THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE CLEAVAGE OF RABBIT IMMUNOGLOBULIN G AT CONSTANT TRYPTOPHAN RESIDUES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

DALE LYNN CAMPBELL

Oklahoma City, Oklahoma

1971

THE CLEAVAGE OF RABBIT IMMUNOGLOBULIN G AT

CONSTANT TRYPTOPHAN RESIDUES

APPROVED BY

Robert Delaney

Martin J. Griffin

Jean A. Harbaugh

Jordan J. n. Lang

Leon Unger

DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

I wish to thank the Oklahoma Medical Research Foundation for its support during the last few years.

I would also like to thank the members of the reading committee for their help in the preparation of this thesis.

Also, I would like to express my appreciation to Mr. John Slagle and Mrs. Janet Rogers for their technical help, and many thanks to Dr. Jordan Tang, Neurosciences Department, Oklahoma Medical Research Foundation, and all the people in his lab for the help they have given me.

Especially, I would like to thank my advisor, Dr. Robert Delaney, Department of Biochemistry and Molecular Biology, University of Oklahoma Medical School, for his advice, counsel and friendship over the last few years.

I would also like to thank my wife, Jolene, who helped so much in the preparation of this manuscript.

TABLE OF CONTENTS

	Page
LIST OF TABLES	v
LIST OF ILLUSTRATIONS	vi
Chapter	
I. INTRODUCTION	1
II. METHODS AND MATERIALS	31
III. RESULTS	45
IV. DISCUSSION	76
V. SUMMARY	95
BIBLIOGRAPHY	97

LIST OF TABLES

Table		Page
1.	Properties of Human Immunoglobulins (30)	4
2.	Amino Acid Composition of Light Chains, Heavy Chains and Cl Fraction	53
3.	Relative Composition of Fragments Obtained from NBS Oxidized Light Chain	57
4.	Relative Composition of Fragments Obtained from NBS Oxidized C_l Fractions	64
5.	Composition of Tryptic Peptides Obtained from Cl Fractions	67
6.	Fractions from Light Chains and Cl Expressed as Residues/Mole	75
7.	Comparison of Predicted and Determined Compositions of Cl Fractions	81

LIST OF ILLUSTRATIONS

Figure		Page
1.	The Site of Proteolytic Cleavages in the Heavy Chain (51)	8
2.	The Interchain and Intrachain Disulfide Bonds of Eu (86)	12
3.	The Interchain Disulfide Bonds of the 4 Types of Human IgG (84)	13
4.	The Evolution of Immunoglobulin Chains	22
5.	The Constant Tryptophan Residues in Eu (86)	24
6.	The Reaction of Tryptophan with N-bromosuccinimide to Produce Chain Cleavage	28
7.	DEAE Cellulose Chromatography of Rabbit IgG	46
8.	Polyacrylamide Gel Electrophoresis of IgG Fractions	47
9.	Double Immunodiffusion of DEAE Cellulose Fractionation Peaks	47
10.	The Immunoelectrophoresis of IgG Fractions from DEAE Cellulose Chromatography	48
11.	Chain Separation in Propionic Acid	51
12.	Chain Separation in 8 M Urea	52
13.	The Spectra of Unreacted and Oxidized Light Chains	54
14.	Chromatography of N-bromosuccinimide Treated Light Chains	56
15.	Chromatography of CNBr Treated Heavy Chains	58
16.	Flow Chart of Reactions of the Heavy Chains and the Resulting Fractions	59
17.	The Spectra of Unreacted and Oxidized C1	60
18.	Chromatography of N-bromosuccinimide Treated C1 Fraction on G-100	62
19.	The Rechromatography of Peak D	63
20.	The Peptide Map of Peak C	65

LIST OF ILLUSTRATIONS--Continued

Figure		Page
21.	The Chromatography of a Trypsin Digest of Peak C	69
22.	The Amino Acid Sequence of Peptides Isolated from Trypsin Digests of Peak C	70
23.	The Peptide Map of Peak E	72
24.	The Peptide Map of Peak B	73
25.	The Peptide Map of Peak DII	74
26.	Comparison of the Amino Acid Sequences from Peak C Peptides with the Sequences for Rabbit IgG Reported by Fruchter (79) and with the Myeloma Protein Eu (86) . .	86
27.	The Sequence of Rabbit Cl Fractions from Residue 109 to the C-terminal (79) and the Location of Isolated Peptides	87

THE CLEAVAGE OF RABBIT IMMUNOGLOBULIN G AT CONSTANT TRYPTOPHAN RESIDUES

CHAPTER I

INTRODUCTION

In ancient times it was realized that persons who recovered from a disease such as smallpox never contracted it again (1). Although this knowledge of the existence of immunity has been with us for many centuries, it has only been the last few years that the scientific community has been able to come to grips with the molecular level reactions that enable the immune response to work. Felton (2) first reported that antibodies were protein in nature in 1932. Later, in 1937, Tiselius (3, 4) reported that antibody to egg white lysozyme was associated with the slowest moving or gamma band that he had been able to separate from immune rabbit serum with his new electrophoresis equipment. He also reported a 7 S value for this band or a molecular weight of 167,000. Other experiments showed that when the rabbit was immunized the amount of protein in the gamma band increased.

The next several years saw much time and effort expended on the study of the physical nature (5), the isolation and purification (6, 7, 8) and the chemical nature (9) of gamma-globulin (IgG). Values given for the molecular weight varied somewhat, but the most generally accepted

values are in the range of 150,000. Rabbit IgG contains about 2.4% by weight carbohydrate. The antibody molecule is elongated with dimensions of about 250 x 40 Å and has little or no alpha helix (9).

As the knowledge of immunoglobulins increased, the nomenclature covering the immunoglobulins also developed. All further references to immunoglobulins will be in accordance with World Health Organization conventions (10, 11, 12, 13, 14).

Early indications were that IgG was composed of a single polypeptide chain (15), however it has always proved to be heterogeneous by physical or chemical criteria. It was not until 1959 that Edelman (16, 17) showed the multichain structure of human IgG. He reduced the disulfide bonds in urea and showed by subsequent column chromatography on carboxymethyl cellulose, ultracentrifugation and starch gel electrophoresis that the molecule has multiple polypeptide chains. Franek (18) extended the findings to show a multichain structure for IgG from the pig, cow, horse and rabbit. Porter and his co-workers did much to establish a structure for the rabbit IgG molecule. Fleischman, Pain and Porter (19) proposed that IgG consists of two chains of molecular weight 110,000 and 40,000 respectively or perhaps four chains; two with a molecular weight of 55,000 and two with 20,000. Results obtained by reduction of disulfide bonds followed by chromatography on Sephadex G-75 in propionic acid and determination of the molecular weight of the resulting fractions in the ultracentrifuge led to this conclusion. In addition, the IgG molecule is dissociable into chains by use of SDS (20) and by use of alkaline conditions (21). Porter (22) then proposed a 4 chain structure with five interchain disulfide bonds. The work leading to this conclusion was pub-

lished by Fleischman et al. (23). This work contained N-terminal analysis and carbohydrate content of the different polypeptide chains. Pain (24) added the molecular weight data in the same series of publications. He found the heavy chain had a molecular weight of 50,000 and the light chain 20,000. Crumpton and Wilkerson (25) also added the amino acid composition. Fleischman et al. (23) concluded that the 4 chain structure accounted for all these observed data. Small (26) and Cohen (27) published confirmation of these conclusions a few months later.

Many early attempts at proteolytic digestion of IgG were made, but little structurally useful data preceded Edelman's work on the multichain structure. Porter (28, 29) made a great breakthrough when he found the digestion of rabbit IgG by papain yielded three fragments of almost identical molecular weight upon carboxymethyl cellulose chromatography. The first two of these fragments (Fab) retained the ability to combine with antigens and were very similar in chemical and biological properties while the third (Fc) crystallized easily and contained most of the antigenic specificity of the intact molecule. The common properties of immunoglobulins for which Fc is responsible, and other properties as summarized by Cohen and Milstein (30) are shown in Table 1. The Fab fragment is composed of the light chain plus a fragment of the heavy chain (Fd). Porter then proposed that the rabbit IgG molecule is formed of two sections with similar structure joined to a third section of a different character. He also proposed that Fc was identical for all IgG and that Fab was different for each specific antibody. Nisonoff proved by measurement of antigen-binding capacity (31) that the Fab fragments contained single combining sites (were "univalent"), and that papain digestion does

TABLE 1

PROPERTIES OF HUMAN IMMUNOGLOBULINS (30)

Properties	IgG	IgA	IgM	IgD
Biological				
Serum conc. (mg %)	800-1680	140-420	50-190	0.3-40
Synthesis rate (mg/kg/day)	20-40	2.7-55	3.2-16.9	0.03-1.49
Catabolic rate (% I V pool/day)	4-7	14-34	14-25	18-60
Distribution (% in I V pool)	48-62	40	65-100	63-86
Antibody activity	+	+	+	
Complement fixation	+	0	+	
Placental passage	+	0	0	0
Presence in cerebrospinal fluid	+	+	0	
Selective seromucous secretion	0	+	0	
Skin sensitization				
heterologous species	+	0	0	
homologous species	0	?	0	
Immunological				
Light-chain types	kappa, lambda	kappa, lambda	kappa, lambda	kappa, lambda
Heavy-chain classes	gamma	alpha	mu	d
types	4	2	2	
Physicochemical				
$S_{20,w}$	6.5-7.0	7,10,13,15,17	18-20, 30	6.2-6.8
Total Carbohydrate (%)	2.9	7.5	11.8	

not effect the binding capacity (32).

Shortly after Porter's work Nisonoff's group produced fragments (Fab') similar to the Fab fragments by digesting with pepsin in the presence of a reducing agent or by digestion with pepsin followed by treatment with a reagent that reduces disulfide bonds (33). The digestion with pepsin gives a 5 S dimer $(Fab')_2$ that dissociates to a 3.5 S monomer (Fab') upon mild reduction. This reduction splits only one disulfide bond (34). The molecular weight of the dimer is 106,000 and that of the monomer is 56,000 (35). In this method the Fc portion is digested to smaller polypeptides.

Work on the multichain nature of the molecule progressed rapidly using the chain separation and proteolytic digestion methods. Palmer et al. (36) showed that the Fab fragments from a single IgG molecule are either Porter's Fab Fraction I or Fab Fraction II, but never both on the same molecule. To do this they first fractionated rabbit IgG on carboxymethyl cellulose then digested each of the fractions with papain and found that each fraction put most of its Fab in one of the two peaks. Goodman (37) showed that Fab and Fab' were antigenically identical. Fleischman et al. (19) showed that Fab contained the antigenic sites for light and heavy chains, but Fc had only antigenic sites for the heavy chains. Nisonoff's group (38, 39) reported that 50% to 66% of the rabbit IgG molecules could be dissociated into half molecules by the reduction of a single disulfide bond. The other molecules required the reduction of two or more bonds. After further study they reported that the same disulfide bond responsible for this dissociation was also responsible for the dissociation of the peptic dimer into Fab' monomers (40). They re-

ported that Fc fragments had the same characteristics (41).

Cebra and his group (42, 43) did a series of studies by cleaving IgG with insoluble papain. This was done because activation of papain requires a reducing agent, and in this manner he could activate the papain and subsequently remove the reducing agent. The papain was made insoluble by polymerizing it and then he carried out a digestion with stirring to keep it in suspension. The reaction could be quickly stopped by centrifugation. By this procedure he (42) was able to produce a 5 S Fab dimer that resembled the product of pepsin digestion. This Fab dimer would dissociate into 3.5 S monomers with the addition of a reducing agent. This work seems to indicate the papain cleavage of Porter (29) is dependent on a reducing agent but later structural work casts some doubt on this interpretation. It is possible that this new Fab dimer resulted from a slight change in specificity of the enzyme. In further work (43) he compared his dimer with $(Fab')_2$ from pepsin and found them to be quite similar except $(Fab')_2$ was a little heavier and could be reversibly dissociated into 3.5 S subunits by reduction and reoxidization whereas his dimer would not reassociate after the disulfide bond was broken.

With the increase in data more accurate conclusions about the structure of IgG could be drawn. Marler et al. (44) added confirmation of the four chain structure and the location of the papain split in the heavy chain by their work on the ultracentrifuge after reducing and dissociating Fab and Fc fragments. Noelken et al. (45) proposed that IgG contains 3 globular portions joined through two short segments of loosely coiled polypeptide chains that are easily hydrolyzed by enzymes. An interchain disulfide bond connects these two chains adjacent to the loosely

coiled segment.

Further studies on the cleavage of IgG were undertaken. Putnam et al. (46) studied plant proteolytic enzymes and found that those which required reducing agents gave splits much like papain. Givol (47) reported a method of cleavage with similar results with trypsin on aminoethylated IgG. Cahnmann et al. (48, 49) reported a chemical cleavage giving a 5 S dimer with cyanogen bromide. Its activity was much the same as that from pepsin (50). A review of the sites of most of the proteolytic digestion cleavages was presented by Givol et al. (51) (Figure 1).

Another of the most important breakthroughs to determining the structure of immunoglobulins came when Edelman and Gally (52) demonstrated that the proteins described by Dr. Henry Bence Jones in 1847 and found in the urine of multiple myeloma patients was the light chain from an immunoglobulin. Edelman and Gally were unable to show any differences by starch gel electrophoresis, amino acid composition and several other criteria between the Bence Jones protein of one patient and the light chains derived from the myeloma protein of the same patient. Since this protein is homogeneous, it is more easily studied than normal heterogeneous light chains. In later studies Schwartz and Edelman (53) showed that the peptides of Bence Jones proteins and the myeloma protein light chains were also identical by 2 dimensional high voltage electrophoresis.

At about this same time Mannik and Kunkel (54) discovered that 2 antigenic types of light chains existed in human Bence Jones proteins. They later extended this observation to 7 S IgG (55). This was also shown by Potter et al. to be true in the mouse (56, 57) and these two types of chains were called kappa and lambda. Gray and Kunkel (58) also

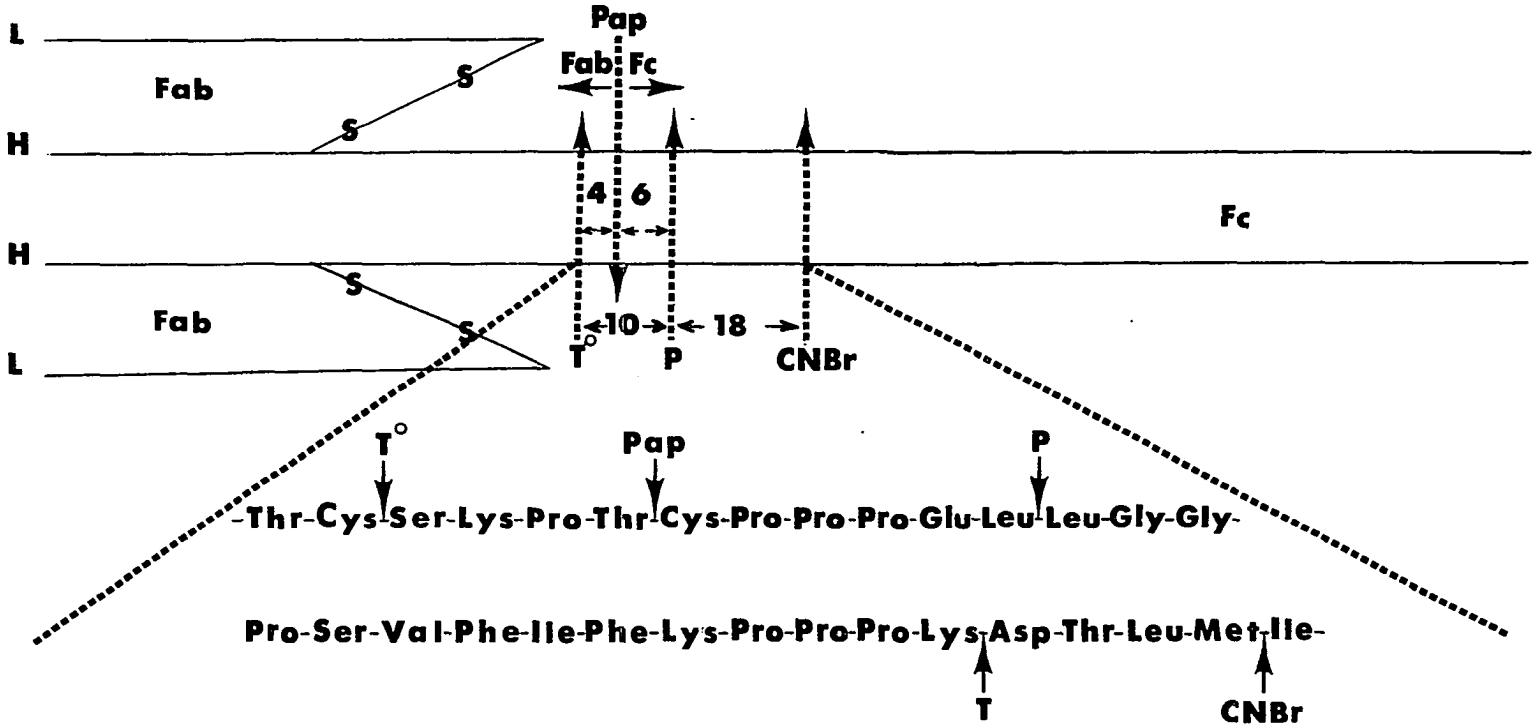

Fig. 1 - The site of proteolytic cleavages in the heavy chain (51).

T stands for a trypsin split, and Pap for papain. T^{O} is a trypsin split at aminoethylcysteine, and P stands for a pepsin cleavage. The numbers indicate the number of residues between each of the splits.

found that IgG could be separated into different antigenic subclasses according to the antigenic determinants on the heavy chain. There are now 4 recognized subclasses of human IgG called IgG1, IgG2, IgG3 and IgG4 (12).

In different developments, Haber (59) and Whitney and Tanford (61) proved that the amino acid sequence alone determined the 3 dimensional structure of the binding site of a immunoglobulin molecule. Working with the Fab fraction of specific antibody they reduced all disulfide bonds in urea or guanidine hydrochloride. They removed the reducing agent and denaturing agent and allowed the polypeptide chains to reoxidize with the solution exposed to air. The binding activity was recovered. Sela (62) managed to accomplish the same thing with whole IgG by reacting the lysine of the IgG with polyalanine before reduction of the disulfide bonds.

The amino acid sequences of the homogeneous Bence Jones proteins began to be published and analyzed. The first complete light chain sequence to be published was by Wikler et al. of the lambda chain Sh (63) but even before this, portions of several chains had been published and a clear pattern was quickly observed. Even in the first comparison of 2 kappa chains (64), Cummings and Roy, it was obvious that the C-terminal half of the two chains were almost identical while the N-terminal fractions showed much variation. Shortly thereafter another partial sequence, Ag, by Titani et al. (65) added to the comparison and the pattern was the same. Further evidence was quickly received from Milstein (66, 67) which confirmed these observations. Milstein also located the half cysteine involved in the interchain disulfide bonds (68) at the C-terminal of the

light chain (see Figure 2 below) and positioned the intrachain disulfide bonds (69) for the light chain.

The story for the heavy chain was more slowly unfolded. Early work by Frangione and Franklin (70) showed the heavy chain to contain a variable and constant portion like the light chain. This was done with peptide maps of the Fd and Fc sections of the chain. Subsequently they proposed the N-terminal half of the Fd region was variable (71). At about the same time, Nelson et al. (72) published the tryptic peptide maps of Fab, Fc, light and heavy chains from specific and nonspecific rabbit IgG. They were able to show heterogeneity in the light chains, but since they did not have enough Fd to properly work with, they only speculated about the existence of heterogeneity in the heavy chain.

Since the heavy chain is twice as long as the light chain, its myeloma proteins are not usually excreted in the urine, so they are more difficult to obtain. Therefore, other approaches to the heavy chain structure were tried. Sequence studies were undertaken on non-myeloma proteins. Early efforts to determine the primary structure came from Givol and Porter (73, 74) who isolated the cyanogen bromide peptides of heavy chain from rabbit IgG and characterized the C-terminal fragment in addition to placing the fragments in order. The following year Hill et al. (75) reported the amino acid sequence for the majority of the Fc fragment and Wilkinson et al. (76) reported the amino acid sequence of the N-terminal fragment that is split from some rabbit IgG molecules by cyanogen bromide cleavage. Since these early advances, Porter's group (77, 78, 79) has nearly completely finished the amino acid sequence of the constant portion of the heavy chain of rabbit IgG. Fleischman (80)

has sequenced the N-terminal 65 residues from the heavy chain variable region of a rabbit with restricted heterogeneity. At this time, only about 30 residues of the rabbit heavy chain are not sequenced.

The intrachain and interchain disulfide bonds came under study also. Pink and Milstein (81) isolated the peptides containing the inter heavy-light chain disulfide bridge from several myeloma proteins and normal pooled IgG. At the time they were unable to position the peptides from the heavy chain except they knew the peptides came from the Fd portion of the molecule. They then reported (82) two intrachain disulfide loops in the Fc portion of the heavy chain that were much like those found on the light chain (69). This same type of structure was then extended to the Fd chain (83). Disulfide bridges between the heavy chains show large differences. There is 1 disulfide bridge in most rabbit IgG (39) and as many as 5 in some human IgG3 molecules (84). It is likely that they are all parallel (84). On the other hand, light-heavy interchain bonds have been located in 2 different positions on the heavy chain (85). Edelman (86) placed the light-heavy interchain bond near the center of the heavy chain for the myeloma protein Eu. O'Donnell *et al.* (87) and Strausbauch *et al.* (88) have placed this bond alongside the first half cysteine residue in the second intrachain disulfide loop for rabbits and goats respectively. O'Donnell (87) also places an extra intrachain disulfide bond in the rabbit IgG. It helps form the second loop in the Fd portion of the molecule. All the disulfide bonds for the IgG1 myeloma protein Eu are shown in Figure 2. Frangione *et al.* (89) have worked out the interchain disulfide bonding for each of the subclasses of human IgG (Figure 3).

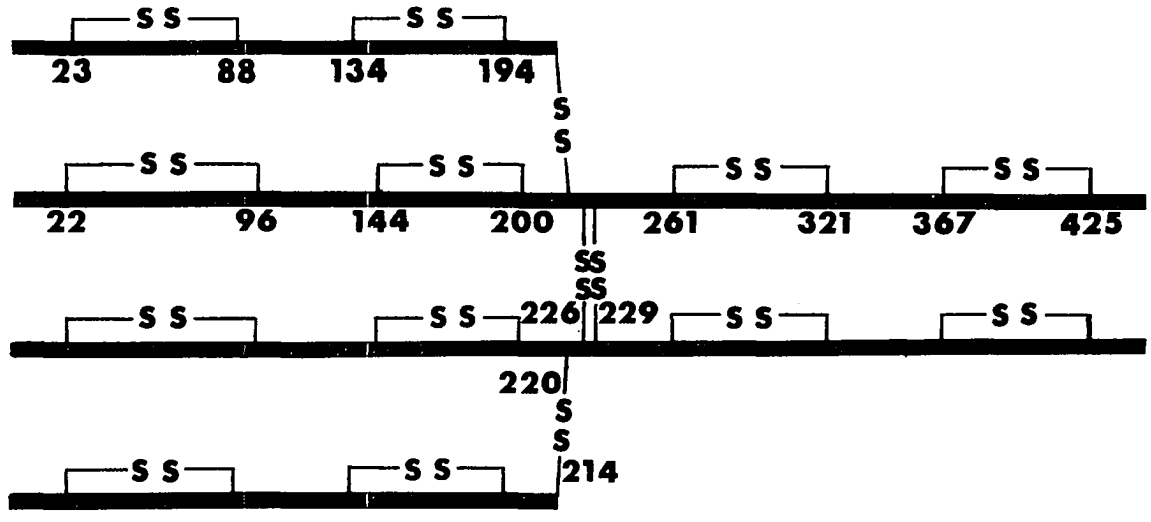

Fig. 2 - The interchain and intrachain disulfide bonds of Eu (86).

The residue number of the intrachain disulfide cysteine residues is indicated in the upper half of the molecule, and the interchain cysteine residues are indicated in the lower half.

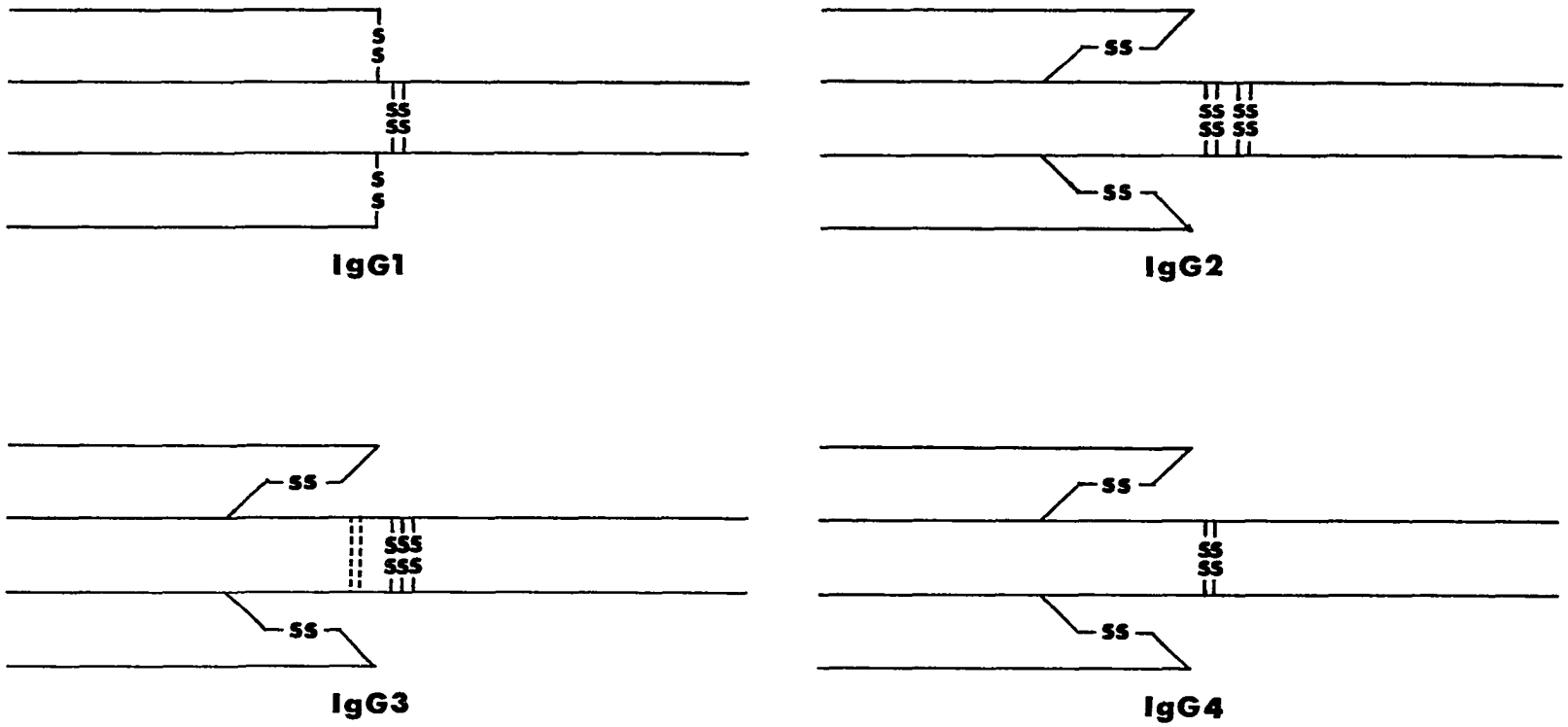

Fig. 3 - The interchain disulfide bonds of the 4 types of human IgG (84).

Dashed lines in IgG3 indicate disulfide bonds that have not been positioned exactly.

One of the advances most important to the understanding of the structure and evolution of immunoglobulins came when Singer (90) saw a relationship between peptides from the C-terminal portion of light chains and the corresponding C-terminal 18 residue sequence from heavy chains. This information was combined with the information from his own work (90) which indicated that a relationship also existed between the variable portions of light and heavy chains. He then proposed that the structural genes that code for light and heavy chains are related. Hill *et al.* (91) were able to expand on this finding as the sequence data for the Fc region of rabbit IgG was established. They were able to show internal homology between the N-terminal and C-terminal halves of the Fc region and from the two halves of Fc with both halves of the light chain Ag. They then proposed a precursor protein of 110 residues that upon gene doubling became a primitive light chain. From this primitive light chain evolved the kappa and lambda chains of today. This primitive gene also went through another doubling to form the heavy chains of today.

Much work was also done on myeloma heavy chains. Porter's group (92, 93) sequenced the N-terminal part of the myeloma heavy chain Daw and several other partial sequences have been reported (94). Another important breakthrough came when Edelman *et al.* (86) reported the sequence of an entire IgG molecule. This included the light and heavy chain of the myeloma protein Eu plus the position of the disulfide bonds. For the first time the relationships of an entire molecule could be studied. The sequence was added proof that Hill's original theory (91) was basically correct. Edelman and Gall (95) further proposed a "domain" theory in which each intrachain disulfide bond helped form a tight conformation or

domain with looser chain connecting the domains. Each domain on either variable or constant regions would contribute an active site to mediate a function of the immunoglobulin. Evidence at the present (96) time seems to uphold this theory.

During the time that much of the structural work was starting, the search to find the binding site of the antigen was also started. Porter's work (29) showed that the binding site was on the Fab fraction. Thus the binding site could be on the N-terminal half of the heavy chain, on the light chain or shared by both. Early evidence showed that antigens would combine with isolated heavy chains but not with isolated light chains (23). A series of experiments on recombination of light and heavy chains was conducted. Edelman et al. (97) showed that the recombination of light and heavy chains of a particular antibody led to a partial recovery of antibody activity. Heavy chains recombined with light chains from different antibodies showed an increase in binding of the heavy chain's antigen but not as great as with the homologous light chain. Light chains against a given antigen mixed with heavy chains against another antigen showed no activity against the light chain's antigen. Later, Yoo et al. (98) were able to demonstrate binding by the light chain by using a fluorescence-enhancement technique to determine the small amount of binding. Roholt et al. (99) showed that the formation of hybrid molecules using light and heavy chains from different antibodies did not involve the formation of disulfide bonds but that the association of the chains was strong enough to hold the molecule together without the formation of covalent bonds. Gray and Mannik (100) showed that heavy chains would bind with light chains from the same antibody in pref-

erence to other light chains even if the other light chains were in great excess. Bridges and Little (101) separated the light and heavy chains from 2 myeloma proteins that bound DNP and studied the recombinants. Light and heavy chain recombinants from the same proteins bound DNP better than either of the 2 hybrids. Dorrington et al. (102) studied recombination events with ORD and found the ORD pattern of the original molecule was recovered when heavy and light chains of the same antibody were allowed to reassociate, but an altered pattern appeared if a hybrid formed.

A different approach to the structure of the antibody combining site was made by Pressman and Roholt (103). A specifically purified antibody was cleaved into Fab fragments. Part of the antibody was incubated with I^{125} and another part with antigen and I^{131}. The two fractions were then combined and digested with pepsin. The resulting peptides were separated by high voltage electrophoresis. The I^{125}-to-I^{131} ratios in some peptides were different from that of the whole digest so these peptides probably came from the combining site of the antibody. In further work Roholt et al. (104) proved that the first two of Porter's fractions (29) were not identical by treating one of the fractions with I^{131} and the other with I^{125}. They then showed that after mixing together and digestion the I^{131}-toI^{125} ratios were not alike. By the same procedure Kitagawa et al. (105) showed that antibodies of the same specificity from different rabbits were different and later Roholt and Pressman (106) concluded from these differences that the antibody from each rabbit is composed of predominantly only a few different types of antibodies.

Another type of approach, affinity labeling, was used by Singer

and his group to try to discover the active site of the molecule. Wofsy et al. (107) proposed that affinity labeling should label the active site and showed by the absorption spectra that the rabbit protein was labeled. Metzger et al. (108) then showed that both light and heavy chains were labeled. Fenton and Singer (109) then reported that the ratio of labeled tyrosine in the heavy chain to the light chain was 2:1 in each of 3 different antibenzoid antibody systems, therefore these tyrosine residues were in relatively consistent positions. The next step was to isolate the peptides containing the labeled tyrosine. Doolittle and Singer (110) reported that the tyrosine residue from both chains were found in highly variable, hydrophobic peptides of about 25 residues in length. They concluded that these peptides were from the variable portion of the antibody chains. A later publication by Good et al. (111) extended these observations to sheep, mice and guinea pig. After further work (112), dipeptides containing much of the labeled tyrosine were isolated and each type of chain showed a predominant dipeptide. For instance a mouse kappa chain showed an aspartic-tyrosine labeled dipeptide. Since none of the constant region tyrosine residues from mouse myeloma proteins has an aspartic residue preceding it, Singer concluded that the residue must come from the variable portion of the chain. The most likely position would be that of residue #86 which is preceded by an aspartic in most mouse light chains. This pattern of a constant tyrosine residue preceded by another constant residue, which differs for each chain and species, holds true for several species. Data obtained by Goetzl and Metzger (113) from labeled DNP binding myeloma protein was interpreted by them to show that tyrosine #34 and not #86 was labeled in this protein.

Sela (62) in different studies was able to show that a correlation exists between the net charge of the antigen and the antibody that binds it. He showed this by making antibodies against several basic and acidic proteins followed by DEAE Sephadex chromatography of the antibodies. The antibodies against the acidic proteins always came off the column in one peak and those against basic proteins were in a different peak. Another experiment showed the size of the binding site was of the order of a tri- or tetra-peptide. Earlier studies by Kabat (114) had shown the upper limit to be the size of a hexasaccharide.

Since Porter's (29) discovery of the Fc portion of the IgG molecule several functions have been associated with this portion of the molecule. It has been indicated to function in complement fixation, skin sensitization, placental transfer, intestinal absorption and control of catabolism (115). Although its amino acid sequence has been published, its function has not been related to the structure. Study of the Fc chains by Goodman (116), Utsumi and Karush (117), Henney et al. (118) and others led to no conclusive proof as to the location of these functional sites. The most informative work along this line was by Prahl (115) who has proposed that the functional site for skin reactivity is on the N-terminal 35-40% of the Fc region. The peptide that represents the rest of the molecule (residue #334-446 on Eu), retains most of its structural integrity as evidenced by its ability to be crystallized, but it has almost no skin reaction. This peptide in high doses was able to cause a slight skin reaction but it was doubtful if the results were meaningful. Any digestion of the N-terminal portion of Fc was effective in causing loss of skin reactions.

Electron microscopy has been helpful in bringing other insights into the structure of immunoglobulins. Feinstein and Rowe (119) working with IgG noticed that the molecule was almost linear when no antigen was bound but that the molecule assumed a "Y" shape when antigen was bound. The angle of the Y varied somewhat. They then proposed a "hinge" section in the molecule. The hinge is the same section of the molecule described by Noelken (45) as loosely coiled and the site of the proteolytic cleavages that enable the large sections of the molecule to be separated. Valentine and Green showed by electron microscopy IgG bound to antigens forming circles consisting of from 3 to 5 IgG molecules. Each IgG molecule was bound to 2 antigens and each antigen was bound by 2 molecules of IgG (120).

It appears that the 3 dimensional structure of the IgG molecule will soon be known. Poljak (121) reports that he has the 6 Å diffraction pattern for an IgG molecule and that the data is good to 2.5 Å. Therefore it will not be long before the 3 dimensional relationships of an entire IgG molecule will be known.

Over the past few years many theories about the evolution of the immunoglobulin structure and mechanism of its synthesis have been proposed. At the present time the most popular theory on its biosynthesis is the "two gene one polypeptide chain" theory. An early exponent was Gray (122); whose theory has been recently reviewed by Williamson (123). The theory basically states that the variable and constant region of a given immunoglobulin polypeptide chain arise from different genes. This phenomenon is probably controlled at the DNA level, so a mechanism must be available to transcribe 2 distinct genes into one mRNA molecule.

This theory further proposes many variable region genes and only a few constant region genes. Also it is thought that light chains have a certain set of variable region genes and heavy chains have a different set.

If the 2 gene 1 polypeptide chain theory is correct, other theories would then have to be modified. Hill's (91) theory of the evolution of the immunoglobulins is one of these. If the genes for the variable and constant regions of the light chain are separate, then doubling of the light chain gene would not have achieved a heavy chain. Litman et al. (124) have proposed some modifications. He proposes that the heavy chain found in existing species comes from 1 (5.75 S immunoglobulins in the duck) or more duplications (2 in IgG, 3 in higher molecular weight heavy chains such as the 7.8 S duck IgM) of the 110 residue segment following the variable region.

A new theory proposed by Delaney (125) goes one step further back in time than Hill's (91) theory. By close examination of the amino acid sequences of Eu (86) he has been able to demonstrate homology between successive segments of the light and heavy chains only 55 residues long. In light of these findings and other theories that have been discussed, a chain of 55 residues doubled to 110 residues. This 110 residue segment duplicated forming the variable and constant precursors. The mechanism for joining variable and constant genes was probably the next factor introduced. The multichain structure appeared presenting 2 chains similar to light chain dimers. Next followed doubling of the constant gene of 1 chain, introducing the heavy chain similar to the heavy chain in duck 5.75 S immunoglobulin. Advent of the 4 chain structure occurred, and further elongation of the heavy chain happened once to produce the

gamma chain, twice to produce the duck 7.8 S immunoglobulin. The evolution of the immunoglobulin chains is shown in Figure 4.

Since it is thought that kappa chains, lambda chains and heavy chains have different sets of variable genes (123) (only 1 set of variable genes for all heavy chains), it is probable that before light and heavy chains diverged there was a doubling of the entire DNA section containing the constant region and all of the variable portions for that early immunoglobulin. The same is true for the divergence of kappa and lambda chains.

As the studies of the primary structure of the IgG molecule continue to be made, most of the attention will be turned to the variable portion of the molecule. A method for the separation of the variable and constant fraction of the molecule would therefore be valuable in these studies. One can see in the future that antibodies may be artificially manufactured to fight one or another disease or infection. The variable region of an antibody against the given antigen could be sequenced. This section could be manufactured and linked to an IgG constant chain giving an effective antibody.

Karlsson *et al.* (126) have proposed a method involving light chains and have used this method to separate the variable and constant regions by limited proteolysis with trypsin, pepsin, or papain. This method results in low yields with only a maximum yield of 22% reported. Therefore a search for a better method would be worthwhile. This dissertation proposes a method that could be more effective in the separation of variable regions from constant regions in both light and heavy chains.

When the internal homology in Edelman's (86) Eu sequence was

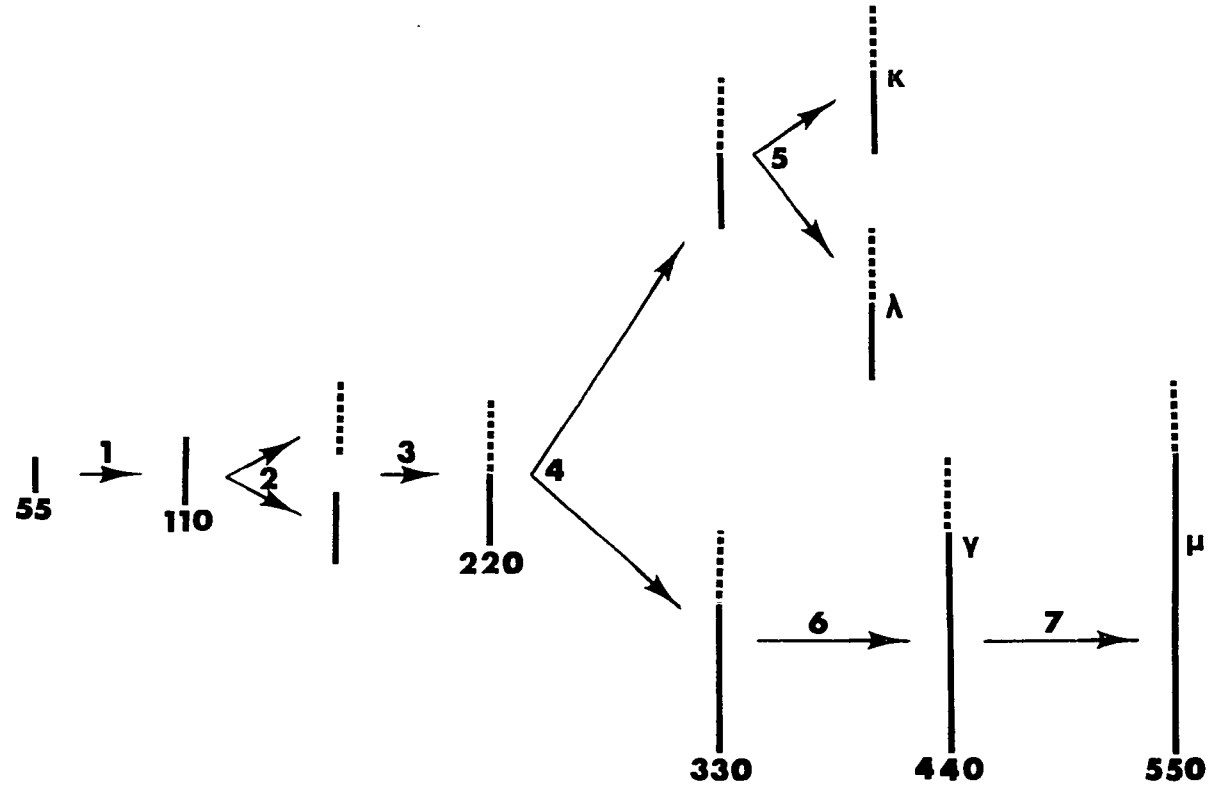

Fig. 4 - The evolution of immunoglobulin chains.

A 55 residue precursor chain doubles to form a 110 residue precursor chain (1). This precursor divides to form variable regions (dashed line) and constant region (2), and a mechanism for joining them is introduced (3). Another doubling occurs (4), to form light and heavy chain precursors. Kappa and lambda chains are formed from the light chain precursor (5), and elongation of the constant regions of the heavy chains forms the present day heavy chains.

examined, it appeared that certain tryptophan residues in the molecule were in invariant positions. One of these residues was close to the variable-constant region dividing line. If cleavage occurred at this residue, it would be possible to obtain a variable-constant region split which would also produce a constant region fragment that would contain much of the Fd sequence unreported at that time. The constant tryptophans in the IgG molecule are shown in Figure 5. These tryptophan residues always occur from 10 to 15 residues after the first half cystine residue in each intrachain disulfide bond. There are also other tryptophan residues in the heavy chain that might be constant. One follows the first tryptophan by about 15 residues and the other 2 are in the Fc region. There have not been enough heavy chains sequenced to be able to state whether these three are constant or not, but they have appeared in all IgG heavy chains sequenced so far. Edelman (96) also noticed the consistency of the tryptophan residues and that the section of the molecule that we were trying to sequence was very hydrophobic.

Dayhoff (127) has shown that tryptophan is the amino acid residue that retains its location in an evolving polypeptide chain better than any other residue. Cysteine is the next most constant residue. Dayhoff (127) has published a matrix indicating the tendency of a given amino acid to remain the same or change to another amino acid within a given accumulated number of mutations. If an amino acid chain of 100 residues undergoes a single point mutation one of the residues in this chain may change. Dayhoff has calculated that after 256 point mutations in a polypeptide chain 100 residues long a tryptophan residue will remain in its original position 48% of the time while cysteine will remain 41%

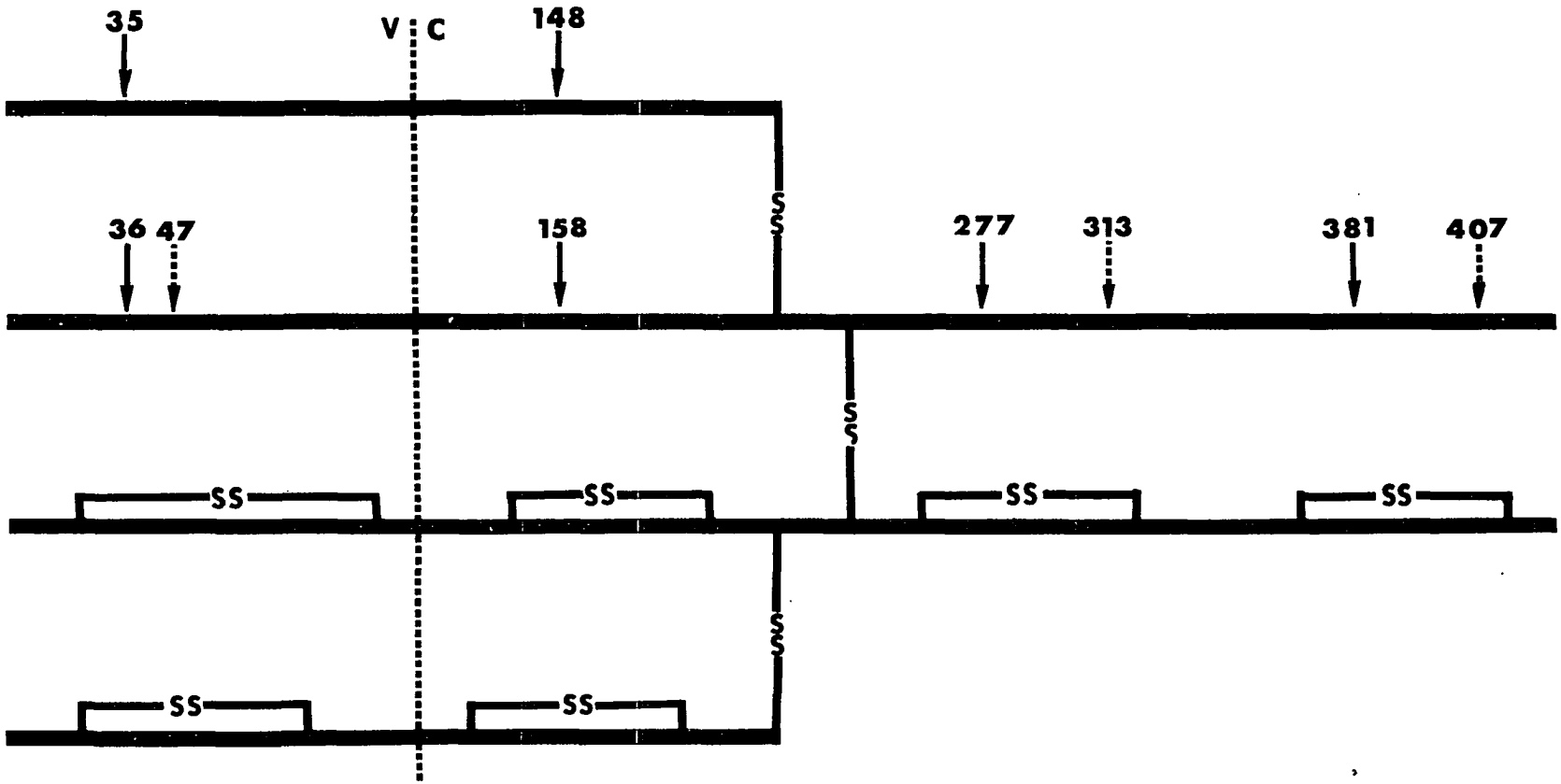

Fig. 5 - The constant tryptophan residues in Eu (86).

The constant tryptophan residues are indicated by arrows and residue numbers in the upper half of the molecule. The dashed arrows indicate residues in the heavy chain that may be constant, but not enough chains have been sequenced to say for sure.

of the time. At the other extreme, methionine will remain only 3% of the time.

It is easy to see why the selective processes of nature help cysteine residues remain in a fixed position so well. Cysteine residues are usually involved in disulfide bonds or in the active site of a molecule. If a mutation occurred that destroyed a cysteine residue, that particular disulfide bond would not be made or the active site would be inactivated. The resulting conformational change of a missing disulfide bond of the protein would probably make it inactive as an enzyme or unsuitable as a structural protein. Therefore such mutations would be highly selected against.

Why tryptophan mutations are so strongly selected against is not so clear. Tryptophan has only one codon, UGG, so one might think that this has something to do with its ability to survive. Although methionine has only one codon, it is more often replaced so this explanation in itself is not satisfactory. It is probable that the disappearance of methionine would be more typical of a single codon, because any mutation of that codon would change the amino acid to something else. The reverse mutation back would be unlikely since only one codon codes for it. Therefore, it is doubtful that the single codon has a great effect on tryptophan immutability.

For an answer to this question, consider the nature of tryptophan. It is the largest of the amino acids, contains a ring structure and is relatively hydrophobic. Dayhoff (127) has noted that the larger amino acids do not mutate very rapidly, so size must have something to do with the tendency of tryptophan mutations to be rejected by natural

selection. The amino acids that replace tryptophan best are tyrosine and phenylalanine (127). Both of these are large, ring-containing amino acids, but the side chain of each only contains 1 ring as opposed to the 2 rings of tryptophan. Since tryptophan is hydrophobic, it would be oriented so that the side chain is pointed into the hydrophobic interior of a folded protein. If a mutation occurred, the tryptophan would be replaced by a smaller residue. This would leave a hole in the interior of the molecule. If the residue replacing it were large and hydrophobic, it would fill most of the void, and possibly, there would be little rearrangement of the tertiary structure of the protein necessary to fill the void. If, on the other hand, the residue was small, a larger hole is left, and rearrangement of the protein would result with some modification of activity. If the residue were hydrophilic, chain rearrangement would almost certainly take place to expose the side chain to the external environment causing loss of biological activity. Therefore, the changing of a tryptophan residue to almost anything else could easily cause loss of biological activity.

If an IgG chain were cleaved by the oxidation of the tryptophan residues, we can predict what should happen. In the case of the light chain there are 2 constant tryptophan residues (Figure 5). One of these is approximately 35 residues from the N-terminal, and the other is approximately 148 residues from the N-terminal. Complete cleavage at these points should give us 3 polypeptides. One of these polypeptides, the N-terminal, should be about 35 residues long. Another, the middle, is about 113 residues long and the last or C-terminal peptide would be around 69 residues in length. Since the molecular weight of the 3 expected pep-

tides is quite different, it should be possible to separate them by molecular sieving. The variable region and about 40 residues from the constant region will be contained in the N-terminal and middle section peptides, while the remainder of the constant region will be contained in the C-terminal peptide.

An analogous situation can be achieved in the heavy chain. Any of the cleavages mentioned earlier (papain, pepsin, trypsin, or cyanogen bromide) to give an Fab type fragment produce a polypeptide chain of approximately the same length as the light chain. The extra constant tryptophan in the heavy chain might not complicate the mixture of products and prevent their resolution by gel filtration since it is close to the first constant tryptophan. The method of cleavage used to obtain the desired fragment was cyanogen bromide cleavage. This method yields a polypeptide of about 245 residues in length from the N-terminal half of the heavy chain.

There are several methods for the oxidization of tryptophan with the resultant cleavage of the adjoining peptide bond. They leave the oxidized tryptophan as the C-terminal end of a new peptide. Figure 6 shows a proposed mechanism for this cleavage. Although N-bromosuccinimide was used in the illustration, all of the reagents react through about the same mechanism. All of the reagents used for this oxidization have some drawbacks. They do not achieve 100% yields of chain cleavage, although all tryptophan molecules are oxidized. Other amino acid residues especially tyrosine and histidine can also be oxidized.

One of the most commonly used reagents for the oxidization of tryptophan and the one used for these experiments is N-bromosuccinimide.

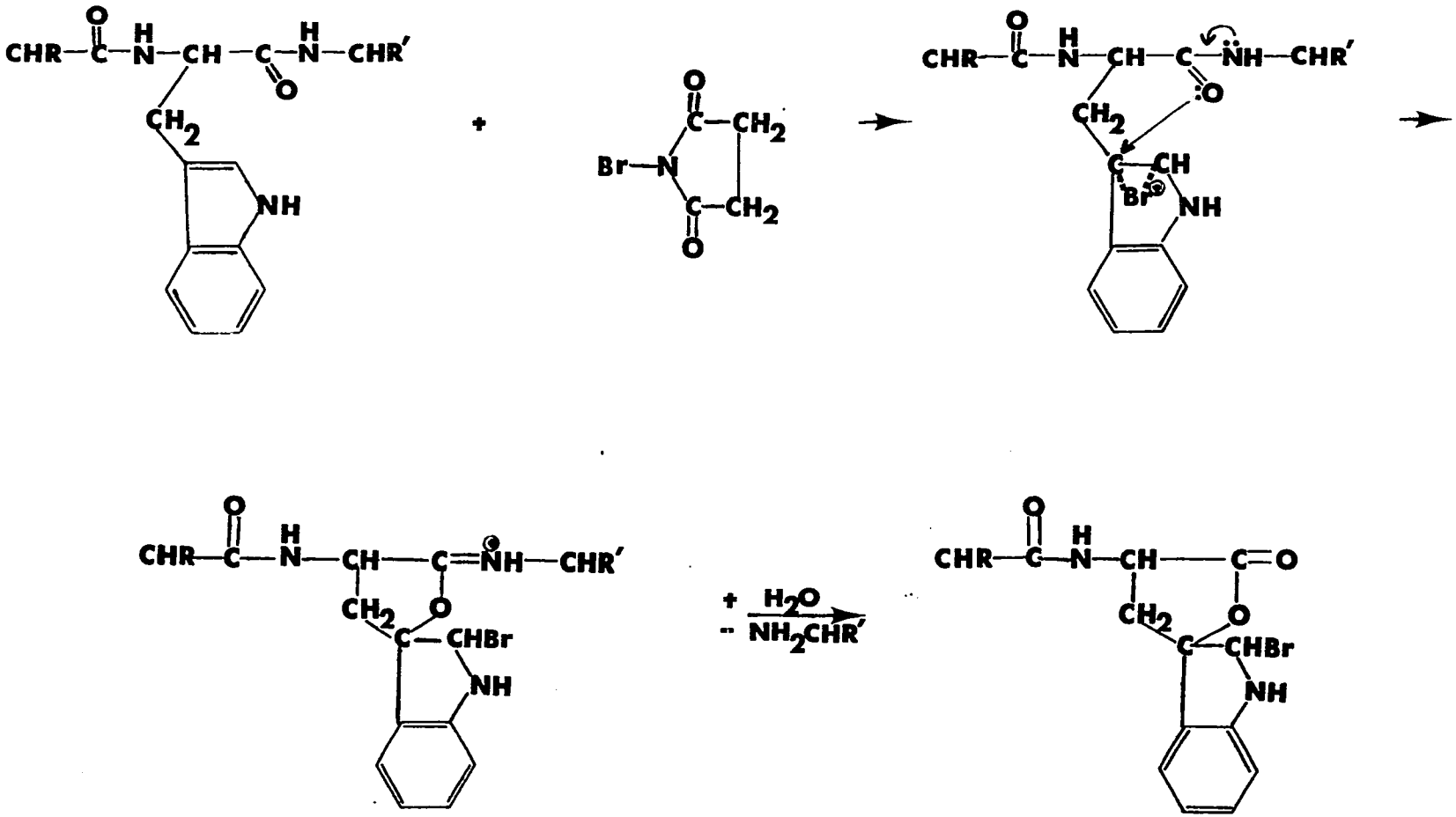

Fig. 6 - The reaction of tryptophan with N-bromosuccinimide to produce chain cleavage.

Witkop (128, 129, 130) has studied this reaction and discussed its application to protein chemistry. The cleavage takes place through the formation of a lactone structure. The formation of the lactone also causes the formation of a Schiff's base involving the nitrogen in the peptide chain. Addition of water to this Schiff's base causes the elimination of the nitrogen and the resulting cleavages of the peptide bond. The method produces a wide range in yield of bond cleavage, generally about 50%. There are several solvent systems that can be used for this reaction. For our purposes, it was decided to use a solvent of 8 M urea titrated to pH 4 with acetic acid. This solvent has two advantages over other solvent systems that were proposed. The 8 M urea serves to destroy tertiary structure exposing any tryptophan residues in the interior of the molecule to the reagent. It also serves to protect the tyrosine from oxidization at the concentration of reagent used to oxidize tryptophan, since the bromo urea intermediate is much slower at oxidizing tyrosine than tryptophan.

Another oxidation procedure for tryptophan has been reported by Atassi (131). In this method oxidation and cleavage are achieved in different steps, so the conditions are adjusted to achieve maximum yields. This procedure may eventually be the best for cleavage of immunoglobulin chains. It has the advantage of being able to achieve almost 100% cleavage of some particular peptide bonds, but on the other hand, some oxidized tryptophan bonds are not hydrolitically cleaved to any appreciable extent. Modification of tyrosine and methionine residues was reported as a side reaction of this procedure, but no chain cleavage occurred at these residues.

A third method for the oxidization of tryptophan has recently been published by Omenn et al. (132). 2-(2-nitrophenylsulfenyl)-3-methylindole was treated with N-bromosuccinimide. The product was then used to treat tryptophan containing peptides and proteins. The cleavage yields of this reaction were on the same order as those for N-bromosuccinimide. The great advantage of this reagent over N-bromosuccinimide is that it is specific for tryptophan and causes no side reactions with other amino acid residues.

The purpose of this work was to first prepare the light chains and N-terminal half of the heavy chains. Secondly, these fractions were to be oxidized by N-bromosuccinimide to produce cleavage at tryptophan residues. Thirdly, the products of this cleavage were to be isolated and characterized by N-terminal analysis, amino acid composition and peptides produced from them by proteolytic enzymes of known specificity. The products of the oxidation of light chains and the N-terminal half of the heavy chain can then be compared to previously published sequences and structure.

CHAPTER II

METHODS AND MATERIALS

Materials

The chemicals used were reagent grade or where necessary, analytical grade, and were purchased from various suppliers and used without further purification with one exception. Pyridine was redistilled after refluxing with ninhydrin to free it from reactive amines. The rabbit serum was purchased from Pel-Freez Bio-Animals, Inc., Rogers, Arkansas.

Methods

Operations, reactions and centrifugation were conducted at room temperature unless otherwise stated. Dialyses and vacuum dialysis were carried out in the cold room at 4°C.

Isolation of IgG

Two methods were used for the isolation of the IgG from the rabbit serum. The first method involved the use of sodium sulfate precipitation (133) followed by DEAE cellulose chromatography (134). The second method utilizes caprylic acid (octanoic acid) to precipitate nearly all the protein in serum except IgG (135).

A volume of rabbit serum was precipitated with 3 volumes of saturated (24%) sodium sulfate. The solution was centrifuged at 4080 x g

for 15 minutes and the supernatant was discarded. The precipitate was redissolved in its original volume of water. Saturated sodium sulfate was added to make the solution 6% with respect to sodium sulfate, and if any precipitate was present in the solution, it was centrifuged and the precipitate discarded. More saturated sodium sulfate was added to make the solution 12% with respect to sodium sulfate and it was centrifuged again. The precipitate was saved. The supernatant was made 18% with respect to sodium sulfate with further addition of saturated solution and was again centrifuged. The precipitate was saved and the supernatant was discarded.

The 6-12% fraction from the sodium sulfate precipitate was redissolved in 0.15 M sodium chloride and dialyzed against several changes of 0.0175 M phosphate buffer at pH 6.3. A column of Bio Rad Cellex-D, an anion exchange cellulose, was poured and equilibrated with 0.0175 M phosphate buffer at pH 6.3. The column bed was 76 x 5 cm. The 6-12% fraction was added to the column and the chromatogram was allowed to develop. After the first 2 peaks had been eluted from the column, the buffer was changed to 0.5 N sodium acetate and 0.5 N sodium chloride at pH 5.5. The column was further developed until no more protein was eluted from it. The separate peaks were pooled, dialyzed, and lyophilized. Polyacrylamide gel electrophoresis, double immuno-diffusion, and immuno-electrophoresis were run on these pools to check for purity.

The 12-18% fraction was dialyzed against water, lyophilized and stored in a freezer at -20°C until used. It was redissolved in 0.0175 M phosphate buffer at pH 6.3 and placed on a Bio Rad Cellex-D column equilibrated in the same buffer. The column was developed with that buffer

until the first peak had been eluted. It was then eluted with 2 M sodium chloride. Only the first peak was used. Disc gels were run to check the purity of the first peak.

The isolation of IgG from serum does not take quite as long with the use of caprylic acid, so this method was adopted after it was published (135). Rabbit serum was diluted with 2 volumes of 0.06 M acetate buffer at pH 4. The pH of the final solution was 4.8 $\pm$ 0.05 or it was adjusted to this pH. Caprylic acid, 6.8 g per 100 ml serum, was slowly added to the solution with vigorous stirring and the mixture was allowed to stir for 30 additional minutes at room temperature. The solution was centrifuged for 15 minutes at 4080 x g. The supernatant was saved, and the precipitate was resuspended in the same volume acetate buffer as originally added. This buffer was 0.015 M acetate at pH 4.8. The resuspension of the precipitate was aided by the use of a blender. The suspension was recentrifuged. The 2 supernatants were combined and filtered with a Buchner funnel and Whatman #1 paper to remove particles, mainly lipid, that had floated while centrifuging. The pH of the solution was adjusted to 5.7 and the ionic strength lowered by dialysis against 0.015 M acetate buffer. The salt concentration of the solution inside the dialysis bags was checked by freezing point depression. After the solution reached the freezing point of -0.09°C, it was batch absorbed with DEAE cellulose to remove small amount of impurities. A quantity of 12 grams of humid pressed DEAE-cellulose in equilibrium with 0.015 M acetate at pH 5.7 was used for every 100 ml of starting serum. The mixture is stirred for 15 minutes and a Buchner funnel was used to remove the DEAE cellulose from the solution. The solution was concentrated by vacuum

dialysis, dialyzed against water and lyophilized. The final product was tested for purity by disc gel electrophoresis.

Chain Separation

Two methods were also used to separate the light from the heavy chains. The first method was used by Fleischman *et al.* (19), in which the chains were partially reduced and blocked and chromatographed in propionic acid. The second method was complete reduction and blocking in 8 M urea.

For partial reduction a 2% solution of protein was made in 0.55 M Tris with 5×10^{-3} M EDTA at pH 8.2. In order to remove oxygen from the solution it was placed in a desiccator and stirred with a magnetic stirrer. The desiccator was evacuated with a vacuum pump until the solution started to foam. The pump was turned off and the solution was allowed to stir for a few more minutes under the vacuum. Nitrogen was then allowed to fill the desiccator and the solution was allowed to stir for another few minutes. The desiccator was reevacuated and refilled with nitrogen again. Mercaptoethanol was added dropwise to a concentration of 0.75 M. The solution was allowed to stand in the covered desiccator for 1 hour. The flask containing the solution was put in a cold water bath and a pH electrode was introduced. Ethylenimine was added slowly to a concentration of 0.75 M and the pH was kept at 8.2 by the addition of 6 N hydrochloric acid. The reaction mixture was allowed to stand in the covered desiccator for thirty minutes. It was then transferred to dialysis tubing and dialyzed against 0.1 M sodium chloride. After 3-4 changes of the dialysis medium the solution was dialyzed against 1 M propionic acid.

A column 5 x 150 cm was poured with Sephadex G-100 in 1 N propionic acid. The reaction mixture was chromatographed on this column to achieve chain separation. The peaks from the column were pooled separately. The chains in solution were precipitated with the addition of solid ammonium sulfate to 50% saturation for heavy chains and 65% saturation for light chains. The protein subfractions were suspended in a smaller volume of 0.1 M sodium chloride solution, dialyzed against water and lyophilized.

The chains were next completely reduced and blocked. They were dissolved in 8 M urea in 0.55 M Tris with 5 x 10^{-3} M EDTA at pH 8.2 giving a 2% solution of protein. Oxygen was removed from the solution as described before. Mercaptoethanol was added to the solution dropwise with stirring to a final concentration of 0.1 M. The solution was allowed to stand overnight. Ethylenimine was added to a final concentration of 0.12 M and the solution was allowed to stand for 15 minutes. More mercaptoethanol was added to make the total mercaptoethanol in the solution 0.13 M. The solution was then dialyzed against water and lyophilized.

Additionally, separation of the chains was achieved by complete reduction and blocking of the whole IgG molecule according to the procedure described for individual chains in the previous paragraph, followed by chromatography on a 5 x 150 cm column poured with Sephadex G-200 in 8 M urea at pH 6. After the chromatogram had developed the peaks were pooled, dialyzed against water and lyophilized.

Cyanogen Bromide Cleavage of Heavy Chains

Cyanogen bromide cleavage of the heavy chain proceeded according

to Givol and Porter (73). The amount of 88% formic acid added was calculated to give a 2% solution of the protein in 70% formic acid. The protein was dissolved in this amount of formic acid. Cyanogen bromide was weighed out in an amount that was twice the weight of the protein. It was dissolved in the amount of water required to make the 88% formic acid a 70% solution. The cyanogen bromide solution was added to the formic acid and protein solution and was allowed to stand in a covered desiccator for 24 hours. An amount of cold water 10 times the volume of the 70% formic acid solution was added and the resulting solution was lyophilized. A small portion of the product was hydrolyzed for amino acid analysis to determine completeness of the reaction. The protein was dissolved in 1 M acetic acid and chromatographed on a Sephadex G-100 column, 4 x 100 cm, equilibrated in 1 N acetic acid. After the chromatogram had developed the peaks were pooled. The first peak, the C1 fraction, was lyophilized and cleaved further with N-bromosuccinimide.

N-Bromosuccinimide Cleavage

N-bromosuccinimide cleavage of the heavy chain C1 fraction and light chains was done as described by Ramachandran and Witkop (129). The amount of N-bromosuccinimide to be used was first determined for both chains by the UV spectrum of the peptides treated with it. A solution of approximately 0.5 mg per ml protein in 8 M urea at pH 4 was scanned from 320 to 240 mμ. Small amounts of N-bromosuccinimide were added to the solution and after each addition of N-bromosuccinimide another scan of the solution was made. The 280 mμ peak decreased and then started to increase again. The point where the 280 mμ reading started to increase was taken as the end point. The amount of N-bromosuccinimide was calcu-

lated. Both chains used approximately 12 moles of N-bromosuccinimide per mole of polypeptide chain.

A 2% solution of the light chains was made in 8 M urea adjusted to pH 4 with acetic acid. N-bromosuccinimide in the amount of 12 moles per mole of protein was dissolved in a small volume of 8 M urea at pH 4. This was added to the protein solution and mixed. After 5 minutes the solution was placed on a Sephadex G-15 column, 2 x 100 cm, in 1 N acetic acid to desalt the protein. After desalting, the protein solution from the column was concentrated by flash evaporation and placed on a Sephadex G-100 column, 195 x 2.5 cm, equilibrated in 1 N acetic acid. After the chromatogram had been developed, the peaks were pooled and the amino acid analysis of the peaks was done.

The heavy chain Cl fraction was treated much the same way. Oxidation took place in the same manner, but desalting was done in 50% acetic acid and the G-100 column was chromatographed in 25% acetic acid. The peaks from the chromatogram were rechromatographed on the same column for further purification. They were flash evaporated to dryness and redissolved in a known volume of acetic acid at a concentration sufficient for complete solution. The peaks were analyzed for their amino acid content and end group.

Trypsin Digests and Peptide Maps

Trypsin digests were done in a 1 - 2% solution of the protein. The polypeptide material was dissolved in water or dilute sodium hydroxide. A few crystals of metacresol purple were added as a pH indicator and the pH of the solution was adjusted to the transition point of the indicator, pH 8.0. Trypsin was dissolved in 10^{-3} M hydrochloric acid and

a ratio of 1 to 50 of trypsin to protein was added to the protein solution. The pH was adjusted to the transition point of metacresol purple with the addition of 10^{-2} M sodium hydroxide. The solution was incubated at 37°C for 2-3 hour and the pH of the solution was kept adjusted to the transition point of metacresol purple with the addition of 10^{-2} M sodium hydroxide. The reaction was stopped by the addition of 0.3 volumes of acetic acid.

Peptide maps were made by Dowex 50 chromatography with a peptide monitoring system developed by Delaney (125). A gradient was developed by two gradient chambers in series. The first chamber contained 500 ml of 0.05 N pyridine in 2 N acetic acid and the second contained 500 ml of 2 N pyridine in 2 N acetic acid. The columns used were 0.9 x 60 cm and 1.2 x 60 cm containing Dowex 50 x 8. They were developed at a flow rate of 60 ml per hour. A portion of the effluent from the column was diverted for peptide analysis and the rest was collected in a fraction collector. The column effluent at 6 ml per hour was added to 6 ml per hour of 3 N sodium hydroxide by a proportioning pump and the solution was pumped for about 1 hour through coils which passed through a 97°C heating bath. A solution of 0.5 M borate containing 0.05% trinitrobenzenesulfonic acid was added at a rate of 18.4 ml per hour to the solution from the heating bath. This solution was pumped through coils at room temperature for about 40 minutes to allow the trinitrobenzenesulfonic acid to develop color. The solution was passed through a flow cell and the optical density at 430 mμ was recorded. The contents of the tubes from the fraction collector representing each peak were pooled and flash evaporated. They were dissolved in a known amount of water or acetic

acid and stored frozen.

Paper Electrophoresis

Paper electrophoresis was used for further purification of peptides and for ionic charge determination (136) of aspartic and glutamic acid containing peptides. The peptides were electrophoresed at pH 6.5, 3.5, and 2.1 for purification and at pH 6.5 and 2.1 to determine the ionic charge. The electrophoreses were run in an electrophoresis tank at 3000 volts for varying periods of time.

N-Terminal Analysis and Sequence Determinations

The N-terminal amino acid of the larger peptides was determined by a dansyl chloride procedure described by Gray (137). Approximately 5 nmoles of the peptide were transferred to a small test tube (6 x 50 mm) and dried. A solution of 0.2 M sodium bicarbonate in the amount of 10 µl and 10 µl of dansyl chloride in acetone was added (2.5 mg dansyl chloride per ml acetone). The pH of the solution was checked on pH paper to make sure it was approximately pH 9. If the pH was too low, it was adjusted by the addition of more sodium bicarbonate. The tubes were covered with parafilm and the reaction was allowed to proceed at 37° for 1 hour. The solutions were dried and 50 µl of constant boiling hydrochloric acid were added. The tubes were sealed and incubated overnight at 105°C. The hydrochloric acid was removed in a vacuum desiccator and the residue was dissolved in 10 µl of 50% pyridine and 50% water.

The end groups were identified by 2 dimensional chromatography on polyamide layer sheets (138). The dansyl hydrolysate was spotted on polyamide layer sheets. Half was spotted on each side of a sheet and on

one of these spots was placed a standard mixture of dansyl amino acids. The chromatogram was developed in the first direction by a water-88% formic acid solvent (200:3, v/v). The second direction was developed in benzene-glacial acetic acid (9:1, v/v). Identification of the end groups was done by comparison under a UV lamp of the spots on the standard containing side of the sheet with the spots on the other side. If the chromatogram contained spots for aspartic and glutamic acid or threonine and serine a third solvent, n-heptane-n-butanol-glacial acetic acid (3:3:1 by volume) was used to separate these pairs of spots.

Sequence determinations were done with 2 methods of Edman degradations described by Gray (139, 140). In the first method a peptide solution was transferred to a small screw cap test tube (3-5 ml). It was dried by flash evaporation. To the dried peptide were added 100 µl of 5% phenylisothiocyanate in pyridine and 100 µl of water. The test tube was flushed with nitrogen and incubated at 45°C for 1 hour. The solution was dried by flash evaporation and 200 µl trifluoroacetic acid was added to the tube. It was incubated for 30 minutes at 45°C and dried by flash evaporation. Water, 250 µl, was added and the solution was extracted 3 times with 250 µl of butyl acetate. An aliquot containing 1-5 nmoles was taken for dansylation and the remainder of the solution was dried by flash evaporation for another cycle of Edman degradation. As more cycles of Edman degradation are run, more of the peptide material is lost by side reactions which form products that do not undergo cyclization and release of the N-terminal amino acid, by extraction into the butyl acetate layer of the extracting system and by use in the dansyl procedure. So each step requires a larger aliquot for the dansyl procedure than the

step before. If a 10 residue peptide were to be sequenced, it would require 0.1 μmole or more of peptide to sequence the peptide. The first aliquot would only require about 1-2% of the sample and later aliquots for dansylation would require high percentages until, in the final step, 100% of the sample is taken. It is also clear that a short degradation by this method requires little material, but longer ones call for much higher amounts.

The aliquot taken for dansylation was placed in a 6 x 50 mm test tube and dried under a vacuum in a desiccator. After 20 μl of 0.05 M sodium bicarbonate was added to the tube, it was dried again to remove ammonia from the sample. Deionized water, 10 μl, was added to the dried residue in the tube and the pH of the resulting solution was checked by spotting a little on pH paper. The pH should be approximately 9. Next, 10 μl of 2.5 mg dansyl chloride per ml acetone were added and the solution was incubated for 1 hour at 37°C. The contents were then dried in the desiccator and 25 μl of constant boiling hydrochloric acid were added to the tube. After it was sealed, the tube was heated at 105°C overnight. The hydrochloric acid was dried and the dansyl amino acid was identified by polyamide layer sheet chromatography as described above.

The second method used for sequencing peptides was a later development by Gray (140) for use on short (5-6 residues or shorter) peptides. A peptide, 5-10 nmoles, was placed in several appropriately labeled 6 x 50 mm test tubes. One test tube was used for each amino acid residue in the peptide to be sequenced. The contents of the tubes were dried in a vacuum desiccator that was heated to 70°C. The tube containing the sample to be analyzed for the N-terminal residue was set aside.

To the other test tubes were added 50 µl of freshly made coupling mixture (600 µl of a 5% solution of phenylisothiocyanate in pyridine and 400 µl of water). The test tubes were covered with parafilm and incubated at 50°C for 45 minutes. The samples were then dried in the desiccator and 50 µl of trifluoroacetic acid was added to them. They were allowed to stand in the heated desiccator for 10 minutes after which they were dried. The tube containing the sample to be analyzed for the second residue was set aside and another Edman cycle was performed on the contents of other tubes. After each cycle, the tube representing the newly exposed residue was set aside. After about 5 cycles, the residue in the test tubes became difficult to dry and this placed a practical limit on the number of residues that could be sequenced by this method. It was possible to obtain a few more cycles by extracting this residue with the water/ethyl acetate extracting system as explained below.

After the Edman degradation cycles were completed, 40 µl of the aqueous phase of a water/ethyl acetate system was added to all the test tubes on which Edman cycles had been performed. These samples were extracted 3 times with 200 µl of the ethyl acetate phase to remove excess reagent and by-products. The contents remaining in the tubes were dried in the heated desiccator.

All samples were then dansylated. After 10 µl of 0.2 M sodium bicarbonate was added to the test tubes, the contents were dried to remove ammonia. Deionized water, 15 µl, was added to the samples and the pH of the solution was checked to be sure it was approximately pH 9. If the pH was correct, 15 µl of a solution of 3 mg dansyl chloride per ml of acetone was added to the sample tubes and they were covered with para-

film. The tubes were incubated at 50°C for 15 minutes. The samples were dried and 100 µl of constant boiling hydrochloric acid was added. The tubes were sealed and incubated at 105°C overnight. The tubes were then dried and extracted twice with 100 µl of the ethyl acetate phase of the water/ethyl acetate system described above. The ethyl acetate was transferred to another properly labeled test tube and dried. The residue was dissolved in 50% water-50% pyridine and run on the polyamide layer sheets as described before.

Protein Hydrolysis

Hydrolysis of proteins and peptides for amino acid analysis was done with constant boiling hydrochloric acid. After the sample to be hydrolyzed was placed in a test tube, it was dried and 1 ml of constant boiling hydrochloric acid was added. The tube was stretched, evacuated, and sealed with an oxygen torch flame. The hydrolysis tubes were placed in a heating block maintained at 110°C for 24 hours. The hydrochloric acid was removed by flash evaporation and the residue was dissolved in 2.5 ml of pH 2.2, 0.2 N citrate buffer. The amino acid analysis was done on a BioCal BioChrom Amino Acid Analyzer with a sample injector using a 2 column system.

Polyacrylamide Gel Electrophoresis

Polyacrylamide gel electrophoresis was run on a Canalco Model 6 apparatus. Standard gels in 4 x 66 mm tubes were used. They were 7% acylamide with a pH of 9.5. The stacking gel has a pH of 8.9 and an acrylamide concentration of 1.25%. A Tris-glycine buffer at pH 8.2 was used in the upper and lower baths. A tracking dye, bromphenol blue, was

placed in the upper bath. After sample application the gels were electrophoresed at 5 ma per tube until the tracking dye nearly reached the bottom of the gel. The gels were stained with aniline black in 7% acetic acid and destaining was done by electrophoresis using 7% acetic acid in both the top and bottom bath of the apparatus.

Immunoelectrophoresis

Immunoelectrophoresis was done using a Gelman apparatus. The electrophoresis was run in a 1% agar gel containing Gelman High Resolution Buffer on 1" x 3" microscope slides. After sample application to the agar, the electrophoresis was run at 360 volts for 45 minutes. After electrophoresis, the center strip was removed from the agar and goat-anti rabbit serum was placed in the center trough. After the precipitate lines had formed in the agar, the gels were rinsed in 0.9% sodium chloride overnight to elute any protein not precipitated. They were rinsed with distilled water and stained with aniline black. The slides were then rinsed with 7% acetic acid to remove excess stain and dried.

Double Immunodiffusion

Double immunodiffusion was done on a 3" x 5" glass slide. A 1% agar solution was poured on the glass about 1/32" thick and allowed to harden. Wells were cut in the agar, and the outside wells were filled with fractions from the DEAE cellulose column. The center well was filled with goat anti-rabbit serum. After bands developed (approximately 2 days) the agar was washed, stained and dried as described above.

CHAPTER III

RESULTS

A total of 25 gm of nonspecific rabbit IgG was isolated from pooled rabbit serum by 2 methods. The first method involved sodium sulfate precipitation followed by DEAE cellulose chromatography. The profile of the DEAE cellulose chromatography is shown in Figure 7. The peaks were pooled and the purity of each peak was checked by polyacrylamide gel electrophoresis shown in Figure 8, double diffusion in Figure 9, and by immunoelectrophoresis in Figure 10.

All of the checks for purity show that peaks A and B are quite clean preparations of IgG. The double diffusion plate and disc gel show only one band for these two peaks and the immunoelectrophoresis also shows them to be a fairly homogeneous preparation of IgG. The immunoelectrophoresis for peak B seems to have a flaw in the agar gel causing the IgG band to be unsymmetrical but the preparation has only 1 band. Peak I is nearly all IgG as shown by the purity checks. The disc gel shows only 1 faint band in addition to the IgG and the double diffusion plate shows 2 faint bands; one in the first half of the peak and a different one in the second half. The immunoelectrophoresis for peak I shows no band in addition to the IgG.

The other peaks were not as pure. Peak II shows 2 extra bands

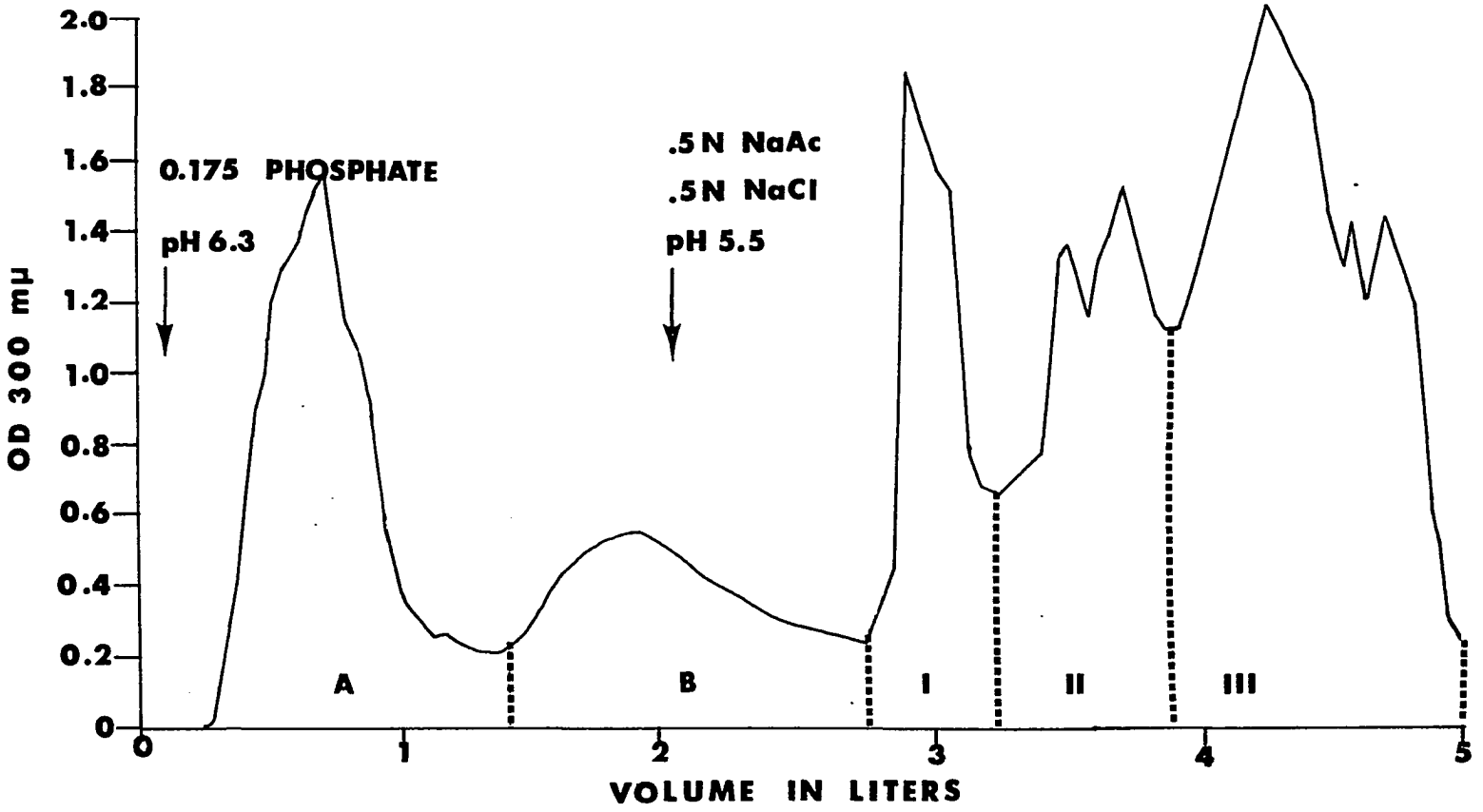

Fig. 7 - DEAE cellulose chromatography of rabbit IgG.

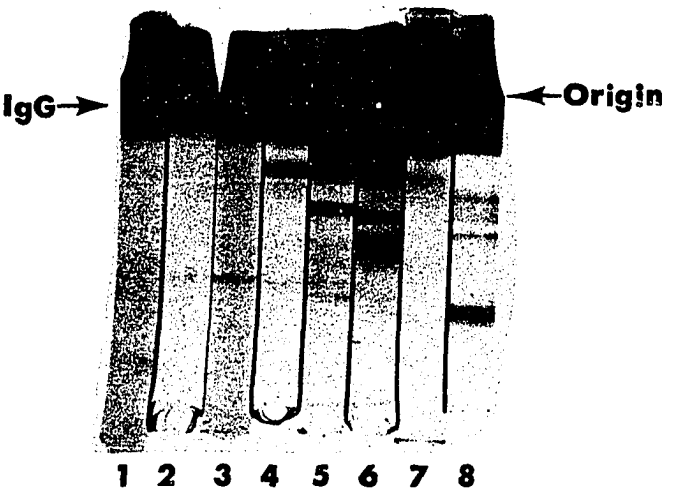

Fig. 8 - Polyacrylamide gel electrophoresis of IgG fractions.

Fractions 1, 2, 3, 4 and 5 are fractions A, B, I, II and III from the DEAE cellulose fractionation of rabbit IgG. Fractions 6 and 7 are the 12-18% sodium sulfate fraction before and after DEAE cellulose chromatography. Fraction 8 is the IgG isolated with caprylic acid.

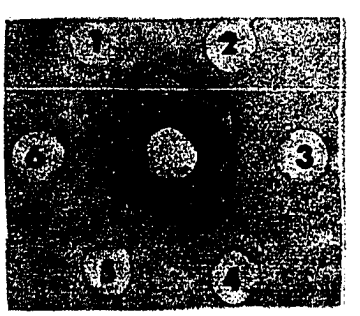

Fig. 9 - Double immunodiffusion of DEAE cellulose fractionation peaks.

The center well contains goat anti-rabbit serum. Wells 1 and 2 contain the center and tail from peak A. Wells 3, 4 and 5 contain the front, center and back of peak I. Well 6 contains peak II.

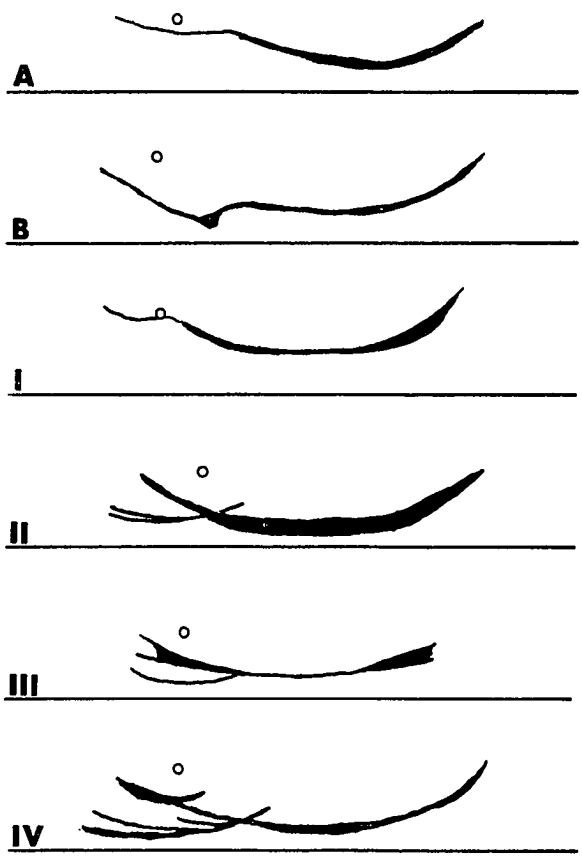

Fig. 10 - The immunoelectrophoresis of IgG fractions from DEAE cellulose chromatography.

The fractions are numbered as they eluted from the column, except fraction IV. Fraction IV is the 12-18% sodium sulfate fraction.

on disc gel and at least 2 faint bands by double diffusion. Immunoelectrophoresis also shows at least 2 extra bands. Peak III shows several bands by both immunoelectrophoresis and disc gel. Peak III is too contaminated to use for further experiments with the IgG but Peak II seems to have very little contamination in relation to the total amount of IgG, so it was used.

The 12-18% sodium sulfate fraction was contaminated with several other proteins as shown by both disc gel and immunoelectrophoresis. Chromatography of this fraction on DEAE cellulose yielded 2 fractions. The first, corresponding in elution to peak A and B, was about as clean as peak I as shown by disc gel electrophoresis and was also used.

The second method for isolation of the rabbit IgG used a caprylic acid precipitation of serum proteins which leaves IgG in solution. This fractionation produced a nearly pure IgG contaminated with small amounts of IgA and ceruloplasmin (127). These contaminants were removed by batch absorption with DEAE cellulose. The purity of these preparations was checked by polyacrylamide gel electrophoresis which is shown in gel 8, Figure 8. This preparation shows at least 4 faint bands in addition to the IgG band, but they are all very faint in comparison to the IgG band. These bands, therefore, constitute only a small portion of the total protein in the preparation. This fraction was also used.

Two methods for the isolation of light and heavy chains from the purified IgG were tried. Whereas the products obtained are of comparable purity, the shorter method of preparation (complete reduction in 8 M urea) has a serious drawback that makes the other method more preferable. The long exposure of protein to urea leads to some reaction of lysine

residues with a breakdown product of urea, cyanate. After 3-4 days in 8 M urea at room temperature, 20-30% of the lysine has carbamylated.

The chain separation by the two methods is shown in Figures 11 and 12. Each chromatogram on Sephadex G-200 and G-100 contains 3 peaks. In each chromatogram, the first peak contains heavy chain dimers and/or aggregates, the second peak contains heavy chain monomers and the third peak contains only light chains. In most reports in the literature (19) this separation is done on Sephadex G-75 which chromatographs all of the heavy chains in the first peak. The weight of heavy chain (peak I and II) to light chain (peak III) in these preparations was always in the range of 2 to 1 which indicates a good yield of chain cleavage. The amino acid composition of the light and heavy chains is given in Table 2.

The light chains were used first for tryptophan oxidation because of several factors. They were pure by-products of our fractionation. The light chain has only 2 constant tryptophan residues and would present a simpler pattern than the C1 fraction which has 3. Finally, the light chain is about 20 residues shorter than the heavy chain, so we would expect a better separation from the C-terminal and middle sections of the molecule.

The UV spectrum of the unoxidized light chains was recorded. N-bromosuccinimide was added until the absorbance at 280 mμ went through a minimum. This occurred at 12 moles N-bromosuccinimide per mole of light chain. The spectra of the unoxidized and oxidized light chains are shown in Figure 13.

After treatment with N-bromosuccinimide, the desalted light chains were then chromatographed on G-100. The results are shown in

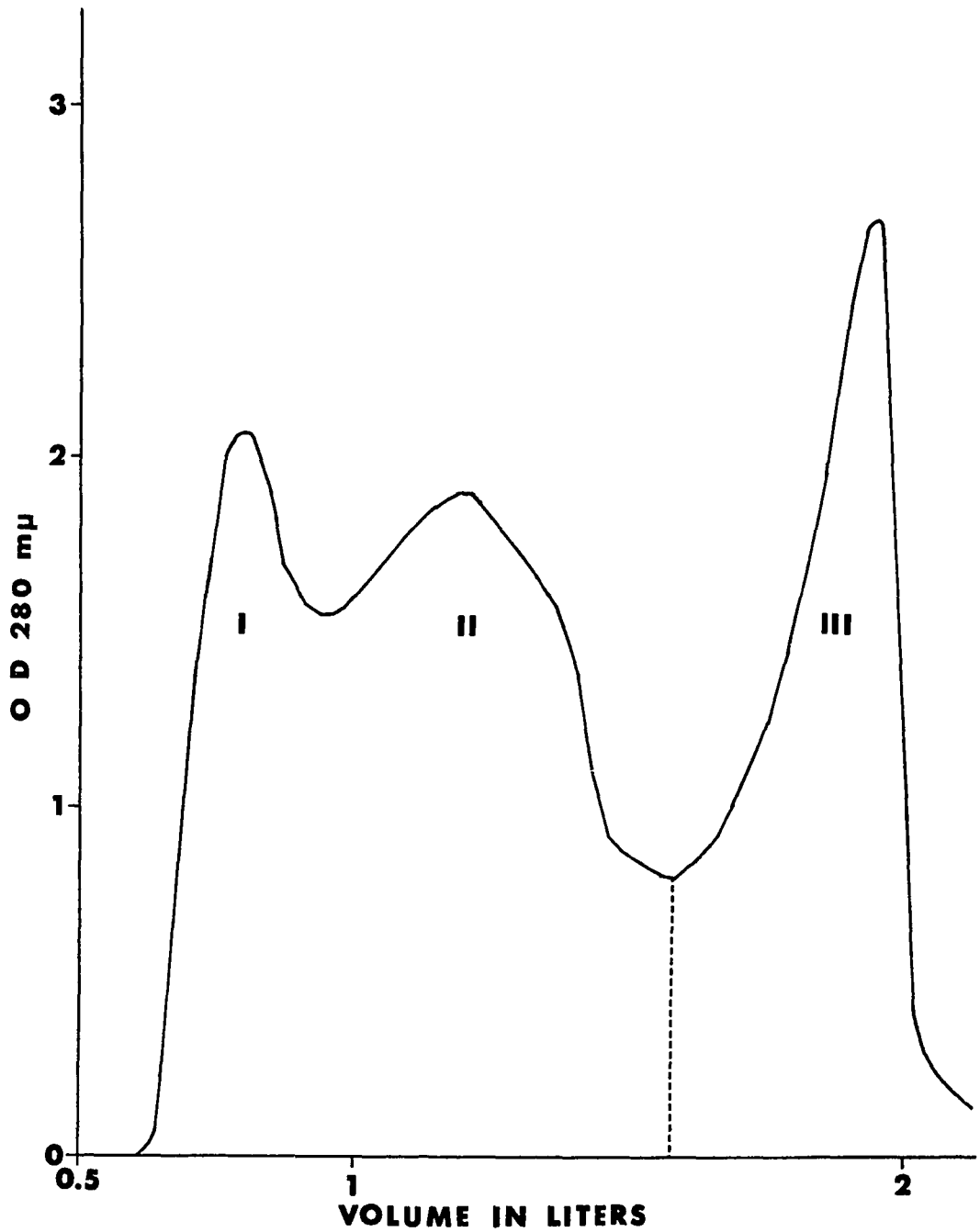

Fig. 11 - Chain separation in propionic acid.

Rabbit IgG was partially reduced and blocked and chromatographed on Sephadex G-100. The column was 5 x 150 cm, and was equilibrated and run in 1 N propionic acid.

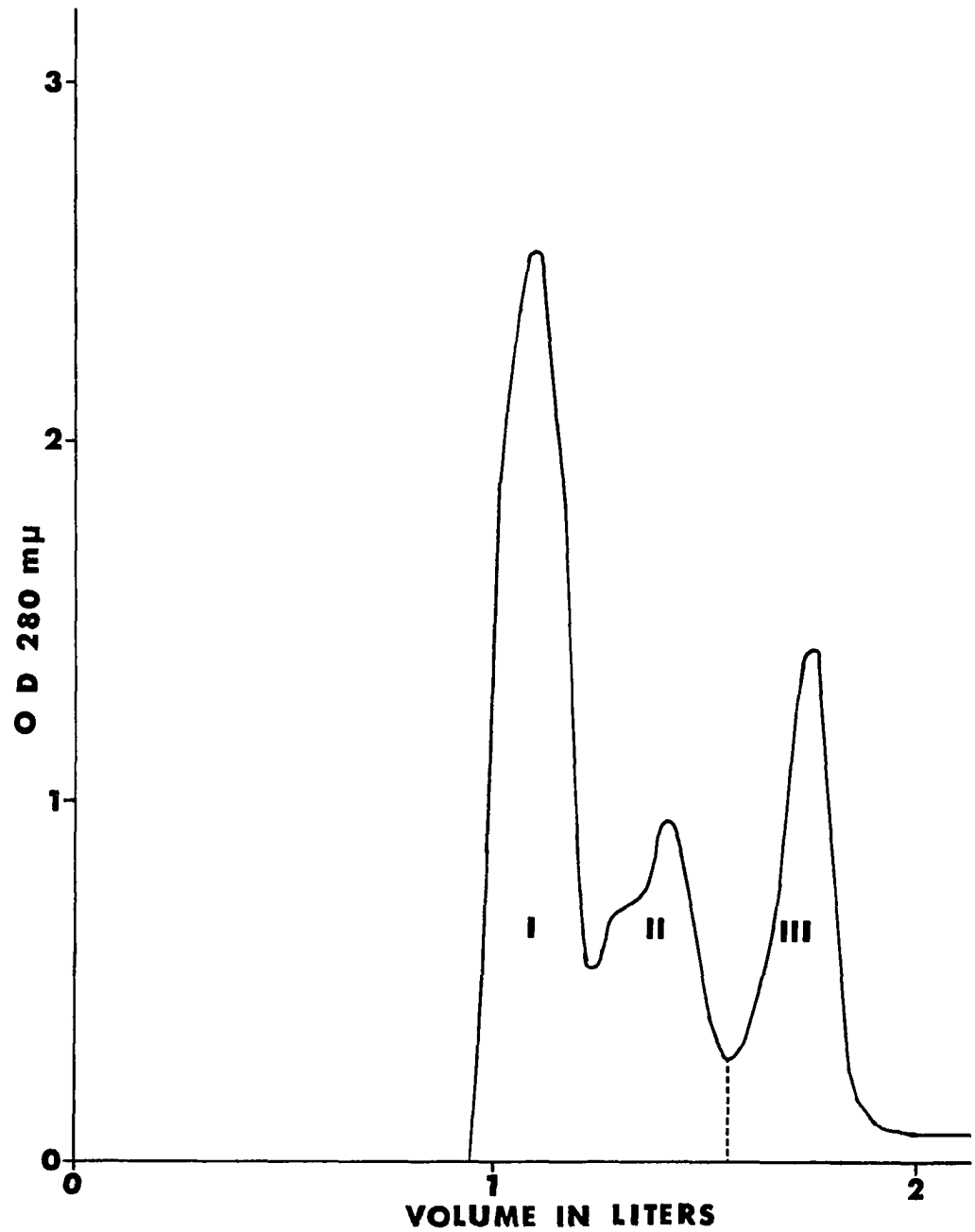

Fig. 12 - Chain separation in 8 M urea.

Rabbit IgG was completely reduced and blocked and chromatographed on Sephadex G-200 in 8 M urea. The column was 5 x 150 cm.

TABLE 2

AMINO ACID COMPOSITION OF LIGHT CHAINS, HEAVY CHAINS AND C_1 FRACTION

Amino Acid	Heavy Chain Residues	Heavy Chain[a]	Light Chain Residues	Light Chain[a]	C_1 Fraction Residues	C_1 Fraction[a]
Aspartic Acid	32.3	7.6	18.3	8.8	13.7	5.9
Threonine	46.2	10.8	22.9	11.0	32.8	14.0
Serine	47.1	11.0	23.9	11.5	30.9	13.2
Glutamic Acid	39.1	9.2	22.4	10.7	16.4	7.0
Proline	41.2	9.7	14.6	7.0	20.3	8.7
Glycine	30.9	7.2	19.6	9.4	22.4	9.6
Alanine	22.1	5.2	13.4	6.4	15.7	6.7
Valine	43.9	10.3	19.6	9.4	24.1	10.3
Cystine[b]	12	-	5	-	8	-
Methionine	6.7	1.6	1.0	0.5	-	-
Isoleucine	15.8	7.3	7.4	3.5	5.8	2.5
Leucine	29.0	6.8	11.3	5.4	17.9	7.6
Tyrosine	16.6	3.9	10.5	5.0	10.4	4.5
Phenylalanine	14.7	3.5	6.7	3.2	8.2	3.5
Lysine	18.3	4.3	11.4	5.5	7.9	3.4
Histidine	6.1	1.4	1.5	0.7	0.9	0.4
Arginine	16.2	3.8	4.6	2.2	6.5	2.8
Tryptophan[b]	8	-	3	-	4	-
Total	446	100	217	100	246	100

[a]Calculated without including tryptophan and cystine. Expressed as residues per 100 residues.

[b]Estimated amounts from literature.

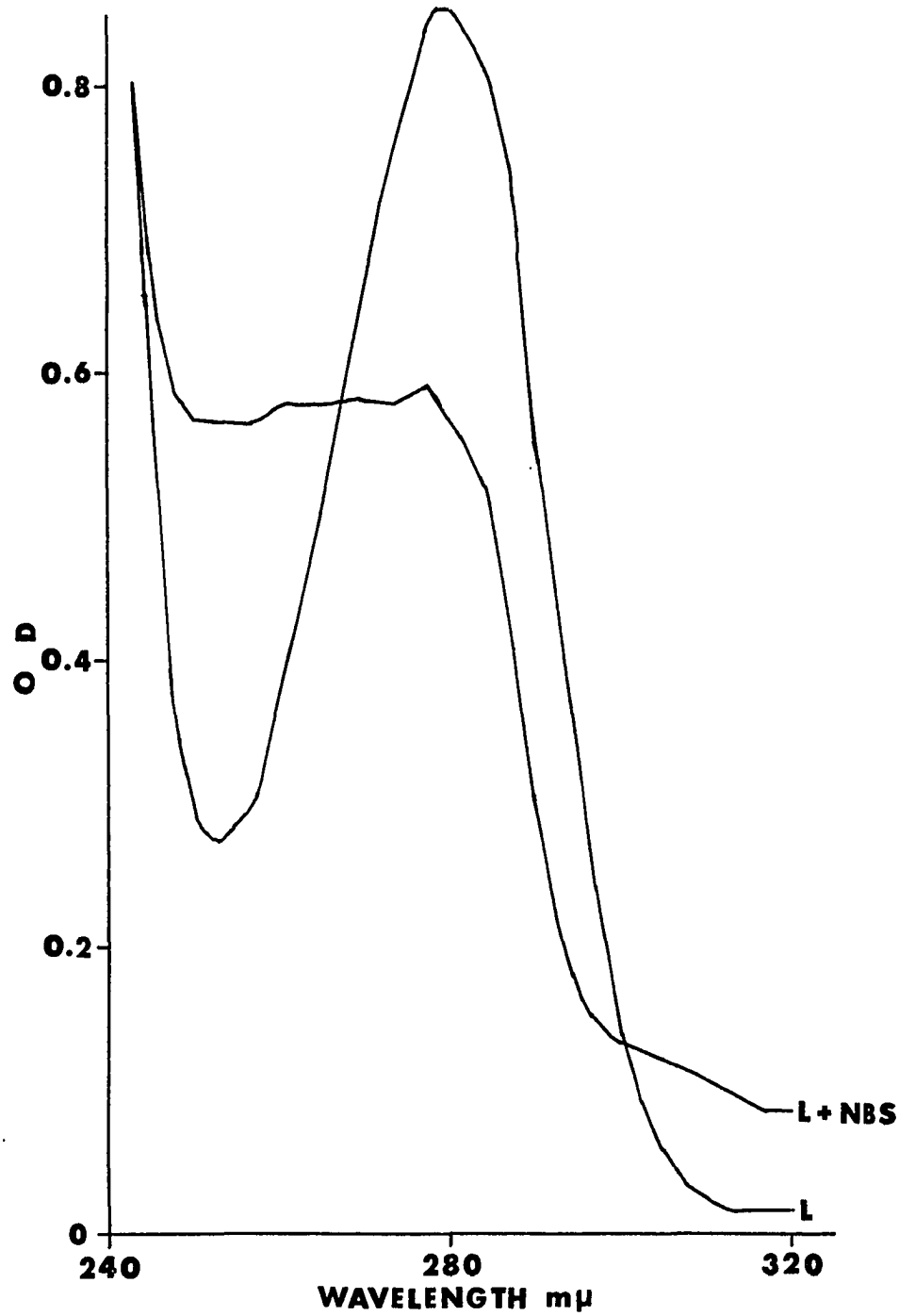

Fig. 13 - The spectra of unreacted and oxidized light chains.

Figure 14. The amino acid composition of the peaks is shown in Table 3. There was very little amino acid material in peak F. It consisted mainly of bromourea which absorbes at 280 mμ.

The N-terminal amino acid for each of the peaks was determined. All of the peaks showed multiple end groups. Peak B had an end group of tyrosine, threonine and serine. Peak C contained tyrosine, aspartic acid, glutamic acid, alanine and serine. Peak D contained lysine, tyrosine and aspartic acid, and peak E contained threonine, serine, glycine, alanine and lysine. These peaks were not rechromatographed.

The heavy chains (500 mg) were treated with cyanogen bromide and the resultant polypeptides were chromatographed on G-100. This chromatogram is shown in Figure 15. The first peak off the column contained the N-terminal half of the heavy chain. The amino acid composition of this peak is given in Table 2.

Work on the C1 fraction from the heavy chain was then started. A flow chart showing the reactions of the heavy chain is shown in Figure 16. The UV spectrum was used as above to determine the correct amount of N-bromosuccinimide to use. As in the light chain, 12 moles of N-bromosuccinimide per mole of C1 were required. The UV spectra are shown in Figure 17.

The C1 chains (250 mg) were treated with 12 moles of N-bromosuccinimide per mole of chain in a final volume of 12.5 ml and were desalted. The resulting peptides proved to be too hydrophobic to remain in solution in 1 N acetic acid. It was finally determined that 25% acetic acid would keep them in solution, so after desalting, they were chromatographed on Sephadex G-100 in that solution. The chromatogram that de-

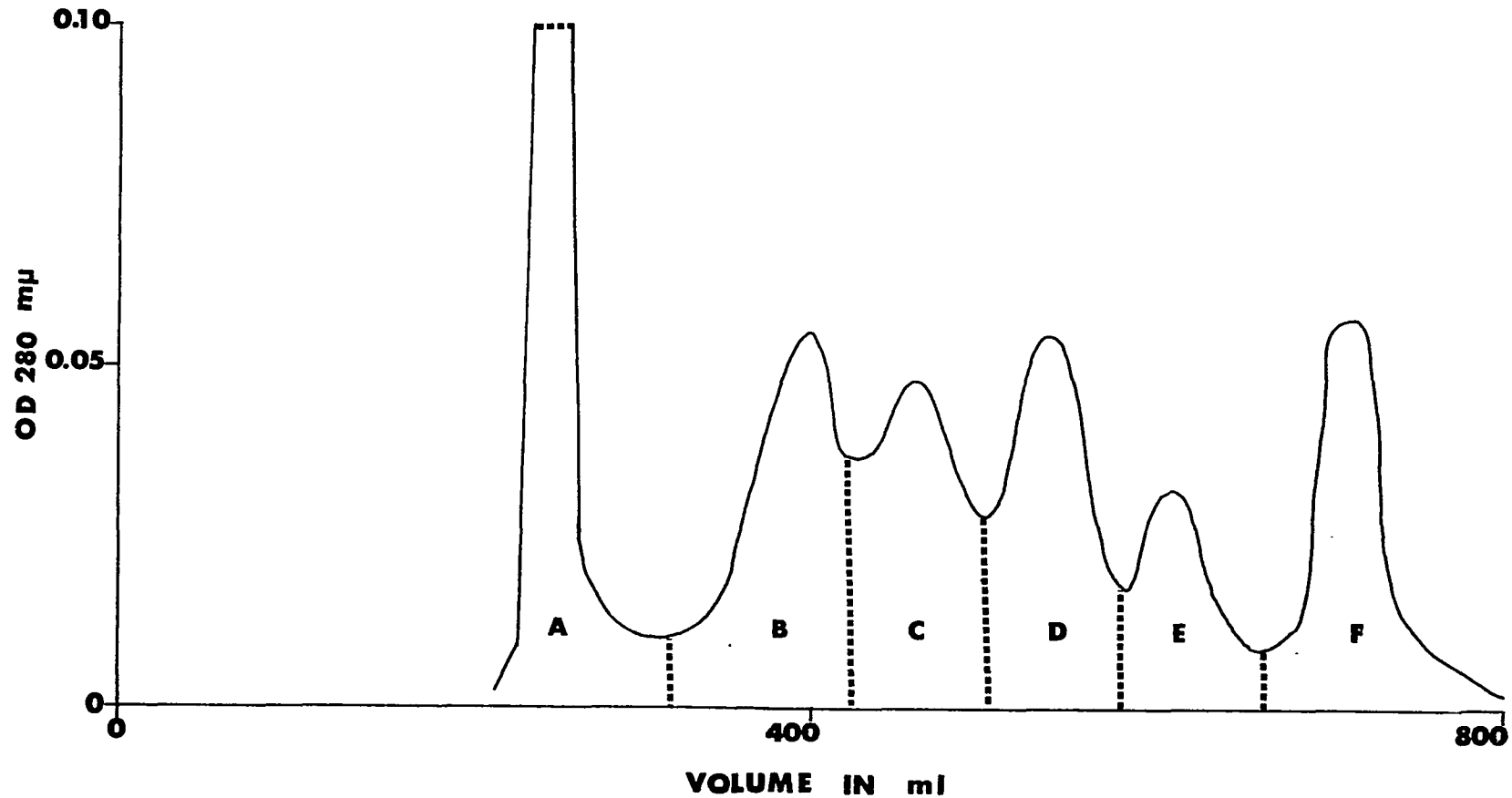

Fig. 14 - Chromatography of N-bromosuccinimide treated light chains.

The column was Sephadex G-100, 2.5 x 195 cm, and it was equilibrated with 1 N acetic acid.

TABLE 3

RELATIVE COMPOSITION OF FRAGMENTS OBTAINED FROM NBS OXIDIZED LIGHT CHAIN

Amino Acid	Peaks[a] A	B	C	D	E	F
Aspartic Acid	8.2	9.0	8.8	10.1	8.8	9.6
Threonine	12.2	11.8	8.7	16.2	14.4	5.7
Serine	11.7	11.2	7.9	13.9	15.6	6.1
Glutamic Acid	8.3	10.4	9.2	10.8	10.3	8.8
Proline	5.2	6.5	7.3	4.6	3.2	4.1
Glycine	10.9	9.0	9.2	8.8	9.7	9.4
Alanine	8.7	6.9	8.3	4.4	10.0	12.4
Valine	8.7	8.0	9.5	7.3	8.6	6.0
Methionine	0.4	tr	0.7	0.1	0.9	0.8
Isoleucine	3.5	3.0	4.1	2.3	3.6	3.6
Leucine	5.0	5.0	6.5	5.9	4.8	7.5
Tyrosine	4.8	4.2	2.9	3.6	2.1	1.9
Phenylalanine	3.9	3.3	3.5	2.6	1.3	3.2
Lysine	5.9	7.1	8.2	5.3	4.1	5.7
Histidine	0.3	1.8	1.5	1.2	0.8	2.3
Arginine	2.4	2.9	3.8	2.8	1.6	12.7

[a]Residues per 100 residues.

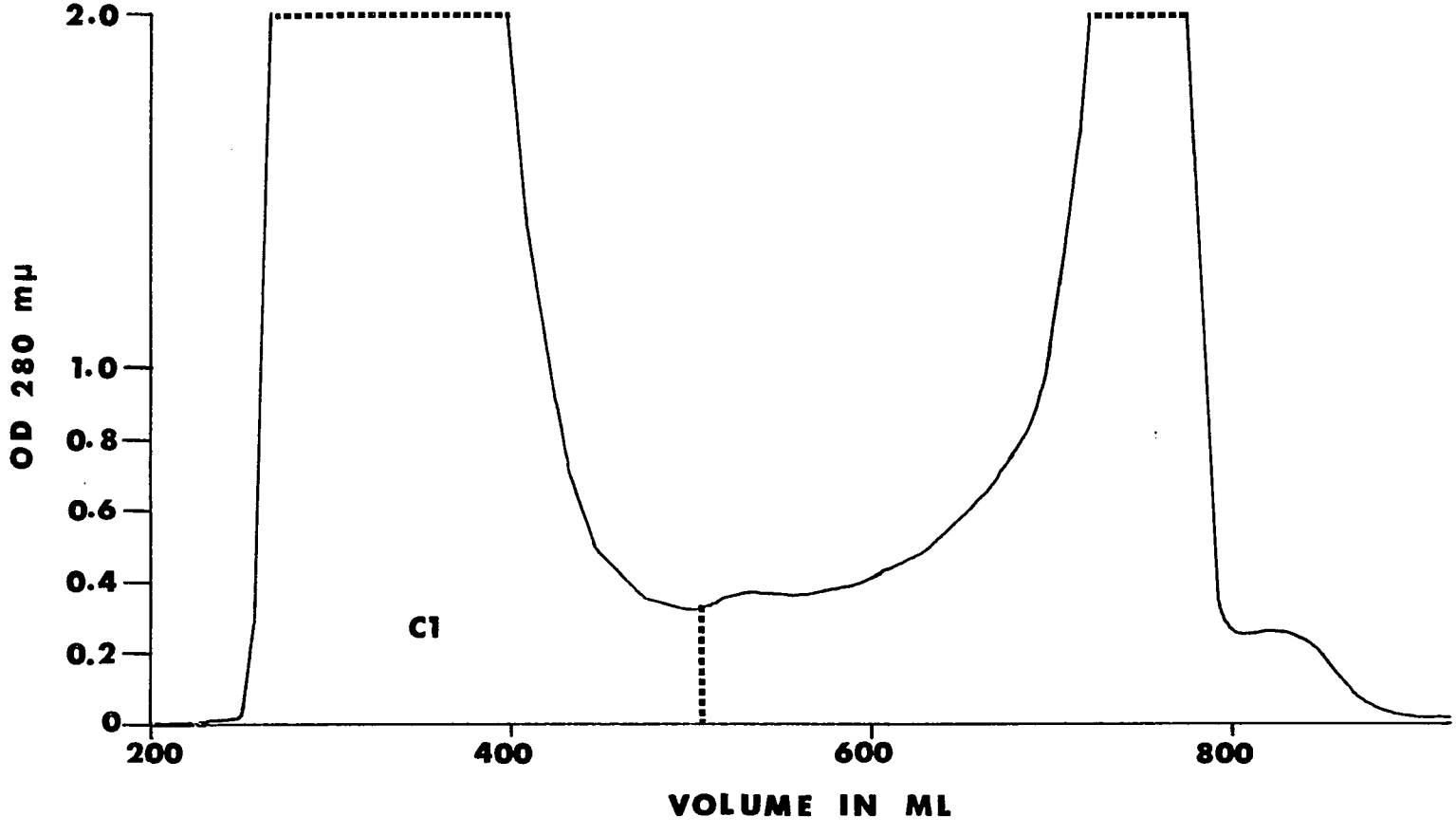

Fig. 15 - Chromatography of CNBr treated heavy chains.

The column was 4 x 100 cm and was run in 1 N acetic acid.

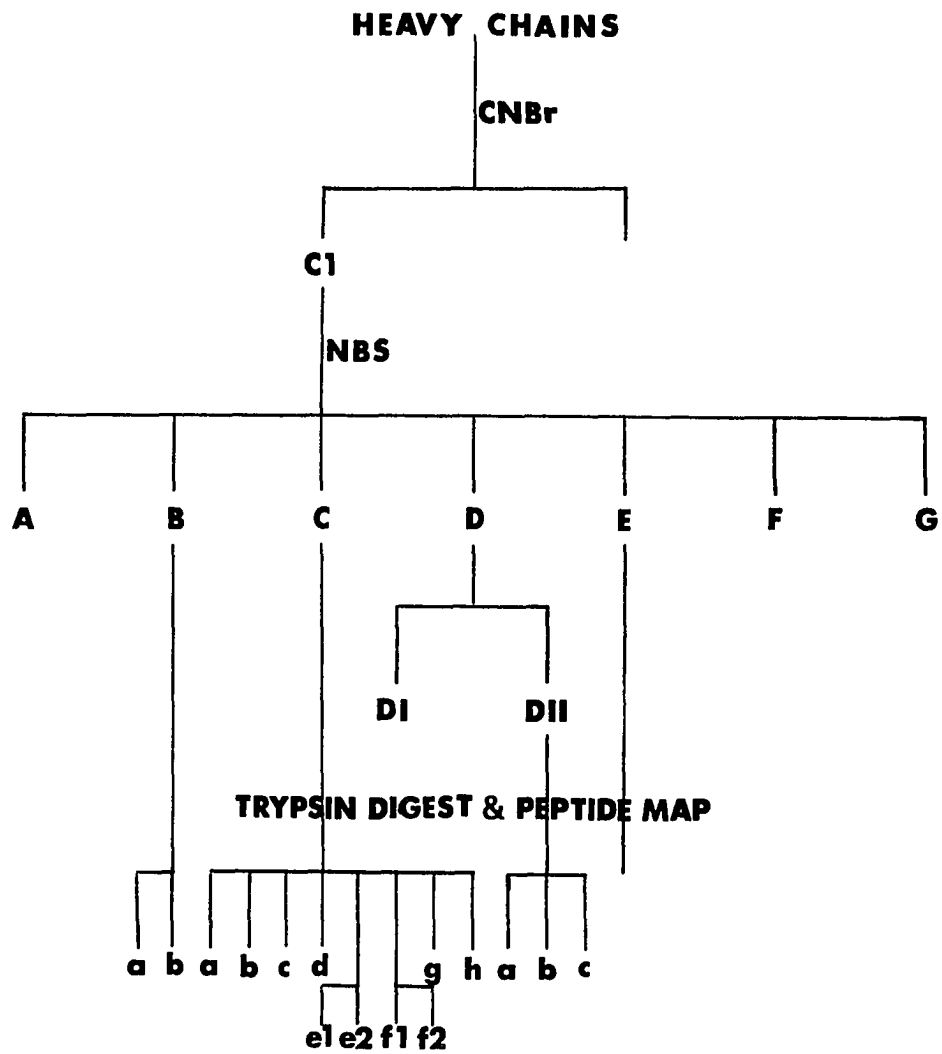

Fig. 16 - Flow chart of reactions of the heavy chains and the resulting fractions.

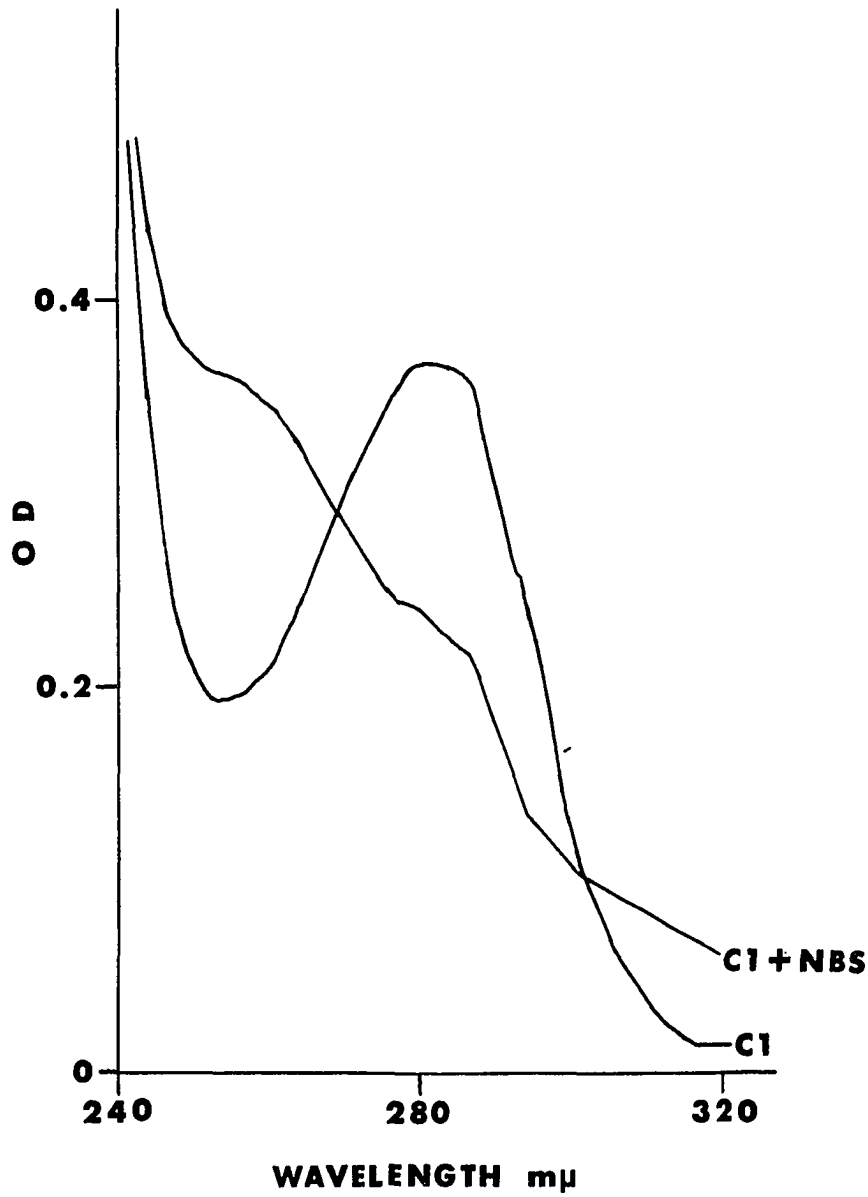

Fig. 17 - The spectra of unreacted and oxidized C1.

veloped (Figure 18) looked somewhat like the chromatogram from the light chain, but on closer examination, it was quite obvious that it was much different. Peaks B, C, D and E were rechromatographed on the same column for further purification. Peak D, on rechromatography (Figure 19), yielded 2 fractions (DI and DII). The amino acid composition of the peaks (B, C, DI, DII and E after rechromatography) were determined and are given in Table 4. Like the last peak from the light chains, there was very little peptide material in peak G.

Amino-terminal analysis of the peaks was carried out. Peak B after rechromatography had end groups of tyrosine, glycine and alanine. Peak C after rechromatography gave a predominant glycine end group with a trace of threonine and serine. Peak E, before rechromatography, gave end groups of glycine, alanine, valine, isoleucine, leucine and tyrosine.

The peaks B, C, DI, DII and E were lyophilized and weighed. From a 300 mg preparation of Cl the approximate yield of the various fractions were 36 mg for B, 30 mg for C, 60 mg for DI, 20 mg for DII and 35 mg for peak E. This corresponds to 12%, 10%, 20%, 7% and 12% of the total starting material for each peak respectively.

Peak C was digested with trypsin and the peptide map was developed from a Dowex 50 column (Figure 20). Pools were made representing each peak. Amino acid analysis of the pools was done. Peptides from peaks ClCb and ClCc were judged homogeneous on the basis of their compositions and were not purified further. Peptide ClCd was further purified by paper electrophoresis at pH 6.5. This peptide gave a yellow cadmium-ninhydrin spot which indicates a glycine end group. ClCe and ClCf were obtained by paper electrophoresis at pH 3.5 of the first 3

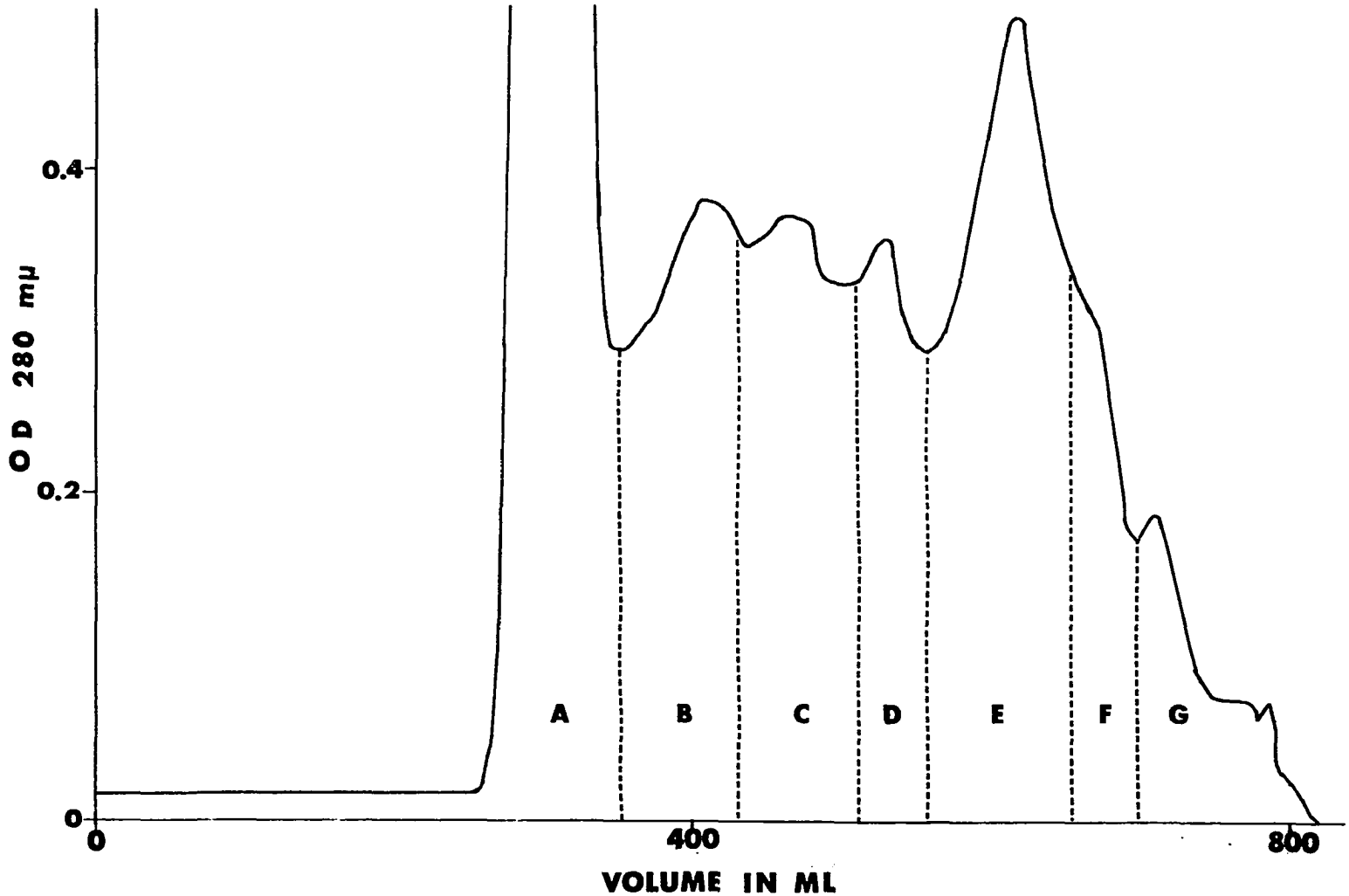

Fig. 18 - Chromatography of N-bromosuccinimide treated C1 fraction on G-100.

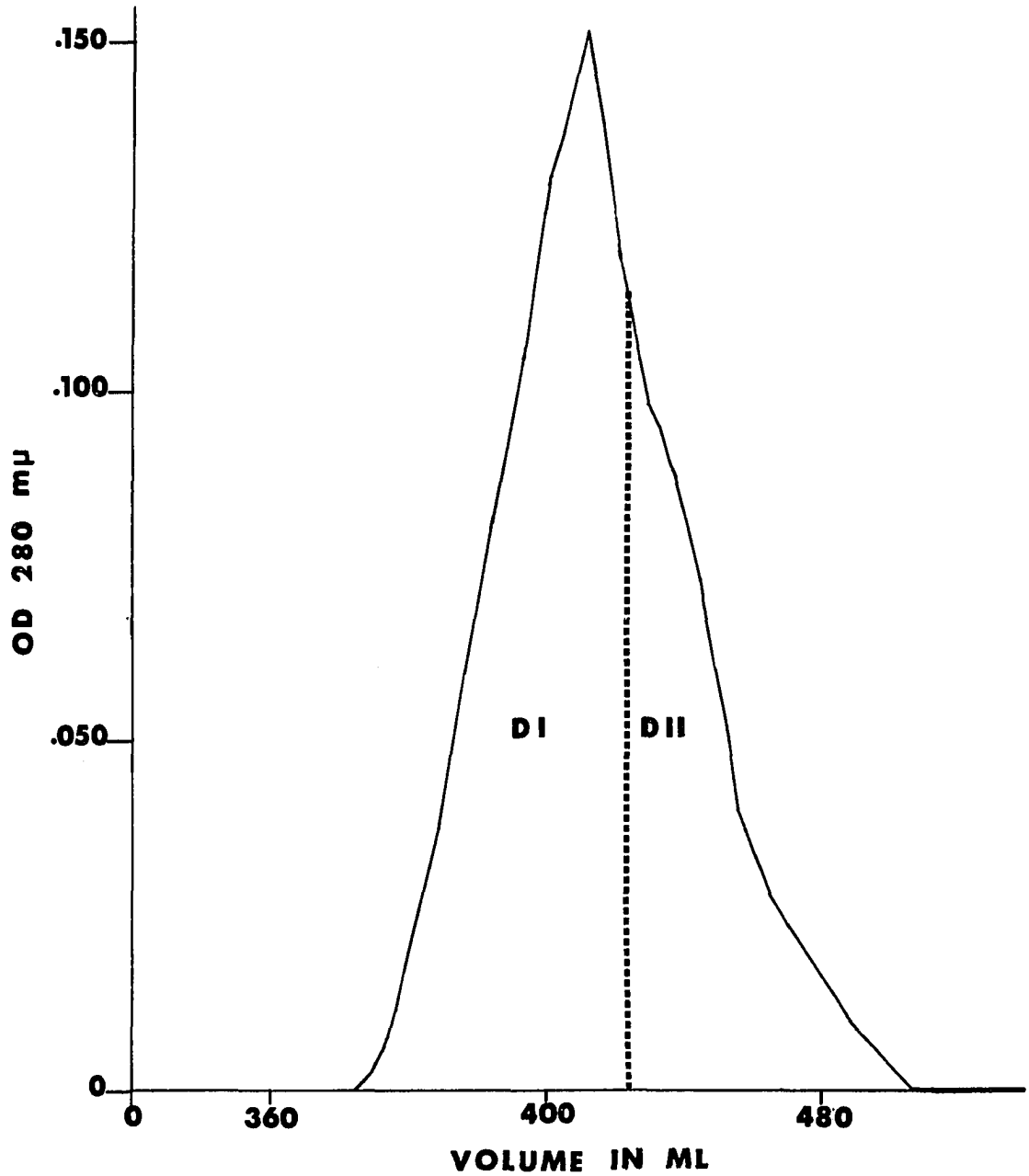

Fig. 19 - The rechromatography of peak D.

TABLE 4

RELATIVE COMPOSITION OF FRAGMENTS OBTAINED
FROM NBS OXIDIZED C1 FRACTIONS

Amino Acid	Peaks[a]							
	A	B	C	DI	DII	E	F	G
Aspartic Acid	8.3	7.0	7.7	8.0	7.2	7.3	6.1	11.0
Threonine	11.8	13.6	12.8	13.2	12.9	13.7	13.6	6.8
Serine	12.4	12.1	14.7	13.2	14.0	11.8	14.5	13.9
Glutamic Acid	7.2	6.1	5.4	7.3	5.9	8.2	8.7	8.3
Proline	5.7	10.9	11.7	9.9	12.6	6.7	5.4	5.4
Glycine	9.7	8.1	7.1	10.6	7.5	9.6	11.2	11.6
Alanine	7.9	5.5	5.5	5.5	5.2	5.7	7.0	8.8
Valine	9.8	10.7	10.0	7.7	10.6	7.7	8.4	10.4
Isoleucine	4.9	2.5	2.0	3.6	2.1	3.6	3.6	5.1
Leucine	7.2	8.2	7.2	6.9	7.4	7.6	8.1	5.0
Tyrosine	4.0	2.8	2.3	3.4	2.0	3.8	3.7	3.6
Phenylalanine	3.5	3.5	3.4	3.1	3.7	3.2	2.9	3.0
Lysine	4.2	5.3	5.4	4.8	5.1	6.3	3.1	2.9
Histidine	0.7	0.8	1.6	0.7	1.3	1.5	0.8	1.1
Arginine	2.7	3.0	3.2	2.0	2.6	3.4	2.9	2.9

[a]Residues per 100 residues.

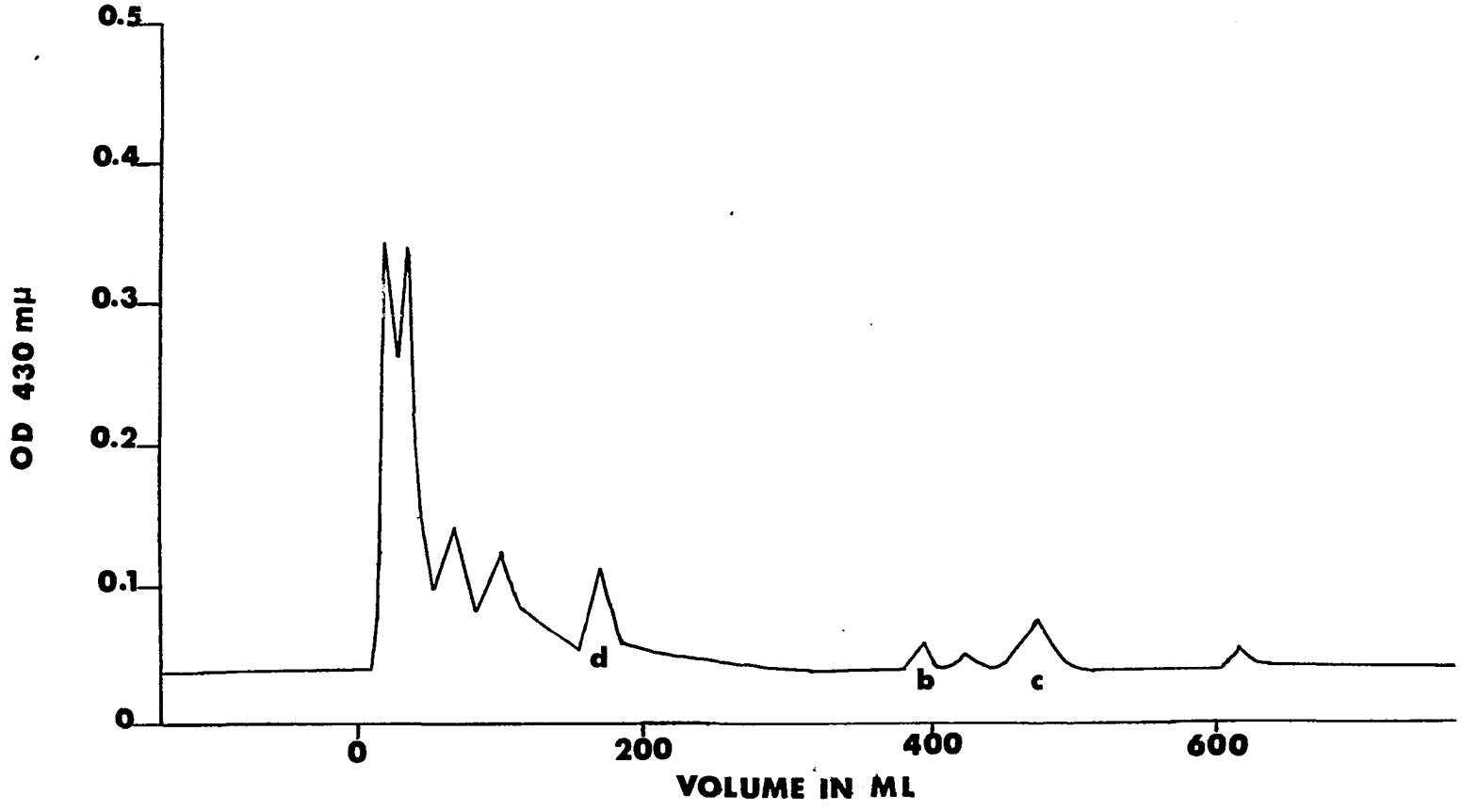

Fig. 20 - The peptide map of peak C.

Peaks b, c and d are listed in Table 5 as C1Cb, C1Cc and C1Cd.

peaks from the Dowex 50 column. C1Ca was purified by Sephadex G-25 chromatography of the section of the chromatogram between C1Cd and the next peak. The amino acid composition of these peptides is shown in Table 5.

Another preparation of peak C was made and digested by trypsin. These peptides were separated on Sephadex G-25 (Figure 21) and were further purified by paper electrophoresis. In addition to the peptides isolated from the Dowex 50 peptide map, 2 other peptides that were the result of tryptic hydrolysis at aminoethylcysteine were also isolated. The composition of the peptides C1Cg and C1Ch are also listed in Table 5.

Peptides C1Ce and C1Cf gave at least 2 ninhydrin spots each. The amino acid compositions of these peptides are given in Table 5. The amino acid compositions of the 2 C1Ce peptides are quite similar, and the compositions of the C1Cf peptides are also similar. Further digestion of the 2 C1Ce peptides was attempted with chymotrypsin and thermolysin, but no meaningful data was obtained from these experiments. Because Cebra (78) had published the sequence of a peptide from the same position as C1Cf, no further work was undertaken at that time on these 2 peptides to find the differences between them. Further work on these peptides is hampered by the difficulty of obtaining enough of a homogeneous peptide to recover further digestion products in good yield.

Peptides C1Cb, C1Cc, C1Cd, C1Cg and C1Ch were sequentially degraded by the Edman method. The results are shown in Figure 22. Electrophoretic mobilities of C1Cd and C1Cg indicated that the acidic residues in these 2 peptides were all in the amide form. The sequence of C1Cd is tentative because the position of the glutamine residue has not been established.

TABLE 5

COMPOSITION OF TRYPTIC PEPTIDES OBTAINED FROM C1 FRACTIONS

Amino Acid	Residues					
	C1Ca	C1Cb	C1Cc	C1Cd[b]	C1Ce1	C1Ce2
Aspartic Acid	0.9	0.9		0.2	1.8	2.2
Threonine	0.8		1.1	1.0	4.0	4.2
Serine			1.1	1.0	6.2	5.2
Glutamic Acid				0.8	2.2	2.6
Proline			1.0	0.9	3.4	2.8
Glycine				1.8	2.4	2.5
Alanine				0.2	1.8	1.8
Valine		0.6	0.9	1.3	4.1	4.5
Aminoethylcysteine[a]					1	1
Isoleucine					0.4	0.7
Leucine	1.0			0.8	2.6	2.2
Tyrosine					0.6	0.8
Phenylalanine			1.0		0.9	0.8
Lysine		1.0		1.1	0.7	0.6
Histidine					tr	tr
Arginine			1.0			
Homoserine[a]	1					

[a]Estimated values.

[b]72-hour hydrolysis.

TABLE 5--Continued

ClCf1	ClCf2	ClCg	ClCh	ClDIIa	ClDIIb	ClDIIc	ClBa	ClBb
1.2	0.4	2.0		1.6	0.8		0.9	
3.0	1.9	1.7	1.6	2.0		1.1		1.2
3.7	3.2		1.0	1.2		1.1		1.1
1.2	0.5							
8.1	9.0	1.1	1.0			1.0		1.0
2.7	3.2			1.7				
1.6	0.5	2.1	1.2	0.4				
2.1	1.6	1.0	1.2	1.0	0.9	1.0	0.9	0.9
2	2		1					
0.6	1.2							
2.6	3.1							
0.2								
1.4	2.3					1.0		0.9
2.7	4.2	1.1		0.3	1.0		1.0	
		1.0						
				0.8		0.8		0.9

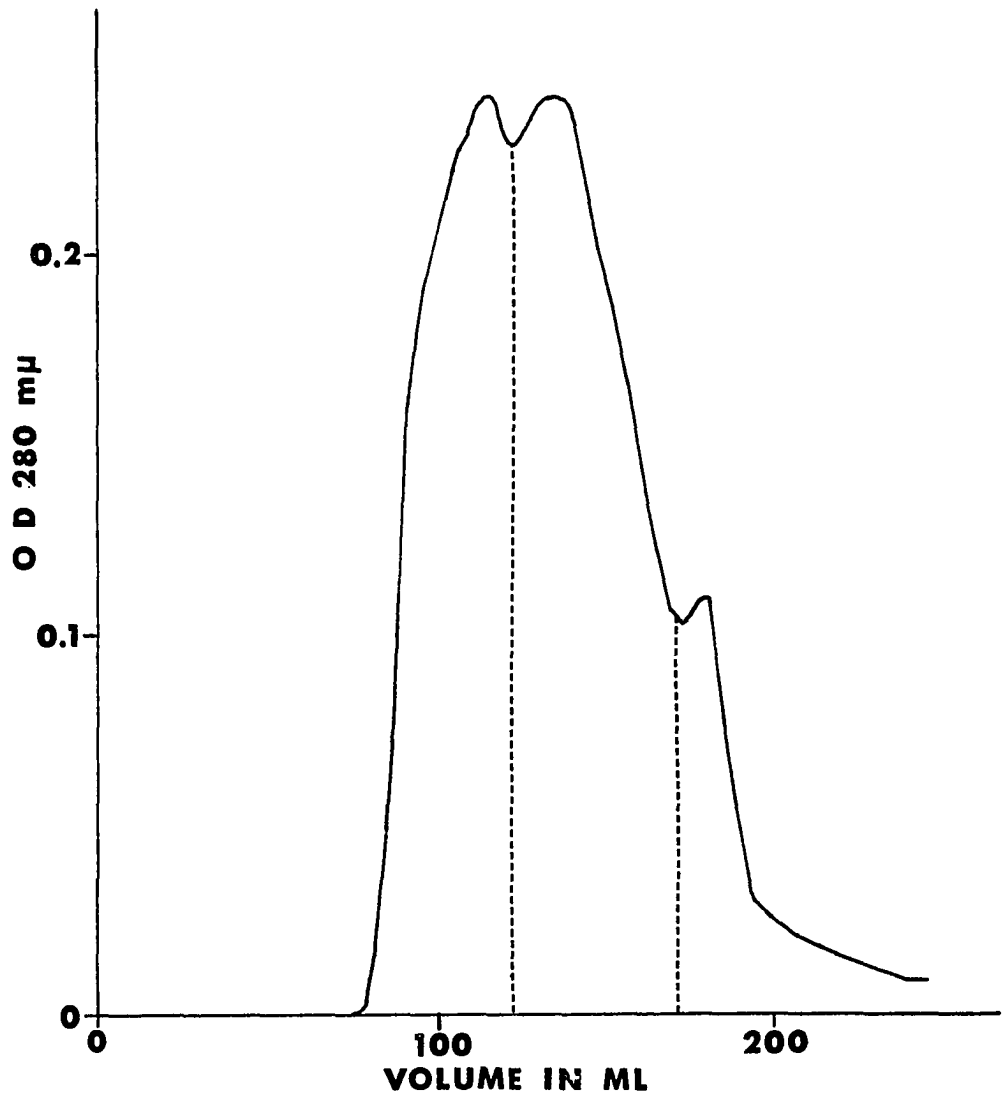

Fig. 21 - The chromatography of a trypsin digest of peak C.

The chromatography was done on a Sephadex G-25 column, 1.5 x 100 cm in 1 N acetic acid.

ClCb

Val-Asp-Lys

ClCc

Thr-Phe-Pro-Ser-Val-Arg

ClCd

Gly-Ser-Gly-Thr-Leu-Val-Pro-Val(Gln)Lys

ClCg

Asn-Val-Ala(His)Pro-Ala-Thr-Asn-Thr-Lys

ClCh

Thr-Val-Ala-Pro(Ser,Thr)AECys

Fig. 22 - The amino acid sequence of peptides isolated from trypsin digests of peak C.

Residues that have not been established are in brackets. No residue could be identified from the 3rd Edman cycle on peptide ClCg. Histidine is placed in this position by homology.

Peak E was digested with trypsin and the peptide map of this peak was developed (Figure 23). Pools were made representing each of the peaks. The pools were analyzed for amino acid content and further purified by paper electrophoresis at various pH's. No homogeneous peptides could be purified from these pools as determined by amino acid ratios.

Peptide maps were also obtained for peaks B and DII (Figures 24 and 25). A few peaks from these chromatograms were subjected to amino acid analysis. The results are given in Table 5.

The peaks from oxidation of light chains and C1 fraction were assigned to particular positions in these polypeptide chains for reasons discussed in the next Chapter. With this knowledge, the molecular weight of each peak can be estimated, and the compositions of these peaks can be expressed as residues/mole. This is done in Table 6. In addition, the yield of each of the C1 fractions can also be calculated, and these are also shown in Table 6. The effective molecular weight of each of the fractions is also given in Table 6. The effective molecular weight on the G-100 column of each fraction is higher than the actual molecular weight because they were chromatographed in high ionic strength solvents.

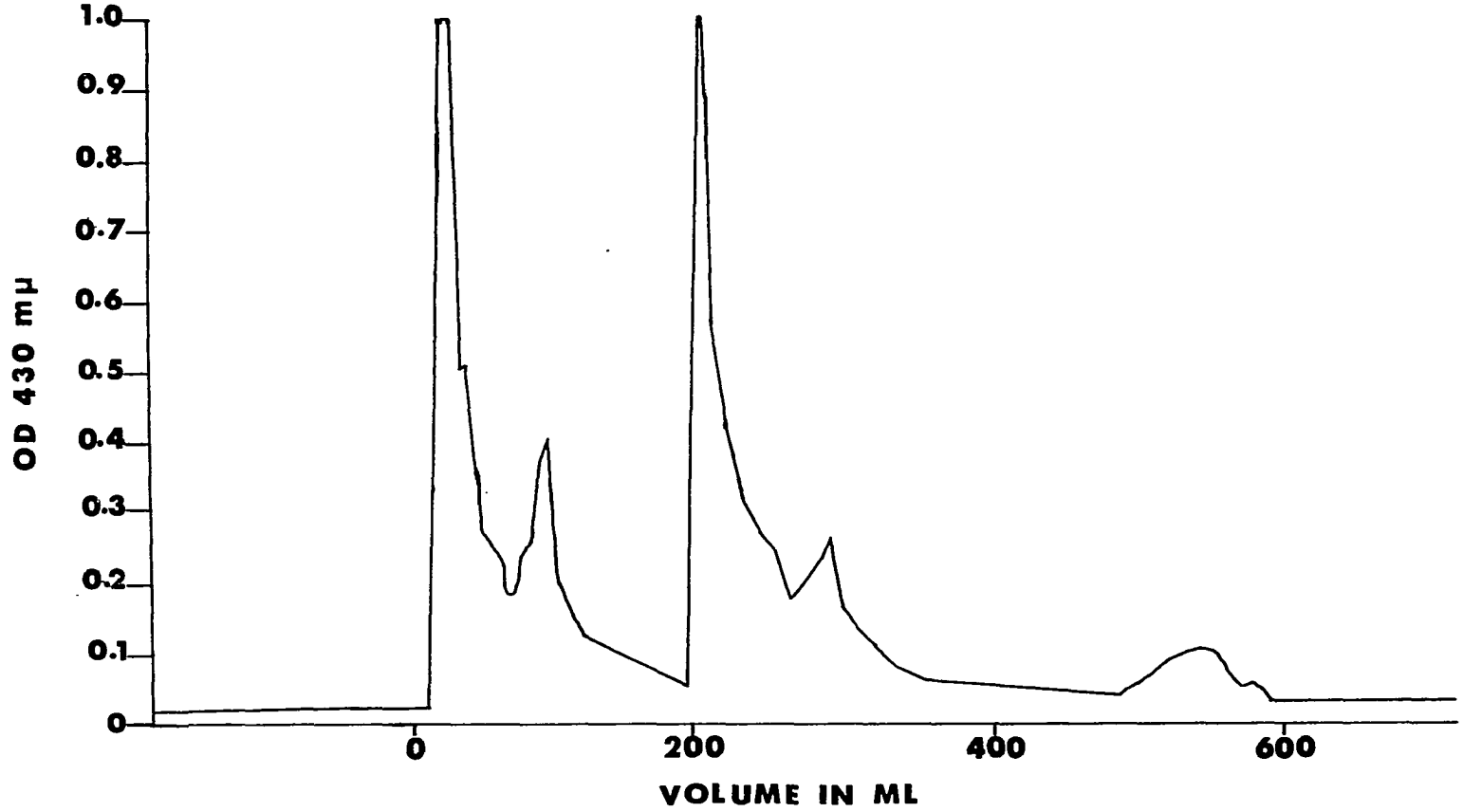

Fig. 23 - The peptide map of peak E.

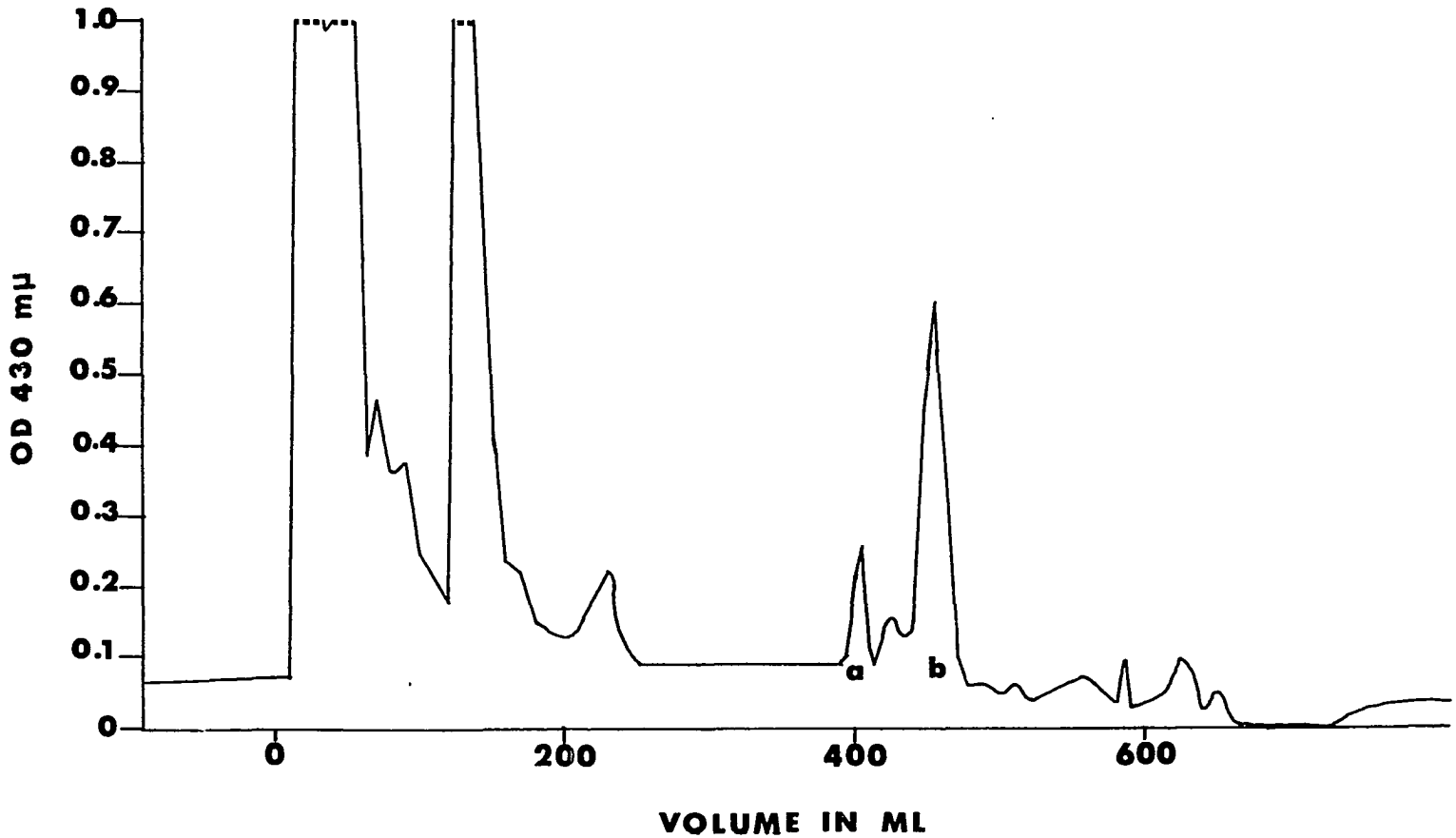

Fig. 24 - The peptide map of peak B.

Peaks a and b are listed in Table 5 as C1Ba and C1Bb.

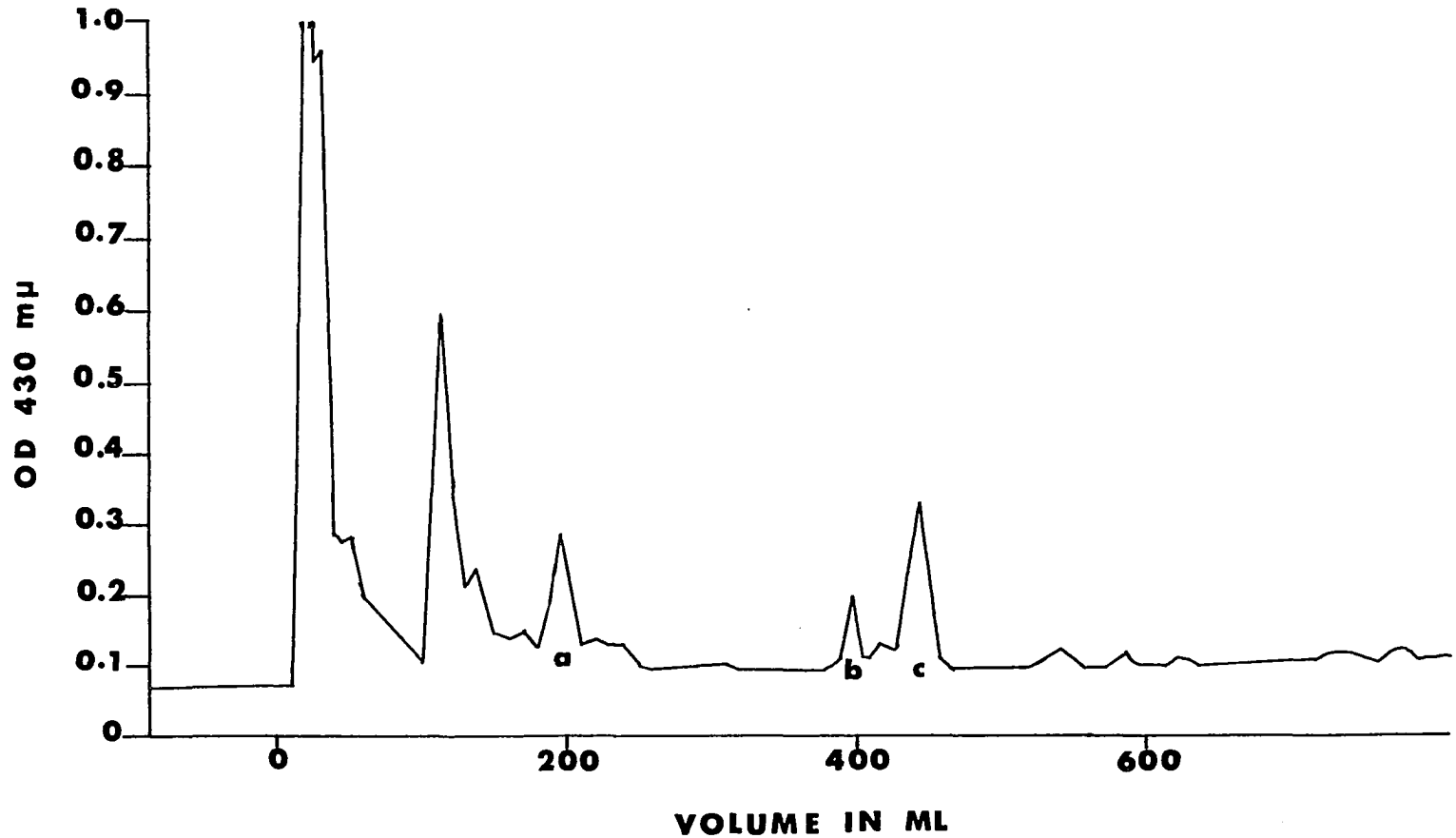

Fig. 25 - The peptide map of peak DII.

Peaks a, b and c are listed in Table 5 as ClDIIa, ClDIIb and ClDIIc.

TABLE 6

FRACTIONS FROM LIGHT CHAINS AND C1 EXPRESSED AS RESIDUES/MOLE

Amino Acid	Light Chains				C1 Fraction					
	B	C	D	E	B	C	DI	DIII	E	F
Aspartic Acid	15.7	9.5	6.8	3.0	14.1	6.5	12.3	6.0	8.6	2.1
Threonine	20.5	9.5	10.9	4.9	27.4	10.8	20.2	10.9	16.0	4.6
Serine	19.6	8.6	9.3	5.3	24.2	12.4	20.2	11.7	13.8	4.9
Glutamic Acid	18.2	10.0	7.2	3.5	12.2	4.5	11.2	4.9	9.6	3.0
Proline	11.4	7.9	3.1	1.1	21.9	9.9	15.1	10.6	7.8	1.8
Glycine	15.8	10.0	5.9	3.3	16.2	5.9	16.2	6.3	11.2	3.8
Alanine	12.0	9.0	3.0	3.4	11.1	4.6	8.5	4.3	6.7	2.4
Valine	13.9	10.3	4.9	2.9	21.6	8.4	11.7	8.9	9.0	2.9
Cysteine[a]	4	2	2	1	7	3	5	3	4	1
Methionine	tr	0.8	tr	0.3	-	-	-	-	-	-
Isoleucine	5.2	4.5	1.6	1.2	5.0	1.7	5.6	1.8	4.2	1.2
Leucine	9.4	7.1	4.0	1.6	16.4	6.1	10.5	6.2	8.8	2.8
Tyrosine	7.8	3.1	2.4	0.7	5.6	1.9	5.2	1.7	4.5	1.3
Phenylalanine	5.8	3.8	1.7	0.4	7.1	2.9	4.7	3.1	3.8	1.0
Lysine	12.4	8.9	3.5	1.4	10.6	4.5	7.4	4.3	7.3	1.0
Histidine	3.1	1.6	0.8	0.3	1.5	1.3	1.1	1.1	1.8	0.3
Arginine	5.0	4.1	1.9	0.5	6.0	2.7	3.1	2.2	3.9	1.0
Tryptophan[a]	1	1	0	1	1	0	1	0	1	1
Total	181	112	69	36	210	88	158	88	122	36
Effective Molecular Weight	34,000	20,000	10,200	6,300	52,000	42,000	27,000		14,000	10,200
% Yield					14	28	20	31	19	

[a]Estimated value

CHAPTER IV

DISCUSSION

One thought that should be kept in mind is that pooled IgG is a heterogeneous protein. Any isolation or chemical reaction of such a protein runs the risk of selecting from the various proteins only a portion of the many types from the whole. A good example of this is the DEAE cellulose chromatography after sodium sulfate precipitation. Although there was much IgG in the peak III and the second peak from the 12-18% fraction, these fractions could not be used because of the impurity. Therefore, any IgG molecule that tended to adsorb to the DEAE cellulose was selected against. The consequence of this is that the complete spectrum of IgG molecules was not achieved. The caprylic acid precipitation method yields nearly 100% IgG in the first step, but the batch absorption on DEAE cellulose removes a small portion of the IgG. This selection is small in comparison to DEAE cellulose chromatography but still exists. The selectivity can appear in nearly every operation carried out on the IgG and can lead to some misinterpretation of the experimental data.

The N-bromosuccinimide cleavage of the light chain yielded the expected results. The first peak, peak A, probably contained uncleaved material and aggregates of light chains and peptides. Its amino acid composition was not identical with that of whole light chains, but it

was not greatly different. Peak B was assigned to the overlap peptide containing the C-terminal and middle sections of the light chain. The amino acid composition is fairly close to that of the light chain, which would be expected, but even more important its composition is close to the average composition between peaks C and D. Peaks C and D are assigned to the middle and C-terminal polypeptides respectively. Peak E is principally the N-terminal peptide. The effective molecular weight data agrees with this assignment. The weight of C plus D (30,000) is near that of B (34,000).

Since very little work has been done on the sequence of pooled rabbit light chains it is difficult to predict with an accuracy the composition of any of the fragments from rabbit light chains. Therefore, the amino acid composition of a peak is, by itself, not strong enough proof to be able to say for certain that a given peak represents a certain fraction. On the other hand indications may be drawn from these compositions. The constant portion of several human kappa chains previously sequenced (127) could give an indication of the possible composition of the C-terminal fragment of the rabbit chain. In some of the amino acid compositions of the various peaks from the oxidized light chains there are large differences in the values of individual amino acids. For instance, in peak D the total threonine + serine content is 30.1% and in peak C it is only 16.6%. Other comparisons of these 2 peaks show proline to be 4.6% for peak D and 7.3% for peak C and alanine to be 4.4% and 8.3%, respectively. Human kappa chains show approximately 27% threonine + serine, 4.5% alanine and 1.5% proline for the C-terminal peptide. Each of these values is much closer to the peak D value than to

the peak C value. The only big exception to these indications is lysine. Peak D has 5.3% lysine and peak C has 8.2% lysine. Human kappa chains have 9% lysine. Some doubt about the correct lysine values exists because of the side reaction of the lysine with cyanate produced by urea. Even with the one exception noted the amino acid composition seems to indicate that peak D is the C-terminal peptide.

There is an even more compelling reason to assign peak D to the C-terminal position. It elutes from the Sephadex column later than peak C, and therefore has a lower effective molecular weight. This is what our model predicted would be the order of elution from the column.

The reasons for assigning peak C to the middle section of the molecule have already been given. It has a higher effective molecular weight than D and its amino acid composition and the composition of peak D average to the composition of peak B. There are 3 exceptions to this average composition. There is more histidine and tyrosine and less leucine in peak B than in either C or D. It is possible that the tyrosine and histidine in the shorter chains were oxidized by the N-bromosuccinimide to a small extent. Also, the small amount of histidine was difficult to calculate on the amino acid analysis because it is incompletely resolved from aminoethylcysteine appearing before it and the ammonia peak appearing after it. This makes it difficult to draw an accurate base line for the calculation. There is no obvious explanation why the leucine value is outside the average.

Peak E was assigned to the N-terminal peptide mainly because of the effective molecular weight of its peak. Since this peptide comes from the variable region of the light chain, the amino acid composition

of other kappa chain peptides can only show tendencies in composition. In human and mouse light chains (127), methionine shows up in this peptide relatively frequently and peak E has the highest methionine content of any of the peptides. Threonine and serine also occur quite often in sequenced kappa chains while proline appears only about once or twice per peptide. Basic amino acids represent only about 2 residues for each N-terminal peptide. All of these tendencies are present in peak E. Amino acid composition supports the assignment of E to the N-terminal peptide.

Although the N-terminal amino acid analysis showed several dansyl amino acids for each of the peaks, the residues that are expected are among these spots. Human and mouse myeloma kappa chains are followed by tyrosine, leucine, or phenylalanine in position 36 and lysine in position 148 (127). Rabbit light chains have aspartic acid and alanine at their N-terminal (23). Peak D contains lysine as an end group and peak B and peak C contain tyrosine. Peak E contains alanine. Unpublished data of Mameesh and Delaney indicate that if peak C is extensively repurified, a homogeneous end group of tyrosine can be obtained. All of the evidence obtained from the chromatography of N-bromosuccinimide treated light chain is consistent with predictions from the model. This information is consistent with the existence of the predicted 2 constant tryptophanes in all rabbit light chains, and probably all immunoglobulin light chains.

In contrast to the light chains, the heavy chain C1 fraction did not yield the expected chromatographic pattern. At first glance, the pattern of this chromatogram appears to be similar to that of the light chain, but there are several important differences.

Unlike the light chain, it was possible to estimate the composition of the N-bromosuccinimide fractions with some degree of accuracy. Cebra (77) has published the composition of the C-terminal 70 residues. By using the sequence of Eu or some other sequenced heavy chain, it was possible to predict the next 15 residues or so up to the tryptophan number 158. Using Wilkinson's data as reported by Porter (141), the composition of the N-terminal 34 residues was given, so it was not hard to estimate the composition for the N-terminal section. The N-terminal and C-terminal composition was subtracted from the complete Cl fraction composition to obtain the expected composition of the middle section. The resultant compositions were converted to mole per cent. These expected compositions and the compositions of the peptides that are assigned to the N-terminal, C-terminal and middle sections are presented for comparison in Table 7.

Since the C-terminal polypeptide chain contains the hinge section of the heavy chain, it should have a high percentage of proline. This one amino acid is probably the best marker in the entire molecule. Peak C contains a high amount of proline. The other amino acid percentages of peak C are fairly close to the predicted values for the C-terminal fraction. By amino acid composition data, peak C is assigned to the C-terminal peptide. This assignment immediately raises the question of why peak C elutes from the column so early? Its predicted molecular weight would indicate that it should elute after peak E. The hinge region, which contains the high proline content, may prevent the polypeptide chain from assuming a globular shape. This would cause the effective molecular weight to be much higher due to a larger Stoke's axis of

TABLE 7

COMPARISON OF PREDICTED AND DETERMINED COMPOSITIONS OF C1 FRACTIONS[a]

Amino Acid	Predicted N-Terminal	Peak F	Predicted C-Terminal	Peak C	Predicted Middle	Peak E
Aspartic Acid	2.9	6.1	6.0	7.7	6.8	7.3
Threonine	14.7	13.6	11.9	12.8	15.4	13.7
Serine	17.6	14.5	14.3	14.7	11.1	11.8
Glutamic Acid	8.8	8.7	3.6	5.4	8.5	8.2
Proline	5.9	5.4	15.5	11.7	4.3	6.7
Glycine	11.8	11.2	7.1	7.1	10.3	9.6
Alanine	2.9	7.0	6.0	5.5	8.5	5.7
Valine	11.8	8.4	11.9	10.0	8.5	7.7
Isoleucine	0.0	3.6	1.2	2.0	4.3	3.6
Leucine	11.8	8.1	8.3	7.2	6.0	7.6
Tyrosine	2.9	3.7	1.2	2.3	6.8	3.8
Phenylalanine	2.9	2.9	3.6	3.4	3.4	3.2
Lysine	2.9	3.1	6.0	5.4	1.7	6.3
Histidine	0.0	0.8	2.4	1.6	0.0	1.5
Arginine	2.9	2.9	1.2	3.2	4.3	3.4

[a]Expressed as residues per 100 residues.

symmetry.

Peak F is assigned to the N-terminal polypeptide of the molecule for several reasons. The amino acid composition is consistent with the predicted value in most cases. Alanine and aspartic acid are the most notable exceptions to this, but even these seemingly large differences are only on the order of one residue from the predicted values. This peak is approximately in the same position on the chromatogram as the N-terminal peak from the light chain and it shows the same amino acid characteristics pointed out for the light chain. It has about 28% threonine plus serine, 2 prolines and 2 basic residues.

Peak E is assigned to the middle section of the C1 fraction. The amino acid composition is about as predicted. One notable exception to this is lysine. The way the predicted value for lysine was reached led to a large error in the value. About 1/3 of the lysine was lost by chemical modification during chain separation. Peaks C and F contained little of these modified residues. Therefore the predicted value for the middle section appears too low. Published values (77) for the composition of the C1 fraction show it to contain about 12 lysine residues or 4 more than shown in this C1 fraction. If these 4 residues are added to the predicted composition of peak E, the actual results are in closer agreement to the predicted result.

The average of peaks C and E seems to approximate the composition of B and the effective molecular weight of C plus E (56,000) is close to that of B (52,000). There are several amino acid mole per cents in B that fall outside the range formed by E and C but all are very close to either E or C. In all but one of these cases, valine, the difference

between E and C is quite small.

The identity of peak DI has not been firmly established. A possibility for the identification of this peak is the overlap of the middle section and the N-terminal section. The predicted molecular weight of this peptide would place it in the chromatogram somewhere between peak B and E, and that is where peak DI is located. Also, the effective molecular weight of E plus F (24,000) is close to that of DI (27,000). Since the middle section would make up about 3/4 of the amino acid residues in this proposed overlap peptide, the amino acid composition would be expected to show a similarity between peak DI and E. DI and E have similar compositions except for proline. Since the peptides on either side of DI, DII and C, both have high proline compositions, a little contamination of DI with either or both of these peaks would raise the proline content markedly. A peptide map of DI could establish whether this polypeptide is the overlap between the N-terminal and middle section of Cl.

Peak DII appears to be a C-terminal polypeptide much like peak C. Their amino acid compositions are quite similar, but some differences between the 2 peptides must exist since they do not chromatograph in the same position. Peak DII also seems to average with peak E to obtain the composition of peak B even better than peak C. Peak A appears to be uncleaved material and/or aggregates of polypeptides. Its amino acid composition is not greatly unlike that of Cl.

N-terminal analysis of the peaks gave a variety of amino acids in each case, but the expected end group was usually among them. Peak B has an expected tryosine or isoleucine end group, and tyrosine was present. The expected end group for peak C was aspartic acid but no aspartic

acid was found. Instead, peak C gave a very predominant glycine end group. Peak E would be expected to contain tyrosine or isoleucine and contained both.

All of the peptides that Cebra's (78) C-terminal polypeptide, T2, contained were isolated from the fractions obtained from the column. They were peptides C1Ca, C1Cb, C1Ce and C1Cf. This was conclusive proof that peak C was the C-terminal peptide. In addition to the T2 peptides, two more peptides were obtained, C1Cc and C1Cd. The peptides C1Cb and C1Cc chromatographed late in the development of the column and therefore make good markers for the C-terminal polypeptide.

The peptide map of peak E showed neither of the two marker peaks from the C-terminal. This peptide map consisted mainly of two large peaks near the start of the run. The peptide map adds more confidence to the assignment of E as the middle section of the C1 fraction.

The peptide map of peak B showed both the characteristic marker peaks from the C-terminal section and the large peptide peaks early in the run characteristic of the middle section, peak E. Amino acid analyses of the suspected C-terminal marker peaks showed compositions identical to C1Cb and C1Cc. This added confirmation to the fact that peak B contains both the C-terminal and middle sections of C1.

Peak DII showed the 2 characteristic peaks from the C-terminal section, and amino acid analysis confirmed that the composition of these peaks is identical to the corresponding peaks from peak C. Total peptide material in the peptide map in comparison to the size of the C1Cb and C1Cc peaks makes it unlikely that large amounts of anything but the C-terminal polypeptide are present. Although this peptide map is somewhat

different from the map of peak C, peak DII is a C-terminal peptide.

The sequences of some of the peptides from peak C were compared to the sequences of the same sections reported by Porter's group (79) and the myeloma protein Eu (86) (Figure 26) and the alignment of peptides is shown in Figure 27. The amino acid sequence of ClCb and ClCc and the composition of ClCa agree with those reported by Hill *et al.* (75) and Porter's group (79). Peptides ClCe and ClCf were not sequenced here, but 2 smaller peptides, ClCg and ClCh, which arose from tryptic hydrolysis at aminoethylcysteine, were partially sequenced. ClCh has the same amino acid composition as reported by Cebra (77) and the first 4 residues are in the order reported. There are probably no differences between ClCh and the remaining sequence of Cebra's. Peptide ClCg shows a proline residue that Fruchter (79) did not include in this peptide. A proline appears in the corresponding peptide from Eu, but it is one residue removed from being in the same position. Electrophoretic mobility of ClCg shows both of the aspartic acid residues are amides.

Peptide ClCd (Figure 26) contains some of the differences between fraction ClC and Fruchter's (79) sequence. The amino acid compositions of ClCd and the corresponding peptide from Fruchter are different, but there is little doubt that the two peptides come from the same location. They both contain the tetrapeptide Ser-Gly-Thr-Leu. The trypsin digest of peak DII yields the peptide, ClDIIa, which has a composition identical to that of Fruchter's sequence. This peptide is one of the obvious differences between peaks C and DII and adds further confirmation to the identity of peak DII. The difference between peaks C and DII is probably greater than allotypic differences and is more likely a subtype

ClCb Val-Asp-Lys

RaIgG Val-Asp-Lys

Eu Val-Asp-Lys

ClCc Thr-Phe-Pro-Ser-Val-Arg

RaIgG Thr-Phe-Pro-Ser-Val-Arg

Eu Thr-Phe-Pro-Ala-Val-Leu

ClCd Gly-Ser-Gly-Thr-Leu-Val-Pro-Val(Gln)Lys

RaIgG Asn-Ser-Gly-Thr-Leu-Thr-Asp-Gly-Val-Arg

Eu Asn-Ser-Gly-Ala-Leu-Thr-Ser-Gly-Val-His

ClCg Asn-Val-Ala(His)Pro-Ala-Thr-Asn-Thr-Lys

RaIgG Asx-Val-Ala-His-Ala - Thr-Asx-Thr-Lys

Eu Asn-Val-Asn-His-Lys-Pro-Ser-Asn-Thr-Lys

ClCh Thr-Val-Ala-Pro(Ser,Thr)Cys

RaIgG Thr-Val-Ala-Pro-Ser-Thr-Cys

Eu Lys-Val-Gln-Pro-Lys-Ser-Cys

Fig. 26 - Comparison of the amino acid sequences from peak C peptides with the sequences for rabbit IgG reported by Fruchter (79) and with the myeloma protein Eu (86).

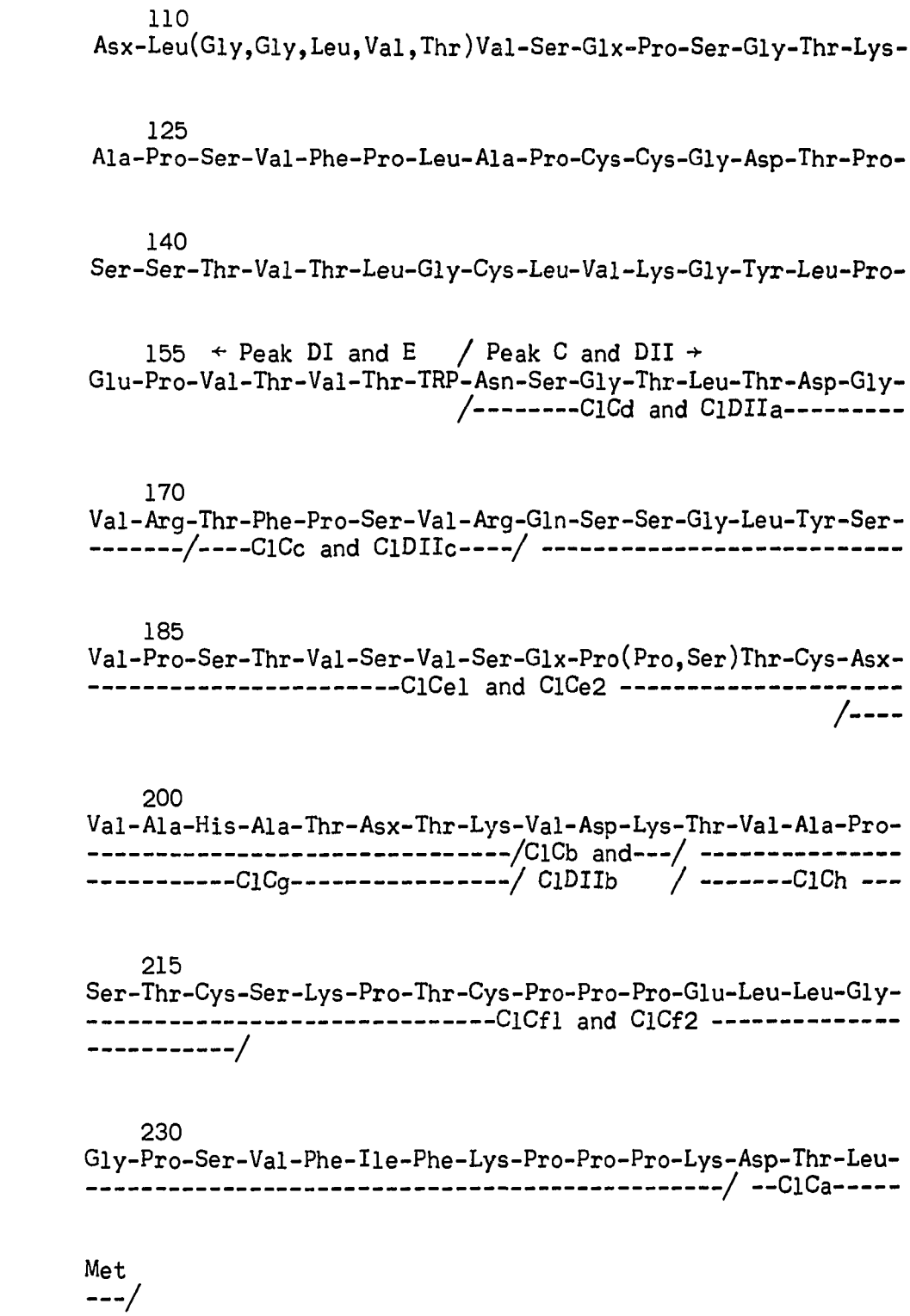

Fig. 27 - The sequence of rabbit C1 fractions from residue 109 to the C-terminal (79) and the location of isolated peptides.

difference as in the human subtypes.

The composition of peak DII is closer to the composition of the polypeptide sequenced by Porter's group than is the composition of peak C. If the IgG sequenced by Porter's group also had sequences like that of my peak C, the method of isolation selectively isolated the peptides that were like peak DII.

If oxidation and cleavage occurred at any residue besides tryptophan, the most likely amino acid to be affected would be tyrosine. Cl chains had 10.4 tyrosines per chain as shown in Table 2. If the tyrosine compositions of the peptides from Table 6 are added together in accordance with the position assignments of the peptides, the totals range from 6.9 (B + F or DI + DII) to a maximum of 7.7 (C + E + F) residues of tyrosine per Cl chain. This represents a loss of about 1/3 of the total tyrosine. An examination of the loss is therefore necessary.

Although peak A (Table 4) has a lower tyrosine percentage than Cl, it is richer in tyrosine than the other fractions. The apparently low recovery of tyrosine could be partially explained by the appearance in peak A of peptides with high tyrosine content. It can not be stated at this time why this happens, but at least three possibilities exist; high tyrosine peptides do not cleave by oxidation, high tyrosine peptides tend to aggregate or peptides that are cleaved tend to lose their tyrosine by some mechanism.

The first question to be considered is, did an actual loss of tyrosine take place or is 10.4 residues of tyrosine an artifically high number for some reason? Fruchter's (79) sequence of rabbit heavy chains shows 2 tyrosine residues from residue 110 to the end of the Cl fraction.

Fleischman's (80) sequence of the first 65 residues shows 4 tyrosines. The myeloma protein Eu shows 5 tyrosines for the intervening 45 residues in the Cl fraction. The loss of tyrosine is most likely an actual loss. On the other hand, Fleischman's Cl fraction contained only 7.4 residues, so individual Cl chains do not all contain the higher number of residues.

There are 3 possible explanations for the loss of the tyrosine; oxidation of specific residues, random oxidation, and/or loss through some other mechanism not involving oxidation, such as destruction during acid hydrolysis.

Fruchter's and Fleischman's sequences indicate that the C-terminal and N-terminal peaks each have 1 residue of tyrosine. Peaks C, DII and E have 1.9, 1.7 and 1.3 residues of tyrosine respectively, therefore any specific oxidation and cleavage at tyrosine residues would most likely be confined to the center section of the Cl chain. This is where all of the tyrosine shortages in Table 6 appear. These facts are inconsistent with random oxidation and cleavage at tyrosine residues, but are, in themselves, not strong enough to rule out the possibility.

If specific oxidation and cleavage at some of the tyrosine residues took place, the peptides that would result would be smaller in molecular weight than E peak peptides and would probably elute from the column in the neighborhood of peaks F and G. G contains little peptide material, so peak F would be expected to contain most of any specific tyrosine cleavage products. Peak F is not big enough to contain a great quantity of these products, and it would also be expected to contain products from cleavages at variable tryptophans. Therefore, cleavage at specific tyrosine residues can not be excluded, but it is unlikely that

it occurs to any great extent.

An argument has already been advanced against random cleavage at tyrosine residues, but further consideration is warranted. If 1/3 of the tyrosine residues were randomly oxidized and 100% cleavage occurred, only a 30% yield would result. Complete cleavage is unlikely, so a much lower yield would probably result. As mentioned in the previous paragraph, these peptides that were released would tend to fall into the chromatogram in the region of peak F. Therefore, random cleavage, if it occurs, does not likely proceed to a very large extent.

If oxidation is the cause of the loss of tyrosine, some predictions can be made concerning it. It is most likely an oxidation of specific tyrosine residues in the middle section of the C1 chain, and the oxidation does not lead to cleavage of the polypeptide chain in the majority of cases.

A third possibility exists that the tyrosine residues were lost by some other mechanism not involving oxidation. One argument for his possibility is that Cebra (77) recovered about 6 of his 9.4 tyrosine residues from his fractionation which did not involve any oxidation. This loss is about the same ratio of loss that tryptophan oxidation shows, so the loss may not be due to oxidation. Cebra's procedures included a fractionation on Sephadex as did these experiments. A very hydrophobic peptide containing most of the tyrosine residues (the middle section peptide is both hydrophobic and high in tyrosine) may have been lost in both fractionations.

Another possibility for loss of tyrosine is during acid hydrolysis. No extra precautions were taken to prevent this, but normal pre-

cautions were taken. Constant boiling hydrochloric acid was used for hydrolysis in an evacuated, Pyrex test tube.

To sum up these arguments, a loss of tyrosine occurred that is confined to the middle section of the Cl chains. Some of this loss can be explained by selective appearance of peptides that contain high amounts of tyrosine in peak A. Other losses of tyrosine besides oxidation have been discussed. It is also possible that oxidation and cleavage of tyrosine residues has occurred, but it is unlikely that this reaction occurred to a large extent.

On the other hand, it is possible to show most of the cleavage occurred at tryptophan residues. The oxidation spectra from the chains indicated tryptophan oxidation (loss of OD at 280 mμ). Unfortunately, calculation of the amount of tryptophan oxidized was not possible because of the 280 mμ absorption of the bromourea complex.

From previously sequenced rabbit chains, a predicted composition of the products was made. The isolated products were compared to these predicted compositions and found to be close except where earlier noted. Among these peptides were found 2 peptides that could only have come from the C-terminal portion of the Cl fragment. This was proved by showing that they contained peptides (ClCb, ClCc, ClDIIb and ClDIIc) which had been previously shown by Cebra (77) to come from this section of the molecule. In addition, other peptides were found (ClCd and ClDIIa) that could have only arisen by chain cleavage of the tryptophan at residue 160 (Fruchter's numbering, Figure 27). Fruchter's sequence from residue 160 to 165 is -Trp-Asn-Ser-Gly-Thr-Leu-. Peak C had a glycine end group and peptide ClCd had a sequence of Gly-Ser-Gly-Thr-Leu-. So, the homol-

ogy between these 2 sequences is very strong. Peptide ClDIIa had a composition that is identical to that of the corresponding peptide from Fruchter's sequence. Peak C and peak DII account for 48% of the total C-terminal sections in the original Cl fraction. From previously assigned positions we saw that peaks DI and E were the peptides split from the C-terminal section, or to put it another way, they also arose from the split at residue 158. Peaks DI and E represent 50% of the total middle section. Therefore, for each C-terminal fraction produced, an intact middle section peptide was also produced. If any appreciable amount of cleavage at tyrosine residues had occurred within either of these sections, these ratios would have varied more widely.

The N-bromosuccinimide oxidation of tryptophan has allowed the isolation of 2 different peptides that come from the same section of the molecule, C and DII. This has shown that in the rabbit, at least, the constant portion of the Fd chain is not as constant as had been thought and may indicate subtype differences. Little work has been done on the DII peptides, but the peak C peptides have been studied enough to show some tendencies. There is an obvious difference between the N-terminal peptides from fractions C and DII, but differences also show up between some peak C peptides studied. Peptides ClCe1 and ClCe2 have different amino acid compositions and electrophoretic mobilities as do peptides ClCf1 and ClCf2. Therefore, differences exist even between peak C peptides which come from the same position. This indicates the nature of the problem that Porter's group overcame in isolating peptides homogeneous enough to sequence from this section of the molecule.

The oxidation of tryptophan by N-bromosuccinimide to produce

chain cleavage has been known for several years, but the method has never been widely used because of the low yields of peptide bond hydrolysis that it produces. Because of this low yield, nearly any other chain cleavage method has been used in preference to it. In this paper, an application has been shown that may be preferable to the method used by Porter's group (77) to obtain the sequence. They obtained their peptides by blocking lysine and digesting with trypsin. Yield data for their method was missing, but visual examination of the chromatograms in their publications indicates that their yields were not any higher than the per cent yields for oxidation with N-bromosuccinimide. Since tryptophan is more constant in a sequence than arginine, its invariability, in this case, makes it as useful as trypsin digestion. In addition, N-bromosuccinimide cleavage is easier to perform.

The C1 chain is a good example of the type of polypeptide chain that can be most effectively cleaved by oxidation. There are no methionine residues in constant locations, so cyanogen bromide cleavage can not be used to cleave the large polypeptide chain into shorter chains. N-bromosuccinimide cleavage is an easy and quick method for achieving the cleavage of a large polypeptide chain into 2 or 3 smaller polypeptide chains. It works best if only 1 or 2 tryptophan residues are present in the polypeptide chain. With more tryptophan residues than this, the peptide patterns can quickly become complicated by all of the different overlapping peptides. In the case of the C1 chain there were 3 constant tryptophan residues, but 2 of them were close enough together that the resulting peptides did not differ greatly in size.

This dissertation has demonstrated that the cleavage of light

chains and heavy chain C1 fractions by tryptophan oxidation is a good approach to the structure of these polypeptides. This method also is an approach to the separation of the constant and variable regions of IgG. N-bromosuccinimide oxidation may not be the most effective method, so work should now be done to examine other methods of oxidation to see if they are more effective. Atassi's method (131) utilizing periodate may prove to be more valuable in this application, and Omenn's method (132) using an oxidizing intermediate between N-bromosuccinimide and the tryptophan, may also have value. In addition chromatography on Sephadex G-100 superfine might improve separations between the various fractions obtained.

CHAPTER V

SUMMARY

Tryptophan appears to be located in relatively constant positions in IgG chains. These positions are so situated that cleavage of these tryptophan residues by oxidation would produce fractions that enable the variable and constant region to be separated. Only about 40 residues of the constant region would be found in the variable fractions.

The constant tryptophan residues are at residue 36 and 148 in the light chain. Cleavage at these positions would result in 3 peptides; an N-terminal peptide of about 36 residues, a C-terminal peptide of about 69 residues and a middle section of about 118 residues. Since oxidation of tryptophan does not produce 100% cleavage, overlap peptides would also result. These peptides are different sizes and Sephadex chromatography should separate them. An analogous situation can be set up in the heavy chain by cyanogen bromide cleavage to produce the N-terminal half of the heavy chain. The constant tryptophan residues in this fraction are at the same approximate locations as those of the light chain.

Rabbit IgG light chains were oxidized by N-bromosuccinimide. The Sephadex chromatography of the resulting polypeptides yielded 6 peaks. The peaks were identified by order of elution and by amino acid analysis. The results were as predicted from position of the tryptophan residues.

The N-bromosuccinimide cleavage of the heavy chain N-terminal fraction produced a chromatography that was different than expected. By amino acid analysis and peptide mapping the identity of the peaks was established. There were 2 different peaks found that each represent the C-terminal polypeptide. Although the order of elution of the peptides from the Sephadex was not as expected, all the peptides that were expected were found.

The usefulness of the oxidation of constant tryptophan residues in structural work on IgG has been demonstrated. Further work will be needed to determine the best method for this oxidation.

BIBLIOGRAPHY

1. Humphrey, J. H., and White, R. G., in Immunology for Students of Medicine, F. A. Davis Company, Philadelphia, Pa., 1964, p. 1.

2. Felton, L. D., J. Immun., 22, 453 (1932).

3. Tiselius, A., Biochem. J., 31, 1964 (1937).

4. Chase, M. W., and Landsteiner, K., Ann. Rev. Biochem., 8, 579 (1939).

5. Marrack, J. R., Ann. Rev. Biochem., 11, 629 (1942).

6. Mayer, M. M., Ann. Rev. Biochem., 20, 415 (1951).

7. Haurowitz, F., Ann. Rev. Biochem., 29, 609 (1960).

8. Porter, R. R., and Press, E. M., Ann. Rev. Biochem., 31, 625 (1962).

9. Cohen, S., and Porter, R. R., Advance. Immunol., 4, 287 (1964).

10. Bull. World Health Organ., 30, 447 (1964).

11. Bull. World Health Organ., 33, 721 (1965).

12. Bull. World Health Organ., 35, 953 (1966).

13. Bull. World Health Organ., 38, 151 (1968).

14. Biochemistry, 9, 3469 (1970).

15. Porter, R. R., J. Biochem., 46, 473 (1950).

16. Edelman, G. M., J. Amer. Chem. Soc., 81, 3155 (1959).

17. Edelman, G. M., and Poulik, M. D., J. Exp. Med., 113, 861 (1961).

18. Franek, F., Biochem. Piophys. Res. Commun., 4, 28 (1961).

19. Fleischman, J. B., Pain, R. H., and Porter, R. R., Arch. Biochem. Biophys., Suppl. 1, 174 (1962).

20. Utsumi, S., and Karush, F., Biochemistry, 3, 1329 (1964).

21. van Eyk, H. G., and Porath, J., Acta Chem. Scand., 21, 866 (1967).

22. Porter, R. R., in Symposium on Basic Problems in Neoplastic Disease, Gellhorn, A., and Hirschberg, E., ed., Columbia University Press, 1962, p. 177.

23. Fleischman, J. B., Porter, R. R., and Press, E. M., Biochem. J., 88, 220 (1963).

24. Pain, R. H., Biochem. J., 88, 234 (1963).

25. Crumpton, M. J., and Wilkerson, J. M., Biochem. J., 88, 228 (1963).

26. Small, P. A., Jr., Kehn, J. E., and Lamm, M. E., Science, 142, 393 (1963).

27. Cohen, S., Biochem. J., 89, 334 (1963).

28. Porter, R. R., Nature, 182, 670 (1958).

29. Porter, R. R., Biochem. J., 73, 119 (1959).

30. Cohen, S., and Milstein, C., Advance. Immunol., 7, 1 (1967).

31. Nisonoff, A., Wissler, F. C., and Woernley, D. L., Arch. Biochem. Biophys., 88, 241 (1960).

32. Nisonoff, A., and Woernley, D. L., Nature, 183, 1325 (1959).

33. Nisonoff, A., Wissler, F. C., Lipman, L. N., and Woernley, D. L., Arch. Biochem. Biophys., 89, 230 (1960).

34. Nisonoff, A., Markus, G., and Wissler, F. C., Nature, 189, 293 (1961).

35. Nisonoff, A., Wissler, F. C., and Lipman, L. N., Science, 132, 1770 (1960).

36. Palmer, J. L., Mandy, W. J., and Nisonoff, A., Proc. Nat. Acad. Sci., 48, 49 (1962).

37. Goodman, J. W., and Gross, D., J. Immun., 90, 865 (1963).

38. Palmer, J. L., and Nisonoff, A., Biochemistry, 3, 863 (1964).

39. Palmer, J. L., Nisonoff, A., and Van Holde, K. E., Proc. Nat. Acad. Sci., 50, 314 (1963).

40. Nisonoff, A., and Dixon, D. J., Biochemistry, 3, 1338 (1964).

41. Inman, F. P., and Nisonoff, A., J. Biol. Chem., 241, 322 (1966).

42. Cebra, J. J., Givol, D., Silman, H. I., and Katchalski, E., J. Biol. Chem., 236, 1720 (1961).

43. Jaquet, H., and Cebra, J. J., Biochemistry, 4, 954 (1965).

44. Marler, E., Nelson, C. A., and Tanford, C., Biochemistry, 3, 279 (1964).

45. Noelken, M. E., Nelson, C. A., Buckley, C. E., III, and Tanford, C., J. Biol. Chem., 240, 218 (1965).

46. Putnam, F. W., Easley, C. W., and Lynn, L. T., Biochim. Biophys. Acta, 58, 279 (1962).

47. Givol, D., Biochem. J., 104, 39c (1967).

48. Cahnmann, H. J., Arnon, R., and Sela, M., J. Biol. Chem., 240, PC 2762 (1965).

49. Cahnmann, H. J., Arnon, R., and Sela, M., J. Biol. Chem., 241, 3247 (1966).

50. Lahav, M., Arnon, R., and Sela, M., J. Exp. Med., 125, 787 (1967).

51. Givol, D., and Lorenzo, F. D., J. Biol. Chem., 243, 1886 (1968).

52. Edelman, G. M., and Gally, J. A., J. Exp. Med., 116, 207 (1962).

53. Schwartz, J. H., and Edelman, G. M., J. Exp. Med., 118, 41 (1963).

54. Mannik, M., and Kunkel, H., J. Exp. Med., 116, 859 (1962).

55. Mannik, M., and Kunkel, H. G., J. Exp. Med., 117, 213 (1963).

56. Potter, M., Dreyer, W. J., Kuff, E. L., and McIntire, K. R., J. Mol. Biol., 8, 814 (1964).

57. McIntire, K. R., Asofsky, R. M., Potter, M., and Kuff, E. L., Science, 150, 361 (1965).

58. Grey, H. M., and Kunkel, H. G., J. Exp. Med., 120, 253 (1964).

59. Haber, E., Proc. Nat. Acad. Sci., 52, 1099 (1964).

60. Whitney, P. L., and Tanford, C., Proc. Nat. Acad. Sci., 53, 524 (1965).

61. Whitney, P. L., and Tanford, C., J. Biol. Chem., 240, 4271 (1965).

62. Sela, M., Proc. Roy. Soc. Ser. B., 166, 188 (1966).

63. Wilker, M., Titani, K., Shinoda, T., and Putnam, F. W., J. Biol. Chem., 242, 1668 (1967).

64. Hilschmann, N., and Lyman, C. C., Proc. Nat. Acad. Sci., 53, 1403 (1965).

65. Titani, K., Whitly, E., Jr., Avogardo, L., Putnam, F. W., Science, 149, 1090 (1965).

66. Milstein, C., Nature, 209, 370 (1966).

67. Milstein, C., Biochem., J., 101, 352 (1966).

68. Milstein, C., Nature, 205, 1171 (1965).

69. Milstein, C., Biochem. J., 101, 338 (1966).

70. Frangione, B., and Franklin, E. C., J. Exp. Med., 122, 1 (1965).

71. Frangione, B., Prelli, F., and Franklin, E. C., Immunochemistry, 4, 95 (1967).

72. Nelson, C. A., Noelken, M. E., Buckley, C. E., III, Tanford, C., and Hill, R. L., Biochemistry, 4, 1418 (1965).

73. Givol, D., and Porter, R. R., Biochem. J., 97, 32c (1965).

74. Givol, D., and Porter, R. R., Israel J. Chem., 4, 68p (1966).

75. Hill, R. L., Delaney, R., Lebovitz, H. E., and Fellows, R. E., Jr., Proc. Roy. Soc. Ser. B., 166, 159 (1966).

76. Wilkinson, J. M., Press, E. M., and Porter, R. R., Biochem. J., 100, 303 (1966).

77. Cebra, J. J., Givol, D., and Porter, R. R., Biochem. J., 107, 69 (1968).

78. Cebra, J. J., Steiner, L. A., and Porter, R. R., Biochem. J., 107, 79 (1968).

79. Fruchter, R. G., Jackson, S. A., Mole, L. E., and Porter, R. R., Biochem. J., 116, 249 (1970).

80. Fleischman, J. B., Biochemistry, 10, 2753 (1971).

81. Pink, J. R. L., Milstein, C., Nature, 214, 92 (1967).

82. Frangione, B., and Milstein, C., Biochem. J., 106, 15 (1968).

83. Frangione, B., and Milstein, C., Nature, 216, 939 (1967).

84. Frangione, B., and Milstein, C., J. Mol. Biol., 33, 893 (1968).

85. Pink, J. R. L., and Milstein, C., Nature, 216, 941 (1967).

86. Edelman, G. M., Cunningham, B., Gall, W. E., Bottlieb, P. D., Rutishauser, U., and Waxdal, M. J., Proc. Nat. Acad. Sci., 63, 78 (1969).

87. O'Donnell, I. J., Frangione, B., and Porter, R. R., Biochem. J., 116, 261 (1970).

88. Strausbauch, P. H., Hurwitz, E., and Givol, D., Biochemistry, 10, 2231 (1971).

89. Frangione, B., Milstein, C., and Pink, J. R. L., Nature, 221, 145 (1969).

90. Singer, S. J., and Doolittle, R. F., Science, 153, 13 (1966).

91. Hill, R. L., Delaney, R., Fellows, R. E., Jr., and Lebovitz, H. E., Proc. Nat. Acad. Sci., 56 (6), 1762 (1966).

92. Porter, R. R., and Press, E. M., Biochem. J., 97, 32p (1965).

93. Press, E. M., Piggot, P. J., and Porter, R. R., Biochem. J., 99, 356 (1966).

94. Press, E. M., and Hogg, N. M., Biochem. J., 117, 641 (1970).

95. Edelman, G. M., and Gall, W. E., Ann. Rev. Biochem., 38, 415 (1969).

96. Edelman, G. M., Biochemistry, 9, 3197 (1970).

97. Edelman, G. M., Olins, D. E., Gally, J. A., and Zinder, N. D., Proc. Nat. Acad. Sci., 50, 753 (1963).

98. Yoo, T. J., Roholt, O. A., and Pressman, D., Science, 157, 707 (1967).

99. Roholt, O., Onoue, K., and Pressman, D., Proc. Nat. Acad. Sci., 51, 173 (1964).

100. Grey, H. M., and Mannik, M., J. Exp. Med., 122, 619 (1965).

101. Bridges, S. H., and Little, J. R., Biochemistry, 10, 2525 (1971).

102. Dorrington, K. J., Zarlingo, M. H., and Tanford, C., Proc. Nat. Acad. Sci., 58, 996 (1967).

103. Pressman, D., and Roholt, O., Proc. Nat. Acad. Sci., 47, 1606 (1961).

104. Roholt, O., Shaw, A., and Pressman, D., Nature, 196, 773 (1962).

105. Kitagawa, M., Yagi, Y., and Pressman, D., J. Immun., 95, 446 (1965).

106. Roholt, O. A., Pressman, D., Immunochemistry, 5, 265 (1968).

107. Wofsy, L., Metzger, H., and Singer, S. J., Biochemistry, 1, 1031 (1962).

108. Metzger, H., Wofsy, L., and Singer, S. J., Proc. Nat. Acad. Sci., 51, 612 (1964).

109. Fenton, J. W., II, and Singer, S. J., Biochem. Biophys. Res. Commun., 20, 315 (1965).

110. Doolittle, R. F., and Singer, S. J., Proc. Nat. Acad. Sci., 54, 1773 (1965).

111. Good, A. H., Ovary, Z., and Singer, S. J., Biochemistry, 7, 1304 (1968).

112. Singer, S. J., and Thorpe, N. O., Proc. Nat. Acad. Sci., 60, 1371 (1968).

113. Goetzl, E. J., Metzger, H., Biochemistry, 9, 1267 (1970).

114. Kabat, E. A., J. Immun., 84, 82 (1960).

115. Prahl, J. W., Biochem. J., 104, 647 (1967).

116. Goodman, J. W., Science, 139, 1292 (1963).

117. Utsumi, S., and Karush, F., Biochemistry, 4, 1766 (1965).

118. Henney, C. S., Jefferis, R., and Stanworth, D. R., Biochim. Biophys. Acta, 154, 295 (1968).

119. Feinstein, A., and Rowe, A. J., Nature, 205, 147 (1965).

120. Valentine, R. C., and Green, N. M., J. Mol. Biol., 27, 615 (1967).

121. Nature, 231, 495 (1971).

122. Gray, W. R., Proc. Roy. Soc. Ser. B., 166, 146 (1966).

123. Williamson, A. R., Nature, 231, 359 (1971).

124. Litman, G. W., Frommel, D., Chartrand, S. L., Fingstad, J., and Good, R. A., Immunochemistry, 8, 345 (1971).

125. Delaney, R., Private communication.

126. Karlsson, F. A., Peterson, P. A., and Berggard, I., Proc. Nat. Acad. Sci., 64, 1257 (1969).

127. Dayhoff, M. O., in Atlas of Protein Sequence and Structure, 4, National Biomedical Research Foundation, Silver Spring, Maryland, 1969.

128. Witkop, B., Adv. Prot. Chem., 16, 224 (1961).

129. Ramachandran, L. K., and Witkop, B., Methods in Enzymology, Hirs, C. H. W., ed., XI, 283 (1967).

130. Spande, T. F., Witkop, B., Methods in Enzymology, Hirs, C. H. W., ed., XI, 498 (1967).

131. Atassi, M. Z., Arch. Biochem. Biophys., 120, 56 (1967).

132. Omenn, G. S., Fontana, A., and Anfinsen, C. B., J. Biol. Chem., 245, 1895 (1970).

133. Kekwick, R. A., Biochem. J., 34, 1248 (1940).

134. Sober, H. A., and Peterson, E. A., Federation Proc., 17, 1116 (1958).

135. Steinbuch, M., and Audran, R., Arch. Biochem. Biophys., 134, 279 (1969).

136. Offord, R. E., Nature, 211, 591 (1966).

137. Gray, W. R., Methods in Enzymology, Hirs, C. H. W., ed., XI, 139 (1967).

138. Woods, K. R., and Wang, K. T., Biochim. Biophys. Acta, 133, 369 (1967).

139. Gray, W. R., Methods in Enzymology, Hirs, C. H. W., ed., XI, 469 (1967).

140. Gray, W. R., and Smith,, J. F., Analytical Biochem., 33, 36 (1970).

141. Porter, R. R., Biochem. J., 105, 417 (1967).