Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 1609–1613, March 1987
Genetics

Structure of a cytotoxic T-lymphocyte-specific gene shows a strong homology to fibrinogen β and γ chains

(cDNA/gene evolution/multigene families/DNA sequence analysis/killer T cell)

TAKUMI KOYAMA*†, LAURIE R. HALL*†, WAYNE G. HASER‡§, SUSUMU TONEGAWA‡§, AND HARUO SAITO*†¶

*Division of Tumor Immunology, Dana–Farber Cancer Institute, and †Department of Biological Chemistry, Harvard Medical School, Boston, MA 02115; and ‡Center for Cancer Research and §Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Communicated by Charles C. Richardson, November 10, 1986 (received for review October 14, 1986)

ABSTRACT **Using a subtractive cDNA cloning method, we isolated a number of T-lymphocyte-specific genes. One of these genes, represented by the cDNA clone pT49, is expressed in cytotoxic T lymphocytes but not in helper T lymphocytes or B lymphocytes. The protein structure deduced from the nucleotide sequence showed a high degree of homology to fibrinogen β and γ subunits. This might indicate that evolutionarily fibrinogen has its origin in lymphocytes. In spite of the strong homology of pT49 protein to the fibrinogen subunits, the positions of the introns in the pT49 gene are totally different from those of the fibrinogen genes.**

Cytotoxic T lymphocytes (CTL), or killer T cells, recognize and lyse target cells bearing specific antigens on the membrane. Their function is critical to the defense of the body against the invading viruses, to the rejection of allografts, and perhaps to the removal of malignant cells. However, the mechanism of the CTL-mediated cytolysis of target cells is poorly understood (1). One way of approaching this problem is to isolate messenger RNAs specifically expressed in CTLs and then identify the function of the corresponding genes. Previously, we have described the isolation of 16 different groups of T-cell-specific cDNA clones (2, 3). During the systematic effort to characterize these cDNA clones, the gene represented by the cDNA clone pT49 (henceforth called the pT49 gene) was found to be highly homologous to the β and γ subunits of fibrinogen. In this report, we describe the initial characterization of the pT49 gene.

MATERIALS AND METHODS

The construction of a T-cell minus B-cell subtractive cDNA library and the isolation of T-cell-specific cDNA clones have been described (2, 3). In short, a cDNA library constructed by subtracting single-stranded cDNA made from the mouse CTL clone 2C (see ref. 4) with B-cell lymphoma (A20-2J) mRNA, was screened for T-cell-specific clones using subtracted T-cell-specific cDNA probe. Preparation of mouse (BALB/c) embryo *Eco*RI partial library by using Charon 4A as a vector has been described (5). These libraries were screened by the methods of Benton and Davis (6) and Hanahan and Meselson (7). Methods for DNA and RNA blot hybridization were according to Southern (8) and Thomas (9), respectively. Approximately 1.5 μg of poly(A)$^+$ RNA was denatured with glyoxal and electrophoresed through a 1.1% agarose gel in 10 mM sodium phosphate buffer (pH 6.5). RNA was transferred to a nitrocellulose membrane filter and hybridized to ^{32}P-labeled nick-translated probe DNA. Hybridization was carried out at 42°C in the presence of 50% formamide and 5× SSC, and the filters were washed at 65°C in 0.2× SSC (1× SSC = 0.15 M NaCl/0.015 M sodium citrate). Nucleotide sequences were determined according to Maxam and Gilbert (10).

RESULTS

The Expression of the pT49 Gene Is CTL Specific. The pT49 cDNA clone was originally isolated as a gene expressed in CTL clone 2C but not in B-cell lines (2, 3). The RNA gel blotting analysis showed that the pT49 gene was expressed in CTL, but not in helper T cells, B cells, or a few other cells tested (Fig. 1*a*). Although the number of T cells tested is rather small, we tentatively consider the expression of the pT49 gene CTL specific. The size of the mRNA detected by RNA blot hybridization [4.2 kilobases (kb)] is much longer than the size of the insert in pT49. Therefore, cDNA libraries made from 2C mRNA were screened to obtain overlapping cDNA clones covering a total of 3.7 kb. The relevant cDNA clones, with the restriction map reconstructed from these clones, are shown in Fig. 2*A*.

The Nucleotide Sequences of the pT49 Gene and Its cDNA. Fig. 3 shows the nucleotide sequence of cDNA, which is 3725 base pairs (bp) long and was determined according to the strategy shown in Fig. 2*A*. The orientation of the cDNA was unambiguously determined due to the presence of the poly(A) tail as well as the signal for the poly(A) attachment site (AATAAA, underlined in Fig. 3). When the nucleotide sequence was translated into the amino acid sequence, only one of the three possible reading frames resulted in a large open reading frame, which is shown in Fig. 3.

The nucleotide sequence of the cDNA apparently lacks the 5′ end of the mRNA, since there seems to be no initiation methionine codon. To obtain the missing sequence at the 5′ end of mRNA, we isolated the genomic DNA clones corresponding to the coding region of the pT49 gene. The Southern blot hybridization of the *Eco*RI-digested mouse (BALB/c) embryonic DNA with the coding-region probe (see Fig. 2*A*), detected a single 10-kb band (Fig. 1*b*). The same probe was used to isolate hybridizing λ phage clones from a mouse embryonic DNA library. One of the clones, T49-ME6, contained the 10-kb *Eco*RI fragment that hybridized to the probe. The comparison of detailed restriction maps of the cDNA and genomic clones indicated that there are two exons in the 10-kb *Eco*RI fragment. Indeed, the nucleotide sequences in and around the exons showed that the protein coding region of the gene is split by an intron of ≈1900 bases (Fig. 4). More importantly, inspection of the 5′ portion of the genomic sequence revealed that the first nucleotide (G) of the cDNA sequence is the third base of the methionine codon ATG. We believe that this codon is used for the initiation of the translation, because there is no other methionine codon or splicing acceptor site (12) between this methionine codon

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: CTL, cytotoxic T lymphocyte(s).

¶To whom reprint requests should be addressed at: Dana–Farber Cancer Institute, 44 Binney Street, Boston, MA 02115.

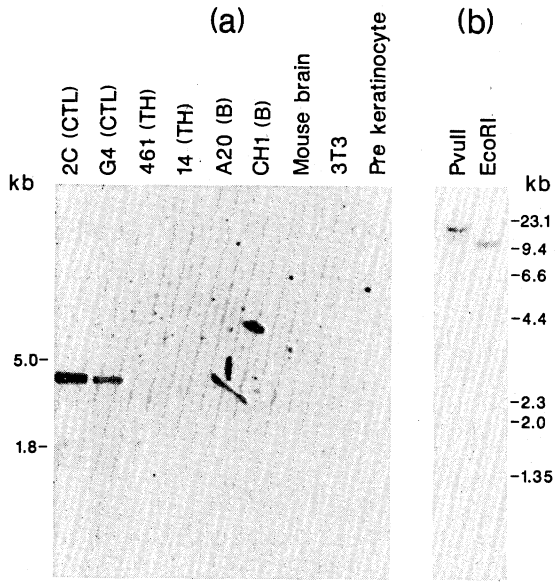

FIG. 1. (*a*) RNA blot analyses of poly(A)$^+$ RNAs. RNA was isolated from two alloreactive BALB.B CTL clones (2C, anti-L^d; G4, anti-D^d; see ref. 4), two T-helper (TH) hybridomas derived from a BW5147 × B10.A spleen cell fusion (14, anti-GAT; 461, anti-dinitrophenyl ovalbumin; see ref. 11); B-cell lymphomas A20-2J and CH1, whole mouse brain, 3T3 preadipocyte cell line, and a prekeratinocyte line. The hybridization probe is shown in Fig. 2*A*. (*b*) DNA blot analysis of BALB/c embryo DNA digested with either *Eco*RI or *Pvu* II. The hybridization probe is the same as in *a*.

and an in-frame termination codon (TGA) present 60 bp upstream.

The Protein Encoded by the pT49 Gene Is Homologous to Fibrinogen Subunits. The predicted amino acid sequence of the pT49 gene was compared with all of the sequences in the database of the Protein Identification Resource, using an Intelli-Genetics computer program, IFIND. Significant degrees of homology were detected only with the fibrinogen β or γ subunits of various animals (13–18) (Fig. 5). The COOH-terminal half of the pT49 protein shows much higher homology to fibrinogen subunits than the NH_2-terminal half. This feature of homology is also noticeable among the fibrinogen subunits themselves, whether the comparison is between the β and γ subunits or between the same subunits of different species, and is interpreted as evidence for the different evolutionary selective pressures on various domains of fibrinogen (19). A more quantitative comparison of these peptides in the COOH-terminal half (amino acids 171–413 of the pT49 protein) showed that the pT49 protein is ≈36% homologous to the β and γ subunits. The β and γ subunits of fibrinogen are also ≈36% homologous to each other in this region. On the other hand, the homology between the NH_2-terminal halves of the pT49 protein and the β or γ subunits is much lower (15–20%).

DISCUSSION

While the relatedness of the pT49 protein to the fibrinogen subunits is not disputable, the possible structure and function of the pT49 protein may not be so obvious. The comparison of functionally important regions of fibrinogen subunits to the corresponding portions of the pT49 protein revealed the following. First, there seems to be the signal peptide necessary for the transport of the protein across the membrane. The calculation according to the method of von Heijne (20) predicts that the most likely cleavage site is between the

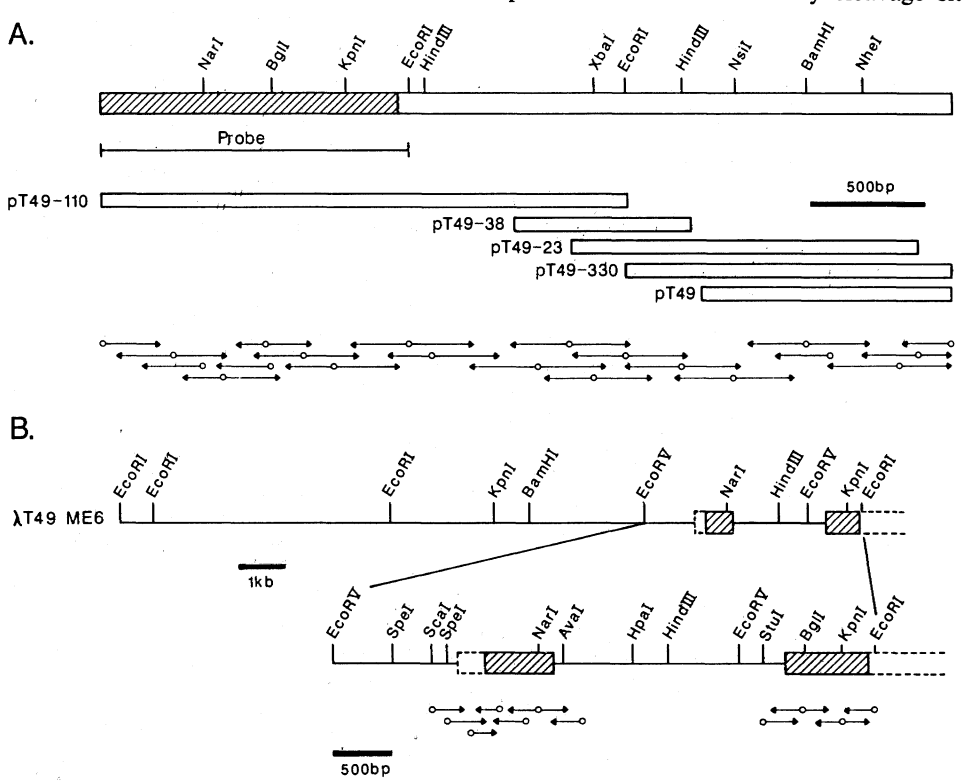

FIG. 2. (*A*) Restriction map of the cDNA of the pT49 gene. The coding and 3′ noncoding regions are shown by hatched bar and open bar, respectively. The individual cDNA clones are indicated by thinner open bars. The hybridization probe used in Fig. 1 *a* and *b* is indicated by a line below the restriction map. Arrows indicate the direction and extent of nucleotide sequence determination. (*B*) Restriction map of the genomic clone of the pT49 gene. Upper portion shows the restriction map of the entire clone, while the bottom portion shows the enlarged map of the protein-encoding region of the pT49 gene. Exons are indicated by boxes. Because the precise 5′ end of the gene is not known, only the approximate position is indicated. Shading and arrows are as in *A*.

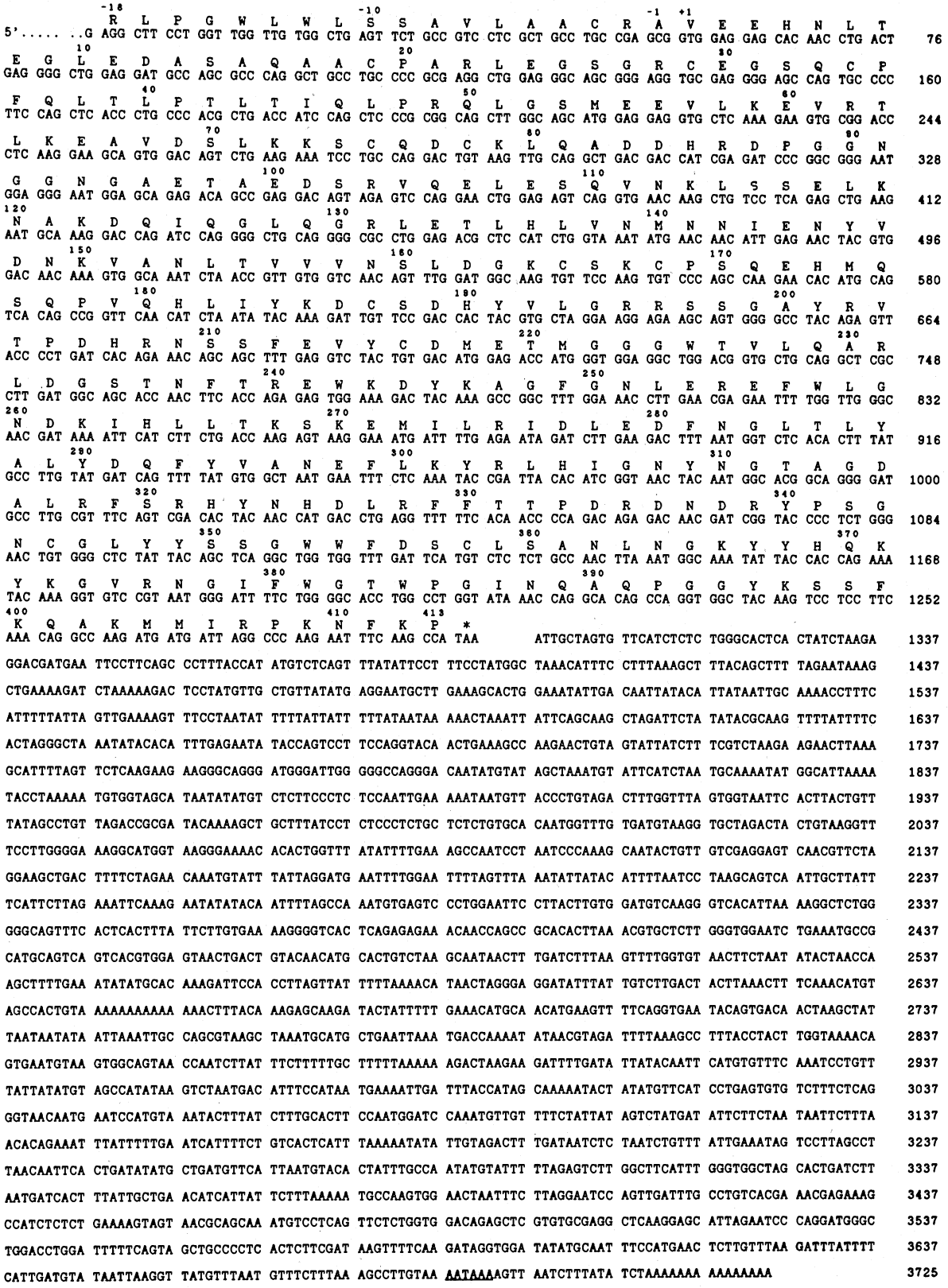

FIG. 3. Nucleotide sequence and predicted amino acid sequence of the pT49 cDNA. Numbers above the amino acid sequence designate amino acid residue position of the putative processed protein, while the numbers on the right show the nucleotide positions. The signal for the poly(A) attachment site is underlined. *, Termination codon. Amino acids are designated by the single-letter code.

alanine and the valine at amino acids −1 and +1, with a score of 9.6. This score is high enough to conclude that there is indeed a signal peptide, although the assignment of the cleavage site must be considered tentative (20). Second, there are two conserved pairs of cysteines, which are indicated by asterisks in Fig. 5. The cysteines are separated by 2 or 4 amino acids rather than 3 amino acids as is the case in the fibrinogen subunits. The algorithm of Chou and Fasman (21)

```
AAGAGTTCACACAAAGTCTAGCCTTCAAAAAGAAAACAGGTTCCCAAACTAGGGAGGAAACAGAATCATTTCCATTTTGGTGACATTTAGTGGGAAGAAGCTCACAGACAT

TTAGACGTTCCAACTCTTTCCCCACTAGTGGACCAAGTATATAATATGGTATCTTTTGGGCACTGGTATTACAACTGTTTTTAAACAAAAGACTTTCCTTGTGCTTTACT

AAAAACCCAGACGGTGAATCTTGAATACAATGCGTGGCACCCACGGCAGGCATTCTATTGTGCATAGTTTTGACTGACAGGAGATGACAGCATTTGGCTGGCTGCGCTTGC

TGAGGACCCTCTCCTCCTGTGTGGCGTCTGAGACTGTGATGCAAATGCGCCCGCCCTTTTCTGGGAACTCAGAACGCC  TGA GTC AGG CGG CGG TGG CTA TTA
                                                -19                             -10
                                                 M   R   L   P   G   W   L   W   L   S   S   A   V   L   A   A
AAG CGC CTG GTC AGG CTG GGC TGC CGC ACT GCA AGG ATG AGG CTT CCT GGT TGG TTG TGG CTG AGT TCT GCC GTC CTC GCT GCC
     -1  +1                                  10                                          170
 C   R   A   V   E   E   H   N   L   T   E   G   L   E                                   P   S   Q   E   H   M
TGC CGA GCG GTG GAG GAG CAC AAC CTG ACT GAG GGG CTG GAG ... ... same as cDNA .. ... ... CCC AGC CAA GAA CAC ATG
            178
 Q   S   Q   P
CAG TCA CAG CCG G / GTAGGTGTAATGAGGGTCATACAGTTTGTTCATGAAAGCTGTATAGCCAGATAGTGGCCATAAACATTAACCCGAGGGAGCATAAGTTAGT

CAGACTTTCACCTGTTAAGTTATGGCAGGAGAAACAAGTGTTTTCTCAAATGAGACAACAGAAATGGTAAATGATCCACGTACAAAAATCCTATTAGTTGTACTCGTTAGA

GACCGTCACTTGCAAGTCTCTAGACCTTCCCTGCTAGGTCGAC ...... about 1500 bases ..... AGTTAGCATCTGTTAGCCTGGTTCAGGAGAGTGTATCA

GAGCCAGGTTCCTCTATCACATAAACTGTAACGCAAGTGAATTGTCCAATTGCTGTTGAGTCTGAGAGTCCTTGAGGTGCATAGCTTTGACTAATAAATCCCCATGCTTTT
                                 180                                     190
                              V   Q   H   L   I   Y   K   D   C   S   D   H   Y   V   L   G   R   R   S   S   G
ATGCTTTTCCTTCCTCCCTCTTCCAG / TT CAA CAT CTA ATA TAC AAA GAT TGT TCC GAC CAC TAC GTG CTA GGA AGG AGA AGC AGT GGG
200                                     210                                                              400
 A   Y   R   V   T   P   D   H   R   N   S                               P   G   G   Y   K   S   S   F   K
GCC TAC AGA GTT ACC CCT GAT CAC AGA AAC AGC ... ... same as cDNA .. ... ... CCA GGT GGC TAC AAG TCC TCC TTC AAA
                                    410         413
 Q   A   K   M   M   I   R   P   K   N   F   K   P   *
CAG GCC AAG ATG ATG ATT AGG CCC AAG AAT TTC AAG CCA TAA ATTGCTAGTGTTCATCTCTCTGGGCACTCACTATCTAAGAGGACGATGAATTC
```

FIG. 4. Nucleotide sequence of the pT49 gene around the protein-encoding region. The coding sequence is identical to that of the cDNA sequence except for the first 2 bases of the initiation codon, which are missing from the cDNA sequence. Amino acids are designated by the single-letter code.

shows that the region between these pairs of cysteines has a very high tendency to assume an α-helical conformation, with an interruption at the middle. In this region of fibrinogen, the three subunits with the α-helical conformation form a three-chain coiled-coil structure interlocked by the disulfide rings (22). While the slightly distorted dispositions of cysteines in the pT49 protein might make the dimeric structure more likely, it may still be possible that the pT49 protein is aggregated with other proteins (or itself) to form a three-stranded coiled coil. Finally, it must be pointed out that the pT49 protein does not have the sequences corresponding to a thrombin-sensitive site for the release of a fibrinopeptide (amino terminus of the β chain) or a site for cross-linking or interaction with platelets (carboxyl terminus of the γ chain) (23). However, if the pT49 protein associates with other proteins, these properties could be attributable to those chains. As a conclusion, while there is much uncertainty it is possible that the pT49 protein is a part of a secreted fibrinogen-like complex.

Although it is premature to speculate about the possible involvement of the pT49 molecule in CTL-mediated cytolysis, we would like to suggest a few of the possibilities. Fibrinogen is activated and degraded by the proteases thrombin and plasmin, respectively, which are in turn activated by other proteases. Therefore, it is interesting to note that three CTL-specific serine protease genes have been isolated by the

```
                                                                                               *    *
Mouse    pT49                                                              VEEHNLTEGLEDASAQAACPARLEGSGR CEGSQCPFQLTLPTLTIQLPRQLGSMEEV
Human    beta   QGVNDNEEGFFSARGHRPLDKKREEAPSLRPAPPPISGGGYRARPAKAAATQKKVERK-PDAGG-LHADPDL-VL-P TG-QL-EA-LQQERPIRNSVDELNNN
Human    gamma                                                                         YVATRDN-CILD-RF-SY-P TT-GIADF-S-YQTKVDKD-Q-L-DI
Rat      gamma                                                                        QYTATRDN-CILD-RF-SY-P TT-GISDF-NSYQTDVDTD-QTL-NI
Lamprey  gamma                                                                         QVRDLKQ-S NDPEF-RY-P TT-GVADV-SKYAKGVDEDSSFIDS-

Mouse    pT49   LKEVRTLKEAVDSLKKSCQDCKLQADDHRDPGGNGGNGAETAEDSRVQELESQVNKLSSE LKNAKDQI QGLQGRLETLHLVNMNNIENYVDNKVANLTVVVN
Human    beta   VEA-SQTSSSSFQYMYLLKDLWQ                           KRQKQ-KDN-NV--EY--- -EKHQLY-DETVNSNIA-NLR-LR S-LENLRS-IQK-ESD-S
Human    gamma  -HQ-ENKTSE-KQ-I-AI-LTYN                           PDE-SKPNMIDAATLK-RKM-EEIMKY  EASILTHDSSIRYLQ E-Y-SNNQ-IV--KEK-A
Rat      gamma  -QRAENRTTEAKE-I-AI-VYYN                           PDQPPKPGMIEGATQK-KKMVEEILKY  EA-LLTH-SSIRYLQ D-YTSNKQ-IT--KQK-A
Lamprey  gamma  -TQLAAKHGI-EGNVNIVNEDVR                           ITRDEA-IIKDSGQ-TVQKI-EEVRIL  EQIGVSHDAQIQELS EMWRVNQQF-TR-QQQLV

                     *    *
Mouse    pT49   SLDGKCSK CPSQEHMQSQPVQHLIYKDCSDHYVLGRRSSGAYRVTPDHRNSSFEVYCDMETMGGGWTVLQARLDGSTNFTREWKDYKAGFGNL
Human    beta   AQMEY-RTP-T    V-CNIPVVSG-E-EEIIRK-GET-EM-LIQ--SSVKPYR-----N-EN-----I-N-Q---VD-GRK-DP--Q----VATNTDGKNYC
Human    gamma  Q-EAQ-QEP-K    DTVQIHDITG---Q-IANK-AKQ--L-FIK-LKA-QQ-L---EIDGS-N----F-K-----VD-KKN-IQ--E---H-SP
Rat      gamma  Q-EAQ-KEP-K    D-VRIYDITG---Q-TANK-AKE--L-SIR-LKATQQSL---ETDGP-N---EFKK-----VD-LKN-IQ--E---H-SP
Lamprey  gamma  DIRQT--RS-QDT   TANKISPITG---QQVVDN-GKD--L-YIK-LKAKQP-L-F-EI- N-N----I-H-H---V----D-VS-RE---Y-AP

Mouse    pT49     EREFWLGNDKIHLLTKSK  EMILRIDLEDFNGLTLYALYDQFYVANEFLKYRLHIG NYNGTAGDALRFS         RHYNHDLRFFTTPDRDNDRY
Human    beta   GLPG-Y-------SQ--RMG  PTE-L-EM--WK-DKVK-H-GG-T-Q--AN--QISVN K-R----N--MDGASQLMGENRTMTI-NGM--S-Y-----GWLT
Human    gamma  TGTT------E----ISTQSAIPYA--VE---W--R-ST-D-AM-K-GP-AD----TYAYFAG-D----FDGFDFGDDPSDKFFTS-NGMQ-S-W-N---KF
Rat      gamma  TGTT------E-N--ISMQSTIPYA---Q-K-WS-R-ST-D-AM-R-GPGSD----TYAYFIG-D----FDGYDFGDDPSDKFFTS-NGMH-S-W-N---KF
Lamprey  gamma  TLTT------E------GQQ  AYR-----T-WENTHR--D-GH-KLTP-SDE---FYSMYLD-D--N-FDGFDFGDDPQDKFYTT-LGML-S--E----K-

Mouse    pT49     PSGNCGLYYSSGWWFDSCLSANLNGKYYH QKYK      GVRNGIFWSTWPGINQAQPGGYKSSFKQAKMMIRPKNFKP
Human    beta   SDPRKQ-SKEDGG---YNR-HA--P--R--WGGQ-TWDMAKH-TDD-VV-MN-K-SWY          -MRKMS-K---FFPQQ
Human    gamma     E---AEQDG----MNK-HAGH---V--QGGT-SKASTPN-YD---I-A--KTRWY          -M-KTT-K-I-F-RLTIGEGQQHHLGGAKQAGDV
Rat      gamma     E---AEQDG----MNK-HAGH---V--QGGT-SKSSTPN-YD---I-A--KTRWY          -M-ETT-K-I-F-RLSIGDGQQHHMGGSKQVGDM
Lamprey  gamma     E-S-AEQDG----MNR-HAGH------FGGN-RKTDVEFPYDD--I-A--HDRWY          -L-MTT-KLL-MGRDLSGHGGQQQSKGN SRGDN
```

FIG. 5. Comparison of the amino acid sequences of pT49 and the fibrinogen β and γ subunits. To increase the identity, some deletions/insertions were incorporated. The amino acids of fibrinogen subunits identical to those of pT49 are indicated by hyphens. The sequences of the fibrinogen subunits are all of the mature proteins (13, 14, 16, 17), while that of pT49 is the proposed processed protein. Two pairs of cysteines, indicated by asterisks, form disulfide rings that link the three subunits of fibrinogen. Amino acids are designated by the single-letter code.

subtractive cloning method (24–26). CTL-specific proteases were also implicated by the effect of protease inhibitors on CTL-mediated cytolysis (27) and by the identification of a serine esterase in these cells (28–30). This led to the suggestion (31, 32) that there is a protease cascade reminiscent of the complement system, with the pore-forming perforin as a substrate (33, 34). The fibrinogen-like complex, however, could be an additional substrate of the protease cascade. The fibrinogen-like complex might work as an adhesion molecule that holds the CTL and the target cells together tightly during the delivery of the lethal hit. Indeed, divalent fibrinogen binds to and aggregates the platelets (23). Alternatively, the fibrinogen-like complex might be a clotting inhibitor. It will prevent undesirable fibrin formation around the CTL or target cells that might be initiated by the release of the CTL-specific proteases into the circulation system. The fragment D of the fibrinogen, which is devoid of the polymerization donor site, is a potent clotting inhibitor.

Perhaps a more immediate contribution of the discovery of the pT49 gene is its implication in the questions about the evolution of the fibrinogen gene family. The first question concerns the origin of the introns in the fibrinogen gene family (35). In spite of their amino acid homology, the number and the location of introns in the β and γ subunits are quite variable: β has 7 and γ has 8 introns, of which only 3 occupy homologous positions. To explain the diversity, one model assumes random intron insertions into an intronless primordial gene and another model assumes selective intron losses from an intron-ridden primordial gene (35). The only intron in the coding region of the pT49 gene does not coincide with any of the introns of the β and γ subunits. This observation is hard to interpret by either model alone. Probably it requires a reconciliatory model that invokes both processes. Another question concerns the origin of the fibrinogen genes themselves (19). As previously mentioned, the COOH-terminal halves of the β and γ subunits show homology of 36%. On the other hand, the human and lamprey γ subunits are $\approx$61% homologous in the same region (17). Thus, the divergence of the β chain of fibrinogen from the γ chain occurred long before the divergence of the lineages that eventually led to human and lamprey. The latter is the lowest animal known to possess plasma fibrinogen. Therefore, paradoxically, it looks as if the genes for the fibrinogen subunits diverged and evolved before fibrinogen came to exist (19). A possible solution suggested by the discovery of the pT49 gene is that the precursor of fibrinogen has evolved as a primitive lymphocyte protein. The development of a completely closed circulation system in primitive vertebrates required a coagulation scheme to protect against the loss of precious iron, and some of the properties of the pT49 protein might have been exploited to form the basis of the blood clotting system.

We are grateful to Dr. Stuart F. Schlossman for his encouragement and generous support. We also wish to thank Drs. Russell F. Doolittle and Herman N. Eisen for valuable comments on the manuscript. This work was supported by an Investigator Award from the Cancer Research Institute and by grants from the Friends of Dana–Farber Cancer Institute and the Whitaker Health Sciences Fund (to H.S.), National Institutes of Health Grants AI-12069-13 (to Stuart F. Schlossman) and CA-38900-06 (to S.T.), and a research grant from Ajinomoto Co. (to S.T.). W.G.H. is supported by a postdoctoral fellowship from the National Institutes of Health (NS07723-02).

1. Henkart, P. A. (1985) *Annu. Rev. Immunol.* **3,** 31–58.
2. Saito, H., Kranz, D. M., Takagaki, Y., Hayday, A. C., Eisen, H. N. & Tonegawa, S. (1984) *Nature (London)* **309,** 757–762.
3. Saito, H., Kranz, D. M., Takagaki, Y., Hayday, A. C., Eisen, H. N. & Tonegawa, S. (1984) *Nature (London)* **312,** 36–40.
4. Kranz, D. M., Sherman, D. H., Sitkovsky, M. V., Pasternack, M. S. & Eisen, H. N. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 573–577.
5. Maki, R., Traunecker, A., Sakano, H., Roeder, W. & Tonegawa, S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 2138–2142.
6. Benton, W. & Davis, R. (1977) *Science* **196,** 180–182.
7. Hanahan, D. & Meselson, M. (1980) *Gene* **10,** 63–67.
8. Southern, E. (1975) *J. Mol. Biol.* **98,** 503–517.
9. Thomas, P. S. (1983) *Methods Enzymol.* **100,** 255–266.
10. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
11. Samelson, L. E., Germain, R. N. & Schwartz, R. H. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 6972–6976.
12. Mount, S. M. (1982) *Nucleic Acids Res.* **10,** 459–472.
13. Watt, K. W. K., Takagi, T. & Doolittle, R. F. (1979) *Biochemistry* **18,** 68–76.
14. Chung, D. W., Que, B. G., Rixon, M. W., Mace, M., Jr., & Davie, E. W. (1983) *Biochemistry* **22,** 3244–3250.
15. Chung, D. W., Rixon, M. W., MacGillivray, R. T. A. & Davie, E. W. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 1466–1470.
16. Crabtree, G. R. & Kant, J. A. (1982) *Cell* **31,** 159–166.
17. Strong, D. D., Moore, M., Cottrell, B. A., Bohonus, V. L., Pontes, M., Evans, B., Riley, M. & Doolittle, R. F. (1985) *Biochemistry* **24,** 92–101.
18. Doolittle, R. F., Watt, K. W. K., Cottrell, B. A., Strong, D. D. & Riley, M. (1979) *Nature (London)* **280,** 464–468.
19. Doolittle, R. F. (1983) *Ann. N.Y. Acad. Sci.* **408,** 13–27.
20. von Heijne, G. (1986) *Nucleic Acids Res.* **14,** 4683–4690.
21. Chou, P. & Fasman, G. D. (1974) *Biochemistry* **13,** 222–245.
22. Doolittle, R. F., Goldbaum, D. M. & Doolittle, L. R. (1978) *J. Mol. Biol.* **120,** 311–325.
23. Doolittle, R. F. (1984) *Annu. Rev. Biochem.* **53,** 195–229.
24. Gershenfeld, H. K. & Weissman, I. L. (1986) *Science* **232,** 854–858.
25. Lobe, C. G., Finaly, B., Paranchych, W., Paetkau, V. H. & Bleackley, R. C. (1986) *Science* **232,** 858–861.
26. Brunet, J.-F., Dosseto, M., Denizot, F., Mattei, M.-G., Clark, W. R., Haqqi, T. M., Ferrier, P., Nabholz, M., Schmitt-Verhulst, A.-M., Luciani, M.-F. & Golstein, P. (1986) *Nature (London)* **322,** 268–271.
27. Chang, T. W. & Eisen, H. N. (1980) *J. Immunol.* **124,** 1028–1033.
28. Pasternack, M. S. & Eisen, H. N. (1985) *Nature (London)* **314,** 743–745.
29. Pasternack, M. S., Verret, C. R., Liu, M. A. & Eisen, H. N. (1986) *Nature (London)* **322,** 740–743.
30. Masson, D., Nabholz, M., Estrade, C. & Tschopp, J. (1986) *EMBO J.* **5,** 1595–1600.
31. Reid, K. B. M. (1986) *Nature (London)* **322,** 684–685.
32. Young, J. D.-E. & Cohn, Z. A. (1986) *Cell* **46,** 641–642.
33. Young, J. D.-E., Cohn, Z. A. & Podack, E. R. (1986) *Science* **233,** 184–190.
34. Tschopp, J., Masson, D. & Stanley, K. K. (1986) *Nature (London)* **322,** 831–834.
35. Crabtree, G. R., Comeau, C. M., Fowlkes, D. M., Fornace, A. J., Jr., Malley, J. D. & Kant, J. A. (1985) *J. Mol. Biol.* **185,** 1–19.