THE PRIMARY STRUCTURE OF A MONOCLONAL λ-TYPE IMMUNOGLOBULIN L-CHAIN OF SUBGROUP II (BENCE-JONES PROTEIN NEI): EVOLUTIONARY ORIGIN OF ANTIBODY VARIABILITY

Fred A. GARVER and Norbert HILSCHMANN

Max-Planck-Institute for Experimental Medicine, 3400 Göttingen, Germany

Received 27 May 1971

1. Introduction

Comparative sequence studies with monoclonal immunoglobulins have contributed significantly to the understanding of antibody structure and variability [1–3]. According to these studies the specificity of the antibody molecule resides in the variable parts of the immunoglobulin L- and H-chains. On the basis of a discriminating chemical homology these variable parts can be arrayed into well defined subgroups:

3 for κ- [4–7], 4 for λ- [8] and 3–4 for H-chains [9, 10]. These subgroups and subdivisions thereof [11–13] strongly support an evolutionary origin of antibody variability [14, 15]. In order to further characterize the nature of antibody variability we present in this communication the amino acid sequence of another human λ-chain of subgroup II.

2. Materials and methods

The NEI λ-chain, a Bence-Jones protein was obtained from the urine of a patient with multiple myeloma by ammonium sulfate precipitation. The protein was purified on CM- or DEAE-cellulose and Sephadex G-100 chromatography. Endgroups were determined by Edman degradation [16] and hydrazinolysis [17]. The hexose content was measured with the anthrone method [18].

After complete reduction, the protein was aminoethylated [19] or carboxymethylated [20] and digested with trypsin or chymotrypsin respectively. The resulting peptides were purified by repeated ion-exchange chromatography, gel-filtration or preparative paper chromatography, as previously described [21]. The amino acid sequence of the tryptic and chymotryptic peptides was determined by the Edman-dansyl technique [22, 23] directly or on the splitting products isolated after digestion with pepsin, chymotrypsin or thermolysin. The amides of glutamic and aspartic acid were localized by high-voltage electrophoresis at pH 6.5 or by amino acid analysis after hydrolysis of the peptide with aminopeptidase M [24] or leucine aminopeptidase [25]. The C-terminal residues of the peptides were established by hydrolysis with carboxypeptidase A and/or B [26] or C [27] or alternatively by hydrazinolysis [17].

The ordering of the tryptic peptides was based on their homology to the λ-chains VIL [28] and BO [29] and confirmed here for the V-region by the isolation of the chymotryptic overlapping peptides. A detailed description of the methods and results will follow in another publication.

3. Results and discussions

The molar amino acid composition of the purified protein is given in table 1. The N-terminal residue was inaccessible to Edman-degradation and the C-terminal residue was found to be Serine. The complete amino acid sequence of the variable part (position 1–109) and the partial sequence of the constant region (110–214) is shown in fig. 1. The numbering system is derived from the reference protein KERN [32]. Therefore, an insertion of three residues is re-

North-Holland Publishing Company – Amsterdam

Table 1
Amino acid composition of Bence-Jones protein NEI.

	Number of residues	
	In the protein [a]	In the sum of tryptic peptides T1–T18
Lys	13.24	13
His	2.01	2
Arg	4.70	5
Asp	13.29	13
Thr [b]	20.60	21
Ser [b]	34.60	35
Glu	22.05	21
Pro	14.17	14
Gly	17.16	17
Ala	17.27	17
Cys [c]	6.33	6
Val [d]	17.07	17
Met	0.78	1
Ile	4.95	5
Leu [d]	11.19	11
Tyr [b]	9.99	10
Phe	5.38	5
Trp [e]	3.24	3
		216

[a] Expressed as number of residues after 24 hr hydrolysis with 6 N HCl at 110° in the monomer (M.W. about 23,500).
[b] Corrected for losses during hydrolysis.
[c] Determined as cysteic acid after oxidation with performic acid [30].
[d] After 72 hr hydrolysis.
[e] Determined spectrophotometrically [31].

quired at position 27 and shown as 27a, 27b and 27c and a deletion of one amino acid occurs at position 96.

The carbohydrate could be located at position 93 of the V-region. The tryptic (T17b) and chymotryptic (Ch-11) peptide are both anthrone positive and yielded hexosamine on amino acid analysis. Neither aspartic acid nor asparagine could be demonstrated as a PTH derivative at this position.

On the basis of the amino acid sequence of its variable part the L-chain NEI clearly belongs to subgroup II of the λ-chains, as well as proteins BO [29] and VIL [28] which also contain 216 residues and begin with an N-terminal α-pyrrolidonecarboxylic acid. As demonstrated in fig. 2, which gives a survey of all known λ-chain sequences, the proteins can be divided into subgroups. Members of one subgroup possess a greater degree of structural homology in comparison to members of other subgroups. Protein NEI exhibits 77% amino acid sequence identity with protein VIL and 73% with protein BO. Comparison of VIL and BO gives 79% residue identity. The homology of these proteins to L-chains of the other λ-subgroups ranges from 53–65%.

Proteins of one subgroup possess unique amino acid residues common only to proteins of that particular subgroup. These linked amino acid residues common only to proteins of subgroup II (VIL, BO, NEI) can be unequivocally identified at positions 3 (ALA), 17 (SER) 23 (THR), 25 (THR), 27a (ASP), 27b (VAL), 27c (GLY), 57 (VAL) and 90 (TYR). Probable subgroup specific residues also occur at positions 8 (ALA), 18 (ILE), 29 (TYR), 38 (HIS), 50 (VAL), 58 (SER) , 88 (SER) and 94 (ASN), but are modified by an individual specific exchange (see fig. 2).

Protein NEI consists of 216 amino acids which is typical for proteins of subgroup II. In contrast proteins of subgroup III contain 213, subgroup IV 211, and subgroup I generally 216 amino acid residues (Ha has 217). In order to maximize the sequence homology between all λ-chains, gaps must be introduced in the alignment. These gaps are, however, subgroup specific as far as the number of deleted amino acids and the position in the chain is concerned. The single amino acid gap in the most hypervariable region (position 91–96) was first described in protein VIL and placed at position 96 because the VAL–VAL sequence at 97–98 is relatively invariant in λ-chains. The presence of this single amino acid deletion at position 96 characterizes all three proteins as members of subgroup II.

Recently five λ-chain subgroups based upon prototype sequence of the N-terminal 20 residues have been proposed [40]. This would result in placing protein VIL and NEI in another subgroup than protein BO. On the basis of total homology, unique deletions and insertions, and subgroup specific residues of these 3 proteins, however, they clearly belong to a single subgroup. The common residues numerically exceed by far the 3 differences at the N-terminal 20 residues between BO and VIL-NEI proteins at positions 8, 10 and 18. The three linked exchanges in the N-termini of proteins BO and HBJ2 could be the beginning of a subdivision of

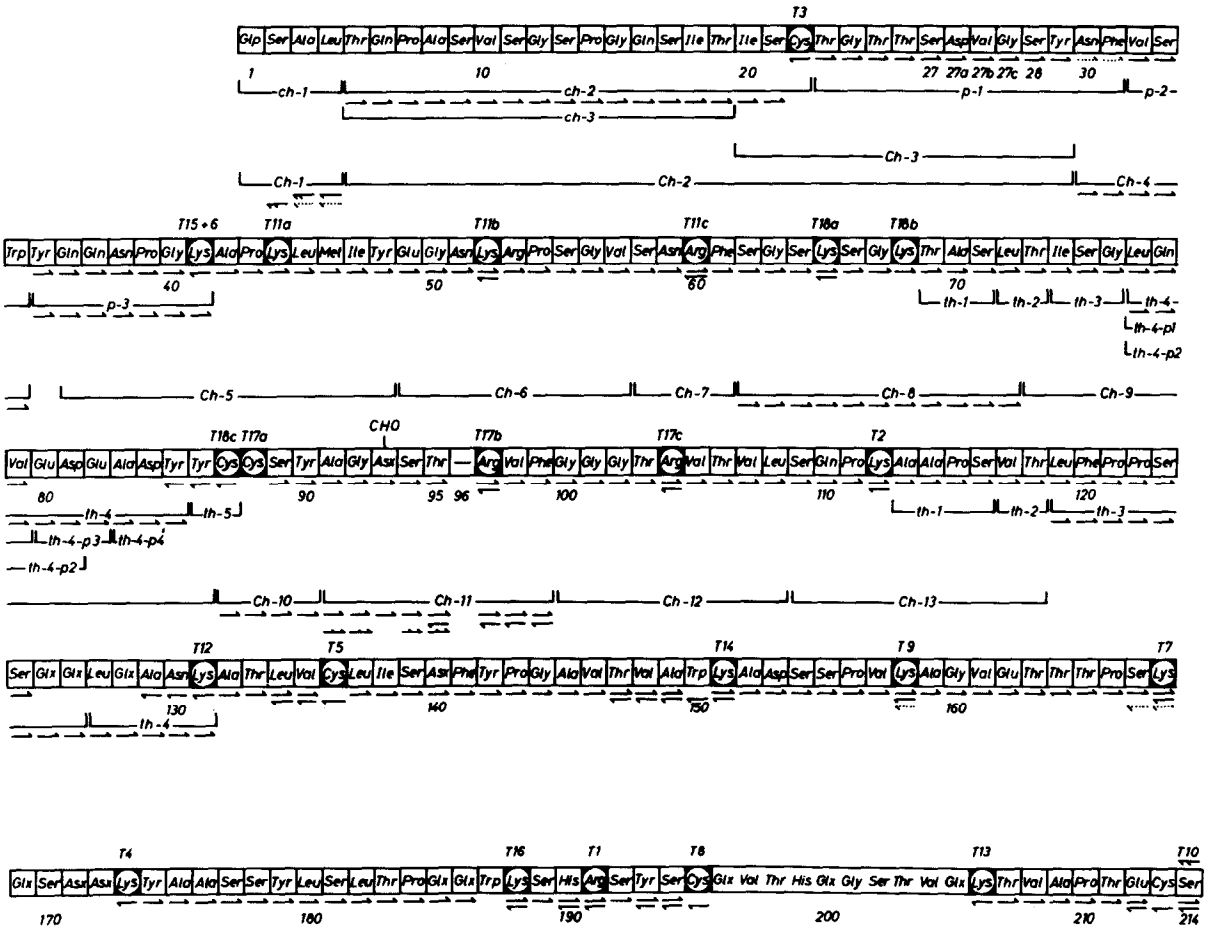

Fig. 1. Amino acid sequence of Bence-Jones protein NEI (λ-type, subgroup II). The protein consists of 216 amino acid residues. The numbering is that of reference protein KERN [32]. Tryptic or chymotryptic splitting products of the protein are indicated as T1 to T18 or Ch-1, Ch-2 respectively. Splitting products of tryptic peptides with chymotrypsin, pepsin or thermolysin are indicated by ch., p- or th-. —↛= Dansyl-Edman degradation [22, 23]; —↠= Edman degradation with identification of PTH-derivatives [16]; ⌇—= Hydrolysis by carboxypepetidase A and/or B [26]; ⌇--= Hydrolysis by carboxypeptidase C [27]; ⌇—= Hydrazinolysis [17].

subgroup II proteins as has been described for κ_I [12–13] and λ_{IV} proteins [11]. Members of one subsubgroup show homology exceeding 85% and represent a subsequent divergence of proteins of that subgroup.

The variability rule of immunoglobulins is in complete agreement with an evolutionary origin of antibody diversity being caused by consecutive gene duplications and subsequent point mutations. The subgroups represent the main branches, the subsubgroups and individual proteins the further ramifications of a phylogenetic family tree which can be constructed by a quantitative comparison of the variable parts. This strict regularity of immunoglobulin variability does not support somatic hypermutation hypotheses, particularly not somatic crossing-over within one subgroup [41]: the linked subgroup specific exchanges are specific for the whole variable part. Special insertion mechanisms to explain the hypervariable regions around positions 27–31, 49–52 and 91–98 are highly improbable [42]. When two proteins are sufficiently similar in their variable parts, the so-called hypervariable sections also show similar if not identical sequences.

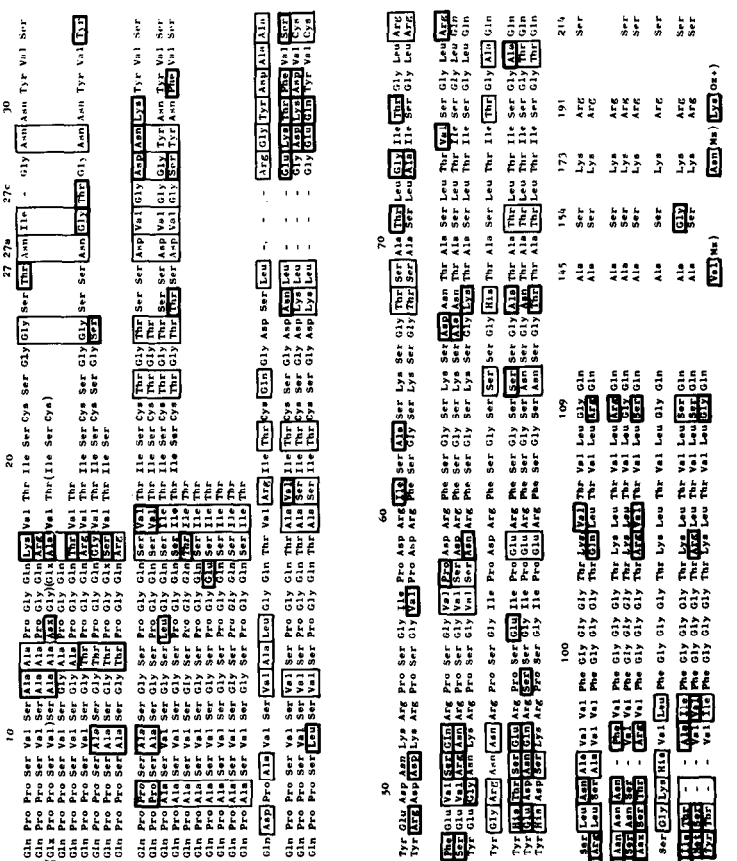

Fig. 2. Comparison of the variable parts of λ-type proteins NEW [8], VIL [28], NEI (this article), KERN [32], BAU [11], 111 [32], 119 [33], BO, HA, SH [29], X [34], HS 94, 78, 2, 86, 68, 77, 70 [35], BJ 98 [36], HBJ 4, 11, 2, 15, 18 [37], KOH [36], PAL [15], RAK [39]. On the basis of their chemical homology the proteins are divided into subgroups I to IV. Subgroup specific exchanges against a basic sequence (identical residues in at least two subgroups) are marked by light boxes, individual specific exchanges by dark boxes, deletions by –. Undetermined sequences are in brackets. Glp = α-Pyrrolidone-carboxylic acid.

Acknowledgements

The generous gift of protein NEI, which was kindly made available to us by Prof. H.E. Bock and Dr. W. Aly, Tübingen, is gratefully acknowledged. We are also grateful to Mr. W. Sinn for the amino acid analyses, and to Mrs. B. Froese and Miss C. Russell for competent technical assistance. This work had been supported in part by the Deutsche Forschungsgemeinschaft.

References

[1] G.M. Edelman and W.E. Gall, Ann. Rev. Biochem. 38 (1969) 415.
[2] C. Milstein and J.R.L. Pink, Progr. in Biophys. Mol. Biol. 21 (1970) 211.
[3] N. Hilschmann, H.U. Barnikol, M. Hess, B. Langer, H. Ponstingl, M. Steinmetz-Kayne, L. Suter and S. Watanabe, in: Current Problems in Immunology, Bayer-Symp. I, eds. O. Westphal, H.E. Bock, and E. Grundmann (Springer-Verlag, Berlin, Heidelberg, New York, 1969) p. 69.
[4] N. Hilschmann, Hoppe-Seyler's Z. Physiol. Chem. 348 (1967) 1718.
[5] H. Niall and P. Edman, Nature 216 (1967) 262.
[6] C. Milstein, Nature 216 (1967) 330.
[7] N. Hilschmann, H.U. Barnikol, M. Hess, B. Langer, H. Ponstingl, M. Steinmetz-Kayne, L. Suter and S. Watanabe, FEBS Symp. 15 (1969) 57.
[8] B. Langer, M. Steinmetz-Kayne and N. Hilschmann, Hoppe-Seyler's Z. Physiol. Chem. 349 (1968) 945.
[9] H. Ponstingl, J. Schwarz, W. Reichel and N. Hilschmann, Hoppe-Seyler's Z. Physiol. Chem. 351 (1970) 1591.
[10] H. Köhler, A. Shimizu, C. Paul, V. Moore and F.W. Putnam, Nature 227 (1970) 1318.
[11] K. Baczko, D. Braun, M. Hess and N. Hilschmann, Hoppe-Seyler's Z. Physiol. Chem. 351 (1970) 763.
[12] S. Watanabe and N. Hilschmann, Hoppe-Seyler's Z. Physiol. Chem. 351 (1970) 1291.
[13] H. Schiechl and N. Hilschmann, Hoppe-Seyler's Z. Physiol. Chem. 352 (1971) 111.
[14] N. Hilschmann, Naturwissenschaften 56 (1969) 195.
[15] L. Hood and D.W. Talmage, Science 168 (1970) 325.
[16] B. Blombäck, M. Blombäck, P. Edman and B. Hessel, Biochem. Biophys. Acta 115 (1966) 371.
[17] S. Akabori, K. Ohno and K. Narita, Bull. Chem. Soc. Japan 25 (1952) 214.
[18] T.A. Scott Jr. and E.H. Melvin, Anal. Chem. 25 (1953) 1656.
[19] M.A. Raftery and R.D. Cole, J. Biol. Chem. 241 (1966) 3457.
[20] C.B. Anfinsen and E. Haber, J. Biol. Chem. 236 (1961) 1361.
[21] H. Ponstingl, M. Hess and N. Hilschmann, Hoppe-Seyler's Z. Physiol. Chem. 352 (1971) 203.
[22] W.R. Gray in: Methods in Enzymology, Vol.11, ed. C.H.W. Hirs (Academic Press, New York, 1967) p. 139 and 469.
[23] K.R. Woods and K.T. Wang, Biochem. Biophys. Acta 133 (1967) 369.
[24] G. Pfleiderer, P.G. Celliers, M. Stanulovic, E.D. Wachsmuth, H. Determann and G. Braunitzer, Biochem. Z. 340 (1964) 552.
[25] E.L. Smith and R.L. Hill in: The Enzymes, 2nd Ed., Vol. 4, eds. P.D. Boyer, H. Lardy and K. Myrbäck (Academic Press, New York, 1960) p. 37.
[26] W. Konigsberg and R.J. Hill, J. Biol. Chem. 237 (1962) 2547.
[27] A. Nordwig, Hoppe-Seyler's Z. Physiol. Chem. 349 (1968) 1353.
[28] H. Ponstingl and N. Hilschmann, Hoppe-Seyler's Z. Physiol. Chem. 350 (1969) 1148.
[29] F.W. Putnam, T. Shinoda, K. Titani and M. Wikler, Science (Washington) 157 (1967) 1050.
[30] C.H.W. Hirs, J. Biol. Chem. 219 (1956) 611.
[31] G.H. Beaven and E.R. Holiday, Adv. Protein Chem. 7 (1952) 319.
[32] H. Ponstingl, M. Hess and N. Hilschmann, Hoppe-Seyler's Z. Physiol. Chem. 349 (1968) 867.
[33] M. Hess and N. Hilschmann, Hoppe-Seyler's Z. Physiol. Chem. 351 (1970) 67.
[34] C. Milstein, J.B. Clegg and J.M. Jarvis, Biochem. J. 110 (1968) 631.
[35] L. Hood and D. Ein, Nature 220 (1968) 764.
[36] C. Baglioni, Biochim. Biophys. Res. Commun. 26 (1967) 82.
[37] L. Hood, W.R. Gray, B.G. Sanders and W.J. Dreyer, Cold Spring Harbour Symp. Quant. Biol. 32 (1967) 133.
[38] A.P. Kaplan and H. Metzger, Biochemistry 8 (1969) 3944.
[39] C.J. Brackenridge, Immunochemistry 4 (1967) 227.
[40] Bull. World Health Org. 41 (1969) 975.
[41] G.M. Edelman and J.A. Gally, Proc. Natl. Acad. Sci. U.S. 57 (1967) 353.
[42] T.T. Wu and E.A. Kabat, J. Exp. Med. 132 (1970) 211.