LAH 306

CLONING, SEQUENCING AND DIFFERENTIAL SPLICING OF THE LYT-2 GENE

Hiromitsu Nakauchi[1]
Garry P. Nolan
Masatoshi Tagawa
Leonard A. Herzenberg

Department of Genetics
Stanford University School of Medicine
Stanford, California

I. INTRODUCTION

Years have passed since a series of monoclonal antibodies was made against important T cell differentiation antigens. These antibodies have been employed in laboratory as well as clinical situations, resulting in a better understanding of many diseases that have immunological abnormalities as an underlying cause. Still the molecular structure and function of these antigens remains ill defined. With the advance of molecular biological techniques, it has become possible to elucidate this information via the process of cloning a gene that codes for a particular cell surface antigen then predicting protein structure by DNA sequencing. In this way various lymphocyte differentiation antigens may be examined at the DNA level. The information generated from this will facilitate understanding of the role these molecules play in the immune response as well as certain diseases in which immunological abnormalities play a role.

Among those molecules that is of interest to the immunologist is the murine lymphocyte differentiation antigen Lyt-2.

---

[1]Present address : Department of Immunology, Juntendo University School of Medicine, 2-1-1, Hongo, Bunkyo-ku, Tokyo, Japan 113

This antigen is expressed on immature T cells as well as certain subpopulations of mature T cells, namely the cytotoxic T cell (CTL) and suppressor subsets (1-3). Human T lymphocytes with analogous functions express the molecule Leu-2 (CD8) (4-6). We had postulated that these molecules are not only functionally but also structurally homologous (7). Functionally, the Lyt-2 and Leu-2 molecules seem to play an auxiliary role in antigen recognition, as monoclonal antibodies against them are known to block cytotoxicity of most CTLs that bear them (8-10). Also the expression of the CD8 molecule correlates well with recognition of class I MHC molecules in both human and murine systems (11-13).

Our research has clarified that the Lyt-2 antigenic determinant is found on 38-kDa (Lyt-2-α) and 34-kDa (Lyt-2-α') polypeptides, usually disulfide-linked to a 30-kDa Lyt-3, primarily as heteromultimers (14) on thymocytes. On peripheral T cells, only 38-kDa Lyt-2-α determinant is found linked to Lyt-3 polypeptide.

To determine the detailed structure of CD8 molecules believed to be important in T cell recognition in the context of class I molecules, we began by isolating molecular clones of cDNA and genomic DNA coding for human Leu-2 (15). This was accomplished using a subtractive cDNA approach with L-cell transfectants selected via the fluorescence-activated cell sorter (FACS) after DNA-mediated gene transfer (15-16). Subsequently, murine Lyt-2 cDNA and genomic clones were isolated (17). In this text we describe the isolation of two types of Lyt-2 cDNA and genomic clones. We also describe the predicted amino acid sequence of the 38-kDa Lyt-2-α and 34-kDa Lyt-2-α' polypeptides and the use of differential splicing as a mechanism for producing two mRNA responsible for them.

II. MATERIALS AND METHODS

A. Cloning of Lyt-2 cDNA and Genomic Gene

A BALB/c thymocyte cDNA library (C.Benoit) was screened with 1.7kb Leu-2 cDNA (15) as a probe in low stringency hybridization conditions. A confirmed Lyt-2 cDNA clone was then used to isolate full-length cDNA clones from a C57BL/6 thymocyte cDNA library (H.Gershenfeld and I.Weissman) and a genomic clone from B10.A liver genomic library (18).

B. DNA Sequence Analysis

DNA sequencing was performed by the dideoxynucleotide chain-termination technique of Sanger et al. (19) after

subcloning restriction endonuclease fragments into M13 mp18 and mp19 phage vectors. Sequences were analyzed on a VAX 11/780 computer using programs described by Staden (20-21). Alignment, translation of cDNA, and the amino acid alignment of Lyt-2 and Leu-2 sequences was accomplished using the BIONET (National Institutes of Health Grant 1U41 RR01681-01) and DB system (22). The Dayhoff Protein sequence bank was searched utilizing the program described by Lipman and Pearson (23).

C. Transfection and Flowcytometric Analysis

Thymidine kinase (TK)-defective L cells were transfected with 20 micrograms of Lyt-2 cDNA plasmid, 1 microgram of chicken TK gene and 10 micrograms of carrier human DNA by the calcium phosphate precipitation method (24). After selection in hypoxanthine/aminopterin/thymidine medium, TK positive cells were stained with monoclonal rat anti-Lyt-2 antibody (53-6.7) (25), and the Lyt-2 positive cells were sorted using a fluorescence-activated cell sorter (FACS) as described (26).

III. RESULTS

A. Cloning of Lyt-2 cDNA and Genomic Gene

By screening 152,000 plaques using 1.7kb Leu-2 cDNA as a probe under conditions of low stringency hybridization, 11 positive phage clones were isolated. After EcoR 1 digestion, four had an insert of 650 bp which hybridized strongly with RNA prepared from thymocytes and the Lyt-2 positive L-cell transfectants but not with L-cell RNA on RNA blots. We subcloned this insert into a pSP65 plasmid (pLY2C-1). Using this pLY2C-1 as a probe, we then screened a C57BL/6 thymocyte cDNA library and a B10.A liver genomic library and obtained two types of cDNAs for Lyt-2 and a genomic Lyt-2 gene.

B. Primary Protein Structure of Lyt-2

The nucleotide and predicted amino acid sequence of Lyt-2-α cDNA is presented in FIGURE 1. The mature Lyt-2-α peptide is 220 amino acids and has a leader peptide of 27 residues. Hydrophobicity analysis reveals a region of strong hydrophobicity near the COOH terminus (residues 157-193) followed by a hydrophilic sequence (residues 194-220 in α, residues 194-195 in α'). We predict these regions to be the transmembrane region and cytoplasmic tail, respectively. Lyt-2-α is distinguished from the second type cDNA (α') in that the α' form lacks 31 base pairs at the end of the predicted trans-

```
                                                                    GGAGAGCACACC

-27                          -20                                        -10
ATG GCC TCA CCG TTG ACC CGC TTT CTG TCG CTG AAC CTG CTG CTG CTG GGT GAG TCG ATT
MET Ala Ser Pro Leu Thr Arg Phe Leu Ser Leu Asn Leu Leu Leu Leu Gly Glu Ser Ile

                        -1  1                                          10
ATC CTG GGG AGT GGA GAA GCT AAG CCA CAG GCA CCC GAA CTC CGA ATC TTT CCA AAG AAA
Ile Leu Gly Ser Gly Glu Ala Lys Pro Gln Ala Pro Glu Leu Arg Ile Phe Pro Lys Lys

                        20                                         30
ATG GAC GCC GAA CTT GGT CAG AAG GTG GAC CTG GTA TGT GAA GTG TTG GGG TCC GTT TCG
MET Asp Ala Glu Leu Gly Gln Lys Val Asp Leu Val Cys Glu Val Leu Gly Ser Val Ser

                        40      ***                                50
CAA GGA TGC TCT TGG CTC TTC CAG AAC TCC AGC TCC AAA CTC CCC CAG CCC ACC TTC GTT
Gln Gly Cys Ser Trp Leu Phe Gln Asn Ser Ser Ser Lys Leu Pro Gln Pro Thr Phe Val

                        60                                         ***
GTC TAT ATG GCT TCA TCC CAC AAC AAG ATA ACG TGG GAC GAG AAG CTG AAT TCG TCG AAA
Val Tyr MET Ala Ser Ser His Asn Lys Ile Thr Trp Asp Glu Lys Leu Asn Ser Ser Lys

                        80                                         90
CTG TTT TCT GCC ATG AGG GAC ACG AAT AAT AAG TAC GTT CTC ACC CTG AAC AAG TTC AGC
Leu Phe Ser Ala MET Arg Asp Thr Asn Asn Lys Tyr Val Leu Thr Leu Asn Lys Phe Ser

                        100                                        110
AAG GAA AAC GAA GGC TAC TAT TTC TGC TCA GTC ATC AGC AAC TCG GTG ATG TAC TTC AGT
Lys Glu Asn Glu Gly Tyr Tyr Phe Cys Ser Val Ile Ser Asn Ser Val MET Tyr Phe Ser

                        120         ***                            130
TCT GTC GTG CCA GTC CTT CAG AAA GTG AAC TCT ACT ACT ACC AAG CCA GTG CTG CGA ACT
Ser Val Val Pro Val Leu Gln Lys Val Asn Ser Thr Thr Thr Lys Pro Val Leu Arg Thr

                        140                                        150
CCC TCA CCT GTG CAC CCT ACC GGG ACA TCT CAG CCC CAG AGA CCA GAA GAT TGT CGG CCC
Pro Ser Pro Val His Pro Thr Gly Thr Ser Gln Pro Gln Arg Pro Glu Asp Cys Arg Pro

                        160                                        170
CGT GGC TCA GTG AAG GGG ACC GGA TTG GAC TTC GCC TGT GAT ATT TAC ATC TGG GCA CCC
Arg Gly Ser Val Lys Gly Thr Gly Leu Asp Phe Ala Cys Asp Ile Tyr Ile Trp Ala Pro

                        180                                        190
TTG GCC GGA ATC TGC GTG GCC CTT CTG CTG TCC TTG ATC ATC ACT CTC ATC TGC TAC CAC
Leu Ala Gly Ile Cys Val Ala Leu Leu Leu Ser Leu Ile Ile Thr Leu Ile Cys Tyr His

                        200                                        210
AGG AGC CGA AAG CGT GTT TGC AAA TGT CCC AGG CCG CTA GTC AGA CAG GAA GGC AAG CCC
Arg Ser Arg Lys Arg Val Cys Lys Cys Pro Arg Pro Leu Val Arg Gln Glu Gly Lys Pro

                        220
AGA CCT TCA GAG AAA ATT GTG TAA AATGGCACCGCCAGGAAGCTACAACTACTACATGACTTCAGAGATCT
Arg Pro Ser Glu Lys Ile Val
```

FIGURE 1. Nucleotide and amino acid sequence derived from thymocyte cDNA clones encoding Lyt-2-α polypeptide. Transmembrane region is underlined. Cysteine residues believed to participate in intrachain disulfide bonding are marked by arrows. Possible N-linked glycosylation sites are marked by asterisks. The 27 amino acid leader peptide is marked as -27 to -1. Mature peptide is marked from 1 to 220.

membrane domain. This effects a shift in the reading frame for α' cDNA, so that the peptide encoded by α' cDNA has only 2 intracytoplasmic residues varsus 27 residues for α cDNA. The predicted molecular weights of the two molecules are 24.7 kDa for α and 21.7 kDa for α'.

C. Lyt-2 Is a Member of the Immunoglobulin Gene Superfamily

A computer search of the Dayhoff protein sequence bank for

sequences similar to Lyt-2 indicates that the N-terminal domain of this molecule is homologous to human and mouse kappa light chain variable regions and heavy chain V regions. Clear but lesser homology was observed with mouse Thy-1, lambda light chain V regions as well as the V regions of T cell receptor alpha, beta and gamma chains.

Alignment of the most significant homologies(FIGURE 2.) suggests that Lyt-2 has an Ig V-like domain. Residues regarded as important for Ig domain structure and function such as

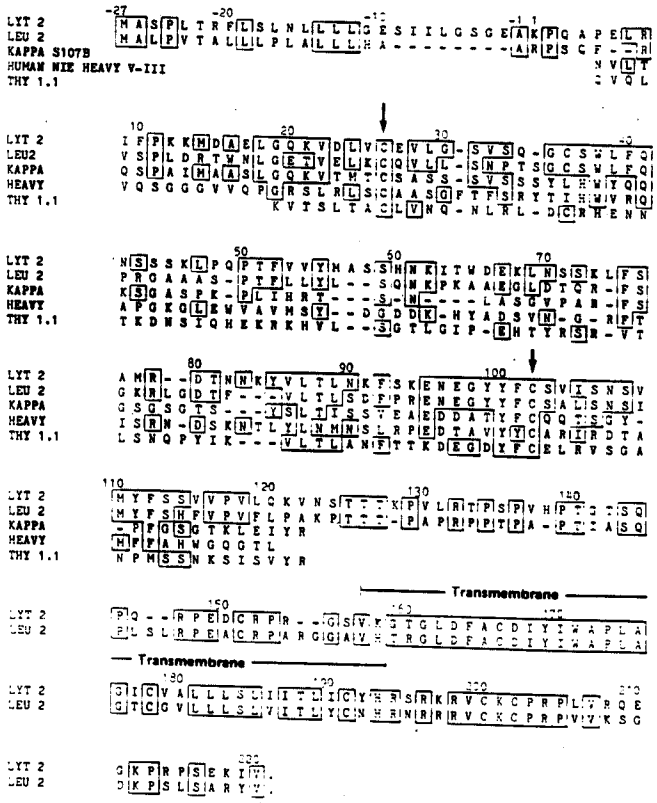

FIGURE 2. Lyt-2 homology to other immunologically relevant proteins. Using the BIONET programs IFIND and ALIGN and the protein homology search program of Lipman and Pearson, the deduced Lyt-2 protein sequence was aligned against Leu-2 and the most significantly homologous V region sequence of mouse kappa light chain, human heavy chain Ig, and the V-like region of Thy1.1. Homologies of these proteins to Lyt-2 are boxed.

the cysteines that form intrachain disulfide bonds, as well as an invariant tryptophan are also found appropriately located, after alignment, in Lyt-2 and Leu-2.

D. Structure of the Lyt-2 Genomic Gene

Comparison of the nucleotide sequence of 5.3Kb Hind-3 cleaved genomic fragment, which includes the entire Lyt-2 gene(16,.27), with that of full length cDNA reveals the intron/exon boundaries shown in FIGURE 3. The Lyt-2 gene has five exons: a fused leader and immunoglobulin variable region like exon, a spacer or a hinge-like region exon, a transmembrane exon and two intracytoplasmic exons. Thus, each exon corresponds approximately to a particular functional domain of the molecule; a similarity to other members of the immunoglobulin superfamily (28), except that Lyt-2 does not have a separate leader exon as do all other immunoglobulin super family members sequenced thus far.

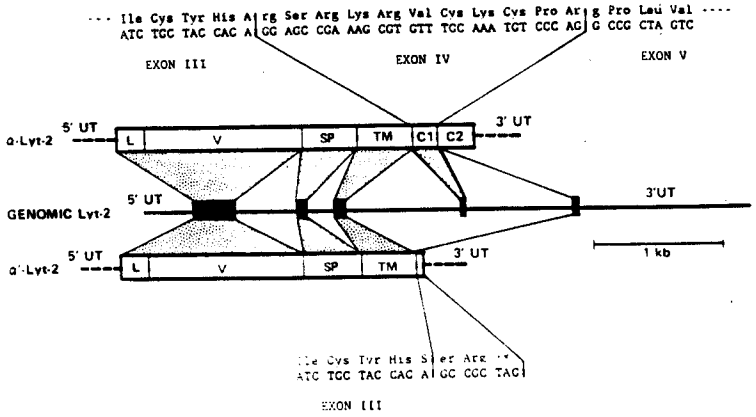

FIGURE 3. Genomic structure of the Lyt-2 gene and comparison of the sequences of α and α' cDNA. Partial sequences of the α and α' cDNA are shown. The splicing pattern, as determined from sequencing of genomic Lyt-2 (27) and the relative location of genomic exons, is shown. There are five exons (I-V) in the Lyt-2 gene; they encode leader/Ig variable-like region (L and V) (exon I), spacer region (SP)(exon II), transmembrane region (TM)(exon III), and cytoplasmic regions C1 and C2 (exons IV and V).

E. Two Types of Lyt-2 cDNA Result from Differential Exon Usage Are Responsible for 38- and 34-kDa Lyt-2 Polypeptides.

In order to confirm that α and α' cDNAs encode the 38- and 34-kDa Lyt-2 polypeptides respectively, we made expression vectors with the two types of cDNA and transfected them into mouse L cells. Immunoprecipitation of radio labeled transfectants showed that α cDNA codes for the 38-kDa peptide, α' cDNA codes for the 34-kDa peptide, and that the genomic clone codes for both 38- and 34-kda peptides.

IV. DISCUSSION

By DNA and amino acid sequence comparison, it became clear that Lyt-2 and Leu-2 are highly homologous to each other and less but clearly homologous with other members of the immunoglobulin gene superfamily (28). These molecules are all related evolutionarily and probably arose from a common ancestral gene. Since Lyt-2 and Leu-2 are related to Ig and Ig-like proteins that function in molecular recognition, and since Lyt-2/Leu-2 restrict CTL recognition of class I expressing targets, it is interesting to speculate that this recognition of class I is mediated by the Ig-like structures of Lyt-2 and Leu-2. The greater divergence in the V-like regions of these two molecules might be due to their coevolution in different major histocompatibility complex counterparts; Lyt-2 evolving to recognize H-2 class I, and Leu-2 evolving to recognize HLA class I.

We also confirm that the Lyt-2 gene is responsible for the production of two differentially-spliced Lyt-2 encoding mRNAs. These RNAs result from alternate usage of the first cytoplasmic exon. α-encoding cDNAs contain each of the five exons of the Lyt-2 gene; whereas α' cDNAs result from the deletion of the fourth exon. Using cDNA expression vectors, it became clear that the α chain (38kDa) of Lyt-2 is produced by mRNA which contains all five exons and correspondingly, the α' chain is encoded by mRNA in which 31 bp from the fourth exon is deleted.

The conversion of α/α' cDNA expression at the polypeptide level in thymocytes to expression of only α in lymph nodes might be related to a T-cell maturation process and be due to a change in the RNA splicing pattern. However, S1 nuclease analysis of lymph node RNA shows that 20% of the Lyt-2 mRNAs are the α' form as compared to 45% in thymocytes (28). This implies that splicing, as well as other mechanisms such as translational control, contribute to regulation of the differential expression of the two Lyt-2 chains.

REFERENCES

1. Cantor, H. & Boyse, E.A. J. Exp. Med. 141:1376 (1975).
2. Cantor, H. & Boyse, E.A. Immunol. Rev. 33:105 (1977).
3. Ledbetter, J.A. et al. J. Exp. Med. 152:280 (1980).
4. Evans, R.L. et al. Proc. Natl. Acad. Sci. USA. 78:544 (1981).
5. Reinherz, E.L. & Schlossman, S.F. Cell 19:821 (1980).
6. Bernard, A. et al. in Leucocyte Typing, eds. Bernard, A., Boumsell, L., Dausset, J., Milstein, C. & Schlossman, S.F. (Springer, Berlin), p.48. (1984).
7. Ledbetter, J.A. et al. J. Exp. Med. 153:310 (1981).
8. Nakayama, E.H. et al. Proc. Natl. Acad. Sci. USA. 76:1977 (1979).
9. Shinohara, N. and Sachs, D.H. J. Exp. Med. 150:432 (1979).
10. Landegren, U. et al. J. Exp. Med. 155:1579 (1982).
11. Swain, S.L. Proc. Natl. Acad. Sci. USA. 78:7101 (1981).
12. Swain, S.L. Immunol. Rev. 74:129 (1983).
13. Krensky, A.M. et al. Proc. Natl. Acad. Sci. USA 79:2365 (1982).
14. Ledbetter, J.A. et al. J. Exp. Med. 153:1503 (1981).
15. Kavathas, P. et al. Proc. Natl. Acad. Sci. USA. 81:7688 (1984).
16. Littman, D. et al. Cell. 40:237 (1985).
17. Nakauchi, H. et al. Proc. Natl. Acad. Sci. USA. 82:5126 (1985).
18. Chien, Y.-h. et al. Nature (London). 309:322 (1984).
19. Sanger, F., Nicklen, S. & Coulson, A. Proc. Natl. Acad. Sci. USA. 74:5463 (1977).
20. Staden, R. & McLauchlan, A.D. Nucleic Acids Res. 10:141 (1982).
21. Staden, R. Nucleic Acids Res. 10:2951 (1982).
22. Staden, R. Nucleic Acids Res. 10:4731 (1982).
23. Lipman, D.J. & Pearson, W.R. Science. 227:1435 (1985).
24. Wigler, M. et al. Cell. 16:777 (1979).
25. Ledbetter, J.A. & Herzenberg, L.A. Immunol. Rev. 47:63 (1979).
26. Kavathas, P. & Herzenberg, L.A. Proc. Natl. Acad. Sci. USA. 80:524 (1983).
27. Nakauchi H. et al. (in press).
28. Williams, A.F. & Gagnon, J. Science 216:696 (1982).
29. Tagawa, M. et al. Proc. Natl. Acad. Sci. USA. 83:3422 (1986).