Attachment of Carbohydrate to the Variable Region of Myeloma Immunoglobulin Light Chains

Harold C. Sox, Jr.,* and Leroy Hood†

IMMUNOLOGY BRANCH, NATIONAL CANCER INSTITUTE, BETHESDA, MARYLAND

Communicated by Elwin A. Kabat, April 10, 1970

Abstract. Approximately 15 per cent of the light chains from homogeneous immunoglobulins in patients with multiple myeloma contain an oligosaccharide group. Five human myeloma light chains that contained carbohydrate were studied. The sequence Asn-__-Ser/Thr was at the site of carbohydrate attachment in all light chains. The carbohydrate group was attached to the asparagine residue of this triplet sequence which in all five light chains was located in the variable region. The occasional occurrence of carbohydrate in myeloma light chains is seen as the consequence of a variable region mutation creating an Asn-__-Ser/Thr sequence to which carbohydrate is attached by an enzyme capable of recognizing the characteristic triplet sequence.

Introduction. Antibody molecules are glycoproteins with a multichain polypeptide structure consisting of two identical light chains and two identical heavy chains. The functional diversity of antibodies is a result of amino acid sequence variability in the amino terminal 107 and 117 residues of light and heavy chains respectively (the variable region). The amino acid sequence of the remainder of heavy and light chains (the constant region) is identical in all molecules of a given class or type.† An oligosaccharide containing approximately 14 sugar residues is covalently bound to the constant region of all heavy chains.‡ Approximately 15% of the light chains isolated from serum myeloma proteins contain an oligosaccharide comparable to that present on all heavy chains.§ In addition, a homogeneous light chain with covalently attached carbohydrate has been isolated from the urine of two patients with multiple myeloma.¶ We have sought to explain why some but not all myeloma light chains contain carbohydrate and report studies of five carbohydrate-containing light chains.

Materials and Methods. Materials: Homogeneous light chains from the urine of patients with multiple myeloma were generously provided by the following persons: Drs. Corrado Baglioni, C. E. Buckley, III, Daniel Ein, Elliott Osserman, and Martin Weigert. Light chains isolated from serum myeloma proteins were the kind gift of Dr. J. T. Harrington, Jr. A kappa light chain from mouse plasmacytoma MPC37 was the gift of Mr. David McKeen and Dr. Richard Asofsky.

Screening procedure for carbohydrate in light chains: Homogeneous light chains from urine (Bence-Jones proteins) were tested for hexosamine either in dialyzed concentrated urine or after purification. Approximately 200 μg light chains were hydrolyzed in 6 N HCl for 15 hr at 110°C and electrophoresed on paper at 7000 V at pH 1.9.¶ The color reaction of hexosamine, migrating between alanine and valine, was maximal about 1 hr after staining with ninhydrin.
Purification of light chains: Unpurified light chains which seem to contain hexosamine by the semiquantitative screening assay were precipitated with ammonium sulfate at 50% saturation and were dialyzed extensively against 0.1 M Tris-Cl, pH 8.0, or 1 N acetic acid. The final step of purification was gel filtration on G-100 Sephadex in 1 N acetic acid or gradient elution from DEAE-cellulose with 0.1-0.3 M Tris-Cl, pH 8.2. The light chains which had been isolated from serum myeloma proteins by gel filtration on G-100 Sephadex required no further purification. All purified light chains were free of serum-protein contamination as evidenced by immunoelectrophoresis at a protein concentration of at least 10 mg/ml.

Isolation of glycopeptides: Purified light chains (4 mg) were prepared for trypsin digestion by aminoethylolation and for pronase or subtilisin digestion by heat denaturation at 100°C for 30 sec. Enzymatic digestion was performed for 3 hr at 37°C by using a 1% enzyme-substrate ratio in 0.1 M ammonium bicarbonate. Peptide maps of tryptic digests were prepared by using two systems previously described. Glycopeptides stay at the origin in the chromatography dimension and migrate no more than 10 cm in the electrophoresis dimension. To isolate glycopeptides from pronase or subtilisin digests, chromatography and electrophoresis were performed for 48 hr and 90 min respectively. Preparative maps were sprayed with 0.03% ninhydrin in ethanol to localize peptides that were eluted with water.

Amino acid composition and sequence of glycopeptides: Semiquantitative amino acid analysis of common region peptides was performed by electrophoresis at pH 1.9. Glycopeptides were hydrolyzed in sealed, evacuated glass tubes in 6 N HCl for 15 hr at 110°C. When tryptophan was suspected, peptides were hydrolyzed in 6 N HCl containing 4% thioglycolic acid. The subtractive Edman procedure was used for peptide sequencing; amino-terminal residues were labeled with dansyl chloride and identified by high voltage electrophoresis or thin-layer chromatography.

Quantitative amino acid analysis was performed using a Beckman model 120C amino acid analyzer.

Carbohydrate analysis: N-acetylglucosamine and N-acetylgalactosamine were quantitated on an amino acid analyzer after hydrolysis with 4 N HCl in sealed, evacuated tubes for 3 hr at 110°C. Destruction during hydrolysis was measured by identical treatment of amino sugar standards. Neutral sugars were measured by the anthrone reaction with mannose as standard. Sialic acid was determined by the theiobarbituric acid assay with N-acetylneuraminic acid as standard.

Results. To obtain material for study, 89 light chains were screened for carbohydrate (Table 1). For screening homogeneous light chains from the urine of patients with multiple myeloma (Bence-Jones proteins), hexosamine was taken as the most reliable indicator of carbohydrate since it is usually the carbohydrate residue linked directly to the polypeptide chain in immunoglobulins and least likely to be removed by glycosidases present in the kidney. Light chains containing hexosamine were purified, peptide maps were prepared, and the yield of glycopeptide was measured by amino acid analysis. In five light chains,

Table 1. Screening of myeloma light chains for carbohydrate.

<table>
<thead>
<tr>
<th></th>
<th>Tested (x/λ)</th>
<th>Positive* (x/λ)</th>
<th>Positive (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bence-Jones proteins†</td>
<td>71 (31/31)</td>
<td>3 (2/1)</td>
<td>4</td>
</tr>
<tr>
<td>Light chains from serum myeloma proteins</td>
<td>18 (12/3)</td>
<td>2 (2/0)</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>89 (43/34)</td>
<td>5 (4/1)</td>
<td>6</td>
</tr>
</tbody>
</table>

* These light chains contained approximately 1 mole of glycopeptide/mole of light chain as determined by N-acetylglucosamine content and by glycopeptide yield from peptide maps.
† Homogeneous light chains isolated from the urine of patients with multiple myeloma.
‡ Twelve light chains were not typed.
TABLE 2. Carbohydrate content of myeloma light chains.*

<table>
<thead>
<tr>
<th>Light Chain</th>
<th>N-acetyl-glucosamine</th>
<th>N-acetyl-galactosamine</th>
<th>Neutral sugar</th>
<th>Sialic acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBJ4†</td>
<td>3.4</td>
<td>0.3</td>
<td>6.5</td>
<td>1.3</td>
</tr>
<tr>
<td>Ful†</td>
<td>3.9</td>
<td>1.8</td>
<td>10.0</td>
<td>1.5</td>
</tr>
<tr>
<td>Mor†</td>
<td>2.9</td>
<td>0.9</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Dup†</td>
<td>2.3</td>
<td>0.3</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>HBJ10†</td>
<td>2.5§</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

* The values are expressed as moles sugar residue/mole light chain.
† Light chain isolated from the urine of a patient with multiple myeloma (Bence-Jones protein).
‡ Light chain isolated from immunoglobulin G serum myeloma protein.
§ This value is based on hydrolysis of the glycopeptide for 15 hr at 110°C in 6 N HCl and is not corrected for hydrolytic losses.
ND, not determined.

the content of amino sugars (three to four residues per light chain molecule, Table 2) was similar to that in heavy chains, indicating that all or nearly all the light chain molecules in these preparations contained a carbohydrate group. Table 1 shows that three of the 71 Bence-Jones proteins and two of 18 light chains isolated from serum myeloma proteins contained carbohydrate.

Most of the common region peptides from tryptic digests of the five myeloma light chains were identified and none contained carbohydrate. Two criteria were used to identify the variable region subgroup14 to which the light chains belonged. The subgroup of two kappa chains, HBJ4 and HBJ10, was determined by the amino acid sequence of the amino terminal 20 residues.15 In a third kappa chain, Mor (a subgroup specific variable region peptide) was isolated. All three kappa chains were subgroup I. By both criteria, the lambda light chain, Ful, was subgroup II.

Table 2 shows the results of quantitative carbohydrate analysis of the five myeloma light chains. The two Bence-Jones proteins contained hexosamines, neutral sugars, and sialic acid in amounts comparable to heavy chains. The other light chains were not completely characterized.

The amino acid sequence of the site of carbohydrate attachment in five myeloma light chains is shown in Figure 1. Carbohydrate is attached to asparagine in each case. In all five proteins, the glycopeptide bond was alkali stable as demonstrated by isolation of glycopeptides in high yield after prior treatment of the proteins with 0.1 N KOH at room temperature for 24 hr. This observation makes it likely that the glycopeptide bond is a glycosylation linkage between the amide nitrogen of asparagine and the C-1 carbon of N-acetyl-glucosamine as.
is the case for many glycoproteins. It is of interest that there is a sequence of amino acids common to all carbohydrate attachment sites: Asn—____—Ser/Thr. The association of this tripeptide sequence with asparagine-linked carbohydrate has been noted for several glycoproteins.

In three of the five myeloma light chains it was possible to position unequivocally the site of carbohydrate attachment to the polypeptide chain. In the kappa chain, HBJ4, the carbohydrate attachment site was found to be Asn-28 by sequence homology between the glycopeptide and light chains of known sequence (Fig. 2) and also by automated sequential Edman degradation from the N-terminal residue to the Ile at position 29. A second kappa chain, Dup, is virtually identical to Ag, Roy, and Eu at positions 63–72. This glycopeptide has two Asx residues that are potential sites for a glycosylamine bond to asparagine. Both aspartyl residues are part of an Asx—____—Ser/Thr triplet sequence. Since extensive enzymatic digestion with large quantities of pronase, subtilisin, and carboxypeptidase failed to yield a glycopeptide containing only one Asx residue, the site of carbohydrate attachment could be to either Asx-65 of Asx-70 or possibly both. Scarcity of this light chain precluded more complete study. The lambda chain, Ful, contained a large tryptic glycopeptide whose amino acid composition was very similar to the segment 24–45, of the reference proteins Ha and New. A glycopeptide isolated from subtilisin digests had an amino acid sequence that was identical to the segment 23–28 of Ha and New except for Asn-26 to which carbohydrate was attached (Fig. 2).

It was not possible to establish definite sequence homology between reference proteins and the glycopeptides from two light chains, HBJ10 and Mor. The tryptic glycopeptide from these proteins clearly was not a common region peptide. The failure to demonstrate definite homology may mean that the site of carbohydrate attachment is in a part of the variable region which shows partic-

<table>
<thead>
<tr>
<th>CHO</th>
<th>CHO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dup</td>
<td>Ser -Gly -Asx -Glx -Ser -Gly -Thr -Asx-Phe-Thr</td>
</tr>
<tr>
<td>65</td>
<td>70</td>
</tr>
<tr>
<td>Ag</td>
<td>Ser -Gly -Ser -Gly -Phe -Gly -Thr -Asp -Phe-Thr</td>
</tr>
<tr>
<td>Roy</td>
<td>Ser -Gly -Ser -Gly -Ser -Gly -Thr -Asp -Phe-Thr</td>
</tr>
<tr>
<td>Eu</td>
<td>Ile -Gly -Ser -Gly -Ser -Gly -Thr -Glx -Phe-Thr</td>
</tr>
<tr>
<td>CHO</td>
<td>CHO</td>
</tr>
<tr>
<td>Ful</td>
<td>Cys -Ser -Gly -Asn -Ser -Thr</td>
</tr>
<tr>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>Ha</td>
<td>Cys -Ser -Gly -Gly -Ser -Ser</td>
</tr>
<tr>
<td>New</td>
<td>Cys -Ser -Gly -Gly -Ser -Thr</td>
</tr>
<tr>
<td>CHO</td>
<td>CHO</td>
</tr>
<tr>
<td>HBJ4</td>
<td>Cys -Arg -Ala -Ser -Glx -Asn -Val -Ser -Asx -Trp -Leu</td>
</tr>
<tr>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>Ag</td>
<td>Cys -Gln -Ala -Ser -Gln -Asp -Ile -Asn -His -Tyr -Leu</td>
</tr>
<tr>
<td>Eu</td>
<td>Cys -Arg -Ala -Ser -Glx -Ser -Ile -Asx -Thr -Trp -Leu</td>
</tr>
</tbody>
</table>

Fig. 2.—Amino acid sequence homology between light chain glycopeptides and light chains of known sequence. The amino acid sequence of the entire variable region is known for Ag, Roy, Eu, Ha, and New. The numbers refer to the position of an amino acid residue in the variable region, counting from the amino terminal residue.
VOL. 66, 1970  
BIOCHEMISTRY: SOX AND HOOD  
979

25 CHO 30

HBJ4 (Human)  
Ala-Ser-Glx-Asn-Val-Ser-Trp Leu

MOPC46 (Mouse)  
Ala-Ser-Glx-Asx-Ile-Ser-Asn-Asp-Leu

MPC37 (Mouse)  
Ala(Ser,Glx,Asx,Val,Ser,Asx)

Fig. 3.—Comparison of amino acid sequence of carbohydrate attachment sites in mouse and human myeloma light chains. The carbohydrate attachment site in the light chain secreted by murine plasmacytoma, MOPC46, is taken from a published sequence of the tryptic glycopeptide. The light chain synthesized by murine plasmacytoma, MPC37, was isolated from a serum myeloma protein (γ-2b) by G-100 Sephadex chromatography following reduction and alkylation. The light chain glycopeptide was isolated from a pronase digest by a preparative peptide map.

ularly extensive amino acid sequence diversity in the myeloma proteins (e.g., positions 25–35 and 89–96). Indeed, the most probable, though tentative, site of carbohydrate attachment in both light chains is at position 34 as determined by comparing the amino acid and codon sequences of the glycopeptides with reference proteins.

The glycopeptides of the human kappa chain, HBJ4, and the mouse kappa chains, MOPC46 and MPC37, are compared in Figure 3. The mouse and human glycopeptides are nearly identical in amino acid composition or sequence, and the site of attachment of carbohydrate appears to be at the same variable region position in all three proteins.

The various sites of carbohydrate attachment to the variable region of the five myeloma light chains are summarized schematically in Figure 4. All of the attachment sites contained the common amino acid sequence Asn-__-Ser/Vuiabe Constant H2N  
1  
214

COOH

Variable

Constant

H2N

1

107

214

CHO

HBJ4 (K)

25

CHO

Dup (K)

65 or 70

CHO

HB10 (K)

Mor (IC)

26

CHO

Ful (I)

Fig. 4.—Schematic summary of sites of carbohydrate attachment to various human myeloma light chains. The heavy lines in the variable region of the prototype light chain indicate the location of regions of extensive amino acid sequence diversity.
Thr to which carbohydrate was linked by a glycosylamine bond. The regions of particularly extensive diversity of the amino acid sequence are depicted in Figure 4 by heavy lines in the prototype light chain; the site of carbohydrate attachment in HBJ4, Ful, and probably in HBJ10 and Mor lies within these regions.

Discussion. The primary finding of this study is that all the human myeloma light chains with covalently linked carbohydrate that were examined shared a common amino acid sequence, Asn-____-Ser/Thr, at the site of carbohydrate attachment. This tripeptide sequence apparently serves as an acceptor for an enzyme (N-acetylglucosamine-asparagine transglycosylase) catalyzing the formation of a glycosylamine bond between asparagine and N-acetylglucosamine.

All human heavy chains apparently have carbohydrate covalently linked to the constant region.2,21 A \( \gamma_1 \) human heavy chain has carbohydrate attached to the Asx residue of an Asx-Ser-Thr triplet at positions 297–299 of the constant region.20 The other \( \gamma \)-chain subclasses also have carbohydrate bound to Asx at apparently homologous common region sites.24 Glycopeptides containing Asn and Ser have been isolated from at least one common region site on a \( \mu \)-chain.25 Although less is known of \( \alpha \)-chain carbohydrate, there is evidence for asparagine-linked carbohydrate.26,27 There are no Asn-____-Ser/Thr triplets in the constant region of the only completely sequenced heavy chain other than the site of carbohydrate attachment.20 The available evidence is consistent with the view that all heavy chain constant regions contain asparagine-linked carbohydrate and thus that all immunoglobulin-forming cells contain the transglycosylase. Thus, light chains are synthesized in cells containing an enzyme for attaching carbohydrate to an appropriate site on a polypeptide chain.

In this study, glycopeptides from the variable region of human myeloma light chains contained carbohydrate associated with the triplet, Asn-____-Ser/Thr in each case. We sought to learn if this triplet in the variable region always was a site of carbohydrate attachment by surveying 2700 overlapping amino acid triplets from complete variable region sequences28 for the characteristic triplet sequence. Three proteins contained the triplet in the variable region but did not contain variable region carbohydrate. Two were Bence-Jones proteins (Ha and Bo)24 and the third was the heavy chain from a serum macroglobulin (Ou).29 There are at least three possible reasons why carbohydrate might not be linked to an Asn-____-Ser/Thr triplet. First, the triplet could be inaccessible to the transglycosylase because of the folding of the polypeptide chain. Second, the cells in which the protein was synthesized could lack the transglycosylase, although this seems unlikely for cells forming complete immunoglobulin molecules. Finally, the oligosaccharide could be removed from the polypeptide by glycosidases in the kidney.2 There are no data allowing a choice among these alternatives.

Two light chains, HBJ4 and Ful, have carbohydrate attached at or near position 28 of the variable region. Furthermore, it seems likely by amino acid and codon sequence homology that proteins Mor and HBJ10 have their carbohydrate attached either in this region or possibly in a second region located between residues 89 and 96. In another light chain, the site of carbohydrate attachment
is in the segment 25–35 of the variable region.30 The segments 25-35 and 89-96 of the light chain variable region are postulated to be directly involved in antigen binding by two kinds of evidence: first, the extensive amino acid sequence diversity of these regions31–33 and, second, affinity-labeling experiments with a mouse myeloma protein34,35 and rabbit antibody,36 both with dinitrophenyl-binding activity. Since the oligosaccharide is hydrophilic, the regions of the light chain variable region that contain carbohydrate are probably on the exterior of the molecule where the carbohydrate can interact with the hydrophilic solvent. In porcine RNase, the oligosaccharide groups do in fact appear to be attached to the exterior of the protein.37

It has been demonstrated that at least ten different species of mammals have Asx-___-Ser/Thr at the site of attachment of carbohydrate to the heavy chain constant region at a site homologous with the human heavy chain.38 Furthermore, two mouse kappa chains have variable region carbohydrate attachment sites virtually identical in position and sequence to the human kappa chain HBJ4 (Fig. 3). Apparently the specificity of the transglycosylase has been preserved throughout much of mammalian evolution.

The function of carbohydrate in immunoglobulins is unknown although the remarkable conservation of heavy chain carbohydrate throughout mammalian evolution suggests an important role for carbohydrate attached to heavy chains. The postulated role for carbohydrate in facilitating secretion of immunoglobulin from the cell could be served by the oligosaccharide present on all heavy chains. The function of variable region carbohydrate is likewise unknown. It seems unlikely that carbohydrate plays an important role in antigen binding since the oligosaccharide is a bulky group made up of a limited variety of carbohydrate residues and cannot be synthesized in a precisely reproducible fashion. In fact, the bulky carbohydrate group attached near the antigen-binding site could interfere with antigen binding. Were this the case, one would predict that the variable regions of light chains from normal serum immunoglobulins should have little or no carbohydrate bound near the antigen-binding site, since interaction of antigen with antibody bound to the surface membrane of immunocyte precursors is apparently necessary for induction of the immune response.39 This possibility is currently being investigated.

This study suggests that if the sequence triplet Asn-___-Ser/Thr is not sterically hindered it can act as a carbohydrate acceptor from a transglycosylase whose primary function is to glycosylate heavy chains and other proteins. This enzyme may respond indiscriminately whenever a polypeptide contains the characteristic sequence triplet. Thus, carbohydrate attached to the light chain variable region may well be an "accident of nature" because of the creation of the sequence triplet Asn-___-Ser/Thr either by the mutational mechanism that is responsible for generating antibody diversity or by some other mutational event.

We wish to thank Dr. Victor Ginsburg for helpful advice and criticism.

* Present address: Department of Medicine, Dartmouth Medical School, Hanover, N.H.
† Present address and to whom requests for reprints should be addressed: Division of Biology, California Institute of Technology, Pasadena, Calif.


The subgroup designations for immunoglobulin variable regions are those proposed by the Conference on Nomenclature for Animal Immunoglobulins, Prague, 1969.

(a) Hood, L. W. R. Gray, B. G. Sanders, and W. J. Dreyer, Cold Spring Harbor Symposia on Quantitative Biology, vol. 32 (1967), p. 133. (b) In glycopeptides from a variety of proteins, many different amino acids have been found immediately C-terminal to the asparagine linked to carbohydrate. Therefore, the characteristic triplet represented in the text as Asn—Ser/Thr in which the 2 represents any amino acid.


Sox, H. C., Jr. and L. Hood, manuscript in preparation.


Melchers, F., Biochemistry, 8, 938 (1969).


In a subclass of α-chains, one of two glycopeptides isolated contained only N-acetylgalactosamine (see ref. 26). Presumably, the oligosaccharide in this glycopeptide is linked O-glycosidically to either serine or threonine (see ref. 2).

Dayhoff, M. O., Atlas of Protein Sequence and Structure, (Silver Spring, Maryland: National Biomedical Research Foundation, 1969).


Goetzl, E. J. and H. Metsger, manuscript in preparation.


