Tanaka, Y., Frank H., and DeLuca, H. F. (1973), *Endocrinology 92*, 417.
Tanaka, Y., Frank, H., DeLuca, H. F., Koizumi, N., and Ikekawa, N. (1975a), *Biochemistry 14*, 3293.
U.S. Pharmacopoeia (1955), 15th Revision, Easton, Pa., Mack Printing Co., p 889.

Primary Structures of N-Terminal Extra Peptide Segments Linked to the Variable and Constant Regions of Immunoglobulin Light Chain Precursors: Implications on the Organization and Controlled Expression of Immunoglobulin Genes†

Yigal Burstein* and Israel Schechter*

ABSTRACT: The mRNA molecules coding for mouse immunoglobulin (Ig) light (L) chains program the cell-free synthesis of precursors in which extra peptide segments precede the N-termini of the mature proteins; i.e., the extra piece is linked to the variable (V) region (the amino terminal half of the L chain). The complete primary structures of the N-terminal extra pieces of three κ-type and three λ-type L-chain precursors were determined. Despite the fact that the extra pieces differ extensively in sequence (up to 73%) they share the following features: a high percentage of hydrophobic residues (69–75%), they are of comparable size (19–22 residues long) and contain fairly long peptide segments (16–19 residues) devoid of any charged residues, a few amino acids are frequently repeated (e.g., five leucines, five serines, and three isoleucines in RPC-20 extra piece), and methionine is the N-terminal residue. These structural features and other experiments indicate that (1) the V region is larger than hitherto realized; (2) the precursor seems to be the immediate translation product in the cell, since the N-terminal Met was identified as the initiator residue; (3) in the cell the precursor is short lived, and the maturation process (cleavage of the extra piece) may regulate secretion of the mature protein; (4) the physiological role of the hydrophobic extra piece may be to favor interaction of the precursor with cell membranes. Two targets for interaction are considered: the endoplasmic membranes and the cell-surface membrane. The N-terminal extra piece can undergo translocation from the V to the C region, as deduced from characterization of the precursor of the κ-type constant (C) region polypeptide fragment (the carboxy terminal half of the L chain) which was found to contain an N-terminal extra piece (17 residues long) with a primary structure identical to that of the extra piece linked to one of the V regions in whole L-chain precursors. These findings can be formally explained by the two genes–one Ig chain hypothesis, if we assume that the DNA coding for the extra piece (Xp DNA) is a constitutive part of the V gene. Alternatively, we raise the speculation that three genes may code for one Ig chain, where the Xp DNA represents a third distinct gene.

It has been recently realized that the mRNA molecules coding for a variety of secretory proteins program the cell-free synthesis of precursors larger than the mature proteins (see Burstein and Schechter, 1977a). Some understanding of the functions of the precursors may be afforded by determination of their structure, i.e., to determine the position (amino- or carboxy-terminal end), size, and sequence of the extra peptide segment. This was first done in 1973 by subjecting to radioactive sequence analysis the precursor of an Ig[1] L chain (Schechter, 1973). The Ig chains comprise a very heterogeneous population of proteins, and they have unique structural features that make the study of their precursors of special interest. In the mature L chain, the V region (the amino-terminal half of the protein) exhibits sequence variability which is responsible for antibody diversity and specificity, and the C region (the carboxy terminal half of the protein) of either κ- or λ-type L chains have a distinct sequence. We have isolated from mouse myeloma tumors L-chain mRNA molecules that program the cell-free synthesis of precursors in which extra pieces (17–22 residues long) precede the N-termini of both the V (precursors of the whole L chain) and C_κ region (precursor of the κ-type C-region fragment). Complete primary structures of the N-terminal extra pieces of the M-41 κ L-chain precursor and M-104E λ_1 L-chain precursor and partial sequences of extra pieces of other L-chain precursors have been reported previously (Burstein and Schechter, 1977a). Here we report the complete primary structures of five additional extra pieces: two of κ L-chain precursors (M-321 and M-63), one of λ_1 L-chain precursor (RPC-20), one of λ_2 L-chain precursor (M-315), and one of the precursor of the C_κ-region polypeptide fragment. The primary structures of these extra pieces and other experiments indicate the function(s) of precursor molecules in general (e.g., interaction with cell membranes, Schechter and Burstein 1976b) and stimulate new ideas concerning the organization and controlled expression of Ig genes (e.g., three genes may control the synthesis of one Ig chain, Burstein et al., 1977a).

† From the Departments of Organic Chemistry and Chemical Immunology, The Weizmann Institute of Science, Rehovot, Israel. *Received October 31, 1977.* This work was supported by Grant CA 20817 from the National Cancer Institute, United States Public Health Service, and by Grant 806 from the United States–Israel Binational Science Foundation, Jerusalem, Israel.

[1] Abbreviations used are: C region, constant region; H chain, heavy chain; Ig, immunoglobulin; L chain, light chain; MOPC-41, MOPC-63, MOPC-104E, MOPC-315 and MOPC-321 are abbreviated to M-41, M-63, M-104E, M-315 and M-321, respectively; V region, variable region; -X-, unknown amino acid residue in a partial sequence; Xp, extra piece; HV, hypervariable.

0006-2960/78/0417-2392$01.00/0

Experimental Section

Materials. Commercial wheat germ was supplied by the "Bar-Rav" Mill, Tel Aviv. [^{3}H]Glutamic acid (16.2 Ci/mmol), [^{3}H]aspartic acid (23.7 Ci/mmol), [^{3}H]threonine (2.1 Ci/mmol), [^{3}H]tryptophan (20 Ci/mmol), [^{35}S]cystine (58 Ci/mmol), and [^{14}C]glycine (92 mCi/mmol) were purchased from New England Nuclear. [^{3}H]Glutamine (26 Ci/mmol), [^{3}H]alanine (31 Ci/mmol), [^{3}H]valine (15.3 Ci/mmol), [^{3}H]isoleucine (26 Ci/mmol), [^{3}H]leucine (53 Ci/mmol), [^{3}H]proline (43 Ci/mmol), [^{3}H]phenylalanine (15.8 Ci/mmol), [^{3}H]tyrosine (22 Ci/mmol), [^{3}H]arginine (16.7 Ci/mmol), [^{3}H]lysine (18 Ci/mmol), [^{3}H]histidine (55 Ci/mmol), and [^{35}S]methionine (240 Ci/mmol) were obtained from The Radiochemical Centre, Amersham. [^{3}H]Asparagine (13 Ci/mmol) was from Schwarz/Mann. The [^{35}S]cystine was reduced to [^{35}S]cysteine as described (Schechter and Burstein, 1976b).

Immunoglobulin L-Chain mRNAs. The mRNA molecules coding for the M-321 κ-type L chain (Schechter, 1973), M-63 κ-type L chain (Burstein and Schechter, 1976), κ-type C-region polypeptide fragment (Burstein et al., 1977a), RPC-20 λ_1-type L chain, and M-315 λ_2-type L chain (Burstein and Schechter, 1977b) have been previously characterized. These mRNAs were prepared from myeloma polysomes specifically precipitated by antibodies to L chain (Schechter, 1973, 1974).

Cell-Free Synthesis of L-Chain Precursors. Translation of the L-chain mRNAs to provide the precursor molecules was carried out in the wheat germ cell-free system (Roberts and Paterson, 1973) as described (Burstein and Schechter, 1976).

Amino Acid Sequence Analyses. The total cell-free products that were labeled by one radioactive amino acid at a time were analyzed in the Beckman Model 890C automatic sequencer, as detailed elsewhere (Schechter and Burstein, 1976a). The [^{35}S]cysteine-labeled product was completely reduced and alkylated prior to sequence analysis (Schechter and Burstein, 1976b). The radioactive samples contained 5 mg of sperm whale apomyoglobin carrier. All samples were sequenced twice. In the duplicate, the pattern of radioactive peaks was identical. Repetitive yields of the radioactive labeled material and of the protein carrier ranged between 91 and 94%. The absolute yields were corrected for background and "out of step" radioactivity (Smithies et al., 1971) and for endogenous activity of the cell-free extract (Schechter and Burstein, 1976a; Burstein et al., 1976a). Absolute yields of the various amino acids studied were: M-321 precursor, 94–103%; M-63 precursor, 67–91% with the exception of [^{3}H]tryptophan and [^{3}H]asparagine where the yields were 35 and 53%, respectively; RPC-20 precursor, 64–73% with the exception of [^{3}H]tryptophan recovered at 42% yield; M-315 precursor, 68–97% with the exception of [^{3}H]tryptophan recovered at 53% yield; C_κ-region precursors, 28–45%.

Results

Primary Structure of the Precursor of M-321 κ-Type L Chain. Analyses of M-321 L-chain precursor labeled with 12 radioactive amino acids have been previously reported, these included [^{3}H]leucine (Schechter, 1973), [^{35}S]methionine, [^{3}H]serine, [^{3}H]proline, [^{3}H]isoleucine, [^{3}H]alanine, [^{3}H]-valine, [^{3}H]threonine, [^{3}H]phenylalanine, [^{3}H]tyrosine, [^{3}H]tryptophan, and [^{35}S]cysteine (Schechter and Burstein, 1976b). The results showed that in the precursor an extra peptide segment (20 residues long) precedes the N-terminus of the mature M-321 L-chain; the partial sequence of the extra piece determined the positions of 16 residues; the amino acids occupying positions 2, 4, 17, and 20 were not identified (Burstein and Schechter, 1977a). In order to determine the entire primary structure of the extra piece, we have synthesized and sequenced M-321 L-chain precursors labeled with the remaining eight amino acids. Flat and low backgrounds of radioactivity were obtained from sequencer runs (24 cycles) of precursors labeled with [^{3}H]asparagine, [^{3}H]arginine, [^{3}H]lysine, and [^{3}H]histidine, thus showing that these amino acids are not present in the extra piece. Results from sequence analyses of precursors labeled with the other four amino acids (Figure 1) show radioactive peaks at the indicated degradative cycles: [^{3}H]glutamic acid, cycle 2; [^{3}H]aspartic acid, cycles 4, and 21; [^{14}C]glycine, cycles 17, 20, and 36; [^{3}H]glutamine, cycles 26, and 37. As expected, after 20 degradative cycles the position of amino acids in the precursor shows perfect homology with the position of these residues in the mature L chain; that is, Asp^{21}, Gln^{26}, Gly^{36}, and Gln^{37} in the precursor match with Asp^1, Gln^6, Gly^{16}, and Gln^{17} in the mature M-321 L chain (McKean et al., 1973). The four residues identified in the extra piece (Glu^2, Asp^4, Gly^{17}, and Gly^{20}) fill all gaps in the partial sequence determined previously (see above). The positions of all residues in the precursor of the M-321 L chain determined here and in earlier studies are given in Figure 2.

Primary Structure of the Precursor of M-63 κ-Type L Chain. Analyses of M-63 L-chain precursor labeled with [^{3}H]isoleucine, [^{3}H]leucine, [^{3}H]alanine, [^{3}H]proline, [^{3}H]serine, and [^{35}S]methionine have shown that the N-terminal extra piece is 20 residues long, with the partial sequence: Met-X-X-X-X-Leu-Leu-Leu-X-X-Leu-Leu-Leu-X-X-Pro-X-Ser-X-X. The partial sequences of the extra pieces of M-63 and M-321 precursors were identical (Burstein and Schechter, 1976). Therefore, we have synthesized M-63 precursors labeled with six radioactive amino acids known to be in the M-321 extra piece: Glu, Thr, Asp, Trp, Val, and Gly. The radioactive peaks derived from sequencer runs of the M-63 percursor (Figure 1) assign these residues in the following positions: Glu^2, Thr^3, Asp^4, Thr^5, Trp^9, Val^{10}, Trp^{14}, Val^{15}, Gly^{17}, Thr^{19}, Gly^{20}, Val^{23}, Thr^{25}. As anticipated for an N-terminal extra piece 20 residues long, Val^{23} and Thr^{25} in the precursor match with Val^3 and Thr^5 in the mature M-63 L chain (McKean et al., 1973). Other residues identified (Glu^2 to Gly^{20}) fill all gaps in the partial sequence of the M-63 extra piece given above.

The N-terminal residue of the mature M-63 L chain was assigned once as asparagine and once as aspartic acid (Figures 1 and 5 in the paper by McKean et al., 1973). Sequence analyses of M-63 precursor labeled with [^{3}H]asparagine yield a distinct radioactive peak at degradative cycle 21, matching with the N-terminal residue of the mature L chain. On the other hand, sequencer runs of M-63 precursor labeled with [^{3}H]aspartic acid yield a single radioactive peak at cycle 4 and none at cycle 21 (Figure 1). These findings establish that asparagine is the N-terminal residue of the mature M-63 L chain. A similar unambiguous discrimination between glutamine and glutamic acid achieved by sequencing of other L chain precursors synthesized in vitro from [^{3}H]glutamine and [^{3}H]-glutamic acid have been reported and discussed elsewhere (Burstein and Schechter, 1977b). The positions of all residues in the precursor of the M-63 L chain determined here and in earlier studies are given in Figure 2.

Primary Structure of the Precursor of RPC-20 λ_1-Type L Chain. Analyses of RPC-20 L-chain precursor labeled with [^{3}H]alanine, [^{3}H]serine, [^{3}H]glutamic acid, [^{3}H]glutamine, [^{3}H]threonine, and [^{35}S]methionine have shown that the N-terminal extra piece is 19 residues long, with the partial sequence: Met-Ala-X-X-Ser-X-X-X-Ser-X-X-Ala-X-Ser-Ser-X-Ala-X-Ser (Burstein and Schechter, 1977b). The ala-

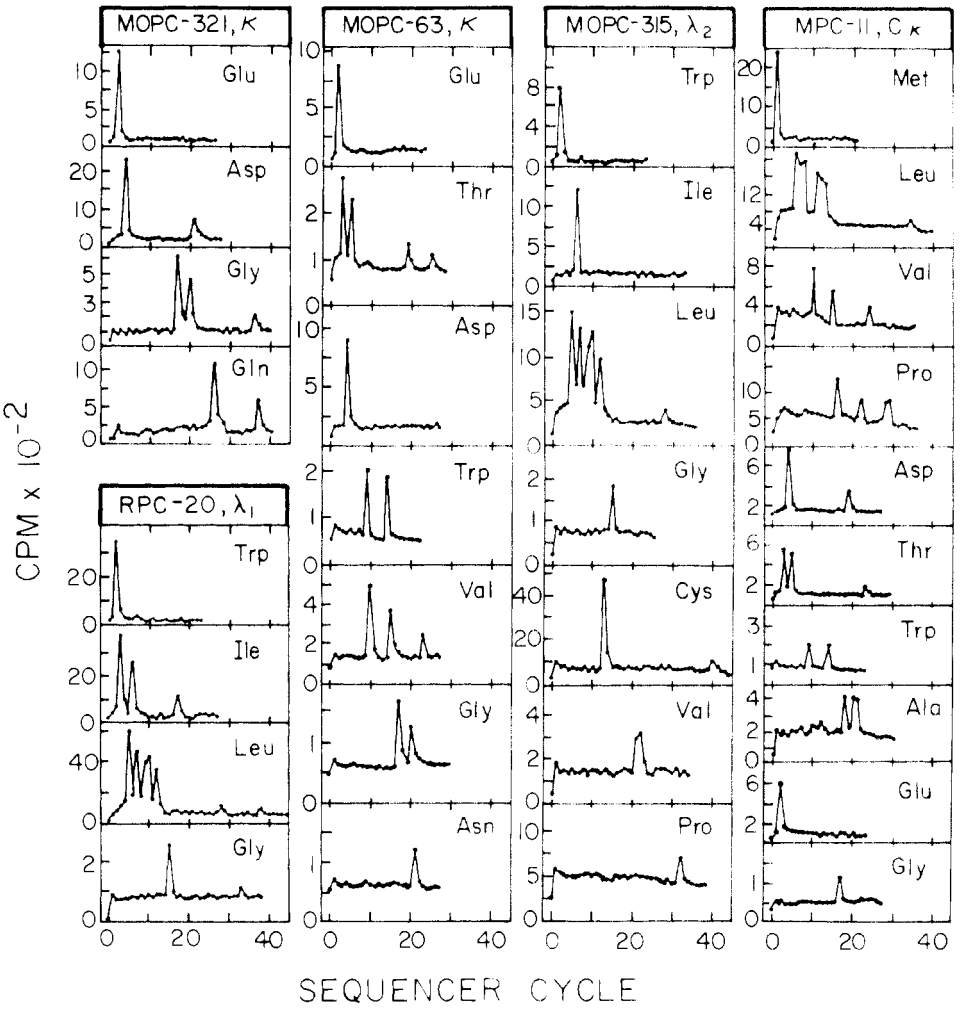

FIGURE 1: Radioactivity recovered at each sequencer cycle from the total cell-free products programmed by L-chain mRNAs. The mRNAs and the radioactive amino acids used to label the products in the wheat germ cell-free system are indicated. The numbers in parentheses represent cpm in the sample analyzed. MOPC-321 (κ): [^{3}H]glutamic acid (11 900), [^{3}H]aspartic acid (32 500), [^{14}C]glycine (20 300), [^{3}H]glutamine (53 200). MOPC-63 (κ): [^{3}H]glutamic acid (22 200), [^{3}H]threonine (11 300), [^{3}H]aspartic acid (28 400), [^{3}H]tryptophan (8900), [^{3}H]valine (27 200), [^{14}C]glycine (10 400), [^{3}H]asparagine (12 200). RPC-20 (λ_1): [^{3}H]tryptophan (65 800), [^{3}H]isoleucine (93 200), [^{3}H]leucine (303 000), [^{14}C]glycine (18 100). MOPC-315 (λ_2): [^{3}H]tryptophan (10 300), [^{3}H]isoleucine (25 700), [^{3}H]leucine (61 600), [^{14}C]glycine (5100), [^{35}S]cysteine (113 000), [^{3}H]valine (31 600), [^{3}H]proline (120 000). C_κ precursor from MPC-11 (N.P.): [^{35}S]methionine (14 600), [^{3}H]leucine (109 000), [^{3}H]valine (48 900), [^{3}H]proline (78 000), [^{3}H]aspartic acid (37 200), [^{3}H]threonine (19 600), [^{3}H]tryptophan (11 700), [^{3}H]alanine (36 000), [^{3}H]glutamic acid (20 500), [^{14}C]glycine (4200). Background radioactivity obtained from sequencer run of the control sample was subtracted. The control sample consisted of wheat germ extract containing each of the labeled amino acids, but without added mRNA (Schechter and Burstein, 1976a). Cycle zero represents a blank cycle (without phenyl isothiocyanate) which was used to wash out potential radioactive contaminants.

nine and serine residues were recovered in high yields (74-78%), while the recovery of the N-terminal methionine was significantly lower (about 10%). By using initiator [^{35}S]-Met-tRNA it was established that this is the initiator methionine (Zemell et al., 1977) which is known to be short lived and rapidly cleaved in other proteins as well (Housman et al., 1970). Accordingly, most of the precursor molecules lacked the N-terminal methionine, and the sequence of the mature RPC-20 L chain was recorded after 18 degradative cycles. The partial sequences of the extra piece of RPC-20 and M-104E λ_1-type L-chain precursors were found to be identical (Burstein and Schechter, 1977b). Therefore, we have synthesized RPC-20 precursor molecules labeled with four radioactive amino acids known to be in the M-104E extra piece: Trp, Ile, Leu, and Gly. The radioactive peaks derived from sequencer runs of the RPC-20 precursor (Figure 1) assign these residues in the following positions: Trp^2, Ile^3, Leu^5, Ile^6, Leu^7, Leu^9, Leu^{10}, Leu^{12}, Gly^{15}, Ile^{17}, Leu^{28}, Gly^{33}, Leu^{38}. In agreement with the previous studies (Burstein and Schechter, 1977b), here again after 18 degradative cycles the position of amino acids in the precursor show perfect homology with the position of these residues in the mature L chain; that is, Leu^{28}, Gly^{33}, and Leu^{38} in the precursor match with Leu^{10}, Gly^{15}, and Leu^{20} in the mature RPC-20 L chain (Weigert et al., 1970; Apella, 1971). The numbering of every degradative cycle from the sequencer run should be increased by one to fit residue position in the precursor prior to removal of the initiator methionine (see above). That is, when the initiator methionine is counted as the first residue, then the tryptophan recovered at cycle 2 fits position 3 in the intact precursor, isoleucine recovered at cycle 3 fits position 4, etc. Applying this correction, it is readily seen that Trp^2 to Ile^{17} recovered from the sequencer analyses fill all gaps in the partial sequence of the RPC-20 extra piece given above. The positions of all residues in the precursor of RPC-20 L chain determined here and in earlier studies are given in Figure 2.

Primary Structure of the Precursor of M-315 λ_2-Type L Chain. Analyses of M-315 L-chain precursor labeled with [^{3}H]alanine, [^{3}H]serine, [^{3}H]glutamic acid, [^{3}H]glutamine, [^{3}H]threonine, and [^{35}S]methionine have shown that the

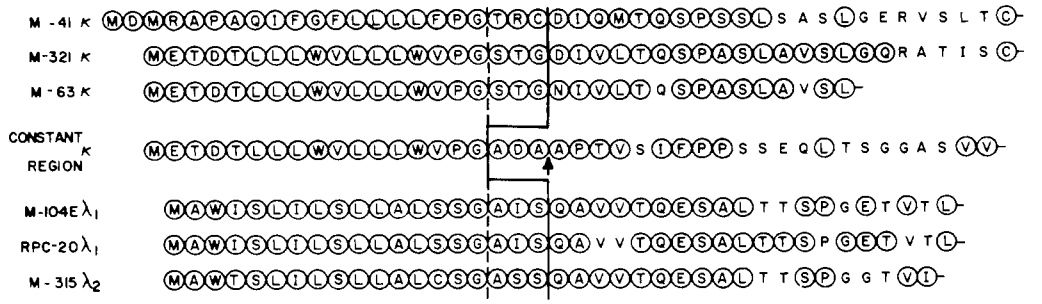

FIGURE 2: Amino terminal sequences of mouse immunoglobulin light chain precursors. Continuous line marks the end of the extra piece and the beginning of the mature L-chain sequence, as well as the point of cleavage with the exception of the C_κ precursor where the cleavage point is indicated by an arrow. In the Discussion we consider the possibility that in all precursors the extra piece may end at the position indicated by the dashed line (see text). Primary structures are based on radioactive sequence analyses of precursor molecules programmed by L-chain mRNAs in the wheat germ cell-free system. Precursors labeled with 20 radioactive amino acids were: M-41 (Burstein and Schechter, 1977a), M-321 (Schechter and Burstein, 1976b, and this report), M-104E (Burstein and Schechter, 1977a), M-315 (Burstein and Schechter, 1977b, and this report). The precursor of M-63 L chain was labeled with 13 radioactive amino acids (Burstein and Schechter, 1976, and this report). The precursor of RPC-20 L chain was labeled with 10 radioactive amino acids (Burstein and Schechter, 1977b, and this report). The precursor of the κ-type constant region was labeled with 11 radioactive amino acids (Burstein et al., 1977a, and this report). Amino acids identified are enclosed within circles; they were assigned in 237 positions. Double occupancy was not observed in any position and sequences of the mature L-chain portions were correctly identified in all cases. The one-letter amino acid code of Dayhoff (1972) is used.

N-terminal extra piece is 19 residues long, with the partial sequence: Met-Ala-X-Thr-Ser-X-X-X-Ser-X-X-Ala-X-X-Ser-X-Ala-Ser-Ser (Burstein and Schechter, 1977b). The alanine, serine, and threonine residues were recovered in high yields (69-97%). On the other hand, the recovery of the N-terminal methionine was low (about 10%). Presumably, this is the short-lived initiator methionine identified in the structurally related λ_1-type M-104E (Burstein et al., 1976a) and RPC-20 (Burstein and Schechter, 1977b) L-chain precursors. Accordingly, most of the precursor molecules lacked the N-terminal methionine, and the sequence of the mature M-315 L-chain was recorded after 18 degradative cycles. The partial sequence of the M-315 λ_2 extra piece differed from that of the M-104E and RPC-20 λ_1 extra pieces (Burstein and Schechter, 1977b). Therefore, in this case we have synthesized and sequenced M-315 L-chain precursor labeled with the remaining 14 amino acids. Flat and low backgrounds of radioactivity were obtained from sequencer runs (24 cycles) of precursors labeled with [^{3}H]phenylalanine, [^{3}H]tyrosine, [^{3}H]aspartic acid, [^{3}H]asparagine, [^{3}H]histidine, [^{3}H]lysine, and [^{3}H]arginine, thus showing that these amino acids are not present in the extra piece. The peaks of radioactivity derived from sequencer runs of precursor molecules labeled with the other seven amino acids (Figure 1) assign residues in the following positions: Trp^2, Leu^5, Ile^6, Leu^7, Leu^9, Leu^{10}, Leu^{12}, Cys^{13}, Gly^{15}, Val^{21}, Val^{22}, Leu^{28}, Pro^{32}. In agreement with previous studies, the positions of amino acids after 18 degradative cycles in the precursor show perfect homology with the positions of these residues in the mature L chain, that is, Val^{21}, Val^{22}, Leu^{28}, and Pro^{32} in the precursor match with Val^3, Val^4, Leu^{10}, and Pro^{14} in the mature M-315 L chain (Dugan-Schulenburg et al., 1973). Similar to the RPC-20 precursor (see above) the number of sequencer cycles is increased by one to fit the position in the precursor prior to the removal of the initiator methionone. Applying this correction, it is readily seen that Trp^2 to Gly^{15} recovered from the sequence analyses fill all gaps in the partial sequence of the M-315 extra piece given above. The positions of all residues in the precursor of M-315 L chain determined here and in earlier studies are given in Figure 2.

Primary Structure of the Precursor of the κ-Type Constant Region Polypeptide Fragment. The mRNA coding for the C_κ-region was purified from two clones derived from the MPC-11 mouse myeloma (Burstein et al., 1977a). Clone 66.2, designated MPC-11(L), has lost the ability to synthesize intact H chain but continues to synthesize and secrete the MPC-11 κ L chain; clone NP-2, designated MPC-11(NP), has lost the ability to synthesize both intact H and L chains; the two clones synthesize small amounts of the mature C_κ-region fragment (Kuehl and Scharff, 1974). The C_κ-mRNA directs the cell-free synthesis of a precursor of the C_κ region in which an extra piece (17 residues long) precedes the N-terminal residue (Ala^{109}) of the mature C_κ region. Sequence analyses of C_κ-precursor molecules labeled with nine radioactive amino acids (Leu, Val, Pro, Asp, Thr, Trp, Ala, Phe, Met) determined the positions of 15 residues in the C_κ-extra piece; the partial sequences of the C_κ-extra pieces in precursors programmed by C_κ-mRNA from both MPC-11 clones were identical; the partial sequence of the C_κ-extra piece matched with the sequence of the extra piece linked to the V region of the M-321 whole L-chain precursor; the amino acids at positions 2 and 17 in the C_κ extra piece were not identified (Burstein et al., 1977a). Since these positions are occupied by glutamic acid and glycine residues in the M-321 extra piece (Figure 2), we have synthesized and sequenced C_κ-precursors labeled with [^{3}H]glutamic acid and [^{14}C]glycine. The results show glutamic acid at position 2 and glycine at position 17 (Figure 1), thus filling all gaps in the partial sequence of the C_κ extra piece. The same results were obtained for C_κ precursors programmed by C_κ-mRNA from both MPC-11 clones. Sequencer data of the C_κ-precursor programmed by C_κ-mRNA from the MPC-11(NP) clone which have not been published previously (Burstein et al., 1977a) are also included in Figure 1. The positions of all residues in the C_κ precursor determined here and in earlier studies are given in Figure 2.

Discussion

The complete sequence data confirm and extend previous suggestions concerning structural features and presumed functions of the L-chain precursor deduced from partial sequences previously reported. These issues are briefly discussed below and new points are raised.

In L-chain precursors, the N-terminal extra piece is linked to the V region of the mature protein. The extra pieces exhibit sequence variability (up to 73%) which follows the pattern of heterogeneity of the V regions of both κ and λ L chains (see lines 1, 2, 3, and 4 in Table I). These findings strongly indicate that the N-terminal extra piece is part of the V region; the implication at the genome level is that the V gene may be larger than hitherto realized (Burstein and Schechter, 1976;

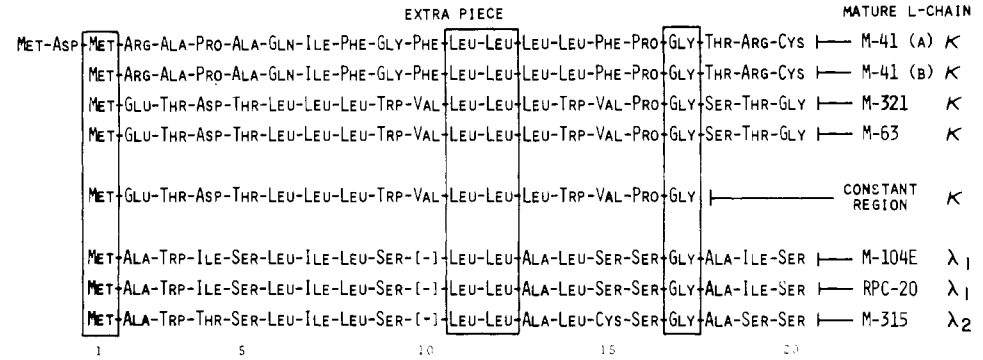

FIGURE 3: Primary structures of the N-terminal extra pieces of immunoglobulin light-chain precursors. The M-41 L-chain mRNA directs the synthesis of two precursors with extra pieces that differ by two residues: the two N-terminal residues (Met-Asp) of the long extra piece (A, 22 residues) are missing in the short extra piece (B, 20 residues); in the other 20 positions both extra pieces have an identical sequence (Schechter and Burstein, 1976c). For maximal sequence homology a gap (indicated by a bar) is inserted in the λ extra pieces. Residues common to all extra pieces are enclosed.

TABLE I: Sequence Heterogeneity at the N-Terminal Extra Piece of Immunoglobulin Light-Chain Precursors.

L chain precursors compared	Amino acid differences in Xp No.[a] (%)	V region No.[b] (%)	HV regions No.[b] (%)
M-41 (κ)/M-321 (κ)	16/22 (73)	53/111 (48)	20/31 (65)
M-63 (κ)/M-321 (κ)	0/20 (0)	8/111 (7)	4/31 (13)
M-104E (λ_1)/M-315 (λ_2)	3/19 (16)	11/110 (10)	4/26 (15)
M-104E (λ_1)/RPC-20 (λ_1)	0/19 (0)	1/110 (1)	0/26 (0)
M-104E (λ_1)/M-41 (κ)	16/22 (73)	64/110 (58)	23/30 (77)
M-104E (λ_1)/M-321 (κ)	14/20 (70)	68/111 (61)	24/31 (77)

[a] Numbers for the extra pieces are based on sequence data given in Figure 3. [b] Numbers for the entire variable and hypervariable regions (Wu and Kabat, 1970) are based on known sequences of the mature L chains (Kabat et al., 1976).

Schechter et al., 1977). The extent of variability of the extra piece is often greater than it is in the entire V region, and it is comparable to the variability found when the hypervariable regions (Wu and Kabat, 1970) of the L chains are examined (lines 1, 3, 5, 6 in Table I). This suggests that the DNA segment coding for the extra piece may be a "hot spot" with an accelerated mutation rate (Schechter and Burstein, 1976b).

Sequence variability of the N-terminal extra piece is also manifested in two protein precursors synthesized by the same cell. The M-315 myeloma secretes IgA (α H chain, λ_2 L chain). The structure of the extra piece of the λ_2 L-chain precursor is known (Figure 3). Recently the partial sequence of the hydrophobic N-terminal extra piece of the M-315 α H chain precursor programmed by mRNA in vitro has been determined (Burstein et al., 1977b). The data show that the extra pieces of the α H-chain and λ_2 L-chain precursors differ extensively in sequence (at least in 7 out of 18 positions). Palmiter et al. (1977) also reported sequence differences in the extra pieces of precursors of three secretory proteins (ovomucoid, conalbumin, and lysozyme) synthesized by chicken oviduct cells.

The finding of N-terminal methionine in all precursors investigated (Figure 2) suggests that these molecules represent the immediate product of mRNA translation within the cell (Schechter and Burstein, 1976a). This is strongly supported by sequence analyses of precursors synthesized in the presence of the initiator [^{35}S]Met-tRNA$^{Met}_1$ or internal [^{35}S]Met-tRNA$^{Met}_2$ as sole source of label. The results showed that the N-terminal methionine, of both κ and λ L-chain precursors, is the initiator residue (Zemell et al., 1977; Schechter et al., 1977). Also, preliminary experiments show N-terminal extra piece sequences on polysome nascent L-chains of intact myeloma cells (Zemell, Burstein and Schechter, unpublished data). The vast majority of precursor molecules, however, should be short lived within the cell and processed to the mature protein (cleavage of the extra piece) rather quickly, because only mature L chains have been isolated from myeloma tumors or body fluids (Svasti and Milstein, 1972; Schechter, 1973).

In a variety of intact cells and cell lysates, the initiator methionine was found to be short lived: it is cleaved within a few minutes after incorporation (Housman et al., 1970) or when the nascent chains attain lengths of 15–30 residues (Jackson and Hunter, 1970). In the wheat germ cell-free system the initiator methionine was quantitatively retained (>90%) at the N-termini of the M-321, M-63, and M-41 precursors (mol wt about 26 000) for at least 4 h (Burstein and Schechter, 1976; Zemell et al., 1977), while in the three λ precursors (Burstein and Schechter, 1977b; Zemell et al., 1977) and in hemoglobin (programmed by hemoglobin mRNA in the same cell-free extract, Burstein et al., 1976b) the initiator methionine was efficiently removed (<10% retention). Cleavage of the initiator methionine may be dependent on both the overall conformation of the chain and the nature of residue X in the Met-X sequence. The data show that residue X is Asp or Arg in the two M-41 precursors, Glu in M-321 and M-63 precursors; in the λ precursors X is Ala (Figure 3) and in hemoglobin it is Val. These findings suggest that when X is a charged residue the Met-X bond is rendered resistant to cleavage by the relevant peptidase.

The ancient evolutionary origin of the N-terminal extra piece has been previously indicated (Schechter et al., 1977) mainly because it is present in the precursors of L and H chains which diverged about 400 million years ago (Dayhoff, 1972), and expression of the Ig gene is contingent on the nucleotide sequence coding for the extra piece that contains the initiator methionine codon (Schechter, 1975; Burstein et al., 1977a). One can speculate that the residues common for all extra pieces (enclosed residues in Figure 3) are from the ancestral gene. These have been preserved through evolution probably because they are functionally significant, e.g., the methionine (when located at the N-terminus) as the initiator residue required for mRNA translation and also the two leucines and the glycine may be relevant for cleavage of the extra piece (see below). These residues have been recently found at the same positions in the extra piece of L chain precursor of rat (manuscript in preparation).

One of the structural features of the extra pieces is the frequent repetition of a few amino acids. Repetition of the Leu-Leu-Leu-Trp-Val sequence in the M-321, M-63, and C_κ-region extra pieces could have originated from duplication of a short DNA segment in the structural gene (Schechter and Burstein, 1976b). This mechanism, however, cannot explain other cases where the residues are randomly distributed in the extra piece: three threonines in M-321 and M-63; three phenylalanines in M-41; five leucines and five serines in the three λ-type extra pieces (see Figure 3). The origin and functional significance of this redundancy is not known; in part it may be due to a strong selective pressure for noncharged amino acid residues. In this connection, we mention that the hydrophobic hexapeptide (Phe-Leu-Leu-Leu-Leu-Phe) present in the M-41 extra piece was also found in the extra piece of preproalbumin (Strauss et al., 1977).

Several lines of evidence indicate that cleavage of the extra piece occurs in conjunction with the endoplasmic membranes; however, attempts to obtain mature L chain by the in vitro incubation of precursors with microsomes or endoplasmic membranes have not been successful (Milstein et al., 1972; Blobel and Dobberstein, 1975). In the intact precursor the hydrophobic extra piece probably interacts with the hydrophobic core of the protein; i.e., the extra piece may be buried inside the protein and it is not accessible to the cleaving enzyme(s). The susceptible substrate may be the shorter incomplete nascent chains where peptide folding is limited, and the extra piece is sterically available for enzymatic attack. The relevant enzyme(s) have not yet been identified, and it is not known whether removal of the extra piece involves one or more cleavage steps. However, structural features of the substrates may indicate their specificity. The sequence variability of the extra piece, including the three positions adjacent to the mature V region (Thr-Arg-Cys, Ala-Ile-Ser, Etc., Figure 3), does not point to a single endoproteolytic event. However, it is seen that all extra pieces have in common one glycine and two leucine residues located, respectively, at position 4 and 9–10, counting from the junction with the mature V region (Figure 3). A unique role was assigned to glycine residues in the structure of the V region because invariant glycines are present in both κ and λ V regions and because of its special steric properties (Wu and Kabat, 1970). Crystallographic studies showed that invariant glycines occur at bends of polypeptide loops of the "basic immunoglobulin fold" (Poljak, 1975). In analogy, the invariant glycine may indicate some common three-dimensional feature of the extra pieces, in which the two invariant leucines may occupy a distinct position in relation to the junction of the extra piece with the mature V region. It is thus possible to speculate on a common folding that may serve as a recognition site for the cleaving enzyme. The invariant glycine (and leucines) may function as a major specificity residue(s) orienting the bound enzyme in a manner that cleavage occurs three residues next to the glycine (in direction of the carboxy terminus). Data of the C_κ precursor support this recognition pattern. The extra piece of the C_κ precursor is 17 residues long with Gly^{17} at the end linked to residue Ala^{109} of the mature C_κ region. The mature C_κ-region fragment recovered from the cell starts at residue Ala^{112} (Kuehl and Scharff, 1974). That is, here again cleavage occurs three residues next to the glycine, causing the removal of three residues of the mature protein (Ala^{109}-Asp^{110}-Ala^{111}; see Figure 2). Sequence alignments show that Gly^{17} of the C_κ extra piece is homologous to the "specific glycine" of the extra pieces linked to the V region and the homology between their pairs of leucines (Figure 3). In this connection, it is worthy to mention that the binding sites of proteolytic enzymes may be quite large (e.g., the active site of papain recognizes simultaneously seven residues in the substrate) and the major specificity region may be removed from the point of cleavage (Berger and Schechter, 1970). Further experiments and sequence data are required to assess the validity of the proposed recognition pattern in the Ig precursor system. For non-Ig precursors it is evident that some details of this recognition pattern are not applicable, because the position homologous to the specific glycine (Figure 3) is occupied by other residues: e.g., alanine in prelysozyme and preovomucoid (Thibodeau et al., 1977); arginine in preproparathyroid hormone (Kemper et al., 1976). It may well be, however, that the recognition site of all precursors is governed by three-dimensional folding of the extra pieces. This folding may have some common conformational features, since all extra pieces are highly enriched with hydrophobic residues.

TABLE II: Hydrophobic Residues in Immunoglobulin Light-Chain Precursors, Membrane Bound Proteins, and Secretory Proteins.

	Residues: hydrophobic/total (%)	
Protein[b]	Extra piece[a]	Mature L chain[b]
M-41 κ precursor (a)	16/22 (73)	97/213 (46)
M-41 κ precursor (b)	15/20 (75)	97/213 (46)
M-321 κ precursor	15/20 (75)	107/218 (49)
M-63 κ precursor	15/20 (75)	107/218 (49)
C_κ precursor	14/17 (82)	50/106 (47)
M-104E λ_1 precursor	13/19 (69)	118/215 (55)
RPC-20 λ_1 precursor	13/19 (69)	118/215 (55)
M-315 λ_2 precursor	13/19 (69)	118/214 (55)
	Hydrophobic domain	Exposed portion
Glycophorin	17/23 (74)	60/124 (48)
Cytochrome b_5	29/40 (72)	40/97 (41)
		Secretory proteins
Lactalbumin (bovine)		60/123 (49)
Trypsinogen (bovine)		110/229 (48)

[a] Numbers for the extra pieces are calculated from sequence data given in Figure 3. [b] Data for the other proteins are from published sequences: mature L chains (Kabat et al., 1976), glycophorin (Segrest et al., 1972, 1973), cytochrome b_5 (Spatz and Strittmatter, 1971), lactalbumin, and trypsinogen (Dayhoff, 1972).

The hydrophobic nature of the extra piece indicated from earlier partial sequence data (Schechter, 1973) was confirmed when the complete structures were determined (Table II). The clustering of hydrophobic residues in a distinct region (i.e., the extra piece) of the precursor is reminiscent of the molecular topography of membrane proteins that contain a hydrophobic domain embedded in the membrane. The complete sequences of the precursors further corroborate this resemblance. No charged residues occur in the hydrophobic domain of glycophorin (Segrest et al., 1972, 1973). Similarly, the entire extra pieces of the three λ precursors and stretches of 16 amino acids in the extra pieces of M-41 (Ala^3-Thr^{18}), M-321, and M-63 (Thr^5-Gly^{20}) κ precursors are devoid of any charged residues (Figure 3). These structural similarities suggest that the role of the hydrophobic extra piece is to favor interaction of the precursor with cell membranes. We proposed two possible targets for interaction: (1) the endoplasmic membranes, where the extra piece is cleaved from the precursor to yield mature protein destined for secretion; (2) the cell-surface membrane, where the extra piece helps anchoring the precursor to serve as the antigen-recognizing receptor (Schechter and Burstein, 1976b). In this process, the hydrophobic extra piece may play

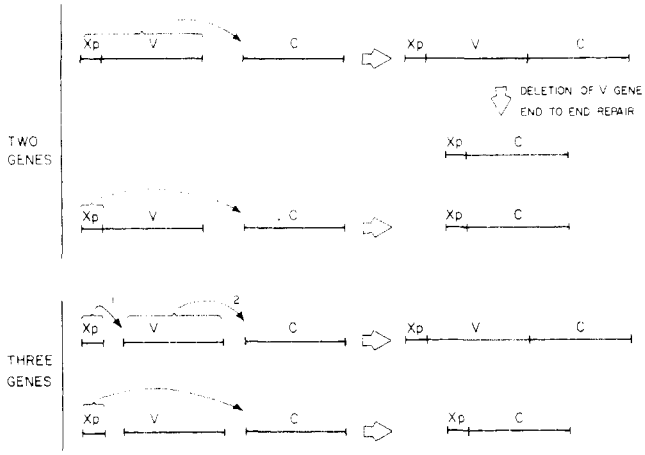

FIGURE 4: Schematic illustration of the DNA segments coding for an Ig L chain according to the two- and three-gene models. Xp, V, and C represent DNA segments coding for the N-terminal extra piece, mature V region and mature C region, respectively. See text for explanations.

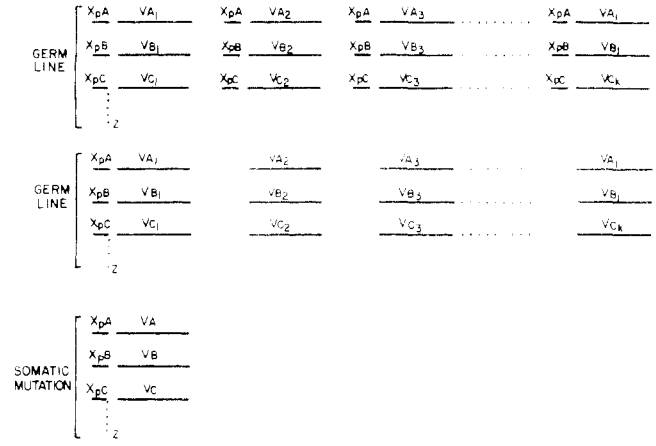

FIGURE 5: Schematic illustration of possible organization of the V region genome according to the germ-line and somatic-mutation theories, assuming that the Xp DNA and mature V gene are not covalently linked. X_P and V represent DNA segments coding for the N-terminal extra piece and mature V region, respectively; subgroups are indicated by small capital letters (A, B, C, . . . Z); members of the same subgroup are indicated by numbers and lower case letters (1, 2, 3, . . . i, j, k).

an important role in translocation of the precursor from the site of synthesis (presumably free polysomes) to the surface membrane, it may promote attachment of the precursor to the plasma membrane, and subsequently the extra piece may be either cleaved or retained. To test the latter stage, we have isolated L chain from purified plasma membranes of M-321 myeloma and subjected it to sequence analyses. The results established that the mature L chain, and not the precursor, is present in the plasma membrane of the myeloma tumor cell (Wolf et al., 1977). Using the same experimental approach we found mature L chains on the surface membrane of spleen lymphocytes (manuscript in preparation). A similar situation has been reported for the membrane lipoprotein of *Escherichia coli*. The protein is synthesized as a prolipoprotein with a hydrophobic N-terminal extra piece (20 residues long), but in the outer membrane it is present as the mature lipoprotein without extra piece (Inouye et al., 1977).

In addition to interaction with membranes, other possible functions of the precursor can be considered. (1) The enzyme(s) cleaving the extra piece may be involved in the regulation of mature protein secretion (Burstein and Schechter, 1976). (2) The metabolism of the extra piece is not yet known. If, however, it is cleaved by one endoproteolytic step, a hydrophobic peptide (about 20 residues long) would be liberated. Such a peptide may react with various components of the protein synthetic machinery and thereby it may be involved in the regulation of protein biosynthesis. (3) Although mature H and L chains recombine to generate tetrameric 7S Ig molecules, it might be that the hydrophobic extra pieces would facilitate this process by serving as a nucleation center for chain association. (4) The DNA segment coding for the extra piece may be involved in the regulation of Ig gene transcription (see below).

The identity of the extra piece of the C_κ precursor with the extra pieces linked to the V regions of M-321 and M-63 L-chain precursors (Figures 2 and 3) suggests translocation of the extra piece from the V region to the C region. This finding can be formally explained by the two genes-one Ig chain hypothesis (Dreyer and Bennet, 1965), but it also raises the possibility that three genes may control the synthesis of one Ig chain (Burstein et al., 1977a). According to the two-genes hypothesis, the DNA coding for the extra piece (Xp DNA) should be a constitutive part of the V gene. The C_κ mRNA could have originated from translocation of the extended V gene (that contains the Xp DNA) to the C gene, deletion of the entire segment of the mature V gene, and "end to end" repair of the remaining Xp DNA to the C gene. Another possibility is that only the Xp DNA portion of the extended V gene was translocated to the C gene, and that this event is sufficient to permit transcription of the C gene.

The N-terminal extra pieces exhibit sequence variability, and in this respect they might be considered as part of the V region. However, this does not necessarily mean that in the genome the Xp DNA is covalently linked to the V gene at all times. Originally, the Xp DNA may be separated from the mature V gene, similar to the DNA of hypervariable regions presumed to be stored in episomes (Wu and Kabat, 1970). This has raised the intriguing speculation that in addition to the mature V and C genes the Xp DNA may represent a third distinct gene, designated Xp gene (Burstein et al., 1977a). The presumed Xp gene may be involved in the regulation of gene transcription. Normally the Xp gene is translocated to the mature V-gene, the spliced genes are then translocated to the C gene to generate a DNA segment (Xp-V-C) from which the whole L-chain mRNA is transcribed. When the Xp gene is accidentally attached to the C gene it leads to its transcription to provide the C-region mRNA. The two- and three-gene models, and the rearrangements in the genome discussed above, are illustrated schematically in Figure 4.

The three-gene model (Figure 4) is consistent with recent studies on the origin of some viral mRNA molecules. It was found that the 5′ end of the mRNAs coding for the coat proteins of adenovirus 2 (Berget et al., 1977) and SV40 (Aloni et al., 1977) is derived from noncontiguous DNA segments of the viral genome. Recently, Tonegawa et al. (1977) have reported on the successful cloning in bacteria of a DNA fragment from mouse embryo that contains the λ_2-type V gene. According to the three-gene model it is predicted that the Xp DNA and the mature V gene are not covalently linked in the noncommitted cell. It is anticipated that the validity of this prediction will be soon subjected to experimental evaluation when the nucleotide sequence of the embryonic λ_2-type V gene will be determined. If it turns out that the Xp DNA and mature V gene are not covalently linked, then it would be necessary to reconsider the organization of the V-region genome in terms of the germ-line theory. Two major possibilities are presented in Figure 5: every mature V gene has an Xp DNA; i.e., the number of Xp DNA

copies is large and equal to the number of all possible regions (Figure 5, top); alternatively, the mature V genes of the same subgroup have only one Xp DNA; i.e., the number of Xp DNA copies is significantly smaller and equal to the number of subgroups (Figure 5, middle). According to the somatic mutation theory, each subgroup is presented by one mature V gene associated with its Xp DNA (Figure 5, bottom).

The C_κ region is commonly assigned from Arg^{108} to Cys^{214} (Svasti and Milstein, 1972; McKean et al., 1973). The finding that in the C_κ precursor the extra piece is linked to residue Ala^{109} suggests that the C_κ region starts at residue Ala^{109}, and residue Arg^{108} may belong to the V region.

The C_κ extra piece matches the sequence of M-321 and M-63 extra pieces, but it is shorter by three residues and ends at the invariant Gly^{17} residue (Figure 3). This suggests that the extra piece may extend from the initiator methionine to the invariant glycine, and the three carboxy-terminal residues found in the other extra pieces may be an integral part of the mature V region which is cleaved in the maturation process (dashed line in Figure 2).

The MPC-11(L) clone synthesizes both the mature MPC-11 L chain and the C_κ region polypeptide fragment (Kuehl and Scharff, 1974). Preliminary sequence data of the precursor of the MPC-11 L chain (programmed by mRNA isolated from this clone) show that it contains an N-terminal extra piece (29 residues long) with an amino acid sequence different from that of the extra piece of the C_κ precursor (unpublished data). Assuming joining of the information for the V and C regions at the DNA level (Hozumi and Tonegawa, 1976) and that the MPC-11(L) clone is homogeneous, then these findings imply that two C_κ genes and two different κ-type V genes are expressed in the same cell.

Note Added in Proof

Since this paper was submitted for publication, Tonegawa and collaborators have determined the nucleotide sequence of an embryonic Ig gene coding for the λ_2 V region (S. Tonegawa, A. Maxam, O. Bernard, and W. Gilbert, personal communication). By matching with the amino acid sequence of the M-315 λ_2 L-chain precursor, it is seen that the DNA coding for the N-terminal extra piece (Xp-DNA) and the DNA of the mature V gene are separated by 93 nontranslatable nucleotides. This finding confirms our prediction that the Xp-DNA may represent a third distinct gene (Xp gene) which is not directly linked to the mature V gene at all times. Considering the precise size of the Xp gene, we predicted that it may extend from the initiator Met codon to the invariant Gly^{17} codon. This is in agreement with the nucleotide sequence data showing that the codons of the three amino acids beyond Gly^{17} are directly linked to the mature V gene.

We thank Dr. S. Tonegawa and his collaborators for providing us with the nucleotide sequence prior to publication.

Acknowledgment

The authors thank Miss Frida Kantor, Mrs. Etty Ziv, and Mrs. Esther Majerovich for excellent assistance.

References

Aloni, Y., Dhar, R., Laub, O., Hurwitz, M., and Khoury, G. (1977), *Proc. Natl. Acad. Sci. U.S.A., 74,* 3686.

Appella, E. (1971), *Proc. Natl. Acad. Sci. U.S.A. 68,* 590.

Berger, A., and Schechter, I. (1970), *Philos. Trans. R. Soc. London, Ser. B, 257,* 249.

Berget, S. M., Moore, C., and Sharp, P. A. (1977), *Proc. Natl. Acad. Sci. U.S.A. 74,* 3171.

Blobel, G., and Dobberstein, B. (1975), *J. Cell Biol. 67,* 835.

Burstein, Y., Kantor, F., and Schechter, I. (1976a), *Proc. Natl. Acad. Sci. U.S.A. 73,* 2604.

Burstein, Y., and Schechter, I. (1976), *Biochem. J. 157,* 145.

Burstein, Y., and Schechter, I. (1977a), *Proc. Natl. Acad. Sci. U.S.A. 74,* 716.

Burstein, Y., and Schechter, I. (1977b), *Biochem. J. 165,* 347.

Burstein, Y., Zemell, R., Kantor, F., and Schechter, I. (1977a), *Proc. Natl. Acad. Sci. U.S.A. 74,* 3157.

Burstein, Y., Zemell, R., and Schechter, I. (1976b), in *Abstracts,* 10th International Congress of Biochemistry, Hamburg, July 1976, p 03-6-144.

Burstein, Y., Ziv, E., and Schechter, I. (1977b), *Isr. J. Med. Sci., 13,* 1048.

Dayhoff, M. O. (1972), Atlas of Protein Sequence and Structure, Vol. 5, Washington D.C., National Biomedical Research Foundation.

Dreyer, W. J. and Bennet, J. C. (1965), *Proc. Natl. Acad. Sci. U.S.A., 54,* 864.

Dugan-Schulenburg, E., Bradshaw, R. A., Simms, E. S., and Eisen, H. N. (1973), *Biochemistry 12,* 5400.

Housman, D., Jacobs-Lorena, M., Rajbhandary, U. L. and Lodish, H. F. (1970), *Nature (London) 227,* 913.

Hozumi, N., and Tonegawa, S. (1976), *Proc. Natl. Acad. Sci. U.S.A. 73,* 3628.

Inouye, S., Wang, S., Sekizawa, J., Halegoua, S., and Inouye, M. (1977), *Proc. Natl. Acad. Sci. U.S.A. 74,* 1004.

Jackson, R., and Hunter, T. (1970), *Nature (London) 227,* 672.

Kabat, E. A., Wu, T. T., and Bilofsky, H. (1976), Variable Regions of Immunoglobulin Chains, Cambridge, Mass., Medical Computer Systems.

Kemper, B., Habener, J. F., Ernst, M. D., Potts, J. T., and Rich, A. (1976), *Biochemistry 15,* 15.

Kuehl, W. M., and Scharff, M. D. (1974), *J. Mol. Biol. 89,* 409.

McKean, D., Potter, M., and Hood, L. (1973), *Biochemistry 12,* 760.

Milstein, C., Brownlee, G. G., Harrison, T. M., and Mathews, M. B. (1972), *Nature (London), New Biol. 239,* 117.

Palmiter, R. D., Thibodeau, S. N., Cagnon, J., and Walsh, K. A. (1977), *FEBS Symp.,* abstracts A6-2,L3,4.

Poljak, R. J. (1975), *Nature (London) 256,* 373.

Roberts, B. E., and Paterson, B. M. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 2330.

Schechter, I. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 2256.

Schechter, I. (1974), *Biochemistry 13,* 1875.

Schechter, I. (1975), *Biochem. Biophys. Res. Commun. 67,* 228.

Schechter, I., and Burstein, Y. (1976a), *Biochem. J. 153,* 543.

Schechter, I., and Burstein, Y. (1976b), *Proc. Natl. Acad. Sci. U.S.A. 73,* 3273.

Schechter, I., and Burstein, Y. (1976c), *Biochem. Biophys. Res. Commun. 68,* 489.

Schechter, I., Burstein, Y. and Zemell, R. (1977), *Immunol. Rev. 36,* 3.

Schechter, I., McKean, D. J. Guyer, R., and Terry, W. (1975), *Science 188,* 160.

Segrest, J. P., Jackson, R. L., Marchesi, V. T., Guyer, R. B., and Terry, W. (1972), *Biochem. Biophys. Res. Commun. 49,* 964.

Segrest, J. P., Kahane, I., Jackson, R. L., and Marchesi, V. T. (1973), *Arch. Biochem. Biophys. 155*, 167.
Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G., and Ballantyne, D. L. (1971), *Biochemistry 10*, 4912.
Spatz, L., and Strittmatter, P. (1971), *Proc. Natl. Acad. Sci. U.S.A. 68*, 1042.
Strauss, A. W., Donohue, A. M., Bennet, C. D., Rodkey, J. A., and Alberts, A. W. (1977), *Proc. Natl. Acad. Sci. U.S.A. 74*, 1358.
Svasti, J., and Milstein, C. (1972), *Biochem. J. 126*, 837.
Thibodeau, S. N., Gagnon, J., and Palmiter, R. (1977), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 36*, 656.
Tonegawa, S., Brack, C., Hozumi, N., and Schuller, R. (1977), *Proc. Natl. Acad. Sci. U.S.A. 74*, 3518.
Weigert, M. G., Cesari, I. M., Yonkovich, S. J., and Cohn, M. (1970), *Nature (London) 228*, 1045.
Wolf, O., Zemell, R., Burstein, Y., and Schechter, I. (1977), *Biochem. Biophys. Res. Commun., 78*, 1383.
Wu, T. T., and Kabat, E. A. (1970), *J. Exp. Med. 132*, 211.
Zemell, R., Burstein, Y., and Schechter, I. (1977), *Isr. J. Med. Sci. 13*, 935.

Phase Behavior of Synthetic Phosphatidylglycerols and Binary Mixtures with Phosphatidylcholines in the Presence and Absence of Calcium Ions†

Earl J. Findlay and Peter G. Barton*

ABSTRACT: Using differential thermal analysis, scanning calorimetry and light scattering, transition temperatures and enthalpy data for the gel to liquid crystalline phase transitions of five synthetic phosphatidylglycerol sodium salts (PG-Na^+) were measured. The values obtained were almost identical with literature values for the corresponding phosphatidylcholines (PC). However, transition temperatures for the fully protonated forms of the saturated phosphatidylglycerols (PG-H^+) were approximately 20 °C higher. For binary mixtures of PG-Na^+ and PC in which the acyl chains of the two species were identical, the width of the thermal transition for the phase change was not appreciably greater than that observed with either of the two components alone. In contrast, mixing of PG-Na^+ and PC with different chain lengths increased the transition width. In the presence of Ca^{2+}, narrow transitions were also observed with mixtures of PG and PC when the chain length of the PG-Ca^{2+} was equal to or two carbons shorter than the PC but the transition width was clearly increased when the chain length of the PG-Ca^{2+} was two carbons longer than the PC. Mixing lipids with greater differences in chain length or mixing saturated lipids with unsaturated lipids in the presence of Ca^{2+} produced two minima in the thermograms, clearly indicative of phase separation. In sum, these results provide evidence for a high degree of miscibility of the phosphoglycerol and phosphocholine head groups, either in the presence or absence of Ca^{2+}, such that the characteristic phase behavior of each mixture is determined primarily by differences in the hydrocarbon chain structure.

Recently, the improved availability of synthetic phosphatidylglycerols (PG) has stimulated examination of the physical properties of the hydrated lipids (Tocanne et al., 1974; Verkleij et al., 1974; Ververgaert et al., 1975; Van Dijck et al., 1975; Jacobson & Papahadjopoulos, 1975; Papahadjopoulos et al., 1973, 1976; Jackson et al., 1974). These studies have utilized mainly differential scanning calorimetry (DSC)[1] and freeze fracture electron microscopy. Complex effects of divalent metal ions on the phase structures formed by dilauroylphosphatidylglycerol (DLPG), dimyristoylphosphatidylglycerol (DMPG), and dipalmitoylphosphatidylglycerol (DPPG) have been reported. The range of existence of the phase structures varied with the metal ion and acyl chain structure of the lipid.

Since biological membrane systems invariably contain mixtures of acidic and zwitterionic phospholipids, it was thought to be useful to characterize the behavior of binary mixtures of phosphatidylglycerols and phosphatidylcholines (PC) in the presence and absence of Ca^{2+} ions. Such mixtures have not previously been systematically investigated as a function of acyl chain length and polar head group composition. We have found that the thermotropic mesomorphism ranges from nearly ideal mixing behavior of the acidic and zwitterionic components at all concentrations examined to distinct phase separations with limited mutual miscibility of the two lipid components. The complex metastable behavior characteristic of the pure PG-Ca^{2+} or PG-Mg^{2+} is not observed with the mixtures.

Materials and Methods

Inorganic chemicals and reagents used in preparation of buffers were obtained from J. T. Baker Chemical Co., or Fisher Scientific Co., and were ACS reagent grade or better. All or-

† From the Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada. *Received October 17, 1977.* This work was supported by the Medical Research Council of Canada (Grant MT 2402) and the Alberta Heart Foundation.

[1] Abbreviations used: DLPG, dilauroylphosphatidylglycerol (1,2-didodecanoyl-*sn*-glycero-3-phosphoglycerol); DMPG, dimyristoylphosphatidylglycerol; DPPG, dipalmitoylphosphatidylglycerol; DSPG, distearoylphosphatidylglycerol; DOPG, dioleoylphosphatidylglycerol; DLPC, dilauroylphosphatidylcholine; DMPC, dimyristoylphosphatidylcholine; DPPC, dipalmitoylphosphatidylcholine; DSPC, distearoylphosphatidylcholine; DOPC, dioleoylphosphatidylcholine; DOPE, dioleoylphosphatidylethanolamine; DPPS, dipalmitoylphosphatidylserine; DTA, differential thermal analysis; DSC, differential scanning calorimetry.

0006-2960/78/0417-2400$01.00/0