Brief Definitive Report

NUCLEIC ACID AND PROTEIN SEQUENCES OF PHOSPHOCHOLINE-BINDING LIGHT CHAINS*

By S.-P. KWAN, S. RUDIKOFF, J. G. SEIDMAN, P. LEDER, and M. D. SCHARFF

From the Department of Cell Biology, Albert Einstein College of Medicine, Bronx, New York 10461; and the Laboratory of Cell Biology, National Cancer Institute, and the Laboratory of Molecular Genetics, National Institute of Child Health and Development, National Institutes of Health, Bethesda, Maryland 20205

Recent studies on the control of the immune response have revealed the importance of idiotypes and their recognition by both B and T cells. Most of these studies have focused on a few well-studied idiotypes, one of which is T15. This is the major idiotype elicited when BALB/c mice are immunized with the hapten phosphocholine (PC), either in its natural form as part of bacterial or parasite cell walls or artificially attached to protein and other carriers (1). Other strains of mice produce lesser amounts of T15 idiotype. The heavy-chain variable region of the T15 myeloma, an IgA κ protein that is the prototype for these antibodies, has been completely sequenced (2, 3). The amino acid sequence of the variable region of the S107 heavy chain, which is also an IgA κ myeloma, is identical to T15 (2, 4). The heavy chain of HOPC8, a third PC-binding IgA κ myeloma bearing the T15 idiotype, has also been completely sequenced and differs by only one amino acid in its third hypervariable region from T15 and S107 (2, 4). Only partial sequences through residues 30–40 of the light chains of these proteins have been reported, and the peptide containing the third hypervariable region has been especially difficult to purify, precluding the immediate completion of the protein sequences (2). It has become necessary to determine the sequence of the third hypervariable region and J segment of the S107 light chain to study hybridomas producing PC-binding antibodies and somatic variants of S107 with changes in antigen binding (5). We have therefore cloned and sequenced the rearranged light-chain variable region gene that codes for the secreted S107 light chain. Here we report the base sequence that codes for amino acid residues 1 through 108 and compare it to the amino acid sequence of the T15 light chain, which has been determined through amino acid residue 88.

Materials and Methods

Cloning of the Light-Chain Gene. The S107 cell line was obtained from the Salk cell bank (The Salk Institute, La Jolla, Calif.). S107 DNA was isolated from cultured cells and the genomic DNA was digested with Eco RI restriction enzyme. Fragments containing the recombined κ gene DNA were identified by Southern blot analysis of S107 compared with embryo DNA with a κ-chain constant-region probe (6). Fragments were then further enriched by preparative agarose gel electrophoresis as described (7). Two distinct fragments of ~18 kilobase (kb) contained rearranged κ-chain DNA. Both were analyzed, but only one, named S107A, is

* Supported by grants AI-10702 and AI-15231 from the National Institutes of Health, grant PLM75-13609 from the National Science Foundation, and grant AP-317 from the American Cancer Society. S.-P Kwan was supported by a fellowship from the Cancer Research Institute, Inc.

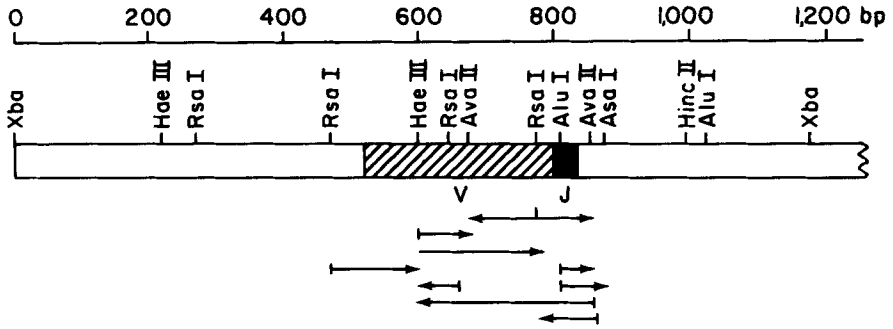

FIG. 1. Physical map of the rearranged S107A light chain DNA. The arrows denote the DNA sequencing strategy. The coding region for the V region is stippled; the J region is solid.

reported here. The DNA from this band was pooled, ligated into the Charon 4A vector, packaged, and transfected into *Escherichia coli* (8). Hybrid phages containing κ-chain DNA were identified by *in situ* hybridization of nitrocellulose filter blots (9) using the C region-specific probe (6) and grown preparatively on *E. coli* LE392. For sequence determination, the S107A clone was restricted with Bam H1 generating a 7-kb fragment which was shown to contain both variable and constant sequences by R loop analysis with S107 light chain messenger RNA. This was subcloned in pBR322 (10). The plasmid was grown and the 7-kb insert was purified as described by Seidman et al. (6). These experiments were carried out in compliance with the National Institutes of Health guidelines on recombinant DNA research.

DNA Sequencing. S107A DNA was digested with Xba I and a 1.2-kb-long fragment was isolated by electrophoresis in 5% polyacrylamide gel. The 1.2-kb fragment was digested with different restriction enzymes as shown in Fig. 1, treated with bacterial alkaline phosphatase (ATP), and labeled at the 5′ end with γ^{32}P-ATP using polynucleotide kinase. These fragments were subjected to second restriction enzyme and single end-labeled DNA were isolated by polyacrylamide gel electrophoresis. Sequencing was performed according to the method of Maxam and Gilbert (11).

Protein Sequencing. The N-terminal sequences of the T15 and S107 κ chains have previously been reported (2). The amino acid sequence of T15 from Trp-35 to Phe-62 (Fig. 2) was determined from a peptide obtained by cleavage at Trp-35 with BNPS-skatole as described (12). This peptide extends from Trp-35 into the C region. The sequence from Phe-62 to Cys-88 was obtained from a tryptic peptide extending from Phe-62 to Lys-103.[1] Sequencing procedures employed have previously been described in considerable detail (12). Positions 61 (Arg) and 85 (His) were not identified in the amino acid sequencing experiments of T15, and are assigned only by nucleotide sequence of S107A.

Results and Discussion

When DNA from the S107 cell line was digested with Eco RI and analyzed for rearranged κ-chain constant-region genes, two bands of ~18 kb were identified and cloned in Charon 4A. A total of 120,000 recombinant phages were analyzed, 10 of which contained κ C-region sequences. On the basis of restriction endonuclease analysis and heteroduplex mapping (not shown), these clones represent two distinct genes. Both genes have been sequenced. One, which we have named S107B, has an unusual V-J recombination and shares only 50% of its coding sequence with the light chain secreted by S107. This will be described in detail elsewhere.[2] A second recombined gene, which we have named S107A, is the subject of the investigations

[1] Rudikoff, S., Y. Satow, E. Padlan, P. Davies, and M. Potter. 1981. Kappa chain structure from a captallized murine Fab′: role of joining segment in hapten binding. Manuscript submitted.

[2] Kwan, S.-P., E. Max, J. G. Seidman, P. Leder, and M. D. Scharff. Manuscript in preparation.

																1			
ATG	TAT	ATA	TGT	TTG	TTG	TCT	ATT	TCT	TAT	TGT	AGG	TGC	CTC	GTG	CAC	GAC	ATT	GTG	ATG
																ASP	ILE	VAL	MET
															20				
ACT	CAG	TCT	CCA	ACT	TTC	CTT	GCT	GTG	ACA	GCA	AGT	AAG	AAG	GTC	ACC	ATT	AGT	TGC	ACT
THR	GLN	SER	PRO	THR	PHE	LEU	ALA	VAL	THR	ALA	SER	LYS	LYS	VAL	THR	ILE	SER	CYS	THR
											30								
GCX	TCT	GAG	AGC	CTT	TAT	TCA	AGC	AAA	CAC	AAG	GTG	CAC	TAC	TTG	GCT	TGG	TAC	CAG	AAG
ALA	SER	GLU	SER	LEU	TYR	SER	SER	LYS	HIS	LYS	VAL	HIS	TYR	LEU	ALA	TRP	TYR	GLN	LYS
	40										50								
AAA	CCA	GAG	CAA	TCT	CCT	AAA	CTG	CTG	ATA	TAC	GGG	GCA	TCC	AAC	CGA	TAC	ATT	GGG	GTC
LYS	PRO	GLU	GLN	SER	PRO	LYS	LEU	LEU	ILE	TYR	GLY	ALA	SER	ASN	ARG	TYR	ILE	GLY	VAL
	60										70								
CCT	GAT	CGC	TTC	ACA	GGC	AGT	GGA	TCT	GGG	ACA	GAT	TTC	ACT	CTG	ACC	ATC	AGC	AGT	GTA
PRO	ASP	ARG	PHE	THR	GLY	SER	GLY	SER	GLY	THR	ASP	PHE	THR	LEU	THR	ILE	SER	SER	VAL
	80										90								
CAG	GTT	GAA	GAC	CTC	ACA	CAT	TAT	TAC	TGT	GCA	CAG	TTT	TAC	AGC	TAT	CCT	CTC	ACG	TTC
GLN	VAL	GLU	ASP	LEU	THR	HIS	TYR	TYR	CYS	ALA	GLN	PHE	TYR	SER	TYR	PRO	LEU	THR	PHE
	100																		
GGT	GCT	GGG	ACC	AAG	CTG	GAG	CTG	AAA	CGT	AAG	TAC	ACT	TTT	CTC	ATC	TTTT			
GLY	ALA	GLY	THR	LYS	LEU	GLU	LEU	LYS	ARG										

FIG. 2. Nucleotide sequence of the active S107A light chain gene and the light chain for which it codes. The numbers indicate the amino acid residue using the Kabat numbering system. The nucleotide sequence was determined as described in the text. The amino acids represent the amino acid sequence deduced from the nucleotide sequence. The solid line represents the identity of amino acid sequence of T15 determined as described in the text.

reported in the present paper. This clone codes for the light chain polypeptide employed in the secreted PC-binding S107 protein. Through the restriction analysis illustrated in Fig. 1, a 1.2-kb Xba-Xba fragment containing all of the V region was obtained and its base sequence was determined. Bases 49–171 code for an amino acid sequence that is identical to that reported for the N-terminal 35 amino acids of the light chain secreted by both the T15 and S107 tumors (2). The sequence of an additional 50 amino acids of the T15 light chain has been determined. As indicated by the continuous line in Fig. 2, this sequence is also identical with that determined from the DNA sequence.

The S107 light chain is of particular interest in that this chain, as well as the idiotype of the native molecule, is highly conserved in all inbred mouse strains (13), which suggests a strong selective value for these gene products. The genetic rearrangement producing the active κ-chain gene involves recombination of the V gene with the J1 gene segment, which is the first J found 5′ from the κ C-region gene (14). It is intriguing that all κ chains analyzed to date from PC-binding proteins employ the J1 sequence, although a number of these light chains (i.e., M603 and M167) differ considerably in amino acid sequence (2, 12).[1] Three-dimensional X-ray analysis of the PC-binding myeloma protein M603 has demonstrated that Leu 96, the first amino acid of the J segment, is a hapten-contacting residue (15).[2] Thus J1, the only J segment encoding Leu at position 96 (15), may be required for PC binding, or these V genes may, due to genetic restrictions, only pair with J1. The nucleotide structure of the S107 κ chain has thus provided important information on the third hypervariable region and J segment from this molecule as well as demonstrating that the T15

and S107 κ chains are identical for at least their N-terminal 88 amino acids. It is therefore probable that these κ chains are identical throughout this entire V region, as has been demonstrated for the heavy-chain V regions.

Summary

An 18-kilobase DNA fragment containing the sequence coding for both the variable and constant regions of the S107 mouse immunoglobulin light chain was cloned from total cellular DNA. The complete nucleotide sequence of the κ-chain variable-region gene is reported. Determination of the amino acid sequence encoded by the DNA is found to be identical to the protein sequence of the T15 light chain through residue 88. Direct sequence analysis confirmed that the J1 joining segment is used in the recombination event producing the active κ light chain gene.

We would like to thank Marion Nau for technical assistance.

Received for publication 9 February 1981.

References

1. Kohler, H. 1975. The response to phosphorylcholine: dissecting an immune response. *Transplant. Rev.* **27**:24.
2. Barstad, P., S. Rudikoff, M. Potter, M. Cohn, W. Konigsberg, and L. Hood. 1974. Imunoglobulin structure: amino terminal sequences of mouse myeloma proteins that bind phosphorylcholine. *Science (Wash. D. C.).* **183**:962.
3. Rudikoff, S., and M. Potter. 1978. Size differences among immunoglobulin heavy chains from phosphorylcholine-binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* **73**:2109.
4. Barstad, P., J. Hubert, M. Hunkapiller, A. Goetz, J. Schilling, B. Black, B. Eaton, J. Richards, M. Wigert, and L. Hood. 1978. Immunoglobulins with hapten-binding activity: structure-function correlations and genetic implications. *Eur. J. Imm.* **8**:497.
5. Cook, W., and M. D. Scharff. 1977. Antigen-binding mutants of mouse myeloma cells. *Proc. Natl. Acad. Sci. U. S. A.* **74**:5687.
6. Seidman, J. G., A. Leder, M. Nau, B. Norman, and P. Leder. 1978. Antibody diversity. *Science (Wash. D. C.).* **202**:11.
7. Polsky, F., M. H. Edgell, J. G. Seidman, and P. Leder. 1978. High capacity gel preparative electrophoresis for purification of fragments of genomic DNA. *Anal. Biochem.* **87**:397.
8. Blattner, F. R., B. G. Williams, A. E. Blechl, K. D. Thompson, H. E. Faber, L. A. Furlong, D. J. Grunwald, D. O. Kiefer, D. D. Moore, J. W. Schuman, E. L. Sheldon, and O. Smithies. 1977. Charon pages: safer derivative of bacteriophage lambda for DNA cloning. *Science (Wash. D. C.).* **196**:161.
9. Benton, W. D., and R. W. Davis. 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. *Science (Wash. D. C.).* **196**:180.
10. Bolivar, E., R. L. Rodriquez, P. J. Green, M. C. Behach, H. L. Heyneker, H. W. Boyer, J. H. Crosa, and S. Falkow. 1977. Construction and characterization of new cloning vehicles II. A multipurpose cloning system. *Gene (Amst.).* **2**:95.
11. Maxam, A., and W. Gilbert. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U. S. A.* **74**:560.
12. Rudikoff, S., and M. Potter. 1978. κ chain variable region from M167, a phosphorylcholine binding myeloma protein. *Biochemistry.* **17**:2703.
13. Claflin, J. L., and J. M. Davie. 1974. Clonal nature of the immune response to phospho-

rylcholine TV. Idiotypic uniformity of binding site associated antigenic determinants among mouse antiphosphorylcholine antibodies. *J. Exp. Med.* **140**:673.

14. Max, E., J. G. Seidman, and P. Leder. 1979. Sequences of five potential recombination sites encoded close to an immunoglobulin κ constant region gene. *Proc. Natl. Acad. Sci. U. S. A.* **76**:3450.
15. Segal, D. M., E. A. Padlan, G. H. Cohen, S. Rudikoff, M. Potter, and D. R. Davies. 1974. The three dimensional structure of a phosphorylcholine binding mouse immunoglobulin Fab and the nature of the antigen binding site. *Proc. Natl. Acad. Sci. U. S. A.* **71**:4298.