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

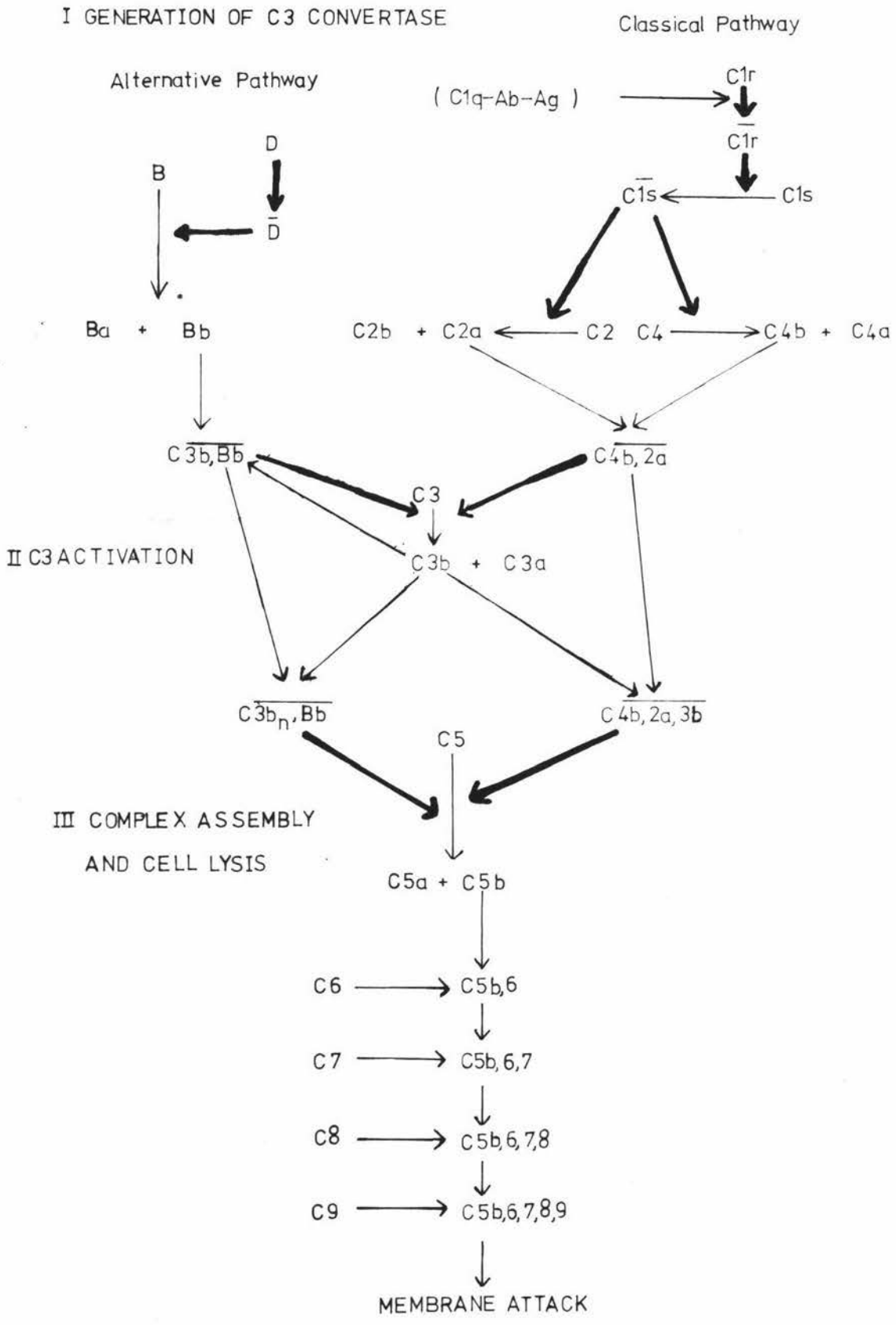

FIGURE 1 The Complement System

↓ 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.