The Molecular Defect in a Protein (CRA) Found in γ1 Heavy Chain Disease, and Its Genetic Implications

E. C. FRANKLIN* AND B. FRANGIONE†

Department of Medicine and Rheumatic Diseases Study Group, New York University School of Medicine, N.Y. 10016

ABSTRACT   A γ1 protein, designated CRA, found in heavy chain disease contains three inter-heavy disulfide bridges instead of the two normally found in γ1 immuno-globulin heavy chains. Almost the entire Fd fragment is missing. The NH_2 terminal region is heterogeneous and contains carbohydrate; after 11 residues that do not resemble any of the known heavy-chain variable-region subclasses, normal synthesis seems to be resumed at the same amino acid residue as in another heavy-chain disease protein (ZUC). This finding raises the possibility that glutamic acid at position 216 represents the beginning of the Fc region, synthesized under the direction of another gene.

To date, 14 cases of "γ heavy chain disease" have been diagnosed, although only 11 of these have been reported (1-3). Studies of the first protein (CRA) found to be characteristic of the disease, and several subsequent studies, have suggested that all patients with the disease produce an abnormal protein which consists of the Fc fragment and piece of the Fd fragment, rather than the entire heavy chain. However, to date it has been possible to define the defect in only one protein (4) and to partially define it in another (5). In the best-studied protein (ZUC), a γ3 protein, there appears (4) to be a deletion of almost the whole Fd fragment of the heavy chain starting at residue 18, with resumption of the sequence normally found in γ3 heavy chains with the glutamic acid at position 216, near the hinge region of the Fc fragment. [Numbering is taken from protein Daw (6) and Eu (7).]

Since study of these proteins promises to provide not only information related to the defect in this disorder, but possibly also a clue as to the genetic control of the constant (c) and variable (v) regions of the heavy chain, it seems worthwhile to report detailed studies on the structure of protein CRA. This protein has been characterized in part previously (8) and shown to consist primarily of the Fc fragment of the γ1 chain plus at least two to three additional peptides, and to possess the two intrachain disulfide loops characteristic of the Fc fragment (9, 10).

In the present study we have chosen cyanogen bromide (CNBr) to cleave the γ1 heavy chain protein CRA in the hope of isolating the amino terminal peptide, which was suspected, on the basis of studies with protein ZUC, to contain the abnormality (4). The amino terminal peptide resulting from cleavage at methionine 252 (11) had approximately 50 residues, and appeared to have three rather than the expected two

inter-heavy chain disulfide bridges. Partial sequence of this fragment indicated that protein CRA resembled ZUC in having a deletion of the Fd fragment of approximately 200 residues, normal synthesis having apparently been resumed at residue 216, in the middle of the heavy chain (4).

MATERIALS AND METHODS

The protein CRA, previously shown to belong to the γ1 subclass and to be Gm (a-), was isolated from the urine by precipitation with 60% saturated ammonium sulfate followed by starch block electrophoresis and filtration on Sephadex G-200 (4). It was slightly contaminated with transferrin on immunoelectrophoresis.

CNBr cleavage

300 mg of protein CRA was dissolved to a final concentration of 30 mg/ml in 70% formic acid. A two-fold (w/w) amount of CNBr was added, and the reaction was allowed to proceed at room temperature for 24 hr. The CNBr was removed by filtration on a Sephadex G-25 column. Amino acid analyses showed that all methionine residues were lost after this treatment.

Complete reduction and alkylation

CNBr-treated freeze-dried protein CRA was dissolved in 15 ml of 0.5 M Tris-6 M guanidine-1 mM EDTA, pH 8.6, and 300 μl of a 10 mg/ml solution of dithiothreitol was added after the tube had been flushed with nitrogen. The reduction was allowed to proceed for 1 hr at room temperature. 600 μl of 0.1 M [¹⁴C]iodoacetic acid (0.5 mCi/mmol) was added and allowed to react for 1 hr at room temperature. A Sephadex G-25 column eluted with 0.3 M acetic acid was used for desalting.

Gel filtration

The reduced and alkylated CNBr cleavage products were separated on a Sephadex G-75 column (140 × 3.5 cm) in 5% formic acid (Fig. 1). Absorbance was measured in a Beckman DU Spectrophotometer at 280 nm. Radioactivity was counted in a Beckman 150 scintillation counter.

Partial reduction and carboxymethylation

These have been described previously (12).

Enzymatic digestion

The radioactive, labeled proteins were digested with trypsin (enzyme/substrate ratio 1:50 in 1% NH_4HCO_3, pH 8.5, for 4-8 hr at 37°C), pepsin (enzyme/substrate ratio 1:20 in 5% formic acid), or both. The digests were passed through a
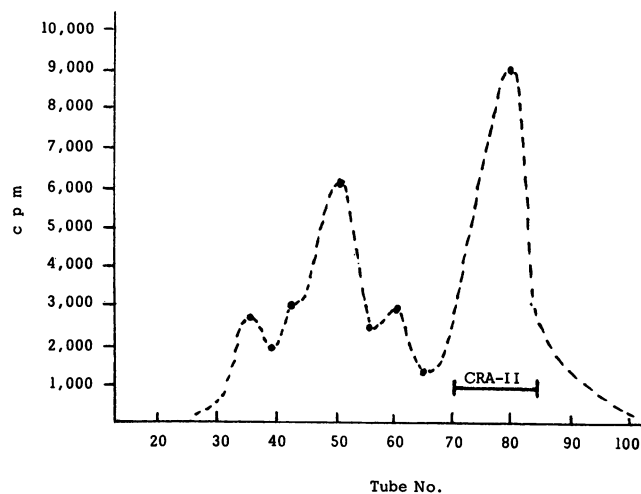
FIG. 1. Separation of CNBr fragments of protein CRA found in heavy chain disease. The totally reduced and [1-^{14}C]carboxy-methylated fragments were applied to a Sephadex G-75 column in 5% formic acid.

Sephadex G-25 column in 0.1 M acetic acid and, when necessary, peptides were purified further by paper electrophoresis. The peptic-tryptic digest was purified directly by paper electrophoresis.

Diagonal maps

Diagonal maps of peptic-tryptic digests at pH 3.5 and 6.5 were made as previously described (10, 13). The number of free sulfhydryl groups was determined by alkylation of the unreduced molecule with [^{14}C]iodoacetic acid (14).

Amino acid analyses were performed in the Beckman 120C analyzer. Samples for amino acid analyses were hydrolyzed in 5.7 N HCl at 110°C under reduced pressure. Reported compositions are based on the values obtained after 22-hr hydrolysis.

Amino acid sequence

The NH$_2$-terminal residue exposed after each cycle of the Edman degradation was identified by dansylation (15), followed by acid hydrolysis. The derivatives were identified on polyamide thin-layer plates (12, 16). The NH$_2$-terminal amino acid residues of protein CRA and of the CNBr fragments were also determined by this technique.

RESULTS

Isolation of CNBr fragments

Protein CRA contains three methionine residues, and the amino acid sequence around them is shown in Table 1. To

TABLE 1. *Sequence around methionine residues in protein CRA*

Peptide	Position	Sequence
	(Ref. 7)	
T1	252	Asp-Thr-Leu-Met
T2	358 (Gm a-)	Glu-Glu-Met-Thr-Lys (ref. 17)
PT3	428	Phe-Ser-Cys-Val-Met (ref. 10)

T, tryptic peptide obtained from CNBr fragment CRA-II (Fig. 1).
PT, peptic-tryptic.

obtain the N-terminal region of protein CRA, the protein was first treated with CNBr so as to cleave the molecule at methionine 252, 358, and 428, and then completely reduced and carboxymethylated with [^{14}C]iodoacetate. The fragments were then separated on Sephadex G-75 in 5% formic acid. Two main radioactive peaks were obtained (Fig. 1). The second radioactive peak (CRA-II) consisted of 46–50 residues. It contained homoserine, carbohydrate, three carboxymethyl-cysteine residues, and had glycine as its major N-terminal residue as well as small amounts of aspartic acid and threonine. The mean of five amino acid analyses is shown in Table 2. Since glycine is the N-terminal residue of the whole molecule, we concluded that peak II contained the N-terminal CNBr fragment of protein CRA.

CRA-II was then digested with pepsin, trypsin, or both under different conditions and the fragments were subjected to filtration on Sephadex G-25 in 0.1 M acetic acid. A peptic peptide eluted in the void volume also had glycine as its major N-terminal residue, together with small amounts of aspartic acid. The mean composition of this peptide (CRA-II P1), based on two separate analyses, is shown in Table 2. Peptide CRA-II P1 was then subjected to trypsin digestion and the resultant peptides were subsequently separated and purified by paper electrophoresis. Of particular interest were several arginine peptides; the composition of the two higher-yield peptides (P1 T4 and P1 T5) is shown in Fig. 2. The major N-terminal residue of peptide P1 T4 was glycine, although heterogeneity was again apparent by virtue of the presence of small amounts of aspartic acid. The partial sequence of peptide P1 derived from CNBr fragment II is shown in Fig. 2. The comparable tryptic peptides derived from CRA CNBr II are also indicated in Fig. 2. The sequence of the tryptic peptide containing the C-terminal methionine is shown in Table 1. Comparison of the sequence of CRA II P1 with the homologous region of three other γ1 myeloma proteins (Fig. 2) clearly indicates that starting with the glutamic acid residue

TABLE 2. *Amino acid compositions of a CNBr fragment (CRA-II) containing the N-terminus of protein CRA, and of a peptic peptide from the same fragment*

	CRA-II	CRA-II-P1
Lys	4.0	2.0
His	1.0	1.0
Arg	1.0	1.0
CMCys	2.4	2.8
Asp	4.4	3.4
Thr	4.4	3.9
Ser	3.1	1.2
Glu	4.0	3.4
Pro	8.2	4.8
Gly	3.0	0.9
Ala	1.1	1.0
Val	1.5	...
Ile	0.4	0.3
Leu	4.4	1.6
Tyr	0.5	...
Phe	3.0	0.7
Homoserine	0.5	...
CHO	+	+
NH$_2$-terminal	Gly ++++	Gly ++++
	Asp ++	Asp ++
	Thr +	

<div style="text-align:center">

		210		216

Daw,Car,γ1,Gm(za)   Asn-His-Lys-Pro-Ser-Asn-Thr-Lys-Val-Asp-Lys-Lys-Val-Glu-Pro-Lys-
Eu   γ1,Gm(fa-)   Arg

CRA   γ1,Gm(a-)   Gly(Asp$_{2.9}$, Leu$_{0.7}$, Ile$_{0.5}$, Phe, CHO) Arg-Thr-Thr-Glu-Glu-Pro-Lys-
          ←——T$_4$ & P$_1$T$_4$————————→   ←———T$_3$ & P$_1$T$_3$———————→
                    (Asp$_3$, Leu$_{1.2}$, Ile$_{0.9}$, Phe, CHO) Arg
          ←——————P$_1$T$_5$——————————→

	220		230

Daw, Car, Eu...   Ser-Cys-Asp-Lys-Thr-His-Thr-Cys-Pro-Pro-Cys-Pro-Ala-Pro-Glu-Leu

CRA   Ser-Cys-Asp-Lys-Thr-His-Thr-Cys-Pro-Pro-Cys-Pro-Ala-Pro-Glu-Leu
          ←—T$_2$ & P$_1$T$_2$—→ ←—————————————P$_1$T$_1$——————————————→
          ←————————————————*P$_1$T$_0$————————————————————→

</div>

Fig. 2. Partial sequence of a CNBr fragment from the N-terminus of protein CRA and the same region of three other γ1 myeloma proteins [Daw (6), Car (17), Eu (7)]. Thick arrow indicates the site at a glutamic acid residue where synthesis resumes normally. *4-hr trypsin digestion.

at position 216 the sequences are identical, while the approximately 11 residues from the amino terminal end to the residue corresponding to number 215 differ from those of any known γ chains.

In an effort to obtain the sequence of the amino terminal region of the molecule, three types of experiments were done, all without success: (1) the whole protein was placed in a Beckman Sequenator; (2) manual Edman degradation of the CNBr fragment CRA-II was attempted; (3) Edman degradation of the N-terminal peptic and an arginine-containing tryptic peptide were performed. In all instances glycine was the major N-terminal residue, after which marked heterogeneity became apparent.

Function of half-cysteine residues

Previous studies using diagonal electrophoresis (10) of peptic digests disclosed the presence of two intrachain disulfide bridges in protein CRA. No other cysteine bridge was detected with this method. When the protein was partially reduced and carboxymethylated with [^{14}C]iodoacetate, subjected to pepsin and trypsin digestion, and run on paper

electrophoresis (11), three main peptides containing carboxymethylcysteine were found, identical to peptides P1 T2, P1 T1 and P1 T0 of Fig. 2.

In order to find out the function of these three half-cysteine residues, we made diagonal maps of peptic-tryptic digests at pH 3.5 and 6.5 (Figs. 3 and 4). Table 3 lists the electrophoretic mobility and sequence of the cysteine-containing peptides. By this method it was possible to isolate and characterize the N-terminal and C-terminal intrachain disulfide loops, as well as the peptides containing the heavy (H–H) bonds and the cysteinyl residue normally involved in the heavy–light (HL) bridge.

At pH 3.5, all the cysteine-containing peptides moved anodally (Fig. 3). Region 1 (with lowest mobility) contained the two peptides (1a, 1b) making up the C-terminal loop. Region 2 contained the N-terminal loop (2a and 2c) as well as peptide 2b. The sequence of peptide 2b is identical to that of peptide P1 T1 (Fig. 2). Peptide 2b was better seen on diagonal maps at pH 6.5, since it was the only one migrating on the anodal side in the first dimension (Fig. 4). It stained yellow with cadmium ninhydrin, and it had the same composition

Figs. 3 and 4. Drawings of the "diagonal maps," at pH 3.5 and 6.5 respectively, of peptic-tryptic peptides subjected to oxidative cleavage of the disulfide bridges. In Fig. 4 (right), only the acidic regions of the first dimension are shown.
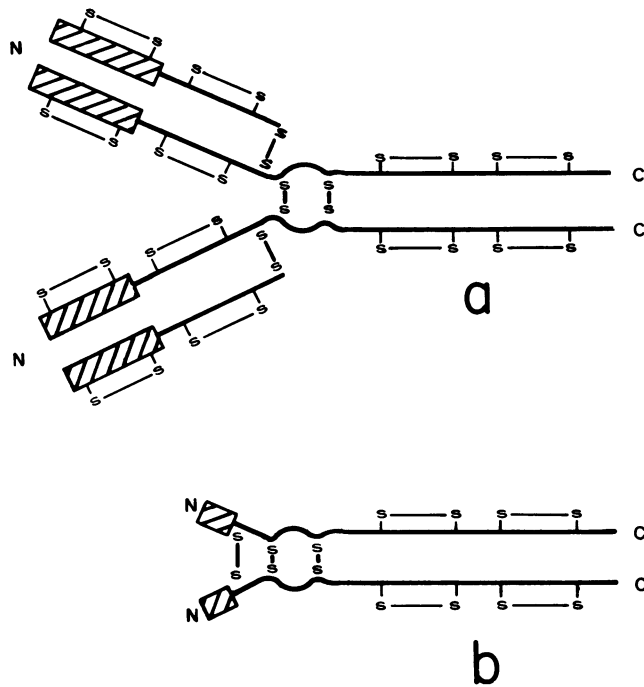
Fig. 5. Schematic diagram of (a) a normal $\gamma 1$ globulin molecule and (b) protein CRA. Most of the Fd fragment is deleted, and there is an additional heavy–heavy disulfide bond. N, NH₂ terminus: C, CO₂H terminus. Hatched area represents variable region.

Table 3. *Properties of peptic-tryptic peptides separated by diagonal electrophoresis (Figs. 3 and 4)*

Peptide	Mobility at pH 6.5	Sequence
1a	0.40	Phe-Ser-Cys-Ser-Val-Met
1b	0.60	Thr-Cys-Leu
2a	0.10	Cys-Lys
2c	0.78	Pro-Glu-Val-Thr-Cys
2b	0.56	Thr-His-Thr-Cys-Pro-Pro-Cys-Pro-Ala-Pro-Glu-Leu
3b	0.50	Ser-Cys-Asp-Lys

and mobility as peptide 2b from Fig. 3. Peptide 2b did not have a mate, indicating that it is joined to itself and must therefore be the peptide bridging both heavy chains. Of particular interest was region 3 (Fig. 3), which contained only a single orange peptide, 3b (serine NH₂-terminal). By its composition, this must be the peptide that is normally involved in the H–L bond. Since at pH 3.5 this peptide does not have a mate it must, in this protein, either contain a free SH group or be joined to the same residue on the other heavy chain via an additional heavy–heavy disulfide bond. To distinguish between these two possibilities, we treated the molecule with [¹⁴C]iodoacetic acid and compared the ratio of ¹⁴C bound (g-atoms/mol of protein) to that for ribonuclease treated similarly. The ratio was <0.3 for CRA, 0.0 for ribonuclease. Since ribonuclease is known not to have a free SH group, this experiment indicates that CRA too lacks such a group and that the half-cysteine at residue 220 must be linked to itself instead of a light chain, which is its usual mate. Confirmatory evidence was obtained by alkylating protein CRA with [¹⁴C]-iodoacetate *without* previous reduction and in the presence of 6 M guanidine. The alkylated protein was subjected to pepsin and trypsin digestion followed by electrophoretic separation and radioautography. No radioactive bands were found, which again indicates the absence of free SH groups.

DISCUSSION

The results of this study are of interest for four reasons. First, it seems that each of the two well-studied γ chain disease proteins has an internal deletion of a part of the Fd fragment rather than, as was originally thought on the basis of immunologic studies, representing the Fc fragment and a contiguous part of the Fd fragment (8). It is particularly striking that

both of the proteins CRA and ZUC, although they belong to different subclasses ($\gamma 1$ and $\gamma 3$ respectively), have a deletion of approximately the same size. In ZUC, the deletion starts at residue 18, whereas CRA has approximately 11 residues before the deletion. Both proteins resume at position 216 with the sequence Glu-Pro-Lys-Ser-Cys. The third protein, Hi, differs in having at least the first 34 residues from the N-terminal end, and, on the basis of a molecular weight of about 40,000, a deletion of only about 100 residues (5). The precise location of the deletion in Hi is not known, nor has the site where the normal sequence resumes been delineated. If it and other heavy chain disease proteins should be shown to resume the proper γ chain sequence at the same point, this may represent the beginning of the sequence coded by the gene that controls synthesis of the Fc fragment constant region. Such a conclusion would give strong support to the concept that several genes control the synthesis of the heavy chain (8). On the basis of prior studies indicating separate genetic control for the v and c regions of the Fd fragment (18), the existence of at least one additional gene for the Fc region would put the total number of genes controlling the synthesis of γ chains at at least three.

The location of the site where synthesis is resumed in the two well-studied proteins is of particular interest since it is two residues removed from the arginine/lysine interchange characteristic of the Gm f and z allotypes of $\gamma 1$ heavy chains (Fig. 2). From their genetic studies and genetic typing of myeloma proteins, Kunkel *et al.* (19) have attempted to explain the linkage of Gm a with z and Gm (non a) with f by postulating the existence of a single gene with two alleles Gm^{az} and Gm$^{non\ a,\ f}$, which expresses itself in both the Fc and Fd fragments. If more direct evidence can be mustered to support the possibility that residue 216 represents the beginning of an Fc fragment gene, an alternative explanation for the observed linkage of the Gm allotypic markers must be considered. It seems possible that an Fd-fragment gene having two alleles Gm^z and Cm^f expresses itself in the constant region of the Fd fragment, while an Fc-fragment gene with two alleles Gm^a and Gm$^{non\ a}$ controls the Fc fragment. Since no recombination has been noted to date, they would have to be closely linked.

A large deletion within a molecule, followed by the synthesis of the product of another closely linked gene, bears a striking resemblance to the effect of certain nonsense mutations which express themselves in one of the genes of a cluster of genes in certain bacteria (20) and suggests that a similar mechanism may be operative in mammalian species.

Secondly, this protein is of interest since it represents an example of a change in the otherwise invariant interchain di-

sulfide bonds. In the absence of light chains, the cysteine residue ordinarily involved in forming the H–L disulfide bond appears to have joined the homologous residue of the other heavy chain to yield an additional inter H–H disulfide bond rather than to remain as a free sulfhydryl group (Fig. 5). Such a possibility is not unlikely, since the two other H–H disulfide bonds probably hold these two half-cysteine residues in close apposition.

Thirdly, the finding of carbohydrate in large amounts close to the N-terminal end is unusual and may explain the difficulties encountered in sequencing that region.

Finally, the observed heterogeneity at the N-terminal region and the finding of glycine as the major N-terminal residue are both distinctly unusual. Glycine has been found once before as the N-terminal residue in a γ chain disease protein–Hi, which has been shown to belong to subclass IV of the variable region (5). In the absence of a precise sequence at the N-terminus it is not possible to determine whether CRA represents yet another variable region subclass or if it commences synthesis at some other site in the Fd fragment. It is unlikely that the observed heterogeneity represents an artefact, since it was found in all preparations of CRA and in all of the peptides examined. Since proteins produced by single clones of plasma cells are usually homogeneous, the observed heterogeneity probably does not reflect the synthesis of different types of molecules. It seems more likely that this may represent degradation of a single molecular species after synthesis, possibly by an amino peptidase similar to that recently described for a nonsense mutant of alkaline phosphatase in *Escherichia coli* (21).

1. Franklin, E. C., J. Lowenstein, B. Bigelow, and M. Meltzer, *Amer. J. Med.*, **37**, 332 (1964).
2. Zawadzki, Z. A., T. G. Benedek, D. Ein, and J. M. Easton, *Ann. Intern. Med.*, **70**, 335 (1969).
3. Franklin, E. C., *New Eng. J. Med.* **282**, 1098 (1970) (editorial).
4. Frangione, B., and C. Milstein, *Nature*, **224**, 597 (1969).
5. Terry, W. D., and J. Ohms, *Proc. Nat. Acad. Sci. USA*, **66**, 558 (1970).
6. Press, E. M., and N. M. Hogg, *Nature*, **223**, 807 (1969).
7. Edelman, G. M., B. A. Cunningham, W. E. Gall, P. D. Gottlieb, H. Rutishauser, and M. J. Waxdal, *Proc. Nat. Acad. Sci. USA*, **63**, 78 (1969).
8. Franklin, E. C., *J. Exp. Med.*, **120**, 691 (1964).
9. Frangione, B., E. C. Franklin, H. H. Fudenberg, and M. E. Koshland, *J. Exp. Med.*, **124**, 715 (1966).
10. Frangione, B., C. Milstein, and E. C. Franklin, *Biochem. J.*, **106**, 15 (1968).
11. Part of this work was done by Blas Frangione for a doctoral thesis, Cambridge, England (1968).
12. Frangione, B., and C. Milstein, *J. Mol. Biol.*, **33**, 893 (1968).
13. Brown, J. R., and B. S. Hartley, *Biochem. J.*, **89**, 59 (1963).
14. Miller, F., and H. Metzger, *J. Biochem.*, **240**, 4740 (1965).
15. Gray, W. R., *Methods Enzymol.*, **11**, 469 (1967).
16. Woods, K. R., and K. T. Wang, *Biochim. Biophys. Acta*, **133**, 369 (1967).
17. Frangione, B., C. Milstein, and J. R. L. Pink, *Nature*, **221**, 145 (1969).
18. Dreyer, W. J., W. R. Gray, and L. Hood, *Cold Spring Harbor Symp. Quant. Biol.*, **32**, 353 (1967).
19. Natvig, J. B., and H. C. Kunkel, *Ser. Haematol.*, **1**, 66 (1968).
20. Garen, A., *Science*, **160**, 149 (1968).
21. Natori, S., and A. Garen, *J. Mol. Biol.*, **49**, 588 (1970).