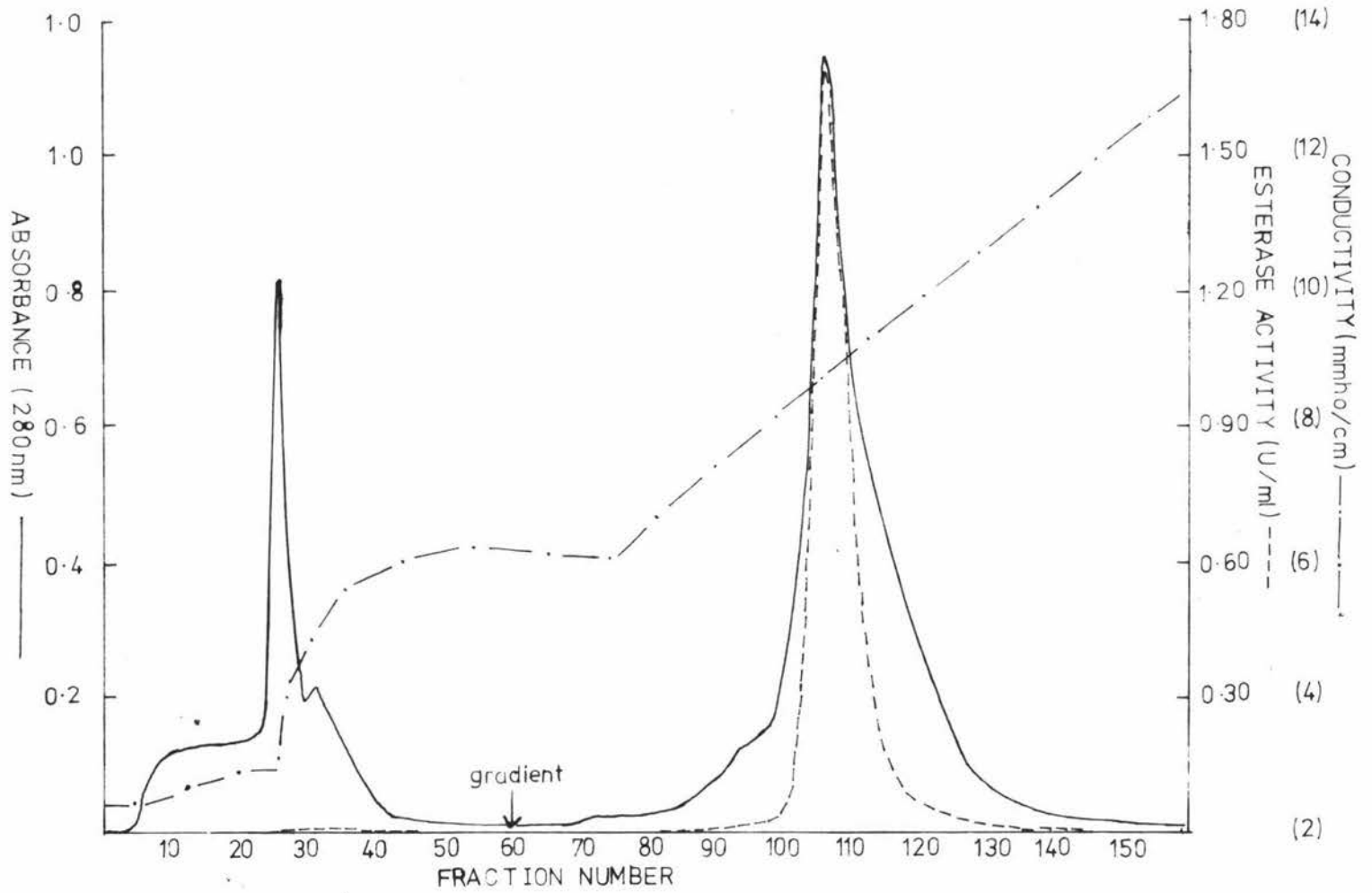


FIGURE 14 DE 52 Chromatography of $\overline{C1s}$ (flow rate 40ml/hr, collecting 8ml fractions)
The gradient consisted of 400ml I=0·2 buffer plus 400ml I=0·4 buffer.

activity. $C\overline{1}$ inhibitor remains in the pseudoglobulin fraction (Davie et al, 1976) and so allows the first esterase assays to be performed on the resuspended euglobulin precipitate (Step 2, Table 1). A yield of 50 - 80% has been obtained for the recovery of C1 in the euglobulin precipitate using haemolytic assay procedures (Gigli et al, 1976). The $C\overline{1s}$ recovery at this stage was unable to be determined using esterase assays and so the amount obtained was arbitrarily set at 100%. The precipitate which formed on prolonged overnight dialysis prior to DE viscose chromatography appeared to contain $C\overline{1s}$ as indicated by a slight decrease in specific activity (Step 2, Table 1).

A 73% recovery of esterase activity accompanied by an eleven-fold increase in specific activity was obtained for DE viscose ion exchange chromatography. Greater than 50% losses of esterase activity were consistently experienced with a two-fold increase in specific activity for DE 52 chromatography (Step 4, Table 1). The removal of a contaminant esterase also possessing cbz-Lys-ONp esterase activity cannot be discounted. $C\overline{1s}$ was the last protein peak to be eluted from DE 52 using the 0.2 - 0.4M sodium chloride gradient, but however a small protein peak containing no cbz-Lys-ONp esterase activity was eluted using a 1M sodium chloride column wash. In both ion exchange chromatographic procedures no other esterase-active peaks were obtained except for the single $C\overline{1s}$ - containing peak. $C\overline{1s}$ was eluted at I=0.25 (8-9 mmho/cm) on the salt gradient.

A final yield of 9 mg $C\overline{1s}$ per litre of serum was obtained. This compares favourably with $C\overline{1s}$ yields of 5-10 mg/litre of serum obtained by workers using variations of the euglobulin precipitation - ion exchange isolation procedure (Sim et al, 1977) (Gigli et al, 1976) (Barkas et al, 1973). The specific activity of purified $C\overline{1s}$ was 1.85 μmoles/min/mg. A value of 2.7 - 3.1 μmoles/min/mg was obtained by Sim and coworkers for $C\overline{1s}$ using the same substrate but under different assay conditions (Sim et al, 1977). Sim and coworkers performed their $C\overline{1s}$ assays at pH 6.0 with a cbz-Lys-ONp concentration of 0.14 mM at 25°C to obtain a Km for this substrate of 0.8 mM

(Sim et al, 1977). The $\overline{C1s}$ assays were performed in this project at pH 6.0 but at only 0.09 mM substrate concentration and at 19 - 20°C. Both of these factors, but mainly temperature, may account for any slight discrepancy in specific activity of $\overline{C1s}$.

The $\overline{C1s}$ obtained was homogeneous as determined by SDS-7.5% polyacrylamide gel electrophoresis (PAGE) (plate 2C) which showed a single protein-staining band having electrophoretic mobility with respect to that of bromophenol blue dye marker of 0.40. Ouchterlony immunodiffusion using anti-$\overline{C1s}$ against purified $\overline{C1s}$ produced a single precipitation line indicating that only one immunological reaction has taken place. (Plate 1).

(ii) Subcomponent C1s Purification

A preparation of unactivated C1s from 2 litres human serum yielded C1s inactive in the cbz-Lys-ONp esterase assay. The entire protein peak eluted from DE viscose with 0.3M sodium chloride (I=0.4) was pooled, dialysed against starting buffer and applied to the equilibrated DE 52 column to give a protein elution profile similar to that as for $\overline{C1s}$ purification (See Figure 14). When the peak height fraction eluted from DE 52 with the salt gradient was assayed using cbz-Lys-ONp an hydrolysis rate of 0.0015 Units/ml was obtained. Incubation of this fraction for four hours with crude $\overline{C1r}$ before assaying resulted in a cbz-Lys-ONp hydrolysis rate of 0.4960 Units/ml. Assuming complete activation to $\overline{C1s}$ the original C1s is therefore 0.3% activated prior to incubation with $\overline{C1r}$. SDS-7.5% PAGE of C1s gave essentially a single protein-staining band.

Unactivated C1s has previously been successfully isolated by the repeated addition of the serine protease and esterase inhibitor, diisopropylfluorophosphate (dip-F) throughout the isolation procedure. (Gigli et al, 1976) (Arlaud et al, 1977a). As the use of this highly toxic inhibitor is actively discouraged for safety reasons the less harmful serine protease inhibitor phenylmethanesulfonylfluoride

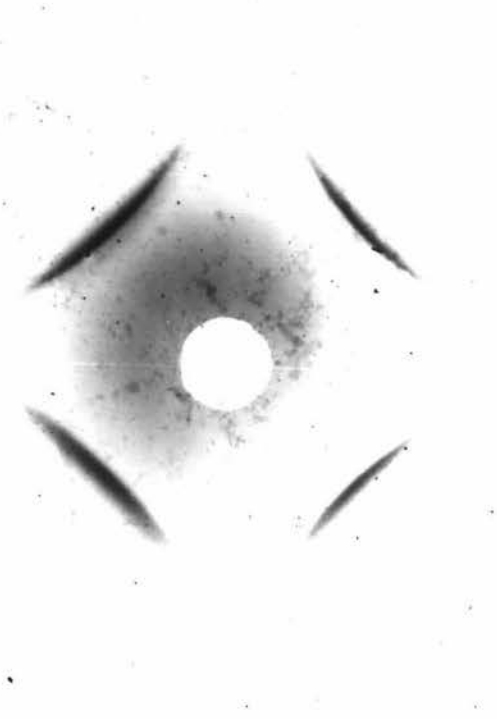


PLATE 1

Immunodiffusion of Purified $C\overline{1s}$

Inner well = anti-$C\overline{1s}$

Outer wells = $C\overline{1s}$

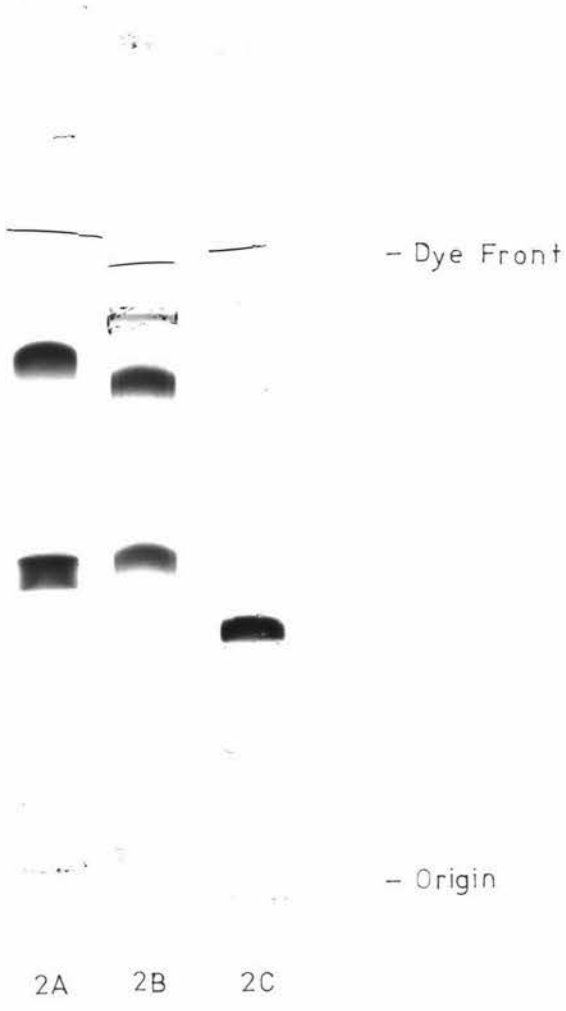


PLATE 2. SDS-Polyacrylamide Gel Electrophoresis of: Reduced and carboxymethylated $C1\bar{s}$ (2A), Reduced $C1\bar{s}$ (2B), $C1\bar{s}$ (2C)

(PMSF) was used to similarly prevent spontaneous activation of C1s by contaminant proteases during isolation. PMSF was equally as effective as dip-F in preventing C1s activation The dip-F final concentration used was 5 mM (Gigli et al, 1976; Arlaud et al, 1977a) and although PMSF at a final concentration of 5 mM was successful in preventing C1s activation, concentrations of 2.5 mM and 1.25 mM were equally as effective. A more specific trypsin-like inhibitor, benzamidine, has also been used successfully. (Campbell et al, 1979). At 4°C and pH 7.4, the temperature and pH used for C1s isolation, the lifetime of PMSF inhibiting ability has been shown to exceed 30 hours. (James, 1978). A single addition of PMSF at a final concentration of 1.25 mM at the C1 solubilization and dissociation step is therefore likely to be able to prevent C1s activation during and after this stage. $\overline{C1}$ inhibitor remaining in the pseudoglobulin would be expected to cause inactivation of C1s-activating proteases during euglobulin precipitation. Strictly performing each stage of the C1s purification procedure at low temperature has been routinely applied (Arlaud et al, 1977a) (Gigli et al, 1976) and would be expected to contribute to zymogen stability.

The specific activity of C1s, activated by incubation with $\overline{C1r}$, was consistently lower than that of $\overline{C1s}$ purified in the absence of PMSF even though both preparations were essentially homogeneous as determined by SDS-polyacrylamide gel electrophoresis. (The Em for C1s was 0.389.) This phenomenon has been described when using dip-F as inhibitor. (Gigli et al, 1976). The reaction of zymogens with irreversible inhibitors (dip-F or PMSF) affects the zymogen's ability to function following activation (Kerr et al, 1978). This is thought to be due to inhibitor reaction with the active site serine of the zymogen thus producing inactive protease or esterase even after normal activation. (Andrews and Baillie, 1979).

(iii) Activation of C1s by $\overline{C1r}$

C1s (0.67 mg/ml) having esterase activity of 0.004 U/ml was incubated with a crude preparation of $\overline{C1r}$ and the esterase

activity generated after 10 minutes and 20 minutes was determined. (Table 2). Correction has been made for autohydrolysis of the ester substrate. Appropriate controls of C1s - and $\overline{C1r}$ - alone showed little esterase activity was produced. $\overline{C1r}$ incubated alone did not yield esterase activity while C1s alone generated only 2% of the esterase activity of C1s incubated with $\overline{C1r}$. As C1s does not self-activate (Gigli et al, 1976), trace contamination by $\overline{C1r}$ would explain the slow generation of esterase activity when C1s - alone was incubated at 37°C. Dodds et al (1978) also found a gradual increase in $\overline{C1s}$ activity on incubation of C1s alone. Up to 1% cross-contamination of C1s by $\overline{C1r}$ has been shown to occur in a preparation of C1s isolated using a similar procedure to that used in these studies. (Arlaud et al, 1977b). Affinity chromatography of C1s using an anti-$\overline{C1r}$ immunoglobulin-Sepharose column was required to completely remove $\overline{C1r}$ to give stable non-selfactivating C1s preparations. When C1s alone was incubated at 37°C for several hours 0.1635 U/ml esterase activity could be generated. However the addition of $\overline{C1r}$ greatly accelerated the rate of generation of $\overline{C1s}$ activity. C1s purified in the presence of PMSF therefore yields C1s which is able to undergo $\overline{C1r}$-catalysed activation to the active form, $\overline{C1s}$.

(iv) Reduction and Carboxymethylation of $\overline{C1s}$

SDS-7.5% PAGE of desalted reduced and carboxymethylated $\overline{C1s}$ (RCM $\overline{C1s}$) gave two protein-staining bands of mobility with respect to bromophenol blue of 0.528 and 0.859 . (Plate 2A). Reduction of $\overline{C1s}$ with 2-mercaptoethanol prior to SDS-7.5% PAGE resulted in two protein bands of mobility 0.525 and 0.842 (Plate 2B). A comparison of polyacrylamide gels following SDS-PAGE of reduced $\overline{C1s}$ and RCM $\overline{C1s}$ was used to determine the extent of disulphide bond reduction before attempts at chain separation were performed (compare Plate 2B with Plate 2A). When complete reduction was achieved the conversion of the single-chain $\overline{C1s}$ protein (Em=0.400) to the constituent a (Em=0.528) and b (Em=0.859) chains could be seen (compare Plate 2C with Plate 2B).

Table 2. Activation of C1s by Incubation with $\overline{C1r}$

Sample	$\overline{C1s}$ activity after 10 min. (U/ml)	$\overline{C1s}$ activity after 20 min. (U/ml)	$\overline{C1s}$ activity generated/min.
C1s + buffer	0.0306	0.0324	0.0002
buffer + $\overline{C1r}$	0.0018	0.0072	0.0005
C1s + $\overline{C1r}$	0.0720	0.1494	0.0077
C1s + $\overline{C1r}_2$	0.0684	0.1422	0.0074

C1s or buffer (0.5 ml) was incubated at 37°C with 0.1 ml of $\overline{C1r}$ or buffer for 10 min. at which time 0.2 ml was assayed for cbz-Lys-ONp esterase activity then incubated for a further 10 min. and the esterase activity of a 0.2 ml aliquot was again determined. An activity of 1 U/ml is equivalent to 1 μmole substrate hydrolysed/min/ml enzyme.

"Buffer" refers to phosphate starting buffer; 40 mM phosphate buffer, pH 7.4 containing 1 mM EDTA. C1s and $\overline{C1r}$ were each dissolved in the above buffer.

$\overline{C1r}$ and $\overline{C1r}_2$ refer to two different preparations of the $\overline{C1r}$-containing peak eluted from DE viscose.

Initial problems of incomplete reduction of $\overline{C1s}$ were due to the small two-fold molar excess of DTT over protein disulphide bonds used to bring about disulphide bond reduction. The 50 mM DTT concentration finally employed in these studies represents a 70-fold molar excess of DTT over protein disulphide bonds using a value of 20 moles cysteine per mole $\overline{C1s}$ (Sim et al, 1977) and a protein concentration of 0.5%. This higher DTT concentration is similar to that used by Sim and coworkers to reduce human $\overline{C1s}$ disulphide bonds (Sim et al, 1977) while an even greater 100 mM DTT concentration with only 0.02% protein concentration was used to reduce factor D of the alternative complement pathway. (Volanakis et al, 1977). Similar harsh reduction methods have been used previously (Gall et al, 1968). Thorough protein denaturation, sometimes even requiring heating, appears to be essential to allow access of reducing agent to effect disulphide bond reduction. The variation in susceptibility of protein disulphide bond reduction is well illustrated in the selective reduction of immunoglobulin interchain disulphides in the absence of denaturant leaving intrachain bonds intact. Harsher reduction methods result in complete reduction. (Konigsberg, 1972).

3.2 Chain Separation

(i) Gel Filtration

All attempts to routinely separate the RCM $\overline{C1s}$ a and b chains using gel filtration techniques were unsuccessful. Gel filtration of RCM $\overline{C1s}$ in the absence of denaturant produced a single protein peak, which, when subjected to SDS-7.5% PAGE contained a mixture of a and b chains. This occurred using Sephadex G100, Sephadex G150 and Sephacryl S200. Gel filtration in the presence of denaturant using Sephadex G100 and Sephadex G200 resulted in partial separation of the a and b chains which were eluted in the expected order but were eluted very close together. SDS-7.5% PAGE of individual fractions showed that that cross-contamination of each chain had occurred and unless many fractions were discarded would not give homogeneous a and b chains.

The elution of the a and b chains as a single peak using gel filtration can be due to three factors:

1. Incomplete Reduction of $\overline{C1s}$ will result in the a and b chains still being covalently linked by disulphide bonds. This problem can be overcome by using the harsher reduction and carboxymethylation procedures described as well as by determining the extent of reduction by performing SDS-PAGE on RCM$\overline{C1s}$ before chain separation is commenced.

2. Noncovalent interactions may cause the a and b chains to adhere. These interactions can be decreased by including a denaturant in the separation buffer. Guanidine hydrochloride (6M) has been used for chain separation by gel filtration on Sephadex G200 (Sim and Porter, 1976; Sim *et al*, 1977). Chain separation by gel filtration using urea or guanidine hydrochloride was however found to be unsuccessful as a routine chain separation technique. Other groups have also found this and some have resorted to preparative gel electrophoresis to separate the a and b chains (Personal Communications). Takahashi *et al* (1975a) used analytical scale PAGE in the presence of 1% SDS to obtain homogeneous

a and b chains for N-terminal sequence determination of each chain. Other functionally related serum serine proteases also require denaturants to bring about chain separation. The two chains of bovine activated Hageman factor (Factor XIIa) of the blood coagulation system can be separated using Sephadex G100 but requires a buffer system of 7% formic acid containing 3M urea (Fujikawa et al, 1977).

3. The selection of a filtration gel having a distribution coefficient incompatible with the proteins to be separated will result in poor separation. The selection of the most suitable gel for a given protein separation is based on the linear distribution coefficient range of that gel. At either end of this range poor separation will be achieved if both proteins are of similar molecular size. As the larger of the two chains of RCM $\overline{C1s}$ is required in a homogeneous state it is undesirable to have this chain eluted in the column void volume as contamination by other larger proteins such as unreduced $\overline{C1s}$ will result.

(ii) Ion Exchange Chromatography

As chain separation by gel filtration was unsuccessful a recent report by Campbell et al (1979) who separated bovine RCM $\overline{C1s}$ a and b chains using DEAE-Sepharose CL-6B in the presence of 8M urea was studied with great interest. This method, with a modification of the ion exchange resin to DEAE-Sephadex A-50, was routinely applied to the separation of human RCM $\overline{C1s}$ a and b chains. The elution profile obtained was the same as for bovine RCM $\overline{C1s}$ chromatography on DEAE-Sepharose CL-6B; the b chain was not bound while the a chain was eluted with approximately 0.25M NaCl, (Campbell et al, 1979). The protein eluted in the unbound peak using DEAE-Sephadex A-50 had an electrophoretic mobility on SDS-7.5% PAGE of 0.830 (Plate 3A) and the bound protein eluted with 0.2M NaCl had an electrophoretic mobility of 0.55 (Plate 3B). These mobilities compare favourably with those found for RCM $\overline{C1s}$ b and a chains of 0.859 and 0.528 respectively and for 2-mercaptoethanol reduced $\overline{C1s}$ of 0.842 and 0.525 respectively for the b and a chains. Fourteen mg (250 nmoles)

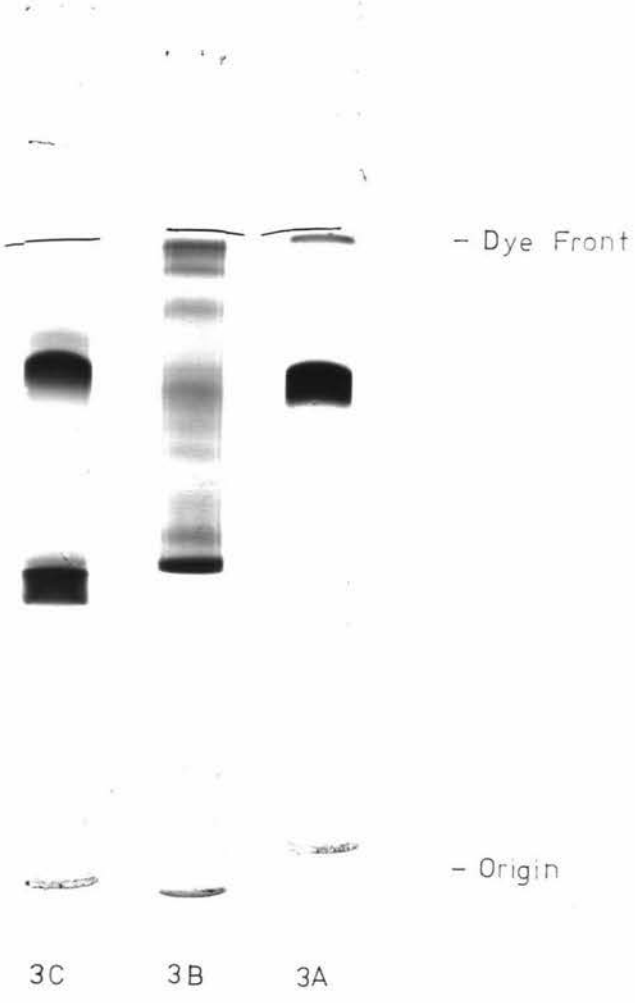


PLATE 3. SDS-Polyacrylamide Gel Electrophoresis of: Reduced and carboxymethylated $C1\bar{s}$ (3C), a chain (3B), b chain (3A)

lyophilized and salt-free RCM $\overline{C1s}$ a chain was obtained from 40 mg $\overline{C1s}$ representing a 52% yield for the reduction and carboxymethylation, ion exchange chain separation and desalting steps. This compares with a bovine RCM $\overline{C1s}$ a chain yield of 60-80% for the ion exchange chain separation step alone (Campbell et al, 1979). Fractions 53-60 (Figure 15) were individually subjected to SDS-7.5% PAGE as a protein contaminant appeared to be eluted slightly before RCM $\overline{C1s}$ a chain. No contaminant was evident in any of the SDS-7.5% polyacrylamide gels of these fractions.

It is interesting to note that even for chain separation by ion exchange methods a denaturant appeared to be required and as well it was thought necessary to introduce a slight charge effect by including 5 mM NaCl in the buffer medium.

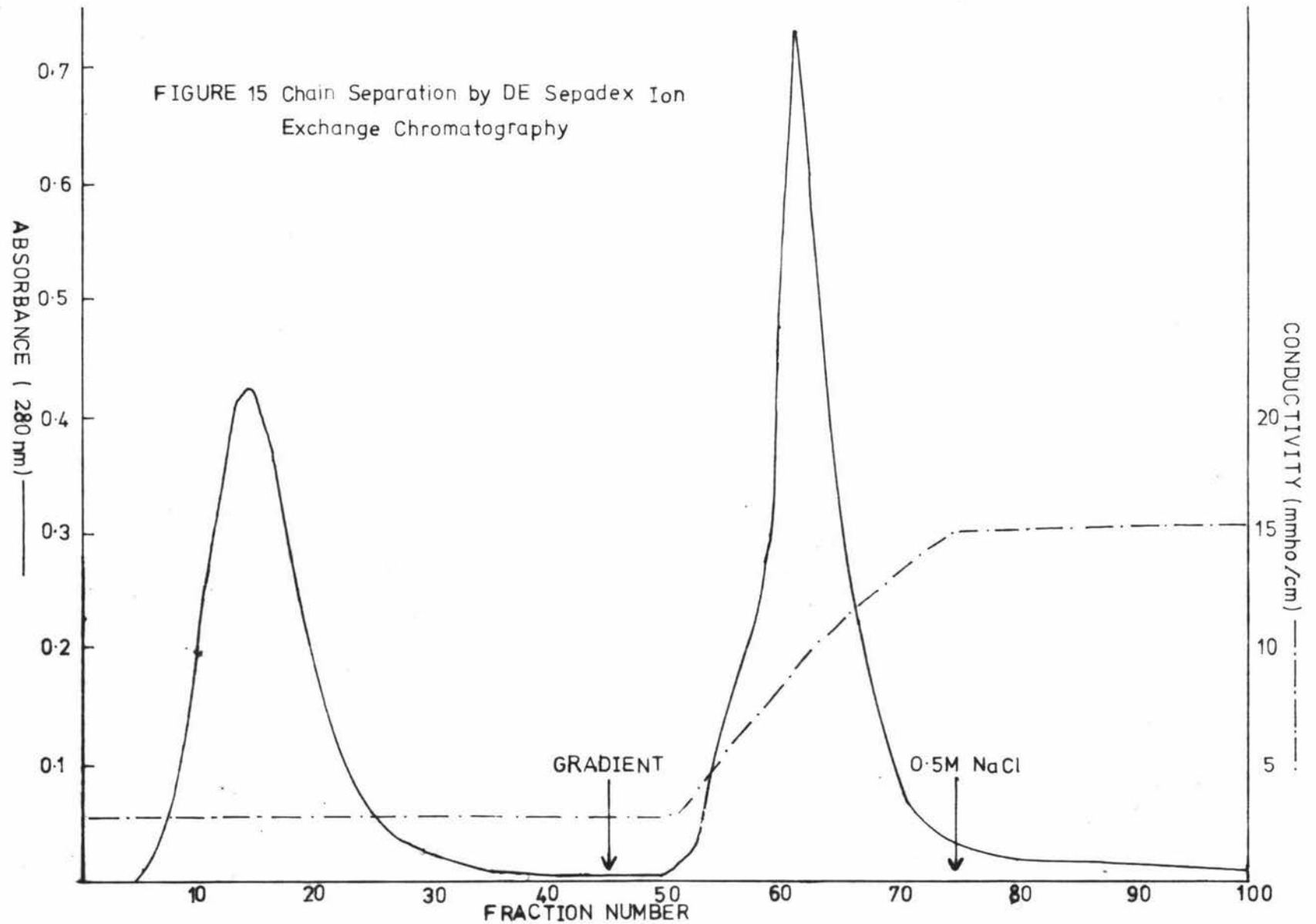


FIGURE 15 Chain Separation by DE Sepadex Ion Exchange Chromatography

3.3 Enzyme Digestions

(i) Carboxypeptidase B Digestion of RCM $\overline{C1s}$

RCM $\overline{C1s}$ (28 mg), twice desalted and lyophilized, was digested by carboxypeptidase B using an enzyme to substrate ratio of 1:60 ($^{w}/_{w}$). Five mg RCM $\overline{C1s}$ was extracted from the digest mixture 0,5,12 and 25 minutes after CPB addition. Each time course sample represents 60 nmoles of RCM a and 60 nmoles of RCM b chains. CPB acted very rapidly to release a quantitative amount of arginine probably well before the first sampling time of 5 minutes as this sample contained 54 nmoles Arginine and remained constant at 57 nmoles for the 12 and 25 minute sampling times (Figure 16). A substrate blank containing 5 mg RCM $\overline{C1s}$ alone was incubated for 25 minutes and resulted in no release of free amino acids. Both the 0 time sample and CPB-alone blank also contained very low levels of free amino acids.

Chymotryptic contamination and CPA-like activity would account for the very high yields of Tyr and Phe and lesser yields of Val and Leu. Such contamination by chymotryptic and possible inherent or trace CPA activity has been known for some time (Folk, 1970; Potts, 1967). The gradual release of most amino acids, excluding Glu, Pro, Gly, CMCys, Asp, Ile, and His suggests both endo- and exo-peptidase activity contamination of this CPB preparation. Inhibitor-treated CPB shows a very restrictive substrate specificity and releases both Arg and Lys from the C-terminus of proteins (Folk, 1970).

The CPB protein substrate consisted of a mixture of RCM $\overline{C1s}$ a and b chains in a 1:1 ratio. While the a chain is predicted to terminate in an Arg or Lys residue on the basis of $\overline{C1r}$ ester substrate specificity (Takahashi *et al*, 1975a) the C-terminal residue of each chain is unknown. Therefore because of the impure CPB and the presence of two protein substrates in the reaction mixture Arg is proposed as the likely C-terminal residue of the RCM $\overline{C1s}$ a chain and

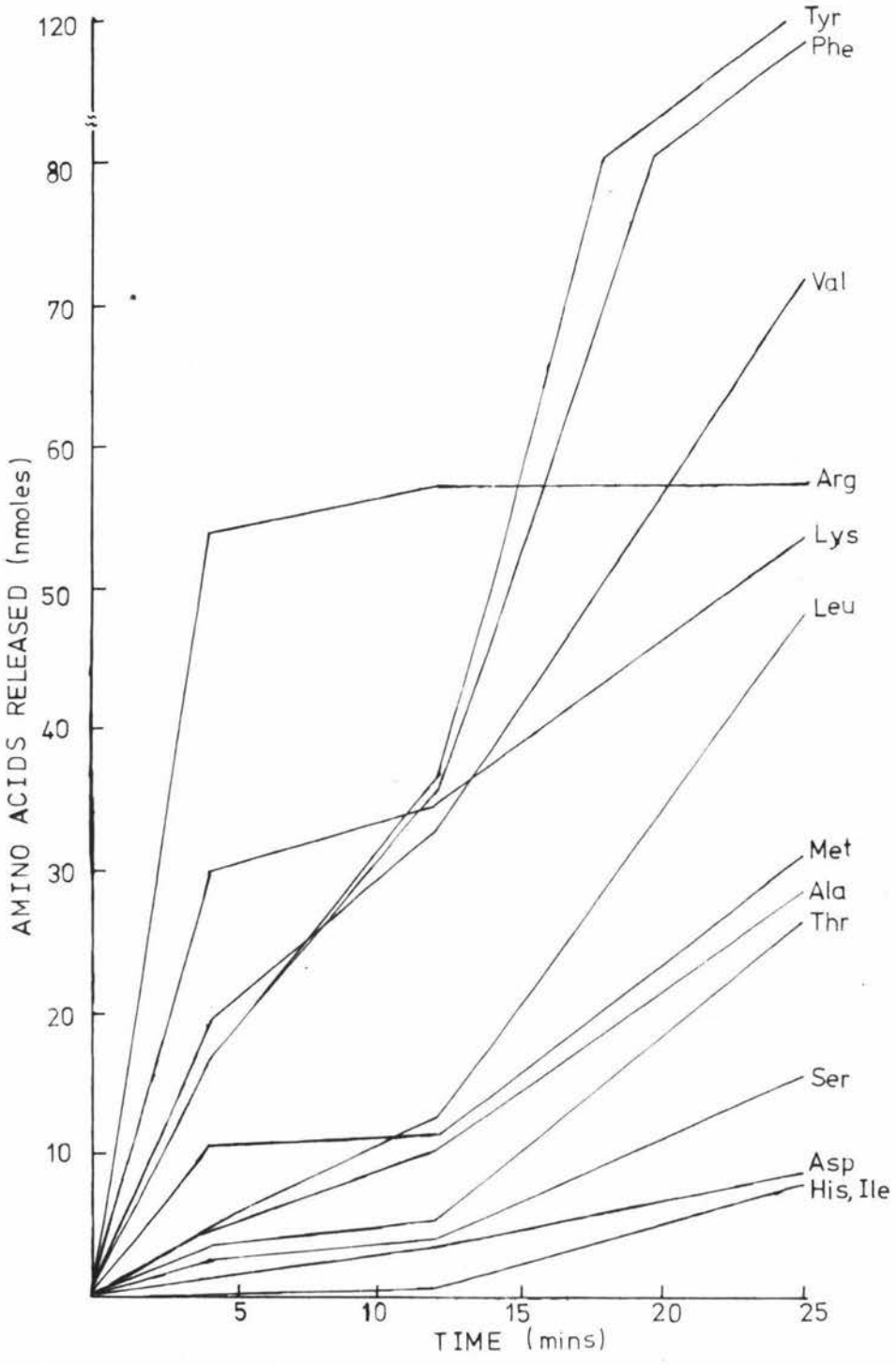


FIGURE 16 Carboxypeptidase B Digestion of $\overline{C1s}$

no other residues can be even tentatively placed.

3.4 Peptide Mapping

(i) Peptide Mapping of Bovine Pancreatic Ribonuclease A

When the endopeptidase-derived peptides of a particular protein are separated by two dimensional electrophoresis and chromatography they form a pattern or "fingerprint" characteristic of that protein. Thus when a protein is modified in some way at a single specific point, for example by carboxypeptidase digestion, a comparison of the peptide maps of the unmodified and modified protein will show which peptide has undergone modification and hence has altered electrophoretic and chromatographic mobilities. Elution of the original peptide from the peptide map of the unmodified protein will give that part of the protein at which the single point modification has occurred; the C-terminal derived peptide in the case of carboxypeptidase digestion. Modification by carboxypeptidase digestion also allows the amino acids cleaved from the C-terminus of the protein to be analysed and these should compare with those present in the isolated C-terminal peptide.

Before this technique was applied to $\overline{C1s}$ a preliminary experiment was conducted using as substrate the much smaller bovine pancreatic ribonuclease A (molecular weight 14,700 daltons). Reduced and carboxymethylated RNase (8 mg) was digested by carboxypeptidase Y for 6 hours and then subjected to a chymotryptic digestion while another 8 mg of RCM-RNase was subjected to a chymotryptic digestion alone. Approximately 30 ninhydrin-staining spots were obtained on the chymotryptic map of each RNase sample.

The map of acidic chymotryptic peptides of unmodified RNase contained an additional red-staining peptide when compared with the acidic chymotryptic peptides of carboxypeptidase-modified RNase. This peptide had an electrophoretic mobility of 0.44 with respect to aspartic acid in the first dimension and a mobility of 0.51 with respect to DNS-Arg in the second chromatography dimension. Amino acid analysis of this peptide following elution from the map as well as

identification of the free amino acids liberated by the 6 hour CPY digestion are shown in Table 3. The amino acid analysis of the eluted peptide gave only three amino acids in approximately equivalent yields of 15 nmoles; these being serine, alanine and valine. It therefore appears that a chymotryptic cleavage at Phe-120 of RCM-RNase has occurred. The low yield of Asp-121 would be due to ninhydrin staining of the peptide and subsequent loss during acid hydrolysis. Ranking of amino acids released by CPY in decreasing yields gives the expected amino acid sequence of RNase of NH_2-Phe-Asp-Ala-Ser-Val-COOH (Smyth et al, 1963) which compares favourably with the composition of the isolated C-terminal peptide of Val,Ser,Ala and (Asp). Table 3. The rapid release of C-terminal aspartic acid by CPY at pH 5.5 and the presence of three valine residues near to the C-terminus of RNase (Val-124, Val-118, Val-116) contributed to the difficulty in interpretation of the results obtained. This tetrapeptide, (molecular weight 390 daltons) having a negative charge of -1 and Em -0.44 at pH 6.5 compares well with a required molecular weight by Offord mobility plot (Offord, 1977) of 440 daltons. This is evidence for the involvement of aspartic acid in an otherwise uncharged peptide at pH 6.5.

Modifications to the Peptide Mapping Technique as a result of Ribonuclease Peptide Mapping.

From the results of the RNase peptide mapping experiment it was decided to slightly modify the procedure to obtain more definitive maps for the much larger $C\overline{1s}$. The use of only very specific enzymes, both the carboxypeptidase and endopeptidase, would result in more definitive "fingerprints" of proteins. The use of CPB to remove only the postulated basic residue from the C-terminus of the $C\overline{1s}$ a chain would give rise to a single specific modification of the C-terminal derived peptide. The loss of a positively-charged residue would be expected to significantly alter the mobility of such a peptide. DCC- or TPCK-treated trypsin or highly purified preparations of chymotrypsin in the presence of a tryptic inhibitor would also reduce the probability of

Table 3. A Comparison of Residues Released by Carboxypeptidase Y and Composition of Carboxyl Terminal Peptide with the Ribonuclease Sequence

Ribonuclease Sequence[a]	Residue	Yield of Residues[b] Released by CPY	Composition of C-terminal Peptide[c]
Val-124	Val	1.77	17.4
Ser-123	Ser	0.85	15.2
Ala-122	Ala	0.76	12.7
Asp-121	Asp	0.71	1.2
Phe-120	Phe	0.63	0.0
His-119	Pro	0.57	0.9
Val-118	His	0.48	0.0
Pro-117	Tyr	0.45	0.0
Val-116	Thr	0.33	0.1
Tyr-115	Leu	0.22	0.1
	Ile	0.18	0.2
	Met	0.15	0.0
	Glu	0.14	0.6
	CMCys	0.13	0.0
	Gly	0.12	0.0
	Lys	0.04	0.0
	Arg	0.04	0.0

Note a. from Smyth et al (1963) Val-124 is the Carboxyl Terminal residue of Ribonuclease

b. moles amino acid released per mole Ribonuclease

c. nmoles of amino acids in eluted peptide, blank values have been subtracted.

secondary fragmentation of peptides. Repeated staining of the peptide maps with very dilute ninhydrin solutions (0.025%) by spraying on one side of the paper only would minimize losses of the N-terminal residue of peptides (Bennett, 1967).

(ii) Peptide Mapping of $\overline{C1s}$

RCM $\overline{C1s}$ a chain (3.6 mg/2 ml 0.1M ammonium bicarbonate, pH 8.0) was digested by CPB for 16 hours while a second a chain sample was incubated alone as a substrate control. Determination of the amino acids released by CPB was inconclusive due to high levels of free amino acids in the a chain control and also to significant chymotryptic and CPA contamination of CPB. Very high levels of free Tyr and Phe and to a lesser extent of Leu and Val were generated in the digestion mixture indicative of chymotryptic contamination of CPB. As well, much Lys and Arg was also produced suggesting that tryptic cleavage of the a chain also took place. As CPB is thought to contain inherent or trace-contaminant CPA activity (Folk, 1970) these amino acids would all be released following endopeptidase cleavage of the $\overline{C1s}$ a chain. Presumably the prolonged 16 hour digestion by CPB caused these lesser side reactions of CPB contaminants to become significant. Lys, Arg, Val, Leu, Tyr and Phe were all released at or greater than three times the theoretical yield expected while His, Asp, Thr, Ser, Ala and Met were released in quantitative mole/mole yields. Glu, Pro, Gly, CMCys and Ile were not released at all

Following CPB digestion both samples of $\overline{C1s}$ a chain were digested with trypsin and a tryptic map of each sample was performed. Approximately 48 peptides were predicted to be obtained based on a value of 4.4 Lys and 4.2 Arg residues/100 residues and an a chain molecular weight of 56,000 daltons (Sim et al, 1977). Only 22 tryptic peptides were identified on the map of unmodified a chain while 17 peptides were identified on the map of CPB - modified a chain. Table 4.

Electrophoresis of the tryptic peptides of unmodified a chain appeared to be affected by the presence of salt with the result that peptide mobilities of modified and unmodified a chain in the first dimension do not compare well. The general patterns of the respective maps were however similar. An additional acidic peptide having an Em-0.50 and chromatographic mobility 0.36 was obtained in the map of unmodified a chain (indicated by * in Table 4).

Less than half of the expected number of tryptic peptides were visualized in the maps of RCM $C\overline{1s}$ a chain therefore although a possible C-terminal peptide had been isolated even if it had been eluted and sequenced its assignment as the C-terminal derived peptide of the a chain could not be proven. The general peptide mapping procedure is applicable to C-terminal peptide isolation from substrates of reasonable size only. Protein substrates of 14,000 daltons (Canfield, 1963) and 30,000 daltons (Carlton and Yanofsky, 1963) have been successfully used. The $C\overline{1s}$ a chain however is much larger with the result that peptide maps are very difficult to clearly interpret making this an unsuitable method for C-terminal peptide isolation from the $C\overline{1s}$ a chain.

Table 4. Mobilities of Tryptic Peptides of $\overline{C1s}$ a chain and CPB-modified $\overline{C1s}$ a chain

	CPB-modified a chain		a chain	
	Mobility (x,y)[a]	colour [b]	Mobility (x,y)[a]	colour [b]
Neutral Peptides	0.60, 0.83	p	0.48, 0.85	p
	0.48, 0.76	p	0.28, 0.85	p
	0.63, 0.61	fp	0.48, 0.62	fp
	0.43, 0.54	fp	0.33, 0.55	fp
	0.72, 0.35	fp	0.62, 0.30	p
	0.83, 0.26	o	0.70, 0.22	o
	0.49, 0.31	fp	-	
	0.53, 0.24	p	0.48, 0.22	p
	0.44, 0.24	fp	0.30, 0.24	p
	-		0.38, 0.40	fp
	-		0.85, 0.14	fp
Acidic Peptides	0.65, 0.20	fp	0.42, 0.30	fp
	0.86, 0.33	fp	0.85, 0.30	fp
	0.93, 0.26	fp	0.88, 0.21	fp
	-		*0.50, 0.36	o/p
	-		0.40, 0.29	fp
	-		0.27, 0.40	fp
Basic Peptides	0.51, 0.44	fp	0.21, 0.44	fp
	-		0.24, 0.54	fp
	0.50, 0.29	p	0.40, 0.41	p
	0.72, 0.22	p	0.65, 0.21	p
	0.36, 0.19	p	0.40, 0.15	o/p
	0.90, 0.17	o/p	0.80, 0.16	o/p

Note a. Mobilities are expressed as (x,y) where x is the electrophoretic mobility (with respect to Asp for acidic and basic peptides; with respect to DNS Arg for neutral peptides) in the first dimension and y is the chromatographic mobility (with respect to DNS Arg) in the second dimension. Corresponding peptides in the a chain and CPB-modified a chain peptide maps are

Table 4.

placed next to each other.

Note b. Colour of peptide spots is:
p = purple, fp = faintly staining purple,
o = orange, o/p = orange/red

* indicates a possible additional peptide not present in the tryptic map of CPB-modified $\overline{C1s}$ a chain.

3.5 Chemical Modification

(i) Protein Carboxyl-blocking

The strategy for labelling a carboxyl group produced by polypeptide chain cleavage was to involve two sequential carboxyl-blocking experiments; the first on zymogen $C\overline{1s}$, and the second using a radioactively-labelled blocking group on activated $C\overline{1s}$. The radioactive label was to provide a convenient method for following the C-terminal amino acid residue of the $C\overline{1s}$ a chain during a chain proteolytic digestion and subsequent C-terminal peptide isolation (Figure 17). A similar simplified procedure has been used previously (Carraway et al, 1968).

Unactivated C1s (15 mg) was reacted with taurine to give 9.5 mg taurine-blocked C1s. A measure of the extent of carboxyl-blocking was made by determining the amino acid composition of C1s after taurine blocking. Seventy-seven moles taurine and 83 moles aspartic acid and 112 moles glutamic acid were obtained per mole C1s representing a 39.5% carboxyl-blocking efficiency. As asparagine and glutamine are degraded during acid hydrolysis to give the respective acidic amino acids Asp and Glu the value of 39.5% carboxyl-blocking represents a minimum value only.

Incubation of taurine-blocked C1s with partially purified $C\overline{1r}$ for several hours at 37°C resulted in no active taurine-blocked $C\overline{1s}$ as determined by the cbz-Lys-ONp esterase assay. Therefore either taurine-blocked C1s is unable to undergo activation by $C\overline{1r}$ and/or the resultant taurine-blocked $C\overline{1s}$ is inactive in the cbz-Lys-ONp assay. To distinguish between these possibilities the carboxyl groups of active $C\overline{1s}$ were blocked using taurine as above.

Acid hydrolysis and subsequent amino acid analysis of the taurine-blocked $C\overline{1s}$ showed that 108 moles taurine had been coupled per mole $C\overline{1s}$, representing a minimum of 54%

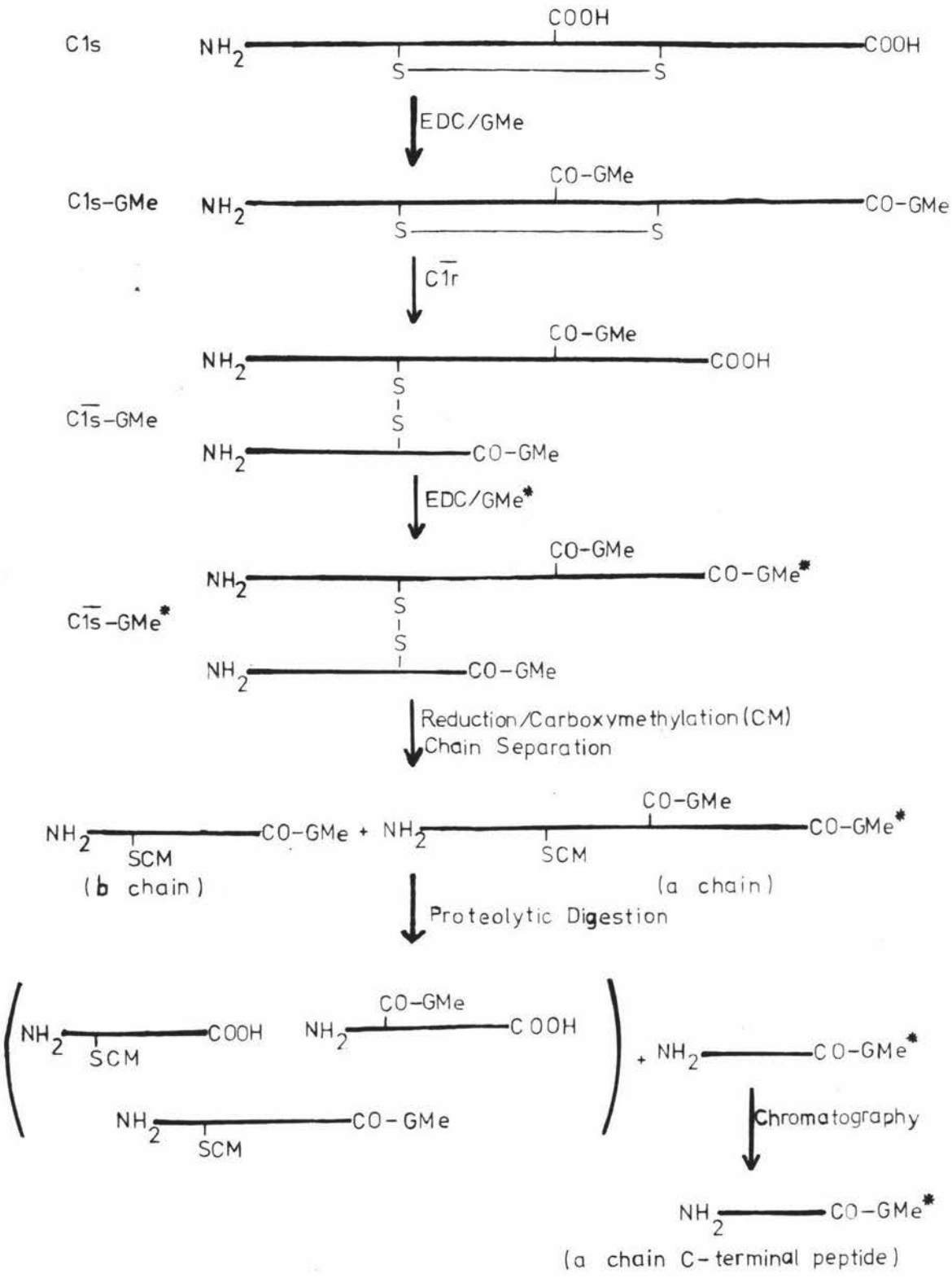


FIGURE 17 The Scheme of the Modification Procedure for the Isolation of the C-terminal Peptide of the C1s a chain

NOTE: GMe = Glycine methyl ester
GMe* = Radioactive glycine methyl ester

carboxyl-blocking. Cbz-Lys-ONp assays of $\overline{C1s}$ and taurine-blocked $\overline{C1s}$ gave hydrolysis rates of 0.687 U/ml and 0.0 U/ml respectively for solutions of similar protein concentrations. The original taurine blocked C1s, after incubation with $\overline{C1r}$ may therefore be in the activated $\overline{C1s}$ state but it is inactive to the cbz-Lys-ONp assay.

As ester hydrolysis appeared to be unsuitable for discriminating between the unactivated taurine-C1s and activated taurine-$\overline{C1s}$ states an alternative method was employed. Carboxyl-blocked C1s or $\overline{C1s}$ was reduced and subjected to SDS-PAGE to give the single and double protein-staining bands respectively. SDS-PAGE of reduced taurine-blocked C1s or $\overline{C1s}$ gave only a single diffuse protein-staining band situated at the origin of the polyacrylamide gel. This is suggestive of crosslinking between different taurine-blocked C1s or $\overline{C1s}$ molecules. Native 7.5% PAGE of taurine-blocked $\overline{C1s}$ or C1s also gave a single diffuse band at the gel origin.

Since protein amino groups are protonated at pH 4.75 and therefore not good nucleophiles carbodiimide-catalysed crosslinking reactions are performed at pH 8. This procedure has been used successfully to crosslink protein molecules or to couple proteins to cell surfaces or insoluble supports. (Carraway and Koshland, 1972). Although C1s and $\overline{C1s}$ carboxyl-blocking reactions were carried out at pH 4.75 it appears that significant protein crosslinking has occurred.

Because taurine is a seldom used blocking group a carboxyl-blocking experiment was performed using glycine methyl ester as the blocking group. $\overline{C1s}$ (7.1 mg) was reacted with glycine methyl ester to yield 6.7 mg glycine methyl ester-blocked $\overline{C1s}$ (94% yield) which like taurine-blocked $\overline{C1s}$, was inactive in the cbz-Lys-ONp assay (0.0 U/ml). Acid hydrolysis and amino acid analysis of glycine methyl ester-$\overline{C1s}$ gave 175.5 moles glycine per mole $\overline{C1s}$ as well as 83.1 moles aspartic acid and 111.4 moles glutamic acid. The glycine content of unblocked $\overline{C1s}$ was determined as

71 moles glycine per mole $\overline{C1s}$. Therefore 104.5 moles glycine methyl ester was coupled per mole $\overline{C1s}$ representing a minimum of 54% blocking of possible C1s carboxyl groups.

Carboxyl-blocking of C1s using taurine gave a minimum of 39% blocking while carboxyl-blocking of $\overline{C1s}$ using taurine or glycine methyl ester both gave a minimum of 54% blocking. Perhaps this difference reflects changes in C1s conformation on activation to $\overline{C1s}$. The amino acid analyses were performed after a single 18 hour acid digestion and tryptophan was not determined at all. It must be emphasised that such calculations as expressed on a molar basis depend on extreme precision of analytical techniques and on an accurate molecular weight value for C1s and $\overline{C1s}$. These percentages then are intended as an indication of the level of carboxyl-blocking only. In the presence of a denaturant all protein carboxyl groups can be blocked while in the absence of a denaturant only those "exposed" carboxyl groups are blocked. The difference gives the number of "buried" or internal carboxyl groups. (Carraway and Koshland, 1972). C1s and $\overline{C1s}$ blocking experiments were performed without denaturant so it is unlikely that 100% carboxyl blocking was achieved. Unfortunately all efforts to produce a $\overline{C1s}$ species having all or most of its carboxyl groups blocked (except the a chain C-terminal carboxyl group) appeared to be unsuccessful. This may have been due to the cross-linking of different protein molecules through amino and carboxyl groups of different C1s molecules as well as irreversible denaturation of C1s caused by carboxyl-blocking or crosslinking thus preventing $\overline{C1r}$-catalysed activation and/or $\overline{C1s}$ catalysis. The failure to bring about the desired peptide bond cleavage of carboxyl-blocked C1s caused the abandonment of this approach to C-terminal peptide isolation.

3.6 Affinity Chromatographic Isolation of a C-terminal Peptide

(i) Preparation of ST-Sepharose 4B

A 1.0 ml settled volume of ST-Sepharose 4B was hydrolysed with 6M HCl and a suitable aliquot ($\frac{1}{10}$) of the freeze dried material was subjected to amino acid analysis to give approximately 1.5 μmoles arginine-containing ligand per ml settled volume of gel. Table 5. ST-Sepharose has previously been prepared containing 1-5 μmoles (Kasai and Ishii, 1975) and 4-11 μmoles (Yokosawa et al, 1976) arginine-containing ligand per ml settled volume of gel. The 1.5 μmoles arginine-containing ligand per ml settled volume gel obtained would represent a minimum value only as acid digestion of the gel appeared to be incomplete. A more suitable method instead of in vacuo hydrolysis would be refluxing enabling hydrolysis to be performed in a much larger volume of acid thereby minimizing losses of gel on the tube walls. Significant losses of the N-terminal residue of peptides, through which the peptide is attached to the gel, can occur on acid hydrolysis (Gabel and Axén, 1976). Protamine sulphate (Salmine) is a highly basic protein from fish sperm and consists of a mixture of several polypeptides, each of about 30 amino acids of which approximately two-thirds are arginine residues. The products of tryptic digestion consist of many di- and tri-peptides containing only one residue of L-arginine at the carboxyl terminus of each peptide. (Kasai and Ishii, 1975).

(ii) Preparation and Purification of Anhydrotrypsin

A summary of anhydrotrypsin preparation is shown in Table 6. The elution profile of anhydrotrypsin purified by affinity chromatography on ST-Sepharose is given in Figure 18. The column-bound anhydrotrypsin peak was eluted with 5 mM HCl only after thorough washing of the ST-Sepharose column. The unbound protein peak still contained anhydrotrypsin as the specific activity of this protein peak was

Table 5. Amino Acid Analysis of ST-Sepharose 4B

Amino Acid	nmoles
arginine	1460
glycine	560
proline	190
histidine	170
serine	150
valine	140
alanine	110
aspartic acid	100

A 1.0 ml settled volume of ST-Sepharose was subjected to acid hydrolysis and an aliquot to amino acid analysis.

Table 6. Anhydrotrypsin Preparation

	Volume (ml)	Protein (mg/ml)	Total Protein (mg)	BAPNA Activity (U/ml)	Specific Activity (U/ml)	% Trypsin Activity
Trypsin	190	2.63	500[a]	223.25	1.175	100
PMS-Trypsin	191	2.17	414.6	0.007	0.003	0.25
Alkaline Trypsin	230.3	1.79	411.6	0.163	0.091	7.74
Affinity Chromatography ST-Sepharose eluate 1	10	6.60	66	0.032	0.005	0.41
ST-Sepharose eluate 2	47	1.15	54	0.008	0.007	0.59
Pooled eluate			120			
			122[a]			

Note a. Protein was determined by the weight of dry material, otherwise by the Coomassie dye binding method using trypsin as the standard protein. ST-Sepharose eluate 2 refers to the 5 mM HCl eluted peak after reapplication of the unbound material from the ST-Sepharose eluate 1 passage.

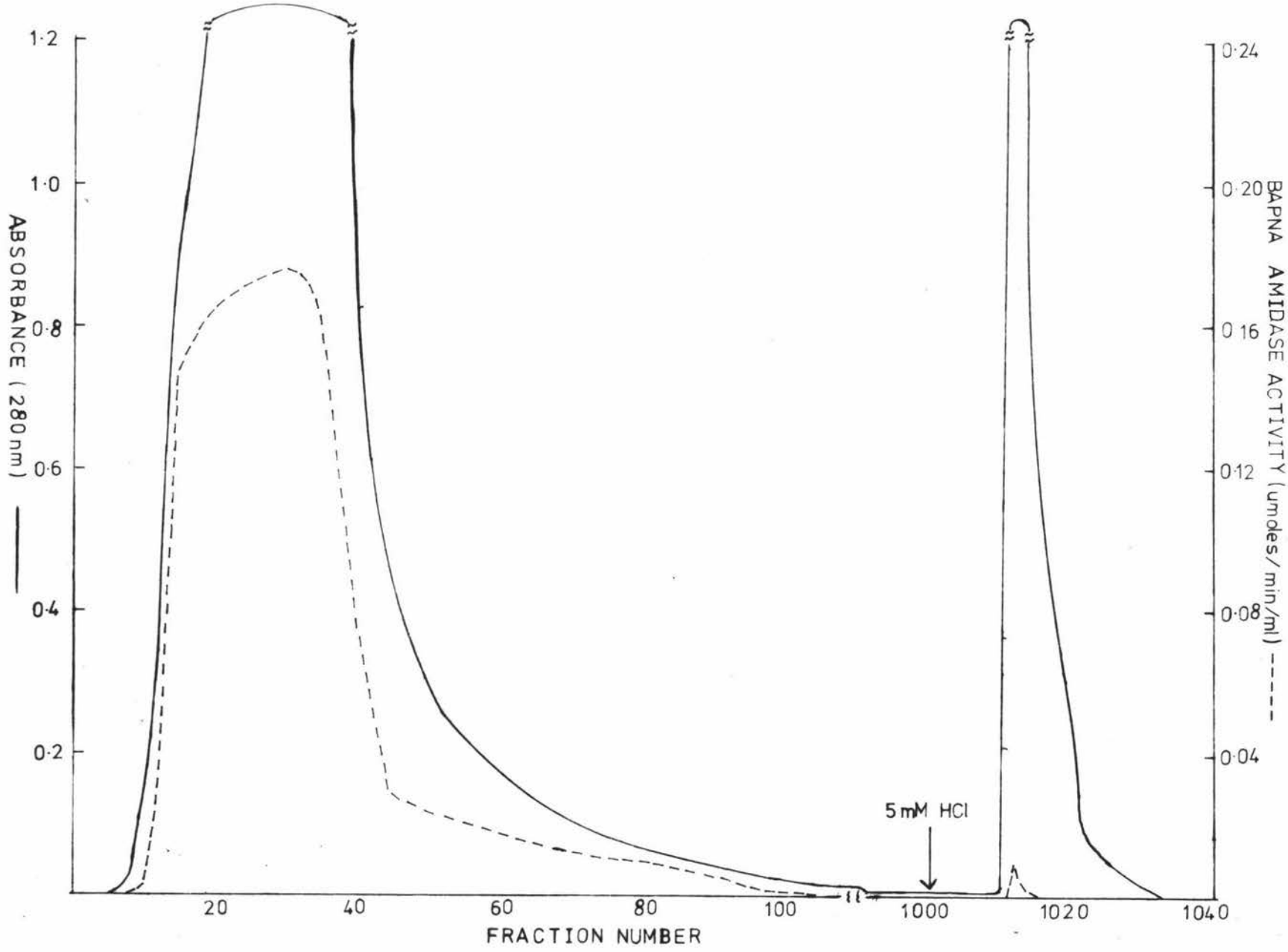


FIGURE 18 Anhydrotrysin Purification using ST-Sepharose

much lower than that of trypsin. As expected the leading edge of this peak contained most of the trypsin because trypsin has a much lower affinity for ST-Sepharose than anhydrotrypsin at pH 5.0. The entire unbound protein peak (see Figure 18) was pooled and after dialysis was reapplied to the re-equilibrated ST-Sepharose column. The column-bound anhydrotrypsin peak eluted by 5 mM HCl was pooled to give ST-Sepharose eluate 2 (see Table 6).

The combined 5 mM HCl eluates contained 120 mg anhydrotrypsin having less than 1% of the specific activity of trypsin representing a 24% yield (by weight). A yield of 24% compares favourably with a yield of approximately 20% obtained by Yokosawa and Ishii (1977). Their preparation was reported as containing less than 0.1% of the specific activity of trypsin. Recently anhydrotrypsin which was prepared in a 24% yield was less than 0.5% active when tested by a proteolytic assay (Sayers and Barrett, 1980). Although the anhydrotrypsin prepared in these studies possessed a specific activity of approximately 0.5% of that of trypsin the errors involved in determining such a low activity and also determining protein concentration would account for any slight differences.

(iii) Preparation of Anhydrotrypsin-Sepharose 4B

The ligand-binding capacity of the anhydrotrypsin-Sepharose column was determined by the application of 2.5mg benzoylglycyl arginine (7,350 nmoles) in 0.1 ml pyridine buffer, pH 5.0. The 280 nm-absorbing peak eluted by 5 mM HCl contained 905 nmoles benzoylglycyl arginine on the basis of glycine and arginine content after acid hydrolysis and amino acid analysis of the anhydrotrypsin-bound material. The anhydrotrypsin-Sepharose column could bind a theoretical maximum of 5,200 nmoles of benzoylglycyl arginine assuming complete coupling of all 120 mg anhydrotrypsin and saturated stoichiometric binding of ligand to immobilized anhydrotrypsin. A final yield of 17.4% (20.9 mg) immobilized anhydrotrypsin capable of a binding ligand was achieved.

(iv) Peptide Isolation by Affinity Chromatography

CT I

Salt-free RCM $\overline{C1s}$ (16 mg) was digested by chymotrypsin, applied to the equilibrated anhydrotrypsin-Sepharose column and eluted with 5 mM HCl. Figure 19. Fractions 112-116 were combined (CT I) and N-terminal end group determination by dansylation indicated a Glu N-terminal residue. As asparagine and glutamine are converted to aspartic acid and glutamic acid respectively on acid hydrolysis the Glu N-terminal residue is more correctly referred to as Glx. Further purification of this peptide by pH 2.1 electrophoresis was then performed as slight contamination by glycine and aspartic acid was evident. Again Glx was determined as the N-terminal residue and slight contamination by glycine was apparent. In both N-terminal dansylations a noncarboxyl terminal ε- DNSLys was also produced. The remainder of fraction CT I was subjected to acid hydrolysis and amino acid analysis. Table 7A. Fractions 31-60, 100-111, 117-120 were individually lyophilised and subjected to pH 2.1 electrophoresis to ensure that no peptides had been eluted other than in the unbound column breakthrough peak (fractions 25-35) or in the 5 mM HCl-eluted peak (fractions 112-116). No significant ninhydrin-positive spots were evident suggesting that specific absorption of a peptide has occurred.

The absence of Ile, Tyr, and Phe in the peptide amino acid composition suggests that a real C-terminal peptide has been isolated. Tryptophan was not determined. Should the isolated peptide contain one of these residues it would be expected to be the carboxyl terminal residue as chymotrypsin only hydrolyses peptide bonds whose carbonyl group is contributed by the above amino acids. The elution profile of the anhydrotrypsin-Sepharose column shows that the 5 mM HCl-eluted peak

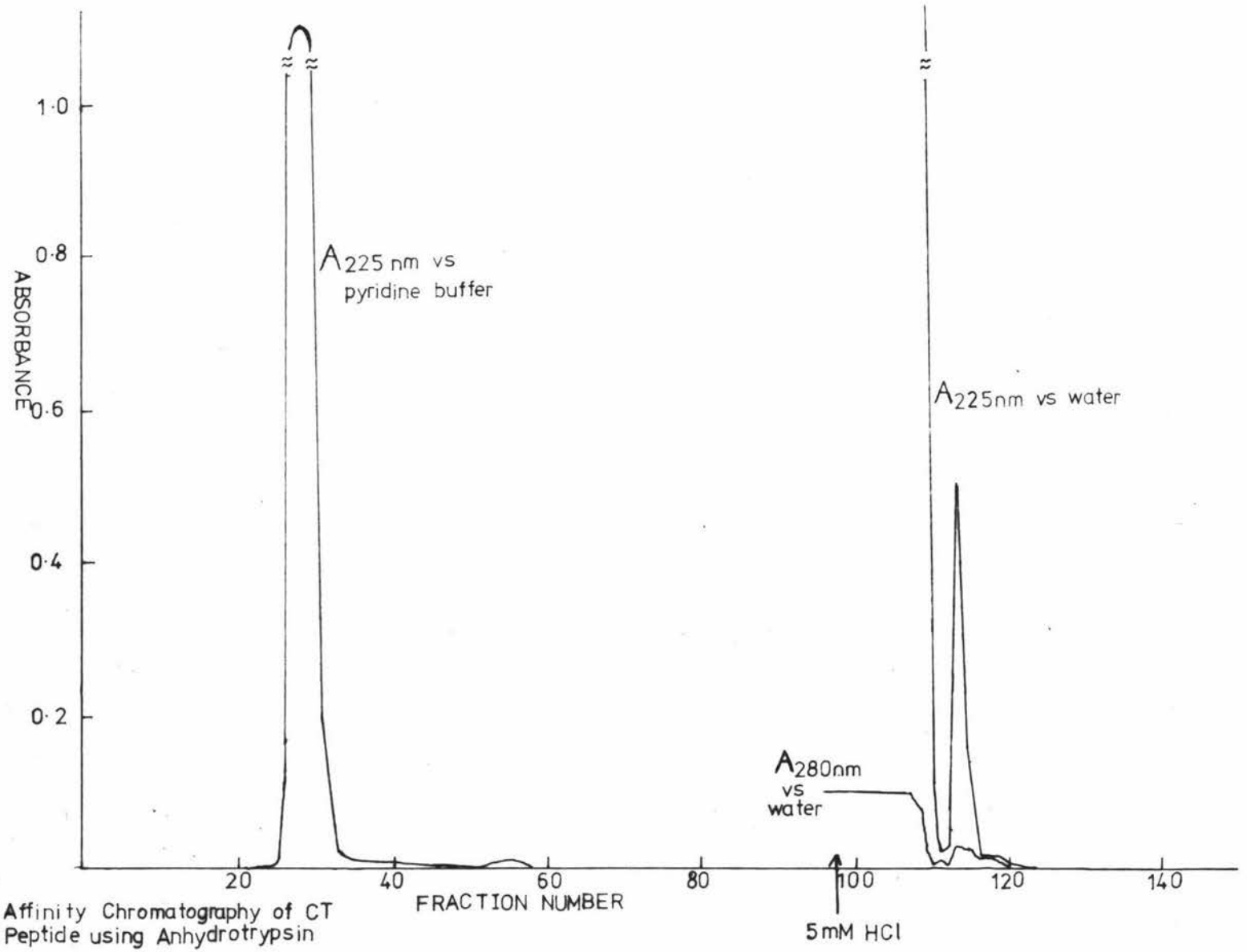


FIGURE 19 Affinity Chromatography of CT Peptide using Anhydrotrypsin

Table 7. Amino Acid Composition of CT I and CT II Peptides

Amino Acid	CT I			CT II		
	Table 7A nmoles	molar[a] equiv.	possible composition.	Table 7B nmoles	molar[b] equiv.	possible composition.
CMCys	0	0		0	0	
Asp	7.6	1.1	1	29.9	1.0	1
Thr	5.4	0.8	1	40.1	1.3	1
Ser	11.7	1.7	2	21.0	0.7	1
Glu	38.0	5.4	5	180.0	6.0	6
Pro	0	0		31.5	1.1	1
Gly	14.2	2.0	2	44.1	1.5	1
Ala	6.9	1.0	1	24.5	0.8	1
Val	7.2	1.0	1	30.6	1.0	1
Met	0	0		0	0	
Ile	3.3	0.5		9.9	0.3	
Leu	4.0	0.6		20.6	0.7	1
Tyr	0.8	0.1		5.2	0.2	
Phe	2.0	0.3		3.3	0.1	
His	1.7	0.2		4.8	0.2	
Lys	17.0	2.4	2	58.5	2.0	2
Arg	18.6	2.6	1-2	64.3	2.1	2

Note a. Molar equivalent is based on a level of 7 nmoles of peptide.

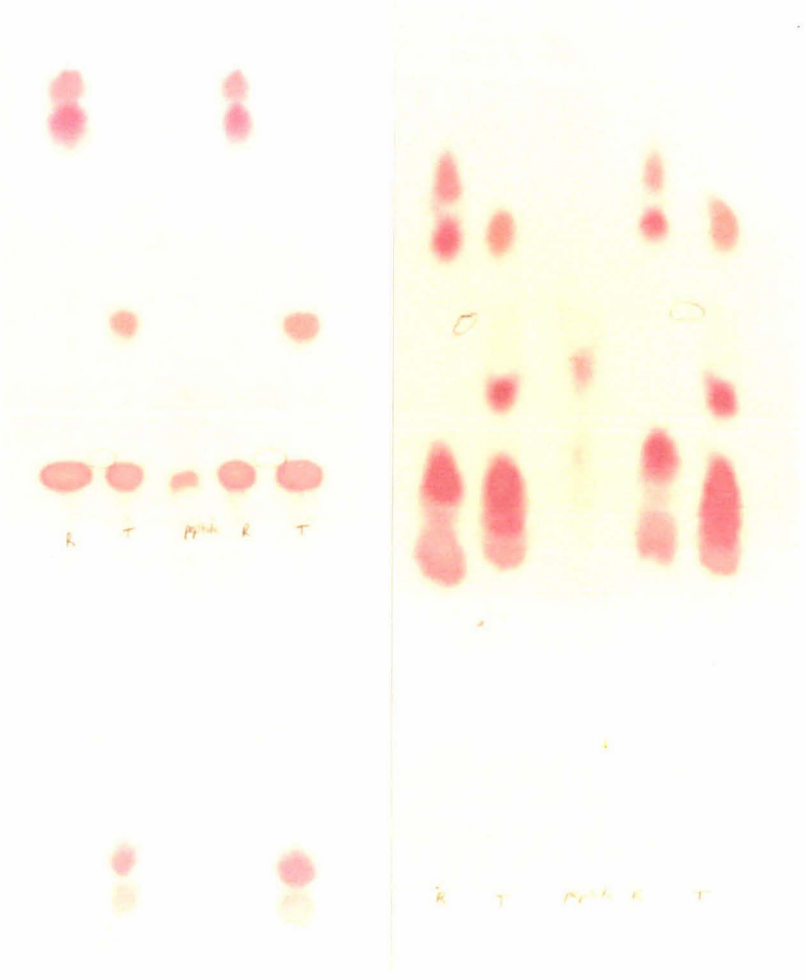
b. Molar equivalent is based on a level of 30 nmoles of peptide.

absorbed only weakly at 280 nm in agreement with the absence of any aromatic residues in the amino acid composition of the peptide. This is evidence that tryptophan is also absent. The peptide having an N-terminal Glx was composed of: 2 Lys, 1 or 2 Arg, 1 Asx, 1 Thr, 2 Ser, 5 Glx, 2 Gly, 1 Ala and 1 Val Table 7A. Because of the low levels of amino acids present it was difficult to differentiate between peptide and background amino acids. No blank spot of equal size and electrophoretic mobility to the peptide on the pH 2.1 electropherogram was eluted and subjected to acid hydrolysis and amino acid analysis. Thus no allowance has been made for contamination by low levels of free amino acids such as Gly, Ala, and Asp etc which are commonly found on electrophoresis paper. Unfortunately arginine was eluted from the amino acid analyser column by the NaOH column wash as a sharp peak overlapping a peak of ninhydrin-positive material consistently eluted at the end of each column run. The arginine content of CT I is therefore thought to be either one or two residues. Electrophoresis at pH 2.1 of CT I gave an Em 1.05 with respect to DNS Arg, which using a Offord plot (Offord, 1977) corresponds to a peptide molecular weight of 350 daltons for charge +2, 760 daltons for charge +3, approximately 1600 daltons for charge +4 or approximately 2300 daltons for charge +5. An approximate molecular weight of CT I is 1800 daltons from the amino acid composition which suggests that the peptide may have a charge of +4 at pH 2.1 requiring the presence of 3 basic amino acids: most probably 2 Lys and 1 Arg residues. Offord plots at pH 2.1 of such highly charged and large peptides give an indication only, as mobilities have been determined only for peptides having a charge of +1 to +3 (Offord, 1977). An average slope of -0.62 (Bailey and Ramshaw, 1973) was used to construct hypothetical mobility plots for charged peptides of +4 and +5.

CT II

RCM-C$\overline{1}$s (30 mg) was digested with chymotrypsin and applied to the anhydrotrypsin column as previously for CT I purification. The 5 mM HCl-eluted peak contained a single peptide having an N-terminal Glx residue in good agreement with CT I. A one-third aliquot of CT II was then subjected to acid hydrolysis and amino acid analysis. Table 7B. The amino acid composition of CT II is in good agreement with that obtained for CT I except for some slight differences which may have been due to the very low amount of CT I peptide available for amino acid analysis. For this reason the CT II amino acid composition is considered to be more accurate than that of CT I. The content of glycine and serine in CT II is less than in CT I while the content of threonine and glutamate has increased slightly. Proline, which was not evident in CT I, was present in CT II at a level of 1 mole/mole of peptide. Arginine was again eluted by the NaOH column wash and so was eluted as a sharp peak superimposed on the usual column wash peak. As for CT I its elution position was displaced significantly from its usual position however an arginine standard was also eluted in the same displaced position. For this reason a peptide containing only one Arg residue is thought to be the more likely case. Such a peptide would have a molecular weight of 1944 daltons and a charge of +4 at pH 2.1. The Em at pH 2.1 of CT II was 0·97 (Plate 4B) and was in excellent agreement with Em 1.05 for CT I. At pH 6.5 the Em 0.0 indicates that CT II is neutral. (Plate 4A). This situation requires equal numbers of acidic and basic residues. The three basic residues of CT I and CT II requires that three of the 6 Glx and 1 Asx are in fact acidic residues with the remainder being amides.

The amino acid sequence of CT II was determined as far as the quantity of material would allow. Glx-Glx-Lys-Glx-Val was the amino acid sequence obtained.



4A 4B

PLATE 4. High Voltage Paper Electrophoresis of CT Peptide at pH 6.5 (4A) and pH 2.1 (4B)

CT III

RCM $\overline{C1s}$ (14 mg) was digested with chymotrypsin and applied to the anhydrotrypsin column as for CT I and II. The 5 mM HCl-eluted peak was freeze dried and used directly for amino acid sequence determination until all the material was consumed to give the following amino acid sequence: Glx-Glx-Lys-Glx-Val-Pro-Glx-Gly. Although the peptide CT III contained some slight impurities, as evidenced by more than one dansyl-amino acid after most cycles, the major DNS-amino acid from the $\overline{C1s}$ peptide was easily identifiable by the intensity of UV absorbance. ϵ-DNS Lys was consistently identified after each Edman cycle. Proline was positioned where a sharp reduction in intensity of UV absorbance occurred as a result of the 16 hour acid digestion of the dansylated peptide. (Gray, 1967). Before and after this residue the DNS-amino acid derivatives were easily identified. Excellent agreement was found for the amino acid sequences determined for CT II and CT III. When the residues of the amino acid sequence of CT III are subtracted from the amino acid composition of CT II the following residues are unplaced: Asp, Thr, Ser, Glx, Glx, Ala, Leu, Lys, Arg and possibly a further Arg.

CT IV

RCM-$\overline{C1s}$ a chain (10 mg) was isolated by DEAE-Sephadex A50 before being thoroughly desalted by dialysis, freeze drying and Sephadex G25 gel filtration. The freeze dried protein was then digested by chymotrypsin and a peptide having an N-terminal Glx residue was isolated by affinity chromatography as before. Approximately 10 nmoles of peptide was digested separately using thermolysin, SV 8 protease, and trypsin at enzyme to substrate ratios of 2.5%, 30% and 2.5% $^{w}/_{w}$ respectively. After freeze drying a one-half aliquot of each digest was subjected to electrophoresis at pH 2.1 and pH 6.5 to determine the extent, if any, of cleavage of

the chymotryptic peptide into smaller fragments. Table 8.

Controls using undigested CT IV and each protease alone were also tested. The protease blanks resulted in no ninhydrin-staining spots while the chymotryptic peptide blank produced a single neutral spot after pH 6.5 electrophoresis. A mobility with respect to DNS Arg of 1.025 was produced on pH 2.1 electrophoresis however two very lightly staining spots having mobility 1.356 (colour purple) and 0.763 (colour orange) were present in all pH 2.1 electropherograms involving the chymotryptic peptide. These contaminants, possibly free Arg or Lys and glycine respectively, were very difficult to accurately locate due to their light staining. The chymotryptic peptide obtained from RCM-$\overline{C1s}$ a chain possessed the same N-terminal Glx residue as the peptide isolated from RCM-$\overline{C1s}$. Both peptides were neutral at pH 6.5 and had an electrophoretic mobility of approximately 1.0 with respect to DNS-Arg at pH 2.1. The presence of other chymotryptic peptides from the $\overline{C1s}$ b chain should have no effect on peptide isolation if the Arg-ending peptide is specifically bound.

Trypsin and possibly also thermolysin appeared to cleave the chymotryptic peptide while SV 8 protease had no effect on CT IV mobility. Thermolytic cleavage of CT IV was very slight only as the majority of the peptide remained intact. This may have been due to using a possibly inactive or partially active thermolysin preparation. Thermolysin cleaves on the N-terminal side of Ile, Leu, Val, Phe, Ala, or Tyr however specificity is not restricted to these residues alone (Heinrikson, 1977). The presence of proline in the P_2' position prevents thermolytic cleavage and so should prevent cleavage of CT IV at valine if the sixth residue is in fact a proline. If Leu is present in the chymotryptic peptide cleavage would be expected to occur. The failure of SV 8 protease to cleave CT IV is

Table 8. Cleavage of the Chymotryptic Peptide by Proteases

Protease	pH 6.5 Electrophoresis			pH 2.1 Electrophoresis		
	EM	Staining Intensity	Colour	EM	Staining Intensity	Colour
Thermolysin	0.0	5+	p	1.041	5+	p
	-0.267	+	p	1.195	3+	p
	+0.504	+	p			
Trypsin	0.0	3+	r	0.948	5+	p
	-0.447	3+	p	1.094	2+	p
	+0.485	2+	p	1.190	2+	p
SV8 Protease	0.0	5+	p	1.040	3+	p
CT IV alone	0.0	5+	p	1.025	5+	p

The chymotryptic peptide was digested by various proteases and the extent of cleavage was determined by electrophoresis of the digest. Staining intensity is based on a scale + (lightly staining) to 5+ (heavily staining). Colour of staining: p represents purple/violet staining spot, r represents red staining spot.

surprising considering the large number of possible glutamic acid residues present. Trypsin cleaved the chymotryptic peptide into what appears to be three fragments. On pH 6.5 electrophoresis these peptides separated into one or more neutral peptides (Tn) plus one basic (Tb) and one acidic peptide (Ta); a situation that can only arise from fragmentation of a neutral peptide such as the chymotryptic peptide isolated from RCM $\overline{C1s}$ a chain. Electrophoresis at pH 2.1 of these peptides produced three spots indicating that only a single peptide which was neutral at pH 6.5 (Tn) was produced.

Tryptic peptides, Ta having pH 6.5 electrophoretic mobility of -0.447 and Tb +0.485 correspond to molecular weights of 440 daltons (charge -1) or 1070 daltons (charge -2) for Ta and 490 daltons (charge +1) or 1120 daltons (charge +2) for peptide Tb. The molecular weight of the neutral tryptic peptide, Tn cannot of course be determined at pH 6.5. The net charge of all tryptic peptides at pH 2.1 should be +2 assuming that the C-terminal residue of the intact chymotryptic peptide is arginine. Peptide molecular weight determinations from pH 2.1 electrophoretic mobilities of 0.9458, 1.094, and 1.190, (all charge +2), correspond to molecular weights of 500, 400 and 300 daltons respectively. Offord points out that more caution is required when estimating molecular weights using Offord Plots than when estimating net charge. (Offord, 1977). Single measurements, especially of peptides having charge +4 or greater, are therefore intended as approximate only. The absence of histidine and cysteine in the chymotryptic peptide reduces the possibility of peptide mobility departing from ideal behaviour because of partial charges (Offord, 1977).

The amino acid composition of these tryptic peptides can be partially deduced from their mobilities at pH6.5 assuming a charge of +2 for each peptide at pH 2.1. Tn

will contain one basic residue plus one acidic residue; Tb will contain one basic residue only and no acidic residues while Ta will contain one basic residue plus either two acidic residues to give net charge -1 at pH 6.5 or three acidic residues to give a net charge -2.

CT V

RCM $C\overline{1s}$ a chain (14 mg) was again isolated by DEAE-Sephadex A-50 and thoroughly desalted before digestion by chymotrypsin and isolation of a chymotryptic peptide possessing an N-terminal Glx residue. The peptide was then digested by trypsin (2.5% w/w) and after freeze drying and dissolving in 0.05M pyridine acetate, pH 5.0, containing 0.02M calcium chloride was reapplied to the equilibrated anhydrotrypsin column. The column elution profile is shown in Figure 20 after first freeze drying column fractions and subjecting a suitable aliquot to pH 6.5 electrophoresis. The column-bound material eluted in fractions 22 and 23 contained two ninhydrin-positive spots: a light-staining neutral peptide(s) and a major peptide having Em-0.429. This agrees well with the mobility of Ta produced by tryptic cleavage of CT IV (Em -0.447). Fractions 22 and 23 were pooled, lyophilised and subjected to Dansyl-Edman amino acid sequencing. The sequence determined was Glx-Glx-(Asx or Gly)-Arg. Two further Edman cycles produced no further DNS-amino acids. As only a single Glycine is thought to be present in the intact chymotryptic peptide the Asx residue is favoured to be the actual third residue of the anhydrotrypsin-bound peptide. The neutral retained peptide also present in fractions 22 and 23 (Figure 20) is thought to be undigested chymotryptic peptide CT V which would be expected to be also bound by anhydrotrypsin. Three ninhydrin-staining peptides appear to have passed unbound through the anhydrotrypsin column;

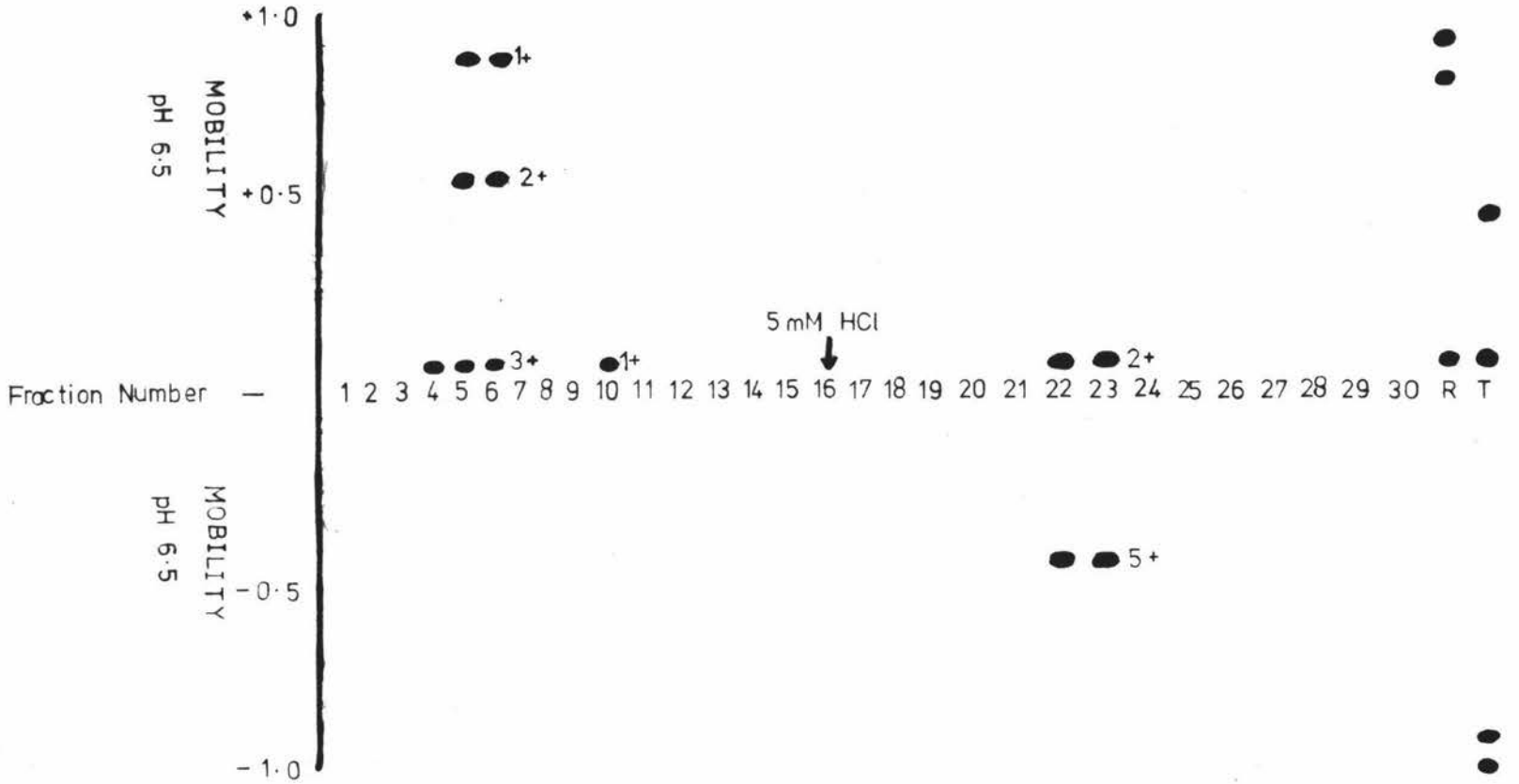


FIGURE 20 Elution Profile of the Tryptic Digest of CT $\underline{\underline{V}}$ Reapplied to Anhydrotrypsin-Sepharose

An aliquot (one-tenth) of each fraction was subjected to pH 6.5 electrophoresis. The tryptic digest of CT V was applied to the Anhydrotrypsin column at Fraction 1. Staining intensity is shown on an increasing scale from +1 to +5.

of these a neutral peptide(s) (Tn) and a basic peptide (Tb) having Em +0.485 had been identified on tryptic digestion of CT IV. The Em obtained in this case was +0.548. The other lighter-staining peptide had mobility +0.885.

3.7 Conclusion

From the results obtained from five chymotryptic digestions of RCM $\overline{C1s}$ or RCM $\overline{C1s}$ a chain a partial primary structure of the chymotryptic peptide bound by anhydrotrypsin is proposed. Figure 21. The neutral chymotryptic peptide specifically bound by anhydrotrypsin was cleaved into 3 fragments Ta, Tb and Tn by trypsin. (Table 9). The alignment of tryptic peptides Ta, Tb and Tn in their correct order in the anhydrotrypsin-bound chymotryptic peptide amino acid sequence was made possible from electrophoretic mobilities and sequence determination of these peptides and the intact chymotryptic peptide. Overlaps in amino acid sequences obtained from CT II and CT III as well as a knowledge of the amino acid composition of the chymotryptic peptide enabled the T1 and T2 structures to be elucidated. The Lysine residue preceding T3 is positioned knowing the specificity of trypsin and the identification of only three tryptic peptides. T3 would be expected to bind to anhydrotrypsin as it is this Arg-ending portion of the chymotryptic C-terminal peptide of $\overline{C1s}$ a chain which is responsible for anhydrotrypsin binding. Merely making the peptide shorter by proteolytic digestion should have no effect on anhydrotrypsin binding by this Arg-ending portion. T3 is an acidic peptide of mobility -0.429 to -0.447 and from the sequence data has a molecular weight of 546 daltons. Assigning a charge of -1 at pH 6.5 this mobility corresponds to a molecular weight of 530 daltons (by Offord plot: Offord, 1977) in excellent agreement. This requires that only two of the Glx-Glx-Asx residues are acidic amino acids while one is an amide. T3 therefore corresponds to Ta the acidic tryptic peptide while T1 and T2 correspond to the neutral and basic peptides Tn and Tb. Both peptides contain 2 Glx residues and T1 contains 1 Lys residue while T2 should also terminate in a Lys residue, being the product of a tryptic digestion. As T1 can only contain one basic residue (see CT IV) the peptide Tb of Em +0.485 to +0.548 at pH 6.5 must have a charge of +1 and therefore molecular weight of 490 daltons (See CT IV).

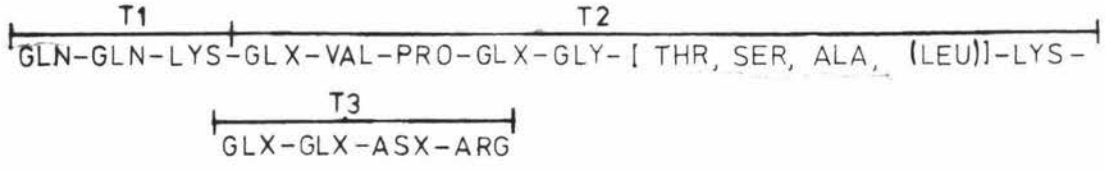


FIGURE 21 The Partial Primary Structure of the C- terminal Chymotryptic Peptide Bound by Anhydrotrypsin

[ ] indicate amino acids present in the amino acid composition of the chymotryptic peptide but not encountered in the amino acid sequence analysis.

( ) indicate amino acids whose presence in the amino acid composition of the chymotryptic peptide is uncertain.

Table 9. Alignment of Tryptic Fragments Ta, Tb and Tn in the CT Peptide

	Em, pH 6.5	Charge pH 6.5	Acidic Residues	Sequence	
Tb	+0.485 to +0.548	+1	0	Gln-Gln-Lys	T1
Ta	-0.429 to -0.447	-1	2	Glx-Glx-Asx-Arg	T3
Tn	0.0	0	1	Glx-Val-Pro-Glx-Gly-[Thr,Ser, Ala,(Leu)]-Lys	T2

The sequence of each tryptic fragment was deduced from its mobility at pH 6.5 (and pH 2.1) as well as the amino acid sequences of the intact chymotryptic peptide and T3 which were determined.

Clearly this peptide must be T1 having the amino acid sequence: Gln-Gln-Lys (molecular weight 410 daltons). T2 must therefore be the large neutral peptide, Tn and should contain 1 Lys and 1 Glu residue with the remaining Glx residues being present as glutamines.

The various peptide molecular weights and charges at pH 6.5 do not correlate well with the mobilities of tryptic peptides at pH 2.1. Molecular weights of the three peptides obtained (assuming a net charge at pH 2.1 of +2) were 300, 400 and 500 daltons. The smaller Ta and Tb peptides can be accounted for however identification of the larger neutral peptide is not possible.

It is interesting to note that the same chymotryptic peptide was isolated by affinity chromatography whether RCM $C\overline{1s}$ or the RCM a chain was the substrate indicating that in both chains only a single C-terminal arginine residue is present.

The presence of an Arginine residue at the C-terminus of the $C\overline{1s}$ a chain as determined by C-terminal peptide sequencing was verified by the very rapid release of arginine from RCM $C\overline{1s}$ by carboxypeptidase B.

Although peptide mapping of RCM $C\overline{1s}$ a chain was unsuccessful, in part due to the low numbers of peptides identified, it is interesting in retrospect that the additional tryptic peptide isolated had an electrophoretic mobility of -0.500 at pH 6.5; similar to the mobility of T3.

A peptide very similar in amino acid composition to T3 has been isolated by J.C. McIntosh. Human RCM $C\overline{1s}$ a chain was isolated by the same methods as used in these studies and a chymotryptic peptide, having a Glx N-terminal residue was isolated by affinity chromatography on the anhydrotrypsin-Sepharose column used in this project. The chymotryptic peptide was further fragmented by trypsin and the digestion mixture was subjected to pH 6.5 high voltage paper electrophoresis and an acidic peptide was isolated by elution from

the paper. A sample of this peptide was subjected to acid hydrolysis and applied to an amino acid analyser capable of determining accurately 5 nmoles of amino acids. Preliminary results suggest an amino acid composition almost identical to that determined by sequence analysis of T3. Two Glx, and one each of Asx, Ser and Arg were thought to be present but what were postulated to be background levels of Gly, Ala, Leu and Val were also found. (Personal Communication).

Further investigation into the primary structure of this C-terminal derived peptide will be required, especially as regards to amide assignment, before an ideal $C\overline{1r}$ model peptide substrate, mimicking the C1s structure at the scissile Arg-Ile bond can be synthesized.

REFERENCES

Akabori, S., Ohno, K. and Narita, K. (1952) Bull. Chem. Soc. Japan, 25, 214

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

Ambler, R.P. (1972) in "Methods in Enzymology" (eds Hirs, C.H.W. and Timasheff, S.N.) Academic Press, New York, 25, 143

Andrews, J.M. and Baillie, R.D. (1979) J. Immunol., 123, 1403

Arlaud, G.J., Meyer, C.M. and Colomb, M.G. (1976) FEBS Lett., 66, 132

Arlaud, G.J., Reboul, A., Meyer, C.M. and Colomb, M.G. (1977) Biochem. Biophys. Acta, 485, 215

Arlaud, G.J., Sim, R.B., Duplaa, A. and Colomb, M.G. (1979) Mol. Immunology, 16, 445

Arlaud, G.J., Villiers, C.L., Chesne, S. and Colomb, M.G. (1980) Biochim. Biophys. Acta, 616, 116

Assimeh, S.N., Bing, D.H. and Painter, R.H. (1974) J. Immunol., 113, 225

Atfield, G.N. and Morris, C.J.O.R. (1961) Biochem. J., 81, 606

Bailey, C.J. and Ramshaw, J.A.M. (1973) Biochem. J., 135, 889

Bailey, J.L. (1955) J. Biochem (Tokyo), 60, 173

Barkas, T., Scott, G.K. and Fothergill, J.E. (1973) Biochem. Soc. Trans., 1, 1219

Bennett, J.C. (1967) in "Methods in Enzymology" (ed. Hirs, C.H.W.) Academic Press, New York, 11, 330

Bergmann, M. and Zervas, L. (1936) J. Biol. Chem., 113, 341

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

Boland, M.J. and Hardman, M.J. (1972) FEBS Lett., 27, 282

Bradford, M.M. (1976) Anal. Biochem., 72, 248

Campbell, R.D., Booth, N.A. and Fothergill, J.E. (1979) . Biochem. J., 183, 579

Canfield, R.E. (1963) J. Biol. Chem., 238, 2691

Carlton, B.C. and Yanofsky, C. (1963) J. Biol. Chem., 238, 636

Carraway, K.L. and Koshland, E. (1972) in "Methods in Enzymology" (eds. Hirs, C.H.W. and Timasheff S.N.) Academic Press, New York, 25, 616

Carraway, K.L., Spoerl, P. and Koshland, E. (1969) J. Mol. Biol., 42, 133

Davie, E.W., Fujikawa, K. Kurachi, K., Koide, T. and Kisiel, W. (1976) Miami Winter Symp., 11, 189

Davis, A.E. (1980) Proc. Natl. Acad. Sci. ., 77,4938

Davis, A.E., Zalut, C., Rosen, F.S. and Alper, C.A. (1979) Biochemistry, 18, 5082

Davis, B.J. (1964) Ann. N.Y. Acad. Sci., 121, 404

Dayhoff, M.O. (1972) "Atlas of Protein Sequence and Structure", National Biomedical Research Foundation, Silver Spring, 5, 56

Dodds, A.W., Sim, R.B., Porter, R.R. and Kerr, M.A. (1978) Biochem. J., 175, 383

Erlanger, B.F., Kokowsky, N. and Cohen, W. (1961) Arch. Biochem. Biophys, 95, 271

Folk, J.E. (1970) in "Methods in Enzymology" (eds. Perlmann, G.E. and Lorand, L.) Academic Press, New York, 19,504

Fothergill, J.E. and Anderson, W.H.K. (1978) Curr. Top. Cell. Regul., 13, 259

Frank, M.M. (1979) Rev. Infect. Dis., 1, 483

Gabel, D. and Axén, R. (1976) in "Methods in Enzymology" (ed Mosbach, K.), Academic Press, New York, 44, 383

Gall, W.E., Cunningham, B.A. Waxdal, M.J., Konigsberg, W.H. and Edelman, G.M. (1968) Biochemistry, 7,1973

Gigli, I., Porter, R.D. and Sim, R.B. (1976) Biochem. J., 157, 541

Goers, J.W.F. and Porter, R.R. (1978) Biochem. J., 175, 675

Gray, W.R. (1967) in "Methods in Enzymology" (ed Hirs, C.H.W.) Academic Press, New York, 11, 469

Gray, W.R. (1972) in "Methods in Enzymology" (eds. Hirs, C.H.W. and Timasheff, S.N.) Academic Press, New York, 25, 121

Harpel, P.C. and Cooper, N.R. (1975) J. Clin. Invest., 55, 593

Hartley, B.S. (1970) Biochem. J., 119,805

Heinrikson, R.L. (1977) in "Methods in Enzymology" (eds. Hirs, C.H.W. and Timasheff, S.N.) Academic Press, New York, 47, 175

Hirs, C.H.W. (1977) in "Methods in Enzymology" (eds. Hirs, C.H.W. and Timasheff, S.N.) Academic Press, New York, 47, 97

Ishizaka, K., Ishizaka, T. and Sugakura, T. (1962) J. Immunol., 88, 690

James, G.T. (1978) Anal. Biochem., 86, 574

Kasai, K. and Ishii, S. (1975) J. Biochem (Tokyo), 78,653

Kassell, B. and Kay, J. (1973) Science, 180,1022

Katz, A.M., Dreyer, W.J. and Anfinsen, C.B. (1959) J. Biol. Chem., 234, 2897

Kerr, M.A., Graham, D.T., Walsh, K.A. and Neurath, H. (1978) Biochemistry, 17, 2645

Khorana, H.G. (1952) J. Chem. Soc., 2081

Knobel, H.R., Villiger, W. and Isliker, H. (1975) Eur. J. Immunol., 5, 78

Konigsberg, W. (1972) in "Methods in Enzymology" (ed. Hirs, C.H.W.) Academic Press, New York, 25,185

Lachman, P.J. and Thompson, R.A. (1970) J. Exp. Med., 131, 643

Lepow, I.H., Naff, G.B., Todd, E.W., Pensky, J. and Hinz, C.F. (1963) J. Exp. Med., 117, 983

Lepow, I.H., Ratnoff, O.D. and Levy, R. (1958) J. Exp. Med., 107, 451

Londsdale-Eccles, J.D., Kerr, M.A., Walsh, K.A. and Neurath, H. (1979) FEBS Lett., 100, 157

Matsuo, H., Fujimoto, Y. and Tatsuno, T. (1966) Biochem. Biophys. Res. Commun., 22, 69

Müller-Eberhard, H.J. (1975) Ann. Rev. Biochem., 44, 697

Naff, G.B. and Ratnoff, O.D. (1968) J. Exp. Med., 128, 571

Nagasawa, S., Takahashi, K. and Koyama, J. (1974). FEBS Lett., 41, 280

Needleman, S.B. (1975) in "Protein Sequence Determination" Springer-Verlag, New York,

Nozaki, Y. (1972) in "Methods in Enzymology" (eds. Hirs, C.H.W. and Timasheff, S.N.) Academic Press, New York, 26, 43

Offord, R.E. (1977) in "Methods in Enzymology" (eds. Hirs, C.H.W. and Timasheff, S.N.) Academic Press, New York, 47,51

Ouchterlony, O. (1953) Acta. Pathol. Microbiol. Scand., 32, 231

Parham, M.E. and Loudon, G.M. (1978) Biochem. Biophys. Res. Commun., 80, 1 and ibid. 80, 7

Parikh, I., March, S. and Cuatrecasas, P. (1974) in "Methods in Enzymology" (eds. Jakoby, W.B. and Wilchek, M.) Academic Press, New York, 34, 77

Porter, R.R. (1977a) Fed. Proc. Fed. Am. Soc. Exp. Biol., 36, 2191

Porter, R.R. (1977b) Biochem. Soc. Trans., 5, 1659

Potts, J.T. (1967) in "Methods in Enzymology" (ed. Hirs, C.H.W.) Academic Press, New York, 11, 648

Ratnoff, O.D. and Lepow, I.H. (1957) J. Exp. Med., 106,327

Ratnoff, O.D., Pensky, J., Ogsten, D. and Naff, G.B. (1969) J. Exp. Med., 129, 315

Reid, K.B.M., Sim, R.B. and Faiers, A.P. (1977) Biochem. J., 161, 239

Sayers, C.A. and Barrett, A.J. (1980) Biochem. J., 189,255

Schechter, I. and Berger, A. (1967) Biochem, Biophys, Res. Commun. 27, 157

Schluck, P. and Kumpf, W. (1926) Hoppe-Seyler's Z. Physiol. Chem., 154, 125

Scott, G.K. and Fothergill, J.E. (1975) Biochem. Soc. Trans, 3, 934

Shotton, D.M. and Hartley, B.S. (1973) Biochem. J., 131,643

Sigler, P.B., Blow, D.M., Matthews, B.W. and Henderson, R. (1968) J. Mol. Biol., 35, 143

Sim, R.B. and Porter, R.R. (1976) Biochem. Soc. Trans., 4, 127

Sim, R.B., Arlaud, G.J. and Colomb, M.G. (1979) Biochem. J., 179, 449

Sim, R.B., Porter, R.R., Reid, K.B.M. and Gigli, I. (1977) Biochem. J., 163, 219

Smyth, D.G., Stein, W.H. and Moore, S. (1963) J. Biol. Chem., 238, 1114

Stark, G.R. (1968) Biochemistry, 7, 1796

Takahashi, K., Nagasawa, S. and Koyama, J. (1975a) FEBS Lett., 50, 330

Takahashi, K., Nagasawa, S. and Koyama, J. (1975b) FEBS Lett., 55, 156

Tanford, C. and Roberts, G.L. (1952) J. Amer. Chem. Soc., 74, 2509

Tang, J. and Hartley, B.S. (1967) Biochem J., 102, 593

Taranta, A. and Franklin, E.C. (1961) Science 134, 1981

Valet, G. and Cooper, N.R. (1974a) J. Immunol., 112, 339

Valet, G. and Cooper, N.R. (1974b) J. Immunol., 112, 1667

Volanakis, J.E., Schrohenloher, R.E. and Stroud, R.M. (1977) J. Immunol., 119, 337

Weber, K. and Osborn, M. (1969) J. Biol. Chem., 244, 4406

World Health Organisation (1968) Bull. W.H.O., 39, 935 or Immunochem., 7, 137 (1970)

Yokosawa, H. and Ishii, S. (1976) Biochem. Biophys. Res. Commun., 72, 1443

Yokosawa, H. and Ishii, S. (1977) J. Biochem. (Tokyo), 81, 647

Yokosawa, H. and Ishii, S. (1979) Anal. Biochem., 98, 198

Yokosawa, H., Hanba, T. and Ishii, S. (1976) J. Biochem (Tokyo), 79, 757

Ziccardi, R.J. and Cooper, N.R. (1976) J. Immunol., 116,496