Immunoglobulin Diversity: Correlation of Non-Allelic Antigenic Markers with the Basic Sequences of the Variable Region of Human Lambda Chains

F. W. Tischendorf, M. M. Tischendorf
Medizinische Universitätsklinik, Tübingen
and
B. Wittmann-Liebold
Max-Planck-Institut für Molekulare Genetik, Abt. Wittmann, Berlin-Dahlem

(Z. Naturforsch. **31 c**, 758–760 [1976]; received August 31, 1976)

Non-Allelic Antigenic Markers, Variable Region, Lambda Chains of Immunoglobulins, Basic Sequences, Subgroups of Lambda Chains

Three forms of the amino terminal half (variable region) of human pathological lambda light chains of immunoglobulins were identified antigenically. By study of all completely sequenced Bence Jones proteins hitherto analyzed and a greater number of proteins subjected to automated sequence determination as well as normal light chains three distinct isotypic basic sequences were identified. The basic sequences are shown to be associated with characteristic antigenic markers representing three V region genes encoding the variable half of lambda chains of immunoglobulins.

The diversity of the amino-acid sequences of human immunoglobulin light chains is considered to be the result of small variations of a few number of basic sequences. Most of the information on the nature of variability of the two main groups of light chains, $\varkappa$ and λ, has been derived from studies of pathological monoclonal light chains, *i. e.* Bence Jones proteins. In 1969 the WHO Commitee proposed that three distinct $\varkappa$ chain variable region basic sequences be designated $V_{\varkappa I}$, $V_{\varkappa II}$, and $V_{\varkappa III}$. A fourth $\varkappa$ chain basic sequence was described recently [1].

Although extensive sequence data on $\varkappa$ chains have been obtained, the data on human λ chains are much more limited. On the basis of homologies among the first 20 to 30 amino acid residues of human λ chains, four [2] or five [3] basic sequences of λ light polypeptide chains have been defined; however, only three subtypes have been recognized immunochemically [4, 5]. The arrangement of λ chains into five prototype sequences as proposed at the 1969 WHO Committee meeting is not entirely convincing, because the variation seen within certain subtypes is not much more than that seen among them [6].

Requests for reprints should be sent to Dr. F. W. Tischendorf, Medizinische Universitätsklinik, *D-7400 Tübingen.*

The availability of representative proteins of different basic sequences has prompted the attempt to establish concordance between immunochemical and structural classification of $\varkappa$ [7, 8] and λ light chains [4, 9]. As more such data accumulate, new ways of arranging the V region sequences into groups become apparent, and revisions of nomenclature need to be made.

From our present immunochemical comparative studies on over 400 λ-Bence Jones proteins and light chains, for which sequence data are available for 177, intergroup similarities between λ chains are evident: The (St^+) determinants [10] were associated with all but one structurally defined $V_{\lambda I}$ proteins, and the (111^+) determinants [4] were found on those proteins with an unblocked amino-terminal residue, *i. e.* tyrosyl and seryl proteins, which do not have pyrrolidone-carboxylic-acid (PCA) at their N-terminal end. Of the 177 Bence Jones proteins tested, both antigenically and structurally, 36 were (St^+) and 53 were (111^+). Polypeptide chains antigenically classified as (St^+) and/or (111^-) exceptionally had blocked N-termini with one exception, protein VIL.

According to the data revealed by Edman degradation the presence of tyrosine in position 2 and a deletion in position 1 is a characteristic of (111^+) polypeptide chains. In 4 proteins (2%) a serine-tyrosine-interchange was found at position 2. Seven proteins (3.5%) began with the NH_2-terminal sequence Ser-1, Tyr-2 and 2 proteins with the sequence Tyr-1, Tyr-2. The antisera used for distinguishing chains of different antigenic specificities were prepared by immunizing rabbits with Bence Jones protein ST (basic sequence $V_{\lambda I}$) and Bence Jones protein 111 (basic sequence $V_{\lambda III}$). Specificity of the antisera was assured by appropriate absorption. These absorbed antisera precipitated only Bence Jones proteins and light chains of the respective basic sequences as demonstrated by studies on 12 completely sequenced λ Bence Jones proteins and 18 proteins with known NH_2-terminal sequences comprising up to 20–40 residues. Antigenic typing was done in the Ouchterlony double diffusion test. In some cases typing was performed by precipitin-inhibition analyses using an intragel absorption technique [11]. The sequence of the NH_2-terminal five amino acids was determined by subjecting native λ chains to sequential amino-terminal (Edman) degradation [12]. The sequence of the first 38 residues (proteins PFA, SG, USH and YO) was determined using a Beckman sequenator with the procedure described by [13].

When V region primary structures of the different antigenic groups are compared with one

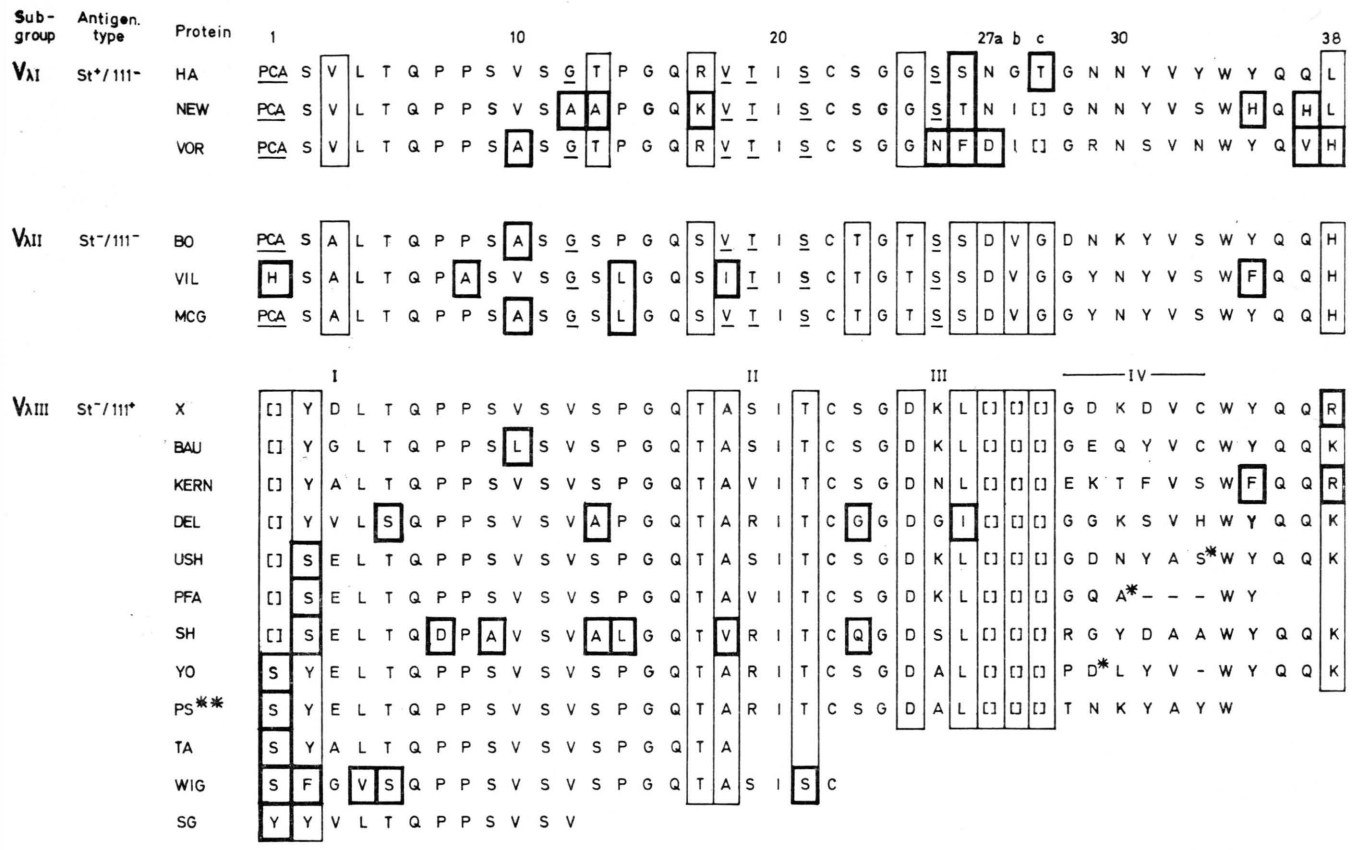

Fig. 1. Comparison of the NH_2-terminal 38 positions of amino acid sequences of human V_λ regions arranged in the three antigenic subgroups discriminated by the variable (St^+) and (111^+) non-allelic antigenic specificities. Amino acid sequences are given in a one-letter notation for amino acid sequences (IUPAC-IUB Tentative rules, Eur. J. Biochem. **5**, 151 [1968]). Positions are numbered from the N-terminus (NEW numbering system). The subgroups are defined by correlated amino acid differences and by amino acid gaps. The subgroup specific positions are indicated by a thin-lined frame. Positions and a region of "hypervariability" in subgroup $V_{\lambda III}$ are marked I—IV. Individually specific differences in non-hypervariable positions are indicated by bold-face type squares, and positions distinguishing $V_{\lambda I}$ and $V_{\lambda II}$ from $V_{\lambda III}$ are underlined. PCA, pyrrolidone-carboxylic acid; * the amino acid residue at this position was not identified; — uncertain results; ** no antigenic typing results available. Sources of data: HA, NEW, BO, VIL, X, BAU, KERN and SH from ref. 16; VOR ref. 17; MCG ref. 18; DEL ref. 19; PS ref. 20; TA ref. 21 and WIG ref. 22.

another as in Fig. 1, it is apparent that these chains can be arranged in subgroups such that the primary structures of the members of any one antigenic subgroup are more similar to one another than they are to those of another subgroup. Chains within a subgroup tend to contain certain amino acid residues at certain positions, and they tend to be of the same chain length. On the average, the V_λ regions of the various λ chain subgroups differ from one another by approximately 40% ($V_{\lambda I}$ versus $V_{\lambda II}$) and 60% ($V_{\lambda I,II}$ versus $V_{\lambda III}$) respectively of their NH_2-terminal sequences. At a number of positions within a subgroup amino acids and gaps can be observed that are subgroup-specific. Certain positions in all λ chain subgroups are highly conserved, *i. e.* few or no variants are found at these positions. Several positions conserved within two subgroups (underlined in the fig.) allow the distinction of $V_{\lambda I,II}$ from $V_{\lambda III}$. Other positions vary to a much greater extent and have been designated "hyper-variable" [14]. The arrangement of the proteins as done in Fig. 1 is in principle agreement with that proposed at the 1969 WHO Committee meeting except that no distinction is made between $V_{\lambda II}$ and $V_{\lambda IV}$ which agrees with the classification of ref. 2 and that protein SH is grouped in the third and not in a separate subgroup (λ_V) because of the close antigenic (and structural) relationship to the "tyrosine" proteins.

The separate grouping of protein SH was recommended at a time where only few data concerning "unblocked" proteins were available. The present data, however, provide evidence that (111^+)

type proteins follow the sequence of protein X. When the NH_2-terminal sequence of (111^+) proteins (except SH) is compared to the corresponding sequence of protein X, a strong homology is observed, since 71% to 87% of the positions are identical. Between SH and X 58% of the positions are identical when the homologous portion of 38 residues is considered. Homology is maximized to 71%, however, when residues adjacent to position 30 ("hypervariable" region) are not taken into consideration, perhaps because of the possible proximity of this area to the antibody active site [6]. Other positions in SH recommended by ref. 2, to be subgroup-specific are identical to homologous positions of several other $V_{\lambda III}$ proteins. Thus, of 18 positions believed to be subgroup specific by these investigators only three remain which may be candidates for subgroup specific residues. These are, however, located in areas (position 7, 9 and 14) highly conserved in light chains of all subgroups, and therefore should not be considered as candidates for subgroup specificity.

From the data resulting from comparison of the known homologous sequences of light chains of different antigenic types, it seems reasonable to propose the following: a) Protein SH does not allow us to define a fourth (fifth) subgroup of variability for the lambda light chains of immunoglobulins. The protein SH was obtained from a Japanese individual and so the question of V_λ polymorphism between races must be considered. It is interesting, however, that the SH constant region is identical to those from several Caucasian individuals and that the USH protein also stems from a Japanese patient. Future antigenic and sequence analyses of λ proteins from different races should resolve this question.

b) The antigenic studies on light chain preparations of γG immunoglobulins of normal individuals which appear to exclude the allelic nature of the three proposed subgroups [11, 5] suggest that the basic sequences $V_{\lambda I}$, $V_{\lambda II}$ and $V_{\lambda III}$ of Bence Jones proteins represent three distinct subgroups of normal light chains rather than pathological proteins.

c) The results indicate further that a strong selection pressure does exist for the considered stretches of sequence which imposes the conservation of subgroup specific segments of the variable region and that at least three germ line genes encode the specificity or V region of λ chains.

Since the (St^+) and (111^+) genetic loci already control $50-60\%$ of λ polypeptide chains, which represent the fairly homogenous subgroups $V_{\lambda I}$ and $V_{\lambda III}$, it appears that not more than a small number of genes encode the V region of immunoglobulin (lambda) light chains.

The authors wish to express their gratitude to Drs. H. F. Deutsch, M. Eulitz, N. Hilschmann, C. Milstein, and F. W. Putnam for their gifts of completely sequenced Bence Jones proteins, to Drs. H. F. Deutsch, C. Tonnelle, and E. M. Press for proteins WIG, TA (both $V_{\lambda III}$), WAL and DAW (both $V_{\lambda I}$), respectively. [Lambda chain DAW showed 89% identy when compared to NEW in fourty-four variable positions [15], *i. e.* residues No. 66-131; personal communication by E. M. Press.] Protein USH was supplied by Dr. H. Ohta. This investigation was supported in part by a grant from the Deutsche Forschungsgemeinschaft.

[1] M. Schneider and N. Hilschmann, Hoppe-Seyler's Z. Physiol. Chem. **355**, 1164 [1974].
[2] B. Langer, M. Steinmetz-Kayne, and N. Hilschmann, Hoppe-Seyler's Z. Physiol. Chem. **349**, 945 [1968].
[3] World Health Organization, Bull. WHO **41**, 975 [1969].
[4] F. W. Tischendorf, Z. Naturforsch. **26 b**, 1292 [1971].
[5] L. Rivat, C. Rivat, and C. Ropartz, Ann. Immunol. (Inst. Pasteur), **125 C**, 697 [1974].
[6] J. A. Gally, Structure of Immunoglobulins, pp. 162–298, in: The Antigens. Vol. I (M. Sela, ed.), Academic Press, New York and London 1973.
[7] A. Solomon and C. L. McLaughlin, J. Exp. Med. **130**, 1295 [1969].
[8] K. Sletten, K. Hannestad, and M. Harboe, Scand. J. Immunol. **3**, 219 [1974].
[9] F. W. Tischendorf and M. M. Tischendorf, Europ. J. Biochem. **13**, 398 [1970].
[10] F. W. Tischendorf and E. F. Osserman, J. Immunol. **102**, 172 [1969].
[11] F. W. Tischendorf, B. Michelitsch, G. Ledderose, and M. M. Tischendorf, J. Mol. Biol. **61**, 261 [1971].
[12] S. Erikson and J. Sjöquist, Biochim. Biophys. Acta **45**, 290 [1960].
[13] B. Wittmann-Liebold, Hoppe-Seyler's Z. Physiol. Chem. **354**, 1415 [1973].
[14] T. T. Wu and E. A. Kabat, J. Exp. Med. **132**, 211 [1970].
[15] N. M. Hogg, Ph. D. Thesis, London 1970.
[16] M. D. Dayhoff, Atlas Prot. Sequence Structure **5** [1972].
[17] M. Engelhardt and N. Hilschmann, Hoppe-Seyler's Z. Physiol. Chem. **356**, 1413 [1975].
[18] J. W. Fett and H. F. Deutsch, Biochemistry **13**, 4102 [1974].
[19] M. Eulitz, Europ. J. Biochem. **50**, 49 [1974].
[20] S. Kochwa, W. D. Terry, J. D. Capra, and N. L. Yang, Ann. N. Y. Acad. Sci. **190**, 49 [1971].
[21] C. Tonnelle, Biochem. Biophys. Res. Comm. **55**, 1112 [1973].
[22] J. W. Fett and H. F. Deutsch, Immunochemistry **13**, 149 [1976].