Disulfide Bridge Peptides and Glycopeptides of a Human IgA1 Myeloma Globulin†

Virginia Moore‡ and Frank W. Putnam

ABSTRACT: An IgA1 human myeloma protein was reduced and alkylated with [^{14}C]iodoacetic acid, and the α1 heavy chain was separated and digested with trypsin to study the disulfide bridge peptides. Four tryptic peptides were isolated and were sequenced by manual procedures and by use of the automatic sequencer. These account for four intrachain half-cystine residues and include a 30-residue peptide and a 27-residue glycopeptide, rich in leucine. The carbohydrate composition and location of the oligosaccharide in the latter were determined. Fragments containing six more half-cystine residues were prepared from other large tryptic glycopeptides and were sequenced. The unblocked amino terminus of the α1 chain was analyzed with the sequencer. Altogether, sequence data were obtained for some 130 residues including the regions around 11 half-cystines and two oligosaccharides. The location of most of these in the α1 chain is unknown, but the probable function of each of the half-cystine residues is deduced by homology to other heavy chains.

Comparative structural study of immunoglobulins of different classes is needed to elucidate the common basis of their antibody function and genetic control and to ascertain their evolutionary relationships. The three major classes of immunoglobulins (IgG, IgM, and IgA) consist of a pair of light chains (either κ or λ) joined to a pair of heavy chains; the latter are characteristic of the class: γ for IgG, μ for IgM, and α for IgA (Putnam, 1969). Unlike IgG immunoglobulins, IgM and IgA have a high carbohydrate content (about 10%), a similar C-terminal nonapeptide sequence (Putnam *et al.*, 1971; Prahl *et al.*, 1971), and an apparently identical joining piece (J chain) that facilitates their polymerization (Halpern and Koshland, 1970; Mestecky *et al.*, 1971). Because of the heterogeneity of normal immunoglobulins and most antibodies of all species, amino acid sequence analysis is most readily done on the homogeneous myeloma globulins and macroglobulins of human patients (Putnam, 1969).

The complete amino acid sequence of several human IgG myeloma globulins is known (Cunningham *et al.*, 1969; Edelman *et al.*, 1969; Press and Hogg, 1970; Ponstingl *et al.*, 1970), and the nearly complete sequence of a human IgM immunoglobulin has been reported (Shimizu *et al.*, 1971a; Putnam *et al.*, 1972), but only partial sequences of some disulfide bridge peptides have been published for human IgA (Prahl *et al.*, 1971; Wolfenstein *et al.*, 1971; Wolfenstein-Todel *et al.*, 1972). The dearth of data is indicated by the fact that no sequence results for α chains of any species are given in the Atlas of Protein Sequence and Structure (Dayhoff, 1972). In our continuing comparative study we have undertaken amino acid sequence analysis of the human myeloma IgA globulin designated Ha. The complete primary structure of the Ha light chain, which is of the λ type, has been reported (Shinoda *et al.*, 1970). In the present work the sequence has been determined around some 140 residues of the Ha α chain including 11 different half-cystine residues comprising the interchain hinge region, the amino terminus, and 6 intrachain bridge peptides including 2 large glycopeptides of different carbohydrate content.

Experimental Section

Materials. The isolation, physicochemical characterization, and antigenic properties of the human myeloma IgA protein Ha have been described and the amino acid composition and peptide maps of the α heavy chain and λ light chain have been reported (Bernier *et al.*, 1965). The carbohydrate composition of the glycopeptides of the α heavy chain was determined by Clamp and Putnam (1967). The complete amino acid sequence of the λ Bence-Jones protein excreted by the same patient has been established (Shinoda *et al.*, 1970). The urinary Bence-Jones protein and the light chain from the serum IgA myeloma globulin are identical in peptide maps and amino acid composition and presumably in amino acid sequence (Bernier *et al.*, 1965). The α1 chain belongs to the V_{HIII} heavy-chain subgroup as shown by sequence analysis of the first 15 residues (Köhler *et al.*, 1970). The antigenic subclass typing as an IgA1 protein was performed by Dr. H. G. Kunkel of the Rockefeller University. Details of experimental procedures in the following section are given by Moore (1972).

Amino Acid Analysis and Sequence Determination. Samples hydrolyzed for 22 hr with 6 N HCl at 110° in an evacuated sealed tube were analyzed with the Beckman Model 120B and 120C automatic amino acid analyzers equipped with high-sensitivity cuvets and recorders. The amino sugars glucosamine and galactosamine were determined from the short-column analyses. For procedures and accuracy of the method see Clamp and Putnam (1964, 1967) and Shimizu *et al.* (1971b). The complete carbohydrate composition of one glycopeptide (T3-Th3) was determined by Dr. John R. Clamp, University of Bristol, England, by gas–liquid chromatography (Clamp *et al.*, 1967; Bhatti *et al.*, 1970). Sequences of small peptides and subpeptides were determined by the dansyl-Edman procedure (Gray, 1967) and the dansyl derivatives were characterized by the method of Woods and Wang (1967). The methods for isolation and characterization of peptides are the same as

† From the Department of Zoology, Indiana University, Bloomington, Indiana 47401. *Received December 1, 1972.* This work was supported in part by grants from the National Institutes of Health (CA 08497), the American Cancer Society (NP-10), the National Science Foundation (NSF GB 29377), and the Damon Runyon Fund for Cancer Research (DRG 1134). Part of this work is taken from the doctoral dissertation of V. M. (1972).

‡ Present address: Division of Natural Sciences, University of California, Santa Cruz, Calif.

those described previously (Shinoda *et al.*, 1970; Titani *et al.*, 1970). Enzymatic digestion with carboxypeptidase A, carboxypeptidase B, and aminopeptidase M were used where appropriate (Shinoda *et al.*, 1970). Several large peptides (T2 and T3) were sequenced by automatic degradation with the Beckman Model 890 Sequencer. For a general reference to procedures see Hermodson *et al.* (1972). The phenylthiohydantoin derivatives were identified by gas chromatography with two columns (Pisano and Bronzent, 1969) with verification by analysis of the regenerated amino acid as needed. The carbohydrate composition of the glycopeptide T3-Th3 was determined by gas–liquid chromatography by Dr. John R. Clamp of the University of Bristol Medical School, Bristol, England.

Results

Purification of Peptides. REDUCTION, ALKYLATION, AND CHAIN SEPARATION. The Ha protein was completely reduced in 7 M guanidine–0.5 M Tris buffer (pH 8.5) and 0.005 M EDTA; 0.1 ml of β-mercaptoethanol was used per 100 mg of protein with reaction under nitrogen for 2 hr at room temperature. The reaction mixture was cooled to 4°, and alkylated with [^{14}C]iodoacetic acid (100 μCi with specific activity of 15.5 Ci/mol) plus enough nonradioactive iodoacetic acid to provide a 10% molar excess over the amount of β-mercaptoethanol used. After 30 min the reaction mixture was dialyzed into 0.15 M NaCl and then into 1 N acetic acid. Completeness of alkylation was checked by amino acid analysis. Chain separations were performed on Sephadex G-100 in 1 N acetic acid (column 5 × 100 cm), with the column effluent being monitored by absorbance at 280 nm. The heavy and light chains were then lyophilized.

TRYPTIC DIGESTION AND SEPARATION OF PEPTIDES. A 1% solution of 1.4 g of carboxymethylated α chain (about 28 μM) was digested for 6 hr with 25 mg of Tos-PheCH_2Cl-trypsin from Worthington Biochemical Corp. The reaction mixture was kept at pH 8.5 by addition of 1 N NaOH using a pH-Stat. After digestion, the pH was adjusted to pH 3 with glacial acetic acid. The insoluble core material was separated by centrifugation. The soluble material was concentrated by rotary evaporation and applied to Sephadex G-25 (two columns, 5 × 150 cm and 5 × 100 cm, linked in tandem) for preliminary fractionation. The column effluent was monitored by absorbance at 280 nm, ninhydrin color, and by radioactivity as measured in a Nuclear-Chicago flow-through scintillation counter. Eight major peptide pools were obtained; the first four pools contained most of the radioactivity. These four were applied to Chromo-Beads[1] type P cation-exchange resin (column 0.9 × 50 cm) connected to a Technicon peptide analyzer and a flow-through scintillation counter. The column was eluted with a concave pyridine acetate gradient from 0.05 (pH 2.55) to 2 M (pH 4.5), followed by a 4 M pyridine wash. The individual peaks were lyophilized. Specific staining techniques described by Easley (1965) were applied for sulfur-positive amino acids, histidine, tyrosine, arginine, and tryptophan.

Purification of Radioactive Tryptic Peptides. The intent was to obtain only the cysteine-containing peptides and to isolate them in high purity regardless of yield. All radioactive peaks were characterized by high-voltage electrophoresis at pH 6.5 and 3.7, and by paper chromatography in 1-butanol–acetic acid–water–pyridine (150:30:120:100, v/v). High-voltage electrophoresis at pH 1.9 was performed on samples that did not move well in the above systems. The system which provided the best resolution was repeated on a preparative scale for each peptide, and the samples were eluted from paper with 1 N acetic acid.

The amino acid composition of the purified tryptic peptides containing [^{14}C]Cm-cysteine is given in Table I. Since some of these peptides approached or exceeded 30 residues in length and also contained carbohydrate, they were very difficult to purify. In several cases enzymatic or partial acid hydrolysates of the large tryptic peptides were made, and only the radioactive subpeptides containing Cm-cysteine were isolated.

THE HINGE PEPTIDE. A large, impure, radioactive peptide with N-terminal leucine was obtained in the first peak from the Sephadex G-25 separation of the tryptic peptides. In amino acid composition this closely resembled the hinge peptide described by Grey *et al.* (1971) and by Wolfenstein *et al.* (1971) because it had a very high content of proline and was rich in Cm-cysteine, serine, and threonine. This made it very difficult to evaluate the stoichiometry. This material contained both galactosamine and glucosamine, whereas the hinge peptide (a tryptic-peptic peptide) contains only galactosamine. The material was assumed to be a mixture of two large glycopeptides because of the presence of the two sugars and the uncertain stoichiometry. Several attempts to separate the glycopeptides failed.

Amino Acid Sequence of the Triacontapeptide T2. PEPTIDE T2. This 30-residue acidic peptide, obtained by tryptic digestion of the α chain, was eluted in the second peak on Sephadex G-25 and was purified by chromatography on Chromo-Beads (Figure 1). The peptide moved as a single spot in paper electrophoresis at pH 1.9, 3.7, and 6.5 and in paper chromatography. Prior to the availability of the automatic sequencer, the sequence of T2 was determined by the manual method on the intact peptide and on its subpeptides. The dansyl-Edman procedure was used to identify the first 8 steps (Figure 2). Most of these steps were verified with the sequenator which extended the sequence considerably.

Separate chymotryptic and thermolytic digestions were done on aliquots of 2 μmol of peptide with each enzyme. The peptides were separated by preparative paper electrophoresis at pH 3.7 and were sequenced by the dansyl-Edman procedure and by digestion with carboxypeptidase A, carboxypeptidase B, and aminopeptidase N where appropriate. The chymotryptic peptides account for all the residues in T2, and the thermolysin peptides provide the necessary overlaps.

PEPTIDE T2-Ch1. Amino acid analysis of T2-Ch1 showed 13 residues. The sequence of the first ten steps was established by the dansyl-Edman method as: Asx-Phe-Pro(Pro)Ser-Glx-Asx-Ala-Ser-Gly. Digestion with carboxypeptidase A for 2 hr revealed only tyrosine, but 67% of the leucine and 80% of the tyrosine were released after 24 hr.

PEPTIDE T2-Ch2. The dansyl-Edman procedure for five steps on T2-Ch2 gave the sequence Thr-Thr-Ser-Ser-Gln-Leu. Glutamine, rather than glutamic acid must be present because the peptide had a neutral mobility at pH 6.5. This is supported by the finding that carboxypeptidase A and aminopeptidase M digestion of T2-Ch2 revealed no glutamic acid but only a large peak at the threonine–serine position, which is where glutamine also is eluted in amino acid analysis. The glutamine could not be quantitated because the peptide contains threonine and serine. Carboxypeptidase A digestion for 2 hr released all the COOH-terminal leucine, and the composite peak increased in size with longer periods of digestion.

[1] Chromo-Beads type P, an ion-exchange resin supplied by the Technicon Corp., Tarrytown, N. Y., for single-column peptide chromatography.

TABLE I: Amino Acid Composition of α-Chain Peptides Containing [^{14}C]Cm-cysteine.[a]

	Tryptic Peptides				Tryptic-Thermolysin Peptides	
	T2	T3	T4[b]	T5[c]	T6-Th2	T7-Th3
Lysine	1.0 (1)		1.2 (1)	1.1 (1)		0.1
Histidine		1.1 (1)		1.1 (1)		
Arginine	0.2	2.4 (2)				
Cm-cysteine	0.85 (1)	0.8 (1)	0.7 (1)	0.8 (1)	2.0 (2)	1.0 (1)
Aspartic acid	3.1 (3)	2.1 (2)	0.1			
Threonine	4.3 (4)	3.1 (3)	3.0 (3)	*1.0* (1)		
Serine	4.2 (4)	1.8 (2)	1.2 (1)	1.1 (1)	0.9 (1)	1.3 (1)
Glutamic acid	3.3 (3)	*2.0* (2)	1.1 (1)			
Proline	3.5 (3)	2.3 (1)	1.1 (1)			
Glycine	2.1 (2)	2.0 (2)	0.2		1.4 (1)	0.4
Alanine	2.8 (3)	1.7 (2)	1.9 (2)			0.2
Half-cystine						
Valine	0.4			2.0 (2)		
Methionine						0.9 (1)
Isoleucine						
Leucine	3.7 (4)	9.5 (9)			*1.0* (1)	
Tyrosine	0.9 (1)		0.9 (1)		1.1 (1)	
Phenylalanine	*1.0* (1)		*1.0* (1)			*1.0* (1)
Carbohydrate		+				
Total residues	30	27	12	7	6	4
Yield (μM)	11	3.7	1.4	2.9	0.5	0.7
Dansyl	Asx	Leu	Thr	Ser	Leu	Phe

[a] Values are given as residues per mole. Stoichiometry is based on italic residues. [b] T4′ obtained in a yield of 1.7 μM had a similar composition but had 0.4 residue of Cm-cysteine and 0.4 residue of half-cystine. [c] T5′ obtained in a yield of 2.9 μM had a similar composition but had 0.5 residue of half-cystine.

Aminopeptidase M digestion for 18 hr released 64% of the leucine, so the enzyme must have freed the glutamine.

PEPTIDE T2-Ch3. The sequence obtained for T2-Ch3 by five dansyl-Edman steps and by carboxypeptidase A digestion is Thr-Leu-Pro-Ala-Thr(Glx,Cys(Cm))Leu. The fact that carboxypeptidase A digestion released only leucine suggests that Cm-cysteine is next to the carboxyl-terminal leucine.

PEPTIDE T2-Ch4. This low-yield peptide was radioactive and sulfur-positive but no CM-cysteine, cysteic acid, or half-cystine were detected on amino acid analysis. However, serine, a possible breakdown product of Cm-cysteine was present as an impurity. The sequence data for the first five steps are identical to T2-Ch3: Thr-Leu-Pro-Ala-Thr. Therefore, these peptides are considered to represent the same segment of T2, differing only in the fact that Cm-cysteine was cyclized, oxidized, or otherwise destroyed during purification of T2-Ch4.

PEPTIDE T2-Ch5. The sequence of this basic tripeptide was established by the dansyl-Edman procedure as Ala-Gly-Lys, which places it at the COOH terminus of T2.

PEPTIDE T2-Th1. Peptide T2-Th1, which was purified by paper electrophoresis at pH 6.5 and 3.7, contains 18 residues. The stoichiometry was good except for the glycine which is a frequent contaminant from the paper. The amino-terminal sequence by the dansyl-Edman method was Asx-Phe. Carboxypeptidase A digests for 2 and 12 hr showed leucine and tyrosine plus a large peak at the composite position for threonine, serine, and glutamine. Hence, the order of the latter three amino acids could not be determined by this procedure. After 12-hr digestion, 34% of the tyrosine and 24% of the leucine were released. These results are compatible with an overlap of T2-Ch1 and T2-Ch2 giving the sequence Leu-Tyr-Thr-Thr-Ser-Ser-Gln.

PEPTIDE T2-Th2. This acidic tetrapeptide was somewhat impure, but the dansyl-Edman procedure yielded the sequence Ala-Thr-Glx-Cys(Cm). This supplements the sequence found for T2-Ch3.

PEPTIDE T2-Th3. The dansyl-Edman procedure on this basic tetrapeptide gave the sequence Leu-Ala-Gly-Lys. This provides the overlap for T2-Ch3 and T2-Ch5 and completes the ordering of the chymotryptic peptides as shown in Figure 2.

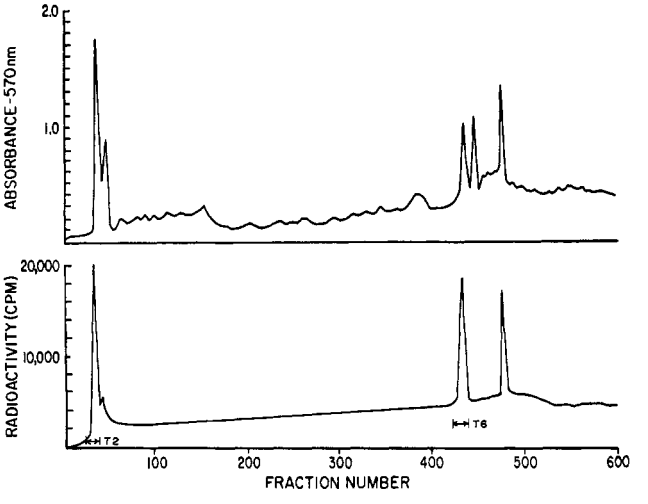

FIGURE 1: Technicon Chromo-Beads purification of peptides T2 and T6 from pool 2 of Sephadex G-25 fractionation.

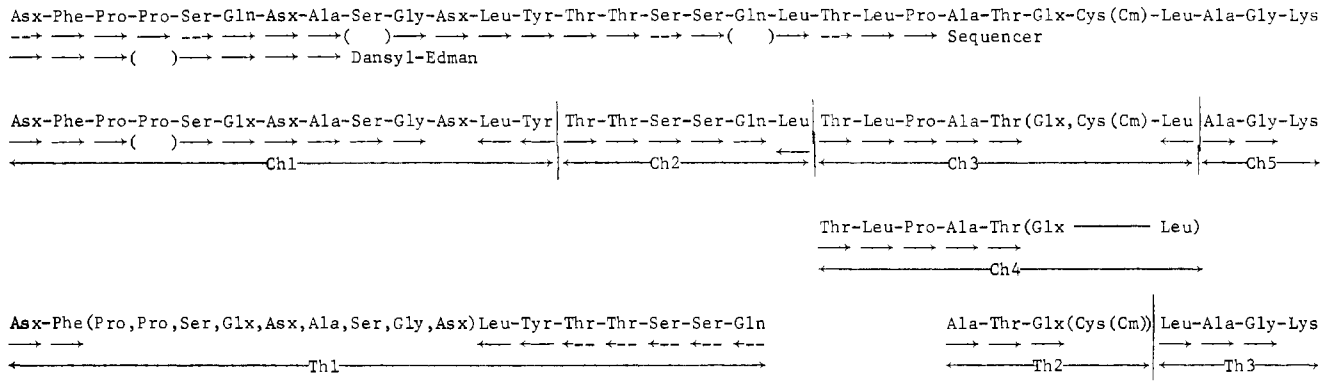

FIGURE 2: Sequence determination of tryptic peptide T2 by manual and automatic methods (→). The chymotryptic (Ch) and thermolytic (Th) peptides were sequenced by the dansyl-Edman method (→) and by use of carboxypeptidase A (←). Unidentified positions are enclosed in parentheses; dash arrows indicate residues that gave a low response in gas chromatography (- - →) cr were ambiguous in enzymatic degradation (←- - -).

SEQUENCER ANALYSIS OF T2. Because the chymotryptic and thermolytic subpeptides of T2 did not provide sufficient overlaps to yield an unambiguous sequence, the whole peptide was twice submitted to automatic sequence analysis with the peptide program. In the first run the peptide was applied to the cup without pretreatment with the Braunitzer reagent 4-sulfophenyl isothiocyanate monohydrate (Braunitzer *et al.*, 1971); this was obtained as Sequenal grade from Pierce Chemical Co., Rockford, Ill. Unclear results were obtained after eight cycles; it was assumed that the peptide had washed out of the cup. The sixth position appeared to be glutamine according to thin-layer chromatography of the phenylthiohydantoin derivative.

The peptide was reisolated by Victor Liu in our laboratory and was treated with the Braunitzer reagent. Automatic sequence analysis was conducted successfully for 22 cycles. At most stages the phenylthiohydantoin derivative was readily identified by gas chromatography. Amino acid analysis was required to verify the dicarboxylic acid residues at positions 6 and 7. The serines at positions 5, 9, and 16 were doubtful in both gas chromatography and amino acid analysis, but the serine at position 17 was clear. No residue was identified for the first position because the amino group is blocked by the Braunitzer reagent. However, the sequencer analysis together with the results on the subpeptides clearly established the structure of the entire peptide (Figure 2).

Sequence of the 27-Residue Glycopeptide T3. PEPTIDE T3. The sequence of the tryptic peptide T3 was very difficult to determine because it is a large glycopeptide and contains an unusual amount of leucine—nine residues out of 27 in the peptide. Peptide T3 eluted in the first peak on Sephadex G-25 and as the sixth peak on Chromo-Beads (Figure 3). The exact number of leucine residues was difficult to determine because the leucine peak on the amino acid analysis chart was off scale or in the inaccurate part of the chart when other amino acids present as a single residue were in the accurate range. This explains why proline which appeared to be present as two residues by amino acid analysis, proved to be present only once according to both manual and automatic sequence analysis.

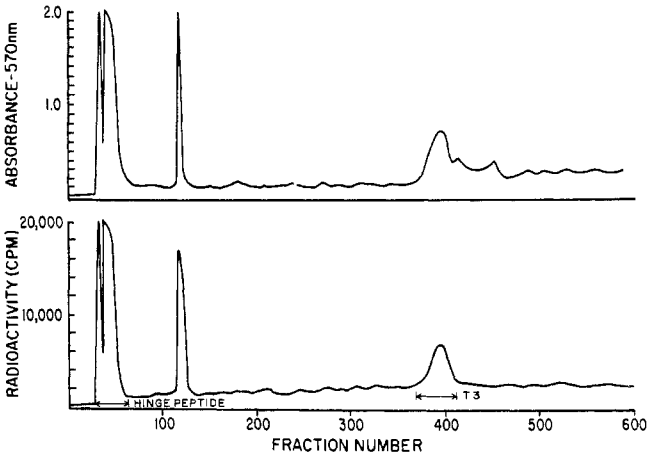

FIGURE 3: Technicon Chromo-Beads purification of the hinge peptide and peptide T3 from pool 1 of Sephadex G-25 fractionation.

Prior to the availability of the automatic sequencer the partial sequence of T3 was determined by the manual degradation method on the intact peptide and on thermolysin subpeptides. The dansyl-Edman procedure was applied to three different samples of T3 to obtain the amino-terminal sequence: Leu-Ser-Leu-His-Arg-Pro-Ala-Leu-Glx-Asx-Leu (Figure 4). The His-Arg-Pro segment presented a great deal of difficulty because histidine is partially destroyed and reacts incompletely in the Edman reaction. This caused considerable overlap on subsequent steps but the sequence was verified from the thermolysin peptides and with the sequencer. Digestion of T3 with a mixture of carboxypeptidase A and carboxypeptidase B for 2 hr showed that the carboxyl-terminal sequence was Thr-Gly-Leu-Arg. The values obtained were: Thr, 0.25; Gly, 0.47; Leu, 1.1, Arg, 1.0. Arginine must be carboxy terminal because this is a tryptic peptide. (The other arginine residue was in an Arg-Pro sequence which is rarely cleaved by trypsin.) Digestion with carboxypeptidase A plus carboxypeptidase B for 8 hr resulted in increased amounts of threonine and leucine and some Cm-cysteine.

Because the sequenator was not available when this peptide was isolated, T3 was subjected to secondary enzymatic cleavage. It was expected that the large number of leucines would result in multiple cleavage by either chymotrypsin or thermolysin. Secondary cleavage on an analytical scale indicated thermolysin gave better results so 1.8 μM of T3 were cleaved preparatively with this enzyme. Several of the eight subpeptides separated were impure because of contamination with submolar amounts of glycine or serine from the paper. Others not illustrated were mixtures of overlapping peptides. Of the subpeptides isolated only T3-Th5 was sulfur positive. On amino acid analysis none of the thermolysin peptides from T3 contained more than trace amounts of Cm-cysteine, even after careful repetition of the acid hydrolysis. However, T3-Th5 and T3-Th4 had 57,000 and 70,000 cpm per min,

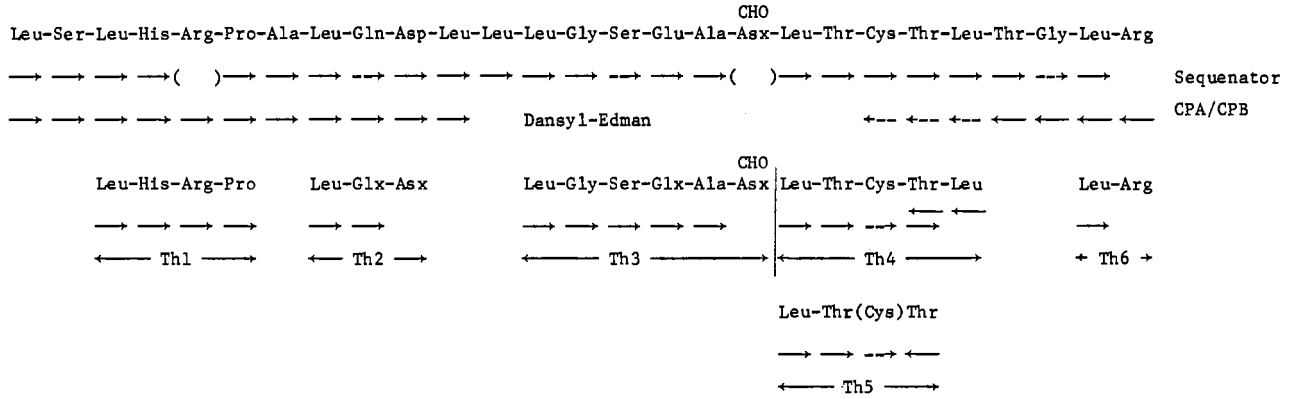

FIGURE 4: Sequence determination of tryptic peptide T3 by manual and automatic methods. Symbols are the same as in Figure 2.

respectively, whereas the other peptides had negligible radioactivity when equivalent aliquots were counted in 6-ml vials in a Beckman LS 230 scintillation counter. Miekka and Deutsch (1970) have also obtained ^{14}C-labeled peptides by reduction-alkylation of disulfide bridges in IgM but without measurable Cm-cysteine on amino acid analysis. They suggested acid degradation of sulfoxide forms of Cm-cysteine to to account for their results, Smyth and Utsumi (1967) have also reported loss of Cm-cysteine on acid hydrolysis, but in their case attributed the disappearance to cyclization of an amino-terminal Cm-cysteine residue.

PEPTIDE T3-Th1. This peptide was recovered in low yield and had considerable contamination with serine and glycine from the paper. However, the dansyl-Edman procedure gave a sequence of Leu-His-Arg-Pro. This peptide fits in the NH_2 terminus of T3. The presence of proline here explains why trypsin did not cleave after the arginine.

PEPTIDE T3-Th2. Two dansyl-Edman steps on this acidic tripeptide gave the sequence Leu-Glx-Asx. This peptide fits within the NH_2-terminal sequence found for T3.

PEPTIDE T3-Th3. This peptide had good stoichiometry and contained all the glucosamine. The sequence was determined by five steps of the dansyl-Edman procedure as Leu-Gly-Ser-Glx-Ala-Asx(CHO). The Asx here is probably asparagine because carbohydrate prosthetic groups containing glucosamine are attached to the amide nitrogen of asparagine in other glycopeptides from immunoglobulin heavy chains (Shimizu *et al.*, 1971b). The carbohydrate in T3-Th3 was analyzed by Dr. John R. Clamp, as is described later.

PEPTIDE T3-Th4. This peptide lacked Cm-cysteine on amino acid analysis but was radioactive (70,000 cpm/min) and showed an unidentified fluorescent dansyl spot after the second Edman step. Dansyl-Edman degradation gave the sequence Leu-Thr(Cys)Thr-Leu. Carboxypeptidase A digestion confirmed the carboxy-terminal leucine; after 24-hr digestion the results were: Leu, 1.3; Thr, 0.85. Serine was present on amino acid analysis of T3-Th4 but was judged to be a breakdown product of Cm-cysteine or a contaminant, for the results of carboxypeptidase A–carboxypeptidase B digestion of the intact peptide T3 also showed there was no serine between the leucine and the Cm-cysteine.

PEPTIDE T3-Th5. The sequence of this radioactive peptide was determined as Leu-Thr(Cys)Thr from two steps of dansyl-Edman degradation and from the action of carboxypeptidase A. The parentheses indicate that the cysteine was not directly confirmed. However, an unidentified fluorescent spot, possibly an oxidized form of cysteine, appeared on the dansyl thin-layer sheet after two steps of Edman degradation. After 24-hr incubation the enzyme liberated 76% of the threonine from T3-Th5 but no other amino acids. This is compatible with Cm-cysteine as the penultimate residue because carboxypeptidase A has little specificity for acidic residues. The sequence Leu-Thr-Cys-Thr is consistent with the dansyl-Edman data on T3-Th4 and with the results of carboxypeptidase A–carboxypeptidase B digestion of T3.

PEPTIDE T3-Th6. Although this dipeptide had background impurities its sequence was clearly Leu-Arg. It represents the C terminus of T3.

OTHER THERMOLYSIN SUBPEPTIDES OF T3. Several other thermolysin fractions of T3 were isolated but appeared to consist of a series of overlapping subpeptides. One peptide with good stoichiometry gave a strong leucine spot on dansylation, but the dansyl map showed four spots after the first Edman degradation, *i.e.*, leucine, serine, glycine, and threonine. This probably represents a series of similar dipeptides and tripeptides all with amino-terminal leucine, *e.g.*, Leu-Ser, Leu-Gly, Leu-Gly-Ser, Leu-Leu-Gly, Leu-Thr, etc. All of these are plausible from the sequence shown in Figure 4 and from the affinity of thermolysin for amino-terminal leucine. This illustrates the difficulty of secondary cleavage of this leucine-rich peptide.

SEQUENCER ANALYSIS OF GLYCOPEPTIDE T3. Because of the difficulty in assembling the sequence of T3 from the thermolysin peptides, the glycopeptide was again isolated from a tryptic digest of the Ha α chain. About 400 nmol of the carboxymethylated peptide were applied to the sequencer and analyzed with the peptide program without modification of the peptide since it had C-terminal arginine. The high leucine content, that had frustrated sequence determination by the manual

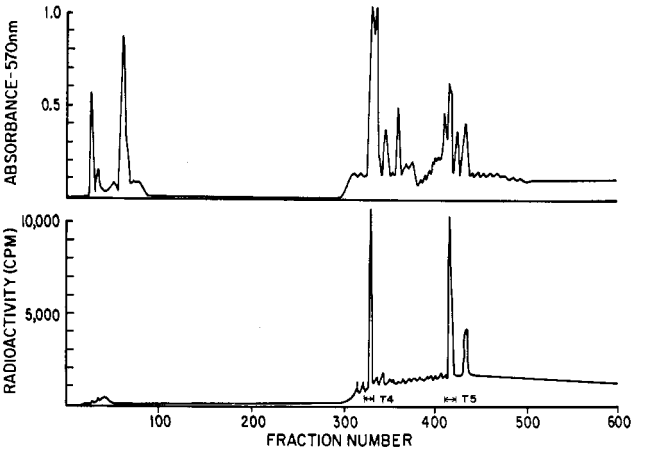

FIGURE 5: Technicon Chromo-Beads purification of tryptic peptides T4 and T5 from pool 4 of the Sephadex G-25 fractionation.

TABLE II: Amino Acid Sequence of Tryptic Peptides T5 and T5′.

T5	Ser-Val-Thr(Cys,His,Val)Lys[a]
T5-Th1	(Ser,Val,Thr,Cys,His)
T5-Th2	Val-Lys
T5′	Ser-Val-Thr-Cys(His,Val)Lys
Sequence	Ser-Val-Thr-Cys-His-Val-Lys

[a] All the half-cystine in T5 was present as Cm-cysteine (0.8 mol/mol of peptide) but 0.5 mol of Cm-cysteine and 0.4 mol of half-cystine were present per mol of T5′.

method, proved to be ideal for the sequencer because of the stability of the thiazolinone derivative of leucine and the high response of leucine phenylthiohydantoin in gas chromatography. Twenty-six cycles were completed successfully. No phenylthiohydantoin derivative was detected at the 18th step; this position must correspond to the asparagine with the carbohydrate group which had been localized in T3-Th3. Nothing was detected by gas chromatography at step 21, but Cm-cysteine was detected with the amino acid analyzer. Unexpectedly, the leucine at position 26 gave a response on gas chromatography equal to 20% of the N-terminal leucine. A complete sequence for T3 could be assigned from a combination of the sequencer results and the previous data on T3 and its thermolysin subpeptides (Figure 4).

Other Tryptic Peptides Containing [^{14}C]Cm-cysteine. PEPTIDE T4. This dodecapeptide eluted in the third peak on the Sephadex G-25 separation and was purified further by chromatography on Chromo-Beads. Preparative paper electrophoresis at pH 6.5 was used as a final step of purification. The dansyl-Edman procedure was performed for the first three steps, giving the partial sequence Thr-Phe-Thr(Cys,Thr,Ala,Ala,-Tyr,Pro,Glu,Ser)Lys. Another peptide, T4′, had similar composition and NH_2-terminal sequence but was incompletely alkylated. Peptide T4′ eluted in the fourth peak on Sephadex G-25 and also was purified further by Chromo-Beads chromatography and preparative electrophoresis at pH 6.5 (Figure 5).

PEPTIDE T5. This tryptic heptapeptide eluted in the fourth peak on Sephadex G-25 separation and was purified by Chromo-Beads chromatography and high-voltage electrophoresis at pH 6.5 (Figure 5). The dansyl-Edman procedure gave a sequence of Ser-Val-Thr. A peptide with similar composition, T5′, was obtained from the third Sephadex G-25 peak and purified further in the same manner as T5. This peptide was incompletely alkylated; it contained 0.5 mol of Cm-cysteine/mol and 0.4 mol of half-cystine. By dansylation the N-terminal sequence of this peptide was Ser-Val-Thr-Cys. Thermolysin was used to cleave T5. When the resulting secondary peptides were purified by preparative electrophoresis at pH 6.5, a pentapeptide was obtained with the composition (Ser,Val,Thr,Cys(Cm),His) and also the dipeptide Val-Lys. Thus, the sequence of T5 can be completed as: Ser-Val-Thr-Cys-His-Val-Lys (Table II).

Thermolysin Peptides Derived from Pool 2 of Sephadex G-25 Separation. Two small radioactive peptides containing Cm-cysteine were obtained by cleavage with thermolysin of a radioactive fraction that eluted as the second peak in the Sephadex G-25 separation of the tryptic peptides. This material was further separated in the chromatography on Chromo-Beads type P (T6 in Figure 1) and appeared to be a large peptide. Although this material moved as a single spot on electrophoresis at pH 1.9, 3.7, and 6.5, and in paper chromatography with butanol–acetic acid–water–pyridine, the stoichiometry and dansylation results indicated the peptide was still impure. Secondary cleavage with thermolysin was performed on half the sample, and two distinct tryptic-thermolysin peptides containing [^{14}C]Cm-cysteine were obtained, *i.e.*, T6-Th2 and T7-Th3.

PEPTIDE T6-Th2. This hexapeptide contains two Cm-cysteine residues. The dansyl-Edman procedure established the amino-terminal sequence as: Leu-Cys(Cm)-Gly-Cys(Cm). Neither of the remaining two amino acids, tyrosine and serine, could be identified as the carboxy terminus by carboxypeptidase A digestion, because both were released in equal amounts even after digestion for only 5 min. Hence, the partial sequence is written as Leu-Cys-Gly-Cys(Tyr,Ser).

PEPTIDE T7-Th3. This radioactive tetrapeptide is of special interest because it contains one of the few methionine residues in the Ha α chain as well as one of the half-cystines. Its sequence was established as Phe-Ser-Cys(Cm)-Met by the dansyl-Edman procedure. The C terminus was confirmed by carboxypeptidase A digestion for 24 hr which released 30 nmol of methionine and 12 nmol of Cm-cysteine; these values represent 90 and 30% recovery, respectively, of the amount of sample used for digestion.

SEQUENCER ANALYSIS OF THE AMINO TERMINUS OF THE α CHAIN. The fully reduced and carboxymethylated α chain was submitted several times to automatic sequence analysis with the protein program. Perhaps because of its poor solubility or high carbohydrate content, this α chain was the most difficult polypeptide to sequence in our experience. Gas and thin-layer chromatography of the phenylthiohydantoin derivatives and amino acid analysis were used to identify the residues. In the most successful run, clear results were obtained up to 20 cycles except for several residues for which the phenylthiohydantoin derivatives are unstable or respond poorly. The results are summarized in Table III. By homology to other unblocked heavy chains of the same subgroup (V_{HIII}), serine is probably present at positions 17, 21, and 25; however, because of a mechanical breakdown in the gas chromatography system the samples after 20 cycles had to be stored under vacuum for 1 week. All positions after Leu-20 are ambiguous.

THE HINGE PEPTIDE. Probably because of the high proline content, this material was resistant to cleavage by chymotrypsin, pepsin, thermolysin, and papain. Partial acid hydrolysis (12 N HCl, 18 hr, 37°) was tried because of the large number of serine and threonine residues. The acid hydrolysate was chromatographed on Chromo-Beads with the Technicon in a pyridine–acetate gradient from pH 1.9 to 5.5. The column was monitored for radioactivity and ninhydrin color. Many peaks were obtained including some with radioactivity. Amino acid analysis and end group analysis showed extensive acid cleavage had occurred, but most of the peaks contained overlapping peptides. Most of the galactosamine was located in a peptide with good stoichiometry (Thr, Ser, Pro_2), but the dansyl-Edman reaction indicated both threonine and serine in the first and second steps.

Three radioactive peptides (TA1, TA2, and TA3) were isolated and were partially sequenced. As shown in Table III, TA1 corresponds to the amino terminus of the tryptic-peptic hinge peptide TPα2 reported by Wolfenstein *et al.* (1971). The partial sequence of TA2 was Ser-Cys(His,Pro,Cys)Arg,

and that of TA3 was His(Pro,Cys)Arg. Together these give Ser-Cys-His(Pro,Cys)Arg. TA2 and TA3 thus fit with the C terminus of the hinge peptide for which no sequence was given by Wolfenstein *et al.* (1971) and for which Grey *et al.* (1971) give Cys-Arg-Leu. Although our data and that of Grey *et al.* permit deduction of a tentative C-terminal sequence of Ser-Cys-His(Pro)Cys-Arg-Leu, this is insecure because of the uncertainty about the number of prolines. Also, this does not accord with the C terminus for the hinge peptide from $\alpha 2$ chains (see Discussion).

The tentative sequence of some 28 residues including at least nine prolines shown for the hinge peptide in Table III represents only the COOH-terminal portion of the tryptic peptide we were unable to isolate in pure form. The latter was obtained in two forms, both of which had N-terminal leucine and appeared to have more than 50 residues. It is possible that both glucosamine and galactosamine are attached at different points in a single large tryptic peptide in the hinge region of the $\alpha 1$ chain. Including the carbohydrate groups this peptide could have a molecular weight close to 10,000. However, we have some evidence from sequencer analysis that the glucosamine and galactosamine are on separate peptides, each of which contains some 30 amino acid residues, and that the two peptides are joined by a disulfide bridge.

Carbohydrate Composition of Glycopeptides. Two large glycopeptides were identified in this work, the hinge peptide, which is characterized by the presence of galactosamine, and the 27-residue tryptic glycopeptide T3 which contains glucosamine. One sample of the large impure tryptic peptide representing the hinge region had 7 residues of galactosamine and 2 of glucosamine per leucine residue, and another had 5 residues of the former and 2 of the latter per leucine residue. Of the 26 peptide fractions obtained after mild acid hydrolysis of this peptide, one contained both amino sugars and another only galactosamine. The latter peptide had N-terminal serine but was impure. However, its composition would place it in the undetermined portion of the hinge peptide.

The carbohydrate in the 27-residue glycopeptide T3 was localized to the Asx residue which is the eighteenth residue in this sequence. The glucosamine was identified in the subpeptide T3-Th3 which has the sequence Leu-Gly-Ser-Glx-Ala-Asx. A minimum of 1.9 mol of glucosamine/mol of peptide was present according to our amino acid analyzer results. However, Dr. J. R. Clamp of the University of Bristol, England, found 4.0 mol of glucosamine/mol of peptide by gas chromatographic analysis. A loss of about 50% of amino sugar during acid hydrolysis (6 N HCl, 105°, 22 hr) has previously been reported by Clamp and Putnam (1964).

In addition to four residues of glucosamine, three of mannose and two of galactose were present in the thermolysin subpeptide (Table IV). Surprisingly, no fucose or sialic acid was present.

Discussion

The molecular weight of the α polypeptide chain less the carbohydrate is reported to be in the range 46,000–52,000, corresponding approximately to 450 residues, and the half-cystine content is given as 16–17 residues in $\alpha 1$ chains by Grey *et al.* (1971). Because of ambiguity about the size of the α chain and the difficulty of precise determination of alkylated derivatives of cysteine the exact number of half-cystines is uncertain. Our analytical data and peptide isolations indicate the number of half-cystine residues is at least 13.

A total of 13 half-cystine residues is the minimum number expected from what is known about the structure of IgA1 proteins and by homology to IgG and IgM, *i.e.*, at least five interchain bridges requiring five half-cystine residues per chain and four intrachain bridges requiring eight. The five interchain bridges include the light–heavy (L–H) bridge, three heavy–heavy (H–H) bridges in the hinge peptide, and one H–H bridge at the COOH terminus. Seven of the eight probable intrachain bridge residues of half-cystine are indicated in Table III. The eighth was not identified but probably is present in the highly conserved prototype sequence Asp-Thr-Ala-Thr-Tyr-Tyr-Cys-Ala-Arg near position 97 just before the hypervariable deletion region characteristic of V_H sequences (Putnam *et al.*, 1972). The Leu-Cys-Gly-Cys-Tyr sequence is puzzling. One of the half-cystines is probably in an intrachain bridge; the other may be involved in the intersubunit linkage responsible for IgA polymers or may be an intrasubunit bridge between heavy chains.

Altogether 13 different half-cystine residues are listed in Table III. Of these 11 are reported in this paper and 9 by Wolfenstein *et al.* (1971). The latter listed two half-cystine residues not yet isolated by us; one of these (TPα1) represents the light–heavy interchain bridge; the other (TPα6) is the COOH-terminal heavy–heavy interchain bridge, which was also described by Prahl *et al.* (1971). We report sequences around four half-cystine residues not identified by Wolfenstein *et al.* (peptides T2, T7-Th3, T5, and the NH_2 terminus). Seven other half-cystine residues were identified by both groups of workers. In separate experiments with several other IgA1 proteins our laboratory has isolated tryptic peptides identical in composition to T2, Tα3, Tα4, and T5. Hence, in Table III the constant region of the human α chain (Cα) is proposed as the probable location of these peptides.

None of the peptides listed in Table III could have been derived from the λ light chain of the IgA1 protein Ha. The complete sequence of the λ Bence-Jones protein from the urine of this patient has been reported (Shinoda *et al.*, 1970). Tryptic peptide maps have indicated the identity of the Ha light chain and the Bence-Jones protein (Bernier *et al.*, 1965). The latter contains five half-cystine residues, which are in the following sequences: Cys-21, Ile-Ser-Cys-Ser-Gly; Cys-86, Tyr-His-Cys-Ala-Ala; Cys-135, Leu-Val-Cys-Leu-Ile; Cys-194, Tyr-Ser-Cys-Gln-Val; Cys-212, Thr-Glu-Cys-Ser. Although some of these sequences have obvious homology to the α-chain peptides, none is identical with those we or Wolfenstein *et al.*, 1971; Wolfenstein-Todel *et al.*, 1972) have reported for either the $\alpha 1$ or the $\alpha 2$ chain. None of the peptides listed in Table III is identical with peptides we have obtained in the complete sequence analysis of the human μ chain (Putnam *et al.*, 1972; F. W. Putnam, unpublished experiments), nor have we identified any of these peptides in preparations of J chain we have prepared from IgA1, IgA2, and IgM proteins.

Because of the frequency of hydrophobic and hydroxy-amino residues on either side of the half-cystine, the sequence homology of light and heavy chains is greatest around the intrachain bridges. These typically involve linkage of an amino-terminal half-cystine in a hydrophobic sequence, such as the prototype Leu-Thr-Cys-Leu-Val, with a carboxyl-terminal half-cystine in an aromatic sequence, such as the prototype Phe-Thr-Cys-X-Val. Phe-Ser-Cys-Met (T7-Th3) and Phe-Thr-Cys-Thr-Ala (T4) fit the latter category, and Val-Thr-Cys-His-Val (T5) and Leu-Thr-Cys-Thr-Leu (T3) fit the former. The sequence Leu-Thr-Cys is present in the constant region of human $\gamma 1$ chains (C3γ) (Edelman *et al.*, 1969) and in the constant region of μ chains (C3μ) (Putnam

TABLE III: Summary of Sequence Data on Human α Heavy Chains.[a]

IgA Subclass	Protein	Peptide	Peptide
IgA1	Pat	TPα1	Ser-Leu-*Cys*-Ser-Thr-Glx-Pro-Asx-Gly-Asx
IgA2	Avi	TPα1	Pro(Asp,*Cys*,Thr,Glu,Glu,Pro,Gly,Gly,Ala,His)
			CHO
IgA1	Pat	TPα2	Val-Thr-Val-Pro-*Cys*-Pro-Val-Pro-Ser-Thr-Pro-Pro-Thr-Pro-Ser-Pro-Ser-
			CHO
IgA1	Ha	TA1-3[b]	Thr-Val-Pro-*Cys*-Pro-Val-Pro(Ser,Thr,Pro,Pro,Thr)————————
IgA2	Avi	TPα2	Val-Thr-Val-Pro-*Cys*-Pro-Val-Pro————GAP————————
IgA1	Pat	TPα6	Ala-Glu-Val-Asp-Gly-Thr-*Cys*-Tyr
IgA2	Avi	TPα6	Ala(Glu,Val,Asp,Gly,Thr,*Cys*)
IgA1	Ha	T2[c]	Asx-Phe-Pro-Pro-Ser-Glx-Asx-Ala-Ser-Gly-Asx- Leu-Tyr-Thr-Thr- Ser-Ser-
IgA1	Ha	T6-Th2	Leu-*Cys*-Gly-*Cys*(Tyr,Ser)
IgA1	Pat	Tα3[c]	Asp-Leu-*Cys*-Gly-*Cys*-Tyr-Ser-Val-Ser(Asp,Ser, Glu,Pro,Pro,Gly, Gly,Ala,
IgA2	Avi	TPα3	Asp(Leu,*Cys*,Gly,*Cys*)
IgA1	Ha	T7-Th3	Phe-Ser-*Cys*-Met
IgA1	Ha	T4[c]	Thr-Phe-Thr(*Cys*,Thr,Ala,Ala,Tyr,Pro,Glu,Ser) Lys
IgA1	Pat	Tα4	Thr-Phe-Thr-*Cys*-Thr-Ala-Ala-Tyr-Pro-Glu-Ser- Lys
IgA2	Avi	TPα4	Thr(*Cys*-Thr-Ala,Ala,His,Pro,Glu,Leu) Lys
IgA1	Ha	T5[c]	Ser-Val-Thr-*Cys*-His-Val-Lys
IgA1	Ha	T3	Leu-Ser-Leu-His-Arg-Pro-Ala-Leu-Glx-Asp-Leu-Leu-Leu-Gly-Ser-Glu-Ala-
IgA1	Pat	TPα5	
IgA2	Avi	TPα5	
			1 10
IgA1	Ha	N-terminus	Glu-Val-Gln-Leu-Val-Glu-Ser-Gly(Gly)Gly-Leu- Val(Lys)Pro-Gly- Gly( )
IgA	For	N-terminus	Glu- Ile-Glu-Leu-Val-Glu-Ser-Gly-Gly-Gly-Leu- Val(Lys)Gly-Gly- Gly-Ser-

[a] Data for IgA1 protein Pat from Wolfenstein *et al.* (1971), for IgA1 protein Ha from this paper, and for IgA2 protein Avi from Wolfenstein-Todel *et al.* (1972). [b] The dash line in TA1-3 indicates a region of undetermined sequence. [c] These peptides were also

TABLE IV: Carbohydrate Composition of IgA Ha and of Glycopeptide T3-Th3.

Monosaccharide	IgA Ha (mol/mol)[a]	Glycopeptide T3-Th3 (mol/mol)[b]
Fucose	3	0
Mannose	11	3.2
Galactose	13	2.0
Glucosamine	18	4.0
Sialic acid	6	0
Galactosamine	Present	0

[a] Taken from Clamp and Putnam (1967). [b] Determined by gas chromatography by Dr. John R. Clamp, University of Bristol, Department of Medicine, Bristol, England (Clamp *et al.*, 1967; Bhatti *et al.*, 1970).

et al., 1972). The sequence Phe-Thr-Cys is present twice in the constant region of μ chains (C2μ and C3μ) (Putnam *et al.*, 1972); in C3μ the sequence is Phe-Thr-Cys-Thr, the same as in peptide T4 in the α1 chain. This tetrapeptide is the longest identical sequence when the 150-odd residues of the α1 chain C region given in Table III are compared with the 1000-odd residues of the C regions of human μ, γ, κ, and λ chains.

The homology of heavy chains appears least in the hinge region. This is the section near the middle of the heavy chain which contains interchain disulfide bridges and is uniquely susceptible to limited proteolysis yielding the Fab and Fc fragments. In the region between the two half-cysteines of the α1 hinge peptide almost half the residues are prolines and all the others but one are threonines or serines. The latter are deleted in the homologous region of the α2 hinge peptide, which has a remarkable pentaproline sequence (Grey *et al.*, 1971; Wolfenstein-Todel *et al.*, 1972). Human and animal α1 chains are also rich in proline in the hinge peptide, but not the human μ chain. Some of these chains have oligosaccharides attached to the hinge area and others do not. However, in all cases the conformation is predictably distorted so that the hinge area is readily accessible to limited proteolytic cleavage.

Unlike γ chains, both μ and α chains have multiple oligosaccharide groups. Dawson and Clamp (1968) have given evidence that there are three different kinds of core units of oligosaccharides attached to the α chain. Type 1, now identified with the hinge peptide, contains only galactosamine and galactose, with the amino sugar attached to a serine through an *O*-glycosidic bond. The second type of core glycopeptide (type 2) is complex and contains mannose, galactose, and glucosamine in the proportion 3:2:3 relative to the aspartic acid (asparagine) to which the glucosamine is attached. Associated with this may be 0–3 residues of sialic acid, 0–1 residue of fucose, and 0–1 additional residue of glucosamine. The third type of oligosaccharide (type 3) is simple and contains only mannose and glucosamine in the ratio 4:3 with the amino sugar attached to aspartic acid.

Type 1, containing galactosamine, is absent in IgG, IgM, and IgA2 proteins and is unique to IgA1. Its absence in IgA2 is explained by the deletion in the α2 chain of the decapeptide segment in the middle of the α1 hinge region which contains this oligosaccharide.

The oligosaccharide attached to our α1 chain peptide T3 is similar to the three complex oligosaccharides in the μ chain

Sequence	Probable Location	Probable Function
	Cα	H–L
	?	?
Thr(Ser,Thr,Pro_2,Ser,*Cys*,His,Pro_2,*Cys*)Arg	Cα1 hinge	H–H
---------Ser-*Cys*-His(Pro,*Cys*)Arg	Cα1 hinge	H–H
——Pro-Pro-Pro-Pro-*Cys*-*Cys*-His -Pro-Arg	Cα2 hinge	H–H
	Cα C terminus	H–H
	Cα C terminus	H–H
Gln-Leu-Thr-Leu-Pro-Ala-Thr-Glx-*Cys*-Leu-Ala-Gly-Lys	Cα	Intrachain
	Cα	Intrachain
Val,Leu,His)Lys	Cα	plus H–H
	Cα	
	Cα	Intrachain
	Cα1	Intrachain
	Cα1	
	Cα2	
CHO	Cα	
Asx-Leu-Thr-*Cys*-Thr-Leu-Thr-Gly- Leu- Arg	Cα	
Thr-*Cys*-Thr-Leu	Cα	Intrachain
Thr(*Cys*,Thr,Leu)	Cα	
20 29		
Leu( )Leu( )(*Cys*)(Ala)(Ala)( )( )(Phe)( )(Phe)	VH	Intrachain
Leu-Arg-Leu(Ser)*Cys*-Ala- Ala (Ser)Gly-Phe	VH	

isolated from a third IgA1 protein (Bur) by Dr. Theresa Low.

(Shimizu *et al.*, 1971) and to the type 2 core unit characterized by Dawson and Clamp (1968). However, according to Dr. Clamp, our oligosaccharide is unique in its absence of both fucose and sialic acid. The type 2 oligosaccharides isolated by Dawson and Clamp (1968) were heterogeneous in their fucose and sialic acid content but all had at least fractional amounts of fucose. The fucose is probably linked in a branched structure to the glucosamine residue that is attached directly to the aspartic acid in the polypeptide chain. The sialic acid is known to be in a terminal branched structure. The absence of fucose in our peptide T3-Th3 suggests that this sugar is added after the main branch structure of the oligosaccharide is laid down.

Two simple oligosaccharides in human IgM that contain only mannose and glucosamine and thus are similar to type 3 have been localized by Shimizu *et al.* (1971a,b) to the Fc region of the μ chain. We have not yet isolated a corresponding oligosaccharide from the Ha α1 chain, but we do have evidence for a second large glycopeptide that contains glucosamine.

Complete sequence determination of the αl heavy chain and assignment of the disulfide bridges will be difficult because many of the half-cystines are located in large tryptic peptides, some of which have unusual composition and contain oligosaccharides. Three half-cystine residues are in the proline-rich hinge peptide (TAl-3), which is a glycopeptide that has some 30–40 amino acid residues. A fourth is in the 27-residue, leucine-rich glycopeptide T3, and a fifth in the 30-residue peptide T2. A seventh half-cystine is in the Phe-Ser-Cys-Met sequence (T7-Th3) which appears to be derived from a large glycopeptide. Furthermore, the half-cystine residues in the light–heavy bridge and at the C terminus are also probably in large tryptic peptides, as is the second half-cystine in the variable region. Hence, for complete sequence determination of the αl-chain aminoethylation of the half-cystines, which yields additional peptides on tryptic digestion, would be more advantageous than carboxymethylation

References

Bernier, G. M., Tominaga, K., Easley, C. W., and Putnam, F. W. (1965), *Biochemistry 4*, 2072.

Bhatti, T., Chambers, R. E., and Clamp, J. R. (1970), *Biochim. Biophys. Acta 222*, 339.

Braunitzer, G., Schrank, B., Ruhfus, A., Petersen, S., and Petersen, U. (1971), *Hoppe-Seyler's Z. Physiol. Chem. 352*, 1730.

Clamp, J. R., Dawson, G., and Hough, L. (1967), *Biochim. Biophys. Acta 148*, 342.

Clamp, J. R., and Putnam, F. W. (1964), *J. Biol. Chem. 239*, 3233.

Clamp, J. R., and Putnam, F. W. (1967), *Biochem. J. 103*, 225.

Cunningham, B. A., Pflumm, M. N., Rutishauser, V., and Edelman, G. M. (1969), *Proc. Nat. Acad. Sci. U. S. 64*, 997.

Dawson, G., and Clamp, J. R. (1968), *Biochem. J. 107*, 341.

Dayhoff, M. O. (1972), Atlas of Protein Sequence and Structure, Vol. 4, Silver Spring, Md., National Biomedical Research Foundation.

Easley, C. W. (1965), *Biochim. Biophys. Acta 107*, 386.

Edelman, G. M., Cunningham, B. A., Gall, W. E., Gottlieb, P. D., Rutishauser, V., and Waxdal, M. J. (1969), *Proc. Nat. Acad. Sci. U. S. 63*, 78.

Gray, W. R. (1967), *Methods Enzymol. 2*, 469.

Grey, H. M., Abel, C. A., and Zimmerman, B. (1971), *N. Y. Acad. Sci. 190*, 37.

Halpern, M. S., and Koshland, M. E. (1970), *Nature (London)*

228, 1276.
Hermodson, M. A., Ericsson, L. H., Titani, K., Neurath, H., and Walsh, K. A. (1972), *Biochemistry 11*, 4493.
Köhler, H., Shimizu, A., Paul, C., Moore, V., and Putnam, F. W. (1970), *Nature (London) 227*, 1318.
Mestecky, J., Zikan, J., and Butler, W. T. (1971), *Science 171*, 1163.
Miekka, S. I., and Deutsch, H. F. (1970), *J. Biol. Chem. 245*, 5534.
Moore, V. (1972), Ph.D. Dissertation, Indiana University, Bloomington, Ind.
Pisano, J. J., and Bronzent, T. J. (1969), *J. Biol. Chem. 244*, 5597.
Ponstingl, H., Schwarz, J., Reichel, W., and Hilschmann, N. (1970), *Hoppe-Seyler's Z. Physiol. Chem. 351*, 1591.
Prahl, J. W., Abel, C. A., and Grey, H. M. (1971), *Biochemistry 10*, 1808.
Press, E. M., and Hogg, N. M. (1970), *Biochem. J. 117*, 641.
Putnam, F. W. (1969), *Science 163*, 633.
Putnam, F. W., Shimizu, A., Paul, C., and Shinoda, T. (1971), *in* Progress in Immunology, Amos, B., Ed., New York, N. Y., Academic Press, p 291.
Putnam, F. W., Shimizu, A., Paul, C., and Shinoda, T. (1972), *Fed. Proc., Fed. Amer. Soc. Exp. Biol. 31*, 193.
Shimizu, A., Paul, C., Köhler, H., Shinoda, T., and Putnam, F. W. (1971a), *Science 173*, 629.
Shimizu, A., Putnam, F. W., Paul, C., Clamp, J. R., and Johnson, I. (1971b), *Nature (London), New Biol. 231*, 73.
Shinoda, T., Titani, K., and Putnam, F. W. (1970), *J. Biol. Chem. 245*, 4475.
Smyth, D. S., and Utsumi, F. (1967), *Nature (London) 216*, 332.
Titani, K., Wikler, M., Shinoda, T., and Putnam, F. W. (1970), *J. Biol. Chem. 245*, 2171.
Wolfenstein, C., Frangione, B., Mihaesco, E., and Franklin, E. C. (1971), *Biochemistry 10*, 4140.
Wolfenstein-Todel, C., Frangione, B., and Franklin, E. C. (1972), *Biochemistry 11*, 3971.
Woods, K. R., and Wang, K. T. (1967), *Biochim. Biophys. Acta 133*, 369.

Enthalpy of Reduction of Disulfide Cross-Links in Denatured Lysozyme†

S. Lapanje‡ and J. A. Rupley*

ABSTRACT: Calorimetric measurements were made at pH 8, 30°, and 6–8 M guanidinium chloride for the reduction of disulfides with dithiothreitol or dithioerythritol (disulfide exchange). The enthalpy changes were −1.5 kcal/mol of disulfide for lysozyme and −1.1 kcal/mol for an exhaustive peptic hydrolysate of lysozyme. These results support the general understanding that the thermodynamics of cross-link formation in a random-coil polymer are dominated by the entropy, and that the cross-linked chain in strong denaturing solvents is essentially a random-coil structure. The enthalpy of disulfide exchange for oxidized glutathione is −0.4 kcal/mol in guanidinium chloride solution and −0.5 kcal/mol in water. Cyclic and noncyclic disulfides have nearly equal enthalpies.

Intramolecular covalent cross-links are generally understood to favor folded over unfolded conformations of protein molecules (Kauzmann, 1959; Scheraga, 1963). This contribution is believed to be an entropy effect which arises from the few configurations accessible to the cross-linked compared with the linear polymer. Because introduction of a cross-link into a random-coil molecule in a good solvent may require that one or more bonds be brought into a conformation with energy higher than obtains for the unconstrained molecule, it is possible that the enthalpy of cross-link formation is not zero. In order to evaluate the enthalpy contribution, the heat of reduction (*i.e.*, of disulfide interchange) of disulfide cross-links in a random coil polypeptide chain was determined in calorimetric experiments that compared in concentrated guanidinium chloride solution the heat of reaction of lysozyme, a peptic hydrolysate of lysozyme, and glutathione. In this solvent, lysozyme has been shown by optical rotatory dispersion (Tanford *et al.*, 1967), viscosity (Tanford *et al.*, 1967), and nuclear magnetic resonance (nmr) experiments (McDonald *et al.*, 1971) to behave as a random coil when the four disulfide cross-links are present or absent.

Materials and Methods

Salt-free hen egg-white lysozyme (six-times crystalline, lot 7102) was obtained from Miles Laboratories; dithiothreitol, dithioerythritol, and oxidized disodium glutathione from Sigma; ultra pure guanidinium chloride (GdmCl)[1] and urea, from Mann Research Laboratories. Other chemicals were analytical grade. Deionized water was used.

Protein solutions were prepared by dissolving weighed amounts of lysozyme in the appropriate solvent. The water content of the protein was taken into account, and it was determined by drying under vacuum at 107° to constant weight and by absorbance of aqueous solutions of weighed

† From the Department of Chemistry, The University of Arizona, Tucson, Arizona 85721. *Received March 12, 1973.* This work was supported by funds from the National Institutes of Health.

‡ On a leave of absence from the Department of Chemistry, University of Ljubljana.

[1] Abbreviation used is: GdmCl, guanidinium chloride.