Partial Duplication in the "Hinge" Region of IgA1 Myeloma Proteins

BLAS FRANGIONE AND CARLOTA WOLFENSTEIN-TODEL

Department of Medicine, New York University School of Medicine, New York, N.Y. 10016

Communicated by Severo Ochoa, October 10, 1972

ABSTRACT The aminoacid sequence of the "hinge" region of two human IgA1 myeloma proteins was found to be identical. A 30-residue fragment contains 3 cysteines, 12 prolines, and carbohydrate, and the sequence of a small segment of the hinge region is repeated at least twice. Since a deletion of 12-13 residues was found in the same region of an IgA1 myeloma protein, it is possible that IgA1 molecules are the result of the insertion of a partially duplicated gene segment in the immunoglobulin genes.

Structural studies of human myeloma proteins belonging to each of the major classes and subclasses of immunoglobulins have clearly defined four or more domains or homology regions that bear striking homologies to each other, suggesting that they have evolved by a series of gene duplications (1-3). An exception to this is the "hinge" region, a fragment of about 20 residues in the center of the heavy chains that bears no homology to any other region of either light or heavy chains. It serves not only as the flexible "hinge" for the two antibody fragments (Fab-fragments), but also links the heavy chains through disulfide bridges to each other, and it may also have other functions. While it is usually rich in proline and cysteine residues, there are striking differences in the aminoacid sequence of the "hinge" region of different classes, and to a lesser extent subclasses, of human heavy chains (2, 4).

In this paper, we present the aminoacid sequence of a fragment that contains the "hinge" region from a human IgA1 myeloma protein and compare it to an extended sequence of another IgA1 protein and to the "hinge" region of an IgA2 myeloma protein (5, 6). Certain unusual structural features of the "hinge" of IgA1 molecules raise several interesting possibilities that may explain the origin and evolution of this region.

MATERIALS AND METHODS

Myeloma protein Oso was purified as described for protein Pat (5). Immunological and chemical typing showed it to be IgA1, k (7).

Enzyme Digestion and Column Chromatography. Protein Oso was digested with t-1-(tosylamido-2-phenyl)ethyl chloromethyl ketone-treated trypsin ( Worthington) in 0.2 M NH4HCO3 (pH 8.3) for 15 hr at 37°, at an enzyme to substrate ratio of 1:50 (w/w). The digest was freeze-dried and was further digested with pepsin (Worthington, twice crystallized) at an enzyme to substrate ratio of 1:50 (w/w) in 5% formic acid for 15 hr at 37°. The digest was fractionated by filtration on a Sephadex G-50 column (3.5 x 180 cm) with 1 M acetic acid as eluant to isolate the hinge peptide. Subtilisin (Nutritional Biochem. Co.) digestion of the hinge peptide was done in 0.2 M NH4HCO3 (pH 8.3) for 15 hr at 37°, at an enzyme to substrate ratio of 1:20 (w/w).

Partial Reduction and Radioactive Alkylation was described (5).

Purification of Radioactive Peptides. The first step in purification of carboxymethylcysteine-containing peptides was paper electrophoresis at pH 3.5, followed by electrophoresis at pH 6.5 and 2.1 (8). The radioactive peptides were localized by autoradiography with Kodak Royal Blue Medical x-ray film. Mobilities at pH 6.5 are expressed as fractions of the distance between ε-Dnp-lysine and aspartic acid (9).

Aminoacid Analysis. Peptides were hydrolyzed at 110° for 20 hr with 6 N HCl, containing 0.1% phenol to prevent destruction of tyrosine (10), in evacuated and sealed tubes. Quantitative aminoacid analyses were performed on a Beckman model 121 automatic aminoacid analyzer equipped with high-sensitivity cuvettes and a recorder. Values are expressed relative to one of the residues taken as 1.0. Amino sugars were qualitatively recognized on the short column of the aminoacid analyzer.

Determination of Aminoacid Sequence. N-terminal residues were identified with dansyl chloride (11), and the derivatives were characterized as described (8). Edman degradation was done manually by dansylation to mark the new amino-terminal residues. The dansyl derivative of carboxymethylcysteine cannot be detected with the solvents used because it does not separate from the 1-dimethylaminonaphthalene-5-sulphonic acid spot. Therefore, when no dansyl amino acid was found, the presence of [14C]carboxymethylcysteine was confirmed by the decrease in negative charge of the peptide after Edman degradation or by the decrease in radioactivity. Carboxypeptidase B (Worthington, diisopropylphosphorofluoridate-treated) digestion was performed in 0.1 M NH4HCO3 for 12 hr at 37° with 5 μg of enzyme per 0.01 μmol of peptide. The digest was freeze-dried and applied to the aminoacid analyzer. The results of the dansyl-Edman procedure and carboxypeptidase B digestion are shown with arrows under the peptide.

Abbreviations: Nomenclature of Immunoglobulins and their chains or fragments follows the recommendations of the World Health Organization [Bull. WHO 30, 447 (1964)]. Myeloma proteins are designated by the first three letters of the patient's name.

Vol. 69, No. 12, pp. 3673-3676, December 1972
RESULTS

Fig. 1 shows the elution pattern obtained by fractionation of a peptic-tryptic digest of myeloma protein Oso on a column of Sephadex G-50. The “hinge” peptide was localized in the first peak after partial reduction and alkylation with 14C-labeled iodoacetic acid, followed by paper electrophoresis at pH 3.5 and radioautography. The peptic-tryptic peptide containing the “hinge” region of IgA₁ molecules runs as a

---

<table>
<thead>
<tr>
<th>Tube No.</th>
<th>A280 nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>0.6</td>
</tr>
<tr>
<td>120</td>
<td>2.2</td>
</tr>
<tr>
<td>160</td>
<td>1.6</td>
</tr>
<tr>
<td>200</td>
<td>0.2</td>
</tr>
</tbody>
</table>

---

**Fig. 1.** Gel filtration of a peptic-tryptic digest of protein Oso (IgA₁) on Sephadex G-50 (3.5 × 180 cm) in 1 M acetic acid. 6-ml fractions were collected at a flow rate of 25 ml/hr. The pooled fractions (Peak 1) are indicated by a bar around fraction 80.

---

**Fig. 2.** Aminoacid sequence of the “hinge” region of two immunoglobulins A₁. Both are carboxymethylcysteine peptic-tryptic (PT) peptides. Protein Oso was further digested with subtilisin (PTS). **Thick lines** indicate duplicated fragment. CHO, carbohydrate fragment.

---

**Fig. 3.** Aminoacid sequence of the “hinge” region of IgA₁ (Oso) and an IgA₂ (Avi) myeloma proteins. Homologous regions are in boxes; **dotted line** indicates gap in A₂ and site of duplication in A₁.
TABLE 1. Aminoacid analyses of a peptic-trypsic carboxymethylcysteine peptide of IgA: myeloma protein Pat that contains the "hinge" region at several steps of an Edman degradation.

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>His</th>
<th>Arg</th>
<th>CMCys</th>
<th>Thr</th>
<th>Ser</th>
<th>Pro</th>
<th>Val</th>
<th>CHO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>1.0</td>
<td>1.0</td>
<td>2.3</td>
<td>4.8</td>
<td>5.0</td>
<td>11.2</td>
<td>2.6</td>
<td>+</td>
</tr>
<tr>
<td>Step 21</td>
<td>0.9</td>
<td>0.8</td>
<td>0.8</td>
<td>1.1</td>
<td>1.8</td>
<td>3.7</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Step 23</td>
<td>1.0</td>
<td>1.0</td>
<td>1.1</td>
<td>2.0</td>
<td>2.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Hydrolysis for 20 hr. Compositions are reported as mol of amino acid per mol of peptide. CMCys, S-carboxymethylcysteine; CHO (+), carbohydrate present.

* For aminoacid analyses of previous steps, see ref. 5.

A series of radioactive bands between ε-Dnp-lysine and the dipeptide glycyll-alanine at pH 3.5 (7). The partial sequence of the "hinge" peptide is shown in Fig. 2; it is compared to the "hinge" peptide obtained from another A1 myeloma protein (Pat). Two other carboxymethylcysteine-containing peptides were also found in the first peak (not shown); their nature and relationship to the "hinge" peptide are under study. The details of the aminoacid sequence of the "hinge" peptide of protein Pat from residues 1 to 18 have been reported (5). We have now extended the sequence to position 22. Aminoacid analysis was performed at step 21 of the Edman degradation without purification and at step 23 after purification by paper electrophoresis. The results are shown in Table 1. In both instances (Fig. 2), the sequence after step 23 (23-25) was inconclusive since small amounts of proline and serine were carried over. Subtilisin was able to split the peptide as indicated in Fig. 2; it was then possible to sequence the region between positions 26 and 30. The aminoacid analysis, the presence of amino sugars, and the NH2-terminal groups of the subtilisin peptides are shown in Table 2. The absence of carbohydrate in peptides PTS 2 and PTS 3 (Table 2) and after step 22 (Table 1) indicates that the carbohydrate moiety is attached either to serine or to threonine residues between position 10 and 22.

**DISCUSSION**

Sequence studies of the "hinge" region of IgA molecules show that like other "hinge" regions, they are rich in cysteine and proline (2); however, the "hinge" region of IgA differs in having at least 15 residues with an unusual sequence that consists of two identical stretches of seven or eight residues and that contains the carbohydrate moiety. It is not known whether there is a single carbohydrate moiety or whether carbohydrate is associated with each of these halves and, therefore, also repeated. It seems likely that this finding is the result of a gene duplication. Comparison of the sequence with that of the "hinge" peptide obtained from an IgA2 protein, a structurally closely related molecule, is particularly informative since these two "hinge" peptides were identical in the first eight residues and in the last five residues (Fig. 3). Nevertheless, the "hinge" region of A2 lacks carbohydrate and has a gap, of about 12-13 residues, just in the place where duplication of a small segment was identified in A1 molecules. At least two possible mechanisms for DNA addition, as described for genetic duplication in bacteriophage lambda (12), can be implicated in explaining this finding. One of these is tandem duplication and unequal crossing-over. The other is DNA insertion at a special sequence and unequal crossing-over at the same region. Both theories can be tested by study of the primary structure of the hinge region of different immunoglobulin heavy chains in man and other species, and particularly by analysis of heavy-chain mutants (heavy-chain disease proteins). Aminoacid sequence studies of three heavy-chain disease proteins (13-15) show that after an internal deletion of variable size (so far always involving the variable (e) and constant (c) region of the Fd-fragment) the same starting point for reinitiation of normal synthesis was found. In all three mutants, the reinitiation occurred at a glutamic acid residue at position 216, just before the "hinge" region. Because of these findings and the lack of homology of this section with the rest of the molecule, several theories were put forward (13-16). One of them suggested that the "hinge" region was under separate genetic control (14). If this were the case, the codon(s) specifying position 216 would represent an attachment site or "site-specific" recombination (12), and one would expect a second site at the other end of the DNA fragment coding for the "hinge" region. The uniqueness of the codon specifying position 216 gained further support with the finding of protein Mγc, which contains a normal sequence through the first 215 residues and has a deletion of 15 residues that commences at position 216 and involves the "hinge" region until position 232 (17). Whether position 232 represents a second "site-specific" recombination remains to be determined. Detailed chemical studies of two recently reported mutants (18, 19), with deletions extending beyond the "hinge" region, may clarify this point.

If the above assumption is correct a short DNA fragment coding for the "hinge" region was integrated into gene(s) coding for immunoglobulins. Indeed, an attractive hypothesis is that the function of this DNA is to fuse duplicated genes. The single chain produced by these fused genes is responsible for distinct biological activities present in each half of the heavy chain and formally associated with separated, although genetically related, molecules. Thus, "gene fusion-junction" is a possible mechanism for the generation of complex molecules in eukaryotic evolution.

We thank Dr. E. C. Franklin for helpful discussions and correction of the manuscript, Ms. F. Prelli for expert technical assistance, and Ms. M. Chavis for secretarial work. This work is supported by USPHS National Institutes of Health Grants.
AM 01431, AM 05064, and AM 02594. B. F. is a Senior Fellow of the Arthritis Foundation, Inc.