# GENOMIC ORGANIZATION OF THE MOUSE PORE-FORMING PROTEIN (PERFORIN) GENE AND LOCALIZATION TO CHROMOSOME 10

## Similarities to and Differences from C9

By JOSEPH A. TRAPANI,* BYOUNG S. KWON,‡ CHRISTINE A. KOZAK,§ CHAYA CHINTAMANENI,‡ JOHN DING-E YOUNG,‖ and BO DUPONT*

*From the *Laboratory of Human Immunogenetics, Memorial Sloan-Kettering Cancer Center, New York, New York 10021; the ‡Department of Microbiology and Immunology and Walther Oncology Center, Indiana University School of Medicine, Indianapolis, Indiana 46223; the §Laboratory of Molecular Microbiology, National Institute for Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892; and the ‖Laboratory of Cellular Physiology and Immunology, The Rockefeller University, New York, New York 10021*

The molecular mechanisms involved in the delivery of cellular injury or death by CTL and NK cells have not been elucidated (1, 2). The granule-exocytosis model of cellular cytotoxicity, the most closely investigated mechanism to date, postulates that an obligate facet of the delivery of the "lethal hit" is the directional release of cytotoxic polypeptides and other molecules (the granule contents) towards the target cell (3–5). Key molecules in this process are a cytolytic protein termed perforin (5, 6), cytolysin (3), pore-forming protein (PFP)[1] (4) or C9-related polypeptide (C9RP) (7), and a family of killer cell–specific, highly homologous serine esterases (SE) (8, 9). PFP is postulated to intercalate the target cell membrane and compromise its osmotic stability; however, the function of SEs in the cytotoxic process is less clear. Intriguing parallels have been drawn between this mechanism and the complement cascade, including the close similarities between the immunological and physicochemical properties of PFP and C9, the ninth component of complement (7, 10–12).

Many of the granule polypeptides have been purified to homogeneity, and cDNA clones encoding PFP (13–16) and several of the SEs (17–25) have recently been characterized. To further analyze the role of PFP in the cytolytic process and the basis for its immunological and evolutionary relationship to C9, we have investigated the

This work was supported in part by U. S. Public Health Service (National Institutes of Health) grants CA-22507, CA-49096, CA-08748, AI-28175, and CA-47307; by a grant from the Lucille P. Markey Charitable Trust; and by a Frances L. and Edwin L. Cummings Memorial Fund Investigatorship from the Cancer Research Institute. J. A. Trapani is supported by a C. J. Martin Travelling Fellowship, National Health and Medical Research Council, Australia. J. D.-E Young is a Lucille P. Markey Scholar. J. A. Trapani's present address is the Department of Pathology, University of Melbourne, Parkville 3052, Victoria, Australia.

Address correspondence to John Ding-E Young, Laboratory of Cellular Physiology and Immunology, Box 37, The Rockefeller University, 1230 York Avenue, New York, NY 10021.

[1] *Abbreviations used in this paper:* PCR, polymerase chain reaction; PFP, pore-forming protein; SP, serine protease; UT, untranslated.

genomic organization of the mouse gene encoding PFP (here designated *Pfp*) and its chromosomal localization.

## Materials and Methods

*cDNA Probes and Library Screening.* Probes used in this study were derived from two overlapping PFP cDNA clones, designated PFP-6 and PFP-15, which together comprise the entire coding sequence for mouse PFP, including ~0.2 kb of 5′ untranslated (5′UT) sequence, but lack most of the 3′UT region (16). PFP-6 lacks the 5′UT sequences and the first 97 bp of the coding region, but extends ~1.3 kb further in the 3′ direction (Fig. 1). The 5′UT region probe used to isolate clones encoding exon I was derived from Pst I digestion of PFP-15 DNA and excision of a 0.19-kb ethidium bromide-staining band from a 1.2% low-gelling temperature agarose gel.

*Screening of Genomic Library.* DNA probes were used to screen a mouse genomic library constructed by partial digestion of high molecular weight DNA with Mbo I and ligation with dephosphorylated Bam HI-digested EMBL 3 arms (Stratagene, San Diego, CA). This library was kindly provided by Dr. Pramod Srivastava, Mt. Sinai Medical School, New York. Phage clones yielding positive hybridization signals were plaque purified, and the insert DNA was subcloned into pBluescript $KS^+$ (Stratagene).

*DNA Sequencing.* DNA sequencing was carried out by the dideoxy chain termination method (26), using either double-stranded plasmid templates (27) or single-stranded templates that were generated from pBluescript phagemids after infection with VCSM13 helper phage (28). Overlapping nucleotide sequences were derived by constructing a series of 17-base oligonucleotides on a DNA synthesizer (381A; Applied Biosystems, Inc., Foster City, CA), which were used as primers for strand extension. This strategy was applied to sequencing both the coding and noncoding DNA strands.

*cDNA Synthesis and Polymerase Chain Reaction (PCR) Amplification.* cDNA fragments encoding the 3′ terminus of mouse PFP mRNA were isolated as follows. First-strand cDNA was synthesized from poly(A)$^+$ RNA derived from the cloned cytotoxic T cell line, CTLL-R8 (29), using the oligonucleotide primer 5′ GCGGCCGC$(T)_{17}$ 3′, exactly as previously described (25). cDNA/RNA molecules were then amplified into double-stranded cDNA by PCR (30), using the above primer and a sense oligonucleotide, whose sequence was derived from the 3′UT region of the mouse *Pfp* genomic clone 64. For PCR, denaturation was at 95°C for 2 min, annealing at 55°C for 2 min, and extension at 74°C for 2 min. The resultant fragment was rendered blunt ended using the Klenow fragment of *Escherichia coli* DNA polymerase I, digested to completion with Not I, then subcloned into pBluescript $KS^+$. Insert DNA was used as a probe in DNA and RNA blotting experiments.

*DNA and RNA Blotting.* These procedures were carried out as previously described (25).

*Chromosomal Assignment of the Mouse Pfp Gene.* High molecular weight DNA was isolated from a previously characterized panel of 18 mouse-Chinese hamster cell lines and from a single mouse-rat hybrid cell line kindly provided by Dr. R. E. K. Fournier (F. Hutchinson Cancer Center, Seattle, WA) (31, 32). The DNA was digested to completion with Eco RI, subjected to electrophoresis through 1% agarose gels, and processed for hybridization as previously described (33). DNA probes were radiolabeled by nick translation to a specific activity of $>10^8$ dpm/μg. Finally, stringent washes were carried out at 65°C in 0.03 M NaCl/0.003 M sodium citrate/0.1% $NaDodSO_4$.

NFS/N mice were obtained from the National Institutes of Health Animal Facility. *Mus musculus musculus* mice originally trapped in Skive, Denmark were provided by Dr. M. Potter (National Cancer Institute, NIH) from his laboratory colony at Hazelton Laboratories, Rockville, MD. NFS/N females were mated with *musculus* males, and the $F_1$ females were mated with *musculus* males. DNA was extracted from the livers of backcross progeny, digested with the restriction enzymes Sca I and Apa I, subjected to electrophoresis on 0.4% agarose gels, and transferred to nylon membranes. The membranes were hybridized with $^{32}$P-labeled PFP-6 or P-*myb* obtained from Oncor, Inc. (Gaithersburg, MD).

## Results

*Isolation of the Mouse Pfp Gene.* A mouse cDNA probe, designated PFP-6 (16), which comprises the entire coding region of mouse PFP cDNA, except for 97 bp at its 5′ end, was used for the initial isolation of *Pfp* genomic sequences from a library constructed in the phage vector, EMBL 3. Of $5 \times 10^5$ clones that were screened, a single reactive clone, designated λ PFP-64 (clone 64), was isolated (Fig. 1). Insert DNA derived from clone 64 was subjected to analysis by Southern blotting, using a variety of restriction endonucleases and the PFP-6 cDNA probe. The resultant hybridization pattern was identical in each case with that observed previously with mouse genomic DNA (16), suggesting that clone 64 contained the entire coding sequence of *Pfp* (not shown).

*Intron-Exon Organization.* Sense and antisense oligonucleotides were synthesized according to the cDNA sequence for mouse PFP (16). These were used to sequence the exons in clone 64 in their entirety and to map intron-exon boundaries (Figs. 1 and 2). *Pfp* was found to consist of three exons, only two of which (exons II and III) encode polypeptide (Fig. 2). Exon I is ~0.3 kb in length and is not translated (see below). Exon II encodes the final 36 nucleotides of the 5′ UT region and the first 536 bases of encoding region. A second intron, ~2.3 kb in length, precedes exon III, which contains the remaining 1,129 bases of coding sequence and the entire 3′UT region.

*Protein Coding Domains of Pfp.* A comparison of the nucleotide and derived amino acid sequences of the *Pfp* gene revealed identity with those predicted by cDNA cloning, as reported by two groups (13, 16). The protein sequence diverges over a stretch of eight residues close to its 3′ end from that reported by Lowrey et al. (15); however, this appears to result from two-frame shifts in the latter cDNA sequence. As predicted from DNA blotting experiments, the coding regions of *Pfp* are entirely contained within a short stretch of the genome (13, 15, 16). The coding information is contained within ~4.8 kb of DNA, a markedly different situation from that found in

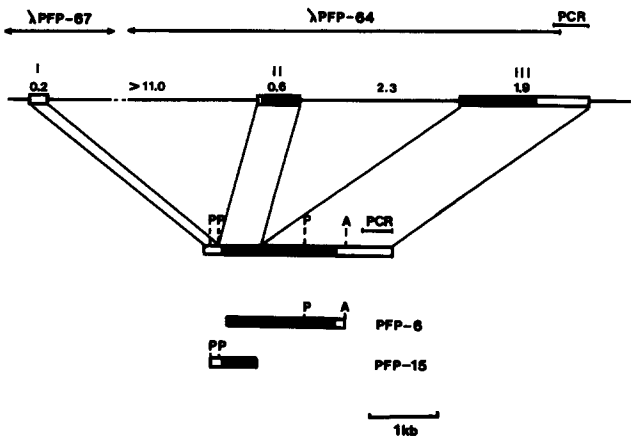


FIGURE 1. Organization of mouse *Pfp*. Exons are shown as boxes and indicated in roman numerals; lengths of exons and introns are shown in arabic numerals (in kilobases). Shaded areas of the exons indicate protein coding domains. P, Pst I; A, a long A-rich tract in the 3′UT region. PFP-6 and PFP-15 are cDNA probes that were used to isolate two reactive phage clones, designated λ PFP-64 and -67. Arrowheads indicate extension of phage inserts into the 5′ intron. The extreme 3′ end of the cDNA and the genomic sequences were determined by PCR amplification (see text).

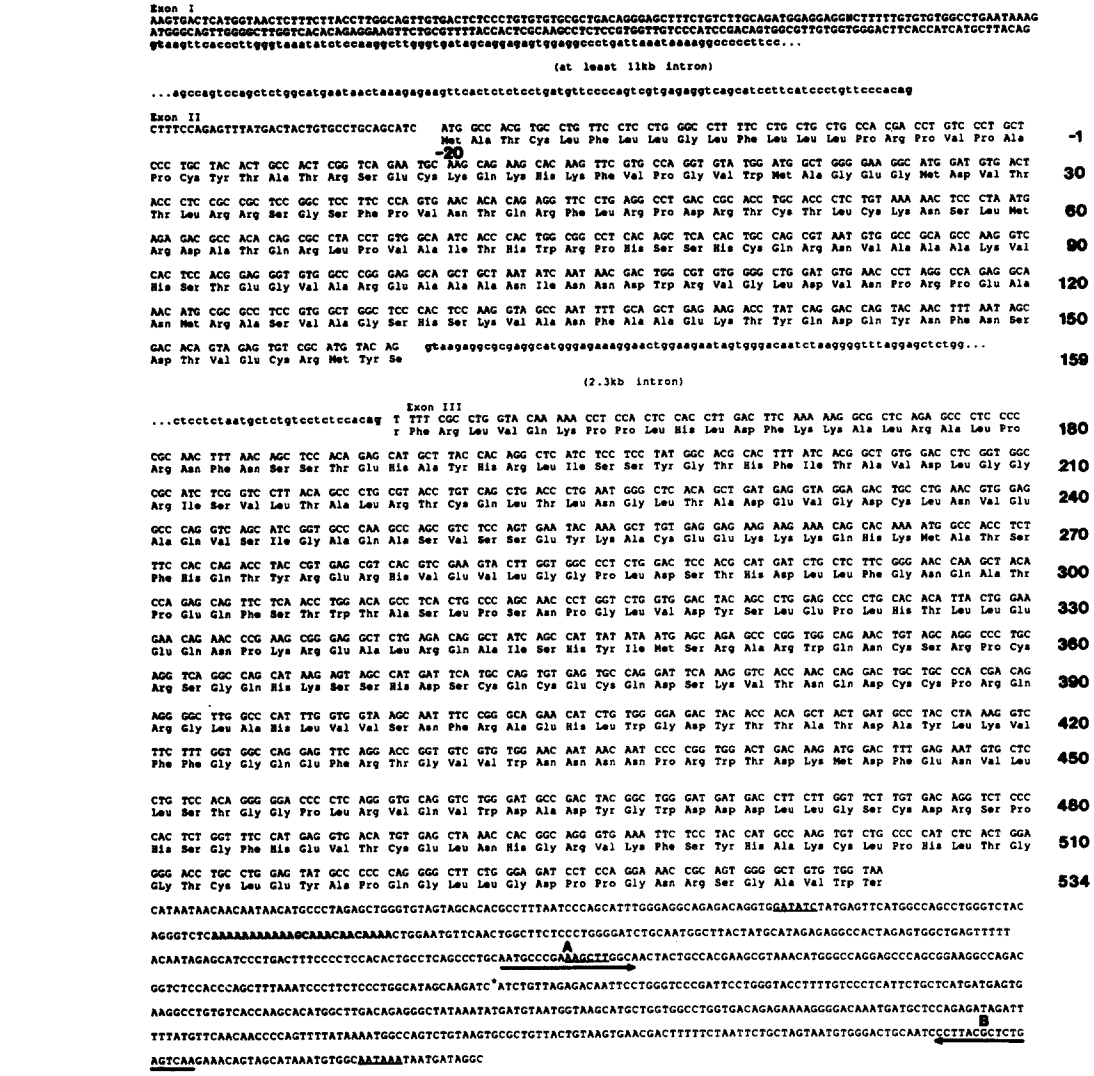

FIGURE 2. Nucleotide sequence of mouse *Pfp*. Exons are shown in upper case, introns (partial sequences) in lower case. Numerals at right and in the text refer to numbered amino acid residues. The first 396 bases of the 3'UT region were sequenced from clone 64, the final 352 bases from a PCR product amplified from CTLL-R8 RNA (see text). Eco RV (gatatc) and Hind III (aagctt) sites and the polyadenylation signal are underlined. An A-rich tract in the 3'UT region is shown in bold characters. The asterisk indicates the position of the Sal I cloning site in EMBL 3 DNA, which was used to generate a 3'UT probe for RNA blot studies (Fig. 4). Sequences that were used to construct sense (*A*) and antisense (*B*) oligonucleotides for PCR amplifications are boldly underlined and arrowed. These sequence data have been submitted to the EMBL/GenBank Data Libraries.

the human *C9* gene (corresponding data are not available for the mouse), which has been shown to be spread over at least 80 kb, with its exons interspersed by several introns as large as 20 kb (34). This radically different configuration is also reflected in the size of the exons in the two genes. The "mosaic" nature of the C9 protein

is more faithfully, though not exactly, represented in the domains encoded by individual exons or clusters of exons (34). The series of 11 exons in the *C9* gene vary in length from only 106 to 247 bp in length, while the two coding exons of *Pfp* are 572 and 1,877 bp long. In contrast to *C9*, the large intron in *Pfp* does not interrupt protein coding sequences and is accommodated at the 5′ end of the gene (Figs. 1 and 2).

*The 5′ Region.* Southern blotting analysis performed with a probe derived from the extreme 5′ end of the published cDNA sequence revealed that clone 64 lacked most of the 5′ noncoding regions of the *Pfp* gene. The probe used for this analysis was a 192-bp fragment derived from Pst I digestion of another cDNA clone, designated PFP-15 (Fig. 1). This 5′ probe was therefore used to identify reactive clones from the same genomic library. Five independent clones (represented by clone 67 in Fig. 1) were isolated, and all five failed to hybridize with the coding region probe (not shown). These observations were consistent with the presence of at least one large intron at the 5′ end of the gene. This contention was further supported by DNA blot analysis of mouse genomic DNA, using 5′ and coding region probes (Fig. 3). DNA that was digested with any of eight restriction endonucleases (seven with a six-base recognition sequence, and Taq I) failed to yield a single common hybridizing band. Identical results were obtained with enzymes that cut DNA less frequently, such as Xho I, as well as with several other enzymes (not shown). Although similarly sized bands were observed when the DNA was restricted with Sac I (Fig. 3, *A* and *B*), these bands were found to be nonidentical in subsequent experiments (not shown). It later became clear that PFP coding regions present in clone 64 were situated at the extreme 3′ end of the DNA insert (see below). As clone 64 contains an insert

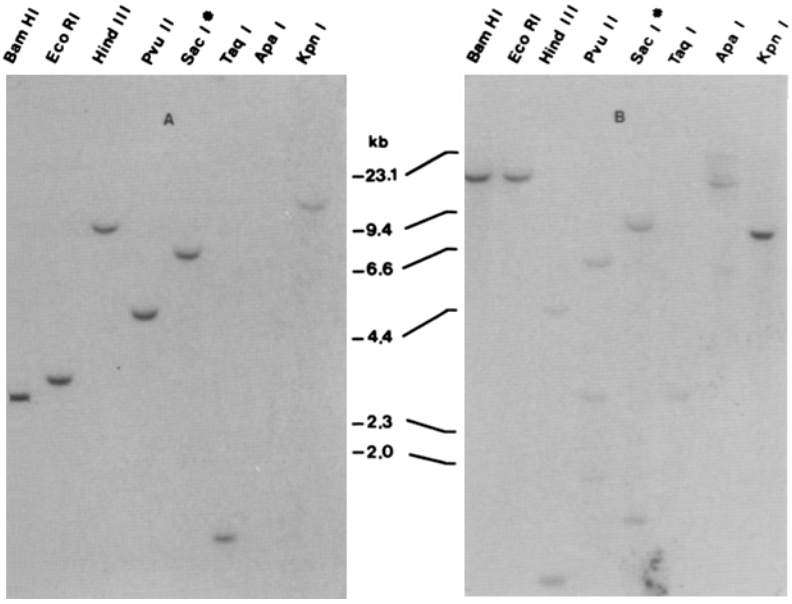


FIGURE 3. Blot analysis of mouse genomic DNA digested with the indicated restriction endonucleases and hybridized with a 5′ cDNA probe (*A*) and with a coding region probe, PFP-6 (*B*). (*) Though the hybridization patterns for Sac I appear similar, the bands were found to be nonidentical (see text).

of at least 16 kb, and the total length of exons II and III and the intron between them is <5 kb, the 5′ intron has a minimum length of 11 kb.

To verify that the 5′ Pst I fragment used in the above analyses did genuinely encode *Pfp* sequences, the fragment was used to probe RNA derived from a cloned cytotoxic T cell line, CTLL-R8 (29). The 5′ probe detected an abundant RNA species of ~2.8 kb (Fig. 4, lane *1*), which migrated identically with that detected by the probe for the coding region (Fig. 4, lane *2*). This result was confirmed using RNA from three other cytotoxic and noncytotoxic cell lines (not shown).

*The 3′ Region.* Analysis of the genomic insert of clone 64 revealed that this clone terminated after the first 0.3 kb of 3′UT sequence, and lacked the putative polyadenylation signal. As attempts at isolating this part of the gene using 3′ probes were fruitless, an alternative strategy was pursued. 3′UT sequences were amplified from CTLL-R8 RNA, using a PCR reaction that was primed with a specific (sense) oligonucleotide primer at only one end (35). This oligonucleotide (Fig. 2 *A*) was constructed from the sequence at the extreme 3′ end of clone 64, while the PCR reaction was anchored at the 3′ end using $(dT)_{17}$, to which a Not I cloning site was attached at the 5′ end, in order to facilitate subcloning of the resultant DNA fragment (see Materials and Methods). A species of ~0.4 kb resulted from this amplification. Substitution of human RNA containing *Pfp* transcripts resulted in no such amplification product (not shown). The fragment was sequenced and demonstrated to encode a classical polyadenylation signal (AATAAA) and poly(A) tail. To determine whether the remainder of the 3′UT region was encoded on exon III or on another exon, a further antisense oligonucleotide was synthesized immediately upstream of the polyadenylation signal (Fig. 2 *B*). Oligonucleotides A and B were then used to amplify intervening sequences in the mouse genome (30). A DNA fragment identical in size to that amplified from RNA resulted, excluding the possibility of an intron at the 3′ end (not shown).

Comparison of the 3′UT sequence derived in this study and that presented previously (15) showed some interesting discrepancies. The 3′UT region of PFP-6 cDNA (16) terminates at a long poly(A) stretch commencing 127 bp beyond the termination codon (Fig. 2). This sequence is not preceded by a polyadenylation signal and is clearly not the genuine poly(A) tail. The sequence presented by Lowrey et al.

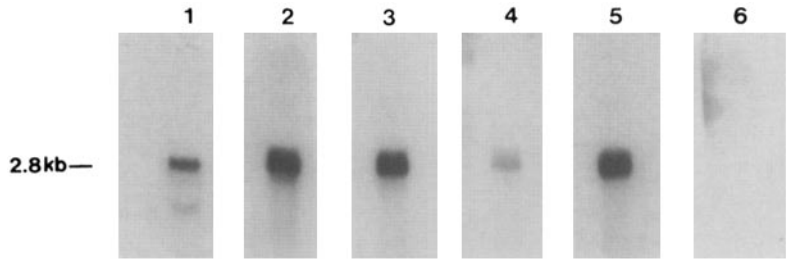


FIGURE 4. Blot analysis of poly(A)$^+$ RNA derived from the mouse cytotoxic cell line, CTLL-R8. RNA (1 μg) was hybridized with a 5′ Pst I fragment from PFP-15 (lane *1*), with a coding region probe, PFP-6 (lane *2*), Eco RV-Hind III (lane *3*), and Hind III-Sal I fragments (lane *4*) from the 3′UT region of clone 64, with a DNA fragment amplified from mouse CTLL-R8 RNA by PCR (lane *5*), and with a 0.3-kb fragment of EMBL 3 DNA (lane *6*).

(15) is virtually identical upstream of this point, but then diverges markedly from that presented in the current study. To check that the extreme 3′ portion of clone 64 (i.e., downstream of the point of divergence) did, in fact, code for *Pfp* exon sequences, three probes derived from the putative 3′UT region of the *Pfp* gene were used in RNA blot analysis. The first two probes consisted of Eco RV–Hind III and Hind III–Sal I fragments from the extreme 3′ end of clone 64 (Fig. 2), and the third was the PCR product derived from amplification of mouse CTLL-R8 RNA (i.e., between oligonucleotides A and B in Fig. 2). Each of these probes detected a 2.8-kb RNA species, which was identical in size to that detected with the coding region probe, PFP-6 (Fig. 4, lanes *2–5*). The reason for the discrepancy between the sequence presented here and that of Lowrey et al. (15) is unclear. One possibility is a polymorphism of the *Pfp* gene, related to the strain of mouse studied in each case. It is conceivable that the A-rich tract at the point of divergence of the two sequences could predispose to an increased recombination frequency, either in vivo, or as an in vitro cloning artifact. The possibility that the previously reported sequence represents an intron seems unlikely, in view of the amplification of genomic sequences presented in this study, and since a typical intron-exon splice consensus motif is not apparent near the point of divergence of the two sequences.

*Chromosomal Localization of the Pfp Gene.* To determine the chromosomal assignment of the *Pfp* gene, the PFP-6 probe was hybridized with DNA derived from a panel of 18 mouse–Chinese hamster hybrid cell lines and a single mouse-rat hybrid cell line (31, 32) (Fig. 5; data for only four hybrid cell lines shown here, two of which hybridized positively for *Pfp*).

Digestion with Eco RI produced hybridizing bands of 8.2 kb in hamster DNA, 16 kb in rat DNA, and 18 kb in mouse DNA. Hybrids were scored for presence or absence of the 18-kb band. This informative band was found to correlate exactly with the presence or absence of chromosome 10 in each of the cell lines (Table I). Included in this set of hybrids were two, HM45 and HM80, that were subclones of the same primary hybrid. These cell lines were karyotypically identical, except

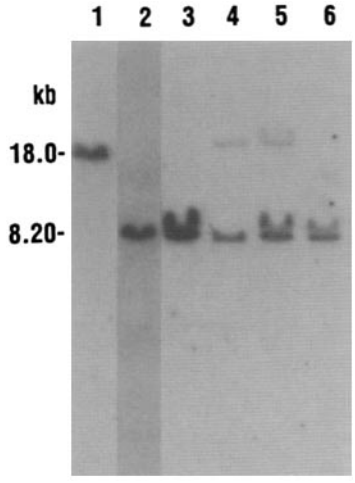


FIGURE 5. Blot analysis of DNA derived from a selected set of mouse–Chinese hamster hybrid cell lines (lanes *3–6*), hybridized with the PFP-6 coding region probe. Lanes *1* and *2* contain control mouse and Chinese hamster liver DNA, respectively. An informative hybridizing band for mouse *Pfp* is indicated at 18 kb.

TABLE I

*Segregation of Mouse Chromosomes with Pfp Crosshybridizing Bands in 19 Somatic Cell Hybrids*

| Mouse chromosome | No. of hybrids (DNA hybridization/chromosome) | | | | Percent discordancy |
|---|---|---|---|---|---|
| | + / + | − / − | + / − | − / + | |
| 1 | 2 | 5 | 1 | 9 | 58 |
| 2 | 3 | 6 | 0 | 9 | 50 |
| 3 | 1 | 6 | 0 | 4 | 36 |
| 4 | 3 | 8 | 0 | 6 | 35 |
| 5 | 0 | 10 | 3 | 3 | 37 |
| 6 | 2 | 6 | 1 | 9 | 56 |
| 7 | 3 | 5 | 0 | 10 | 56 |
| 8 | 2 | 10 | 1 | 4 | 29 |
| 9 | 2 | 10 | 1 | 5 | 33 |
| 10 | 3 | 15 | 0 | 0 | 0 |
| 11 | 0 | 11 | 2 | 1 | 21 |
| 12 | 1 | 3 | 0 | 7 | 64 |
| 13 | 2 | 6 | 1 | 9 | 56 |
| 14 | 1 | 11 | 1 | 4 | 29 |
| 15 | 1 | 1 | 0 | 8 | 80 |
| 16 | 1 | 9 | 1 | 5 | 38 |
| 17 | 2 | 5 | 1 | 8 | 56 |
| 18 | 1 | 6 | 0 | 7 | 50 |
| 19 | 1 | 7 | 1 | 7 | 50 |
| X | 3 | 6 | 0 | 8 | 47 |

This table is compiled from 19 cell hybrids involving either mouse-Chinese hamster or mouse-rat fusions (see Materials and Methods). Symbols indicate the presence ( + /) or absence ( − /) of mouse *Pfp* restriction fragments as related to the presence (/ + ) or absence (/ − ) of a particular mouse chromosome. The number of discordant observations is the sum of the + / − and − / + observations. The lowest discordance is the basis for chromosomal assignment.

that HM80 contains mouse chromosome 10 and HM45 does not. Only HM80 produced the mouse *Pfp* fragment (data not shown). The *Pfp* gene was therefore assigned to chromosome 10.

To confirm and extend the chromosome 10 assignment, progeny from an interspecifies backcross were also typed for *Pfp*. Digestion with Apa I produced 17.8- and 7.2-kb crossreactive fragments in NFS/N mice and 17.8- and 10.2-kb fragments in *musculus* mice. The NFS/N 7.2-kb fragment was detected in 25 of 70 progeny of the backcross (NFS/N × *musculus*)$F_1$ × *musculus*. Analysis of these same mice showed linkage of *Pfp* to the oncogene locus *Myb*, which was detected by the presence of an 8.4-kb Sca I fragment inherited from the NFS parent (recombination = 16/64 = 0.25 +/−0.054). This provides evidence for genetic linkage to *Myb*, which has been mapped to the centromeric region of chromosome 10 (36, 37).

## Discussion

This study reports the genomic organization of mouse *Pfp*, which has been found to be surprisingly uncomplicated. The gene consists of three exons with two inter-

vening introns, and the entire protein coding information is contained within a span of <5 kb. This format shows close similarity with preliminary reports of the structure of human *Pfp*, as recently presented by Lichtenfeld et al. (38). These two studies reveal marked differences from the intron-exon organization of human *C9* (34).

With the isolation of cDNA clones encoding PFP and C9 (13–16, 39, 40), the immunological, physico-chemical, and functional similarities of the two molecules were found to reside over a surprisingly small portion (about one-third) of the two molecules. This dichotomy between structural analogy and overall discordance of the amino acid sequences is extended by the comparison of the respective genes. PFP and C9 polypeptides have two analogous domains. The first is believed to allow membrane insertion and to comprise an amphipathic $\alpha$-helical structure. The second is a contiguous cysteine-rich domain that has homology to that found in the EGF family of cysteine-rich motifs. Unlike C9 and the other members of the membrane attack complex, C8$\alpha$, C8$\beta$, and C7, PFP lacks thrombospondin and LDL receptor-like domains near its $NH_2$ terminus (41). The region of similarity between the two proteins comprises amino acids 160–390 of PFP (approximately the central one-third of the molecule) and 320–560 of C9 (the COOH-terminal one-third) (13, 15, 16). In the *C9* gene, this region is encoded on four exons (exons 7–10), three of which, exons 7, 8, and 9, are relatively closely clustered, but separated from exon 10 by a large intron (34). In sharp contrast, the corresponding regions of *Pfp* are represented in continuo in approximately the 5′ half of the open reading frame of exon III. The PFP-specific $NH_2$- and COOH-terminal portions of the molecule are encoded on exon II and on the remainder of the coding portion of exon III. These two peripheral domains have no similarity to C9 and are presumed to confer structural and functional characteristics unique to PFP.

The marked differences between the *C9* and *Pfp* genes have implications for their evolutionary relationship. Though not reflected with perfect fidelity in their intron-exon boundaries (34), the nature of the genes encoding the late complement components seems generally consistent with the notion that domains of these mosaic proteins correspond to discrete units of genetic information that may have coalesced to form a single gene, after a series of recombinations, duplications, and deletions. The C6, C7, C8$\alpha$, C8$\beta$, and C9 polypeptides all have structures consistent with this "domain" hypothesis (39–46). It is still unclear whether this model is applicable to PFP, as the two peripheral portions of this molecule are of unknown function and genetic origin. Taken at the simplest level of comparison, PFP and C9 share two domains that fulfill roles essential to their overall functions, namely those providing for lipid insertion and polymerization. It is conceivable that these domains arose from a primordial ancestor gene encoding a polypeptide with a more general function in immunity. A series of subsequent intron insertions (*C9*) and deletions (*Pfp*) would then account for the present divergent configurations of the two genes. The mechanisms responsible for these arrangements would presumably be analogous to those responsible for generating other multigene families, such as the Ig superfamily (47). This gene family, members of which are scattered over many chromosomes, encodes a widely divergent group of polypeptides that display one or more copies of a common Ig-like domain, as well as other features that define and differentiate individual family members (47). Alternatively, the markedly differing gene organizations for *Pfp* and *C9* could suggest that their common features may have arisen

independently, by a process of convergent evolution. Precedents for gene duplications and likely instances of convergent evolution in the complement *C2* and *C4* genes have previously been described in mouse and in man (48, 49). These observations would be consistent with the clearly important biological roles played by the complement cascade proteins and by cytotoxic effector molecules.

An important feature of the *Pfp* gene and of other genes involved in cellular cytotoxicity is their inducibility as a function of T cell activation and proliferation, and in response to a wide variety of stimuli, including antibody binding, exposure to certain cytokines such as IL-2, and as a function of the activity of intracellular protein kinase C (50). It is anticipated that examination of 5′ controlling regions of the *Pfp* gene will yield important information regarding the role played by cytotoxic mediators in the normal immune response.

## Summary

Genomic clones encompassing the entire coding region of the mouse lymphocyte pore-forming protein gene (*Pfp*) have been isolated and used to determine its intron-exon organization. In contrast to *C9*, *Pfp* has a simple structure, consisting of only three exons (two of which encode polypeptide), a large 5′ intron, and a single, smaller intron that is situated approximately one-third of the way through the protein-coding portions of the gene. The regions encoding the homologous domains of PFP and C9 are encoded on exons 7, 8, 9, and 10 of *C9*, but form only approximately half of the open reading frame of exon III in *Pfp*. Although encoding polypeptides with related functions, the two genes possess such sharply contrasting structures as to suggest that their analogous regions may have risen independently, by a process of convergent evolution. Using a panel of somatic cell hybrid cell lines, *Pfp* has been mapped to chromosome 10.

We thank Renate Klatt for excellent technical assistance, Dr. Perrin C. White for his critical review of the manuscript, and Drs. Ralph M. Steinman and Zanvil A. Cohn for constant advice and encouragement.

*Received for publication 19 September 1989.*

## References

1. Young, J. D.-E, and C.-C. Liu. 1988. Multiple mechanisms of lymphocyte-mediated killing. *Immunol. Today.* 9:140.
2. Clark, W. R. 1988. Perforin: a primary or auxiliary lytic mechanism? *Immunol. Today.* 9:101.
3. Henkart, P. A. 1985. Mechanism of lymphocyte-mediated cytotoxicity. *Annu. Rev. Immunol.* 3:31.
4. Young, J. D.-E, and Z. A. Cohn. 1986. Cell-mediated killing: a common mechanism. *Cell.* 46:641.
5. Podack, E. R. 1985. The molecular mechanism of lymphocyte-mediated cytotoxicity. *Immunol. Today.* 6:21.
6. Podack, E. R., and G. Dennert. 1983. Assembly of two types of tubules with putative cytolytic function by cloned natural killer cells. *Nature (Lond.).* 302:442.
7. Muller-Eberhard, H. J. 1988. The molecular basis of target cell killing by human lymphocytes and killer cell self protection. *Immunol. Rev.* 103:99.

8. Pasternack, M. S., and H. N. Eisen. 1985. A novel serine esterase expressed by cytotoxic T lymphocytes. *Nature (Lond.).* 314:743.
9. Masson, D., and J. Tschopp. 1987. A family of serine esterases in lytic granules of cytotoxic T lymphocytes. *Cell.* 49:679.
10. Tschopp, J., D. Masson, and K. K. Stanley. 1986. Structural/functional similarity between proteins involved in complement- and cytotoxic T-lymphocyte-mediated cytolysis. *Nature (Lond.).* 322:831.
11. Young, J. D.-E, Z. A. Cohn, and E. R. Podack. 1986. The ninth component of complement and the pore-forming protein (perforin 1) from cytotoxic T cells: structural, immunological and functional characteristics. *Science (Wash. DC).* 233:184.
12. Liu, C.-C., B. Perussia, Z. A. Cohn, and J. D.-E Young. 1986. Identification and characterization of a pore-forming protein of peripheral blood natural killer cells. *J. Exp. Med.* 164:2061.
13. Shinkai, Y., K. Takio, and K. Okumura. 1988. Homology of perforin to the ninth component of complement (C9). *Nature (Lond.).* 334:525.
14. Lichtenfeld, M. G., K. J. Olsen, P. Lu, D. M. Lowrey, A. Hameed, H. Hengartner, and E. R. Podack. 1988. Structure and function of human perforin. *Nature (Lond.).* 335:448.
15. Lowrey, D. M., T. Aebischer, K. Olsen, M. Lichtenfeld, F. Rupp, H. Hengartner, and E. R. Podack. 1989. Cloning, analysis, and expression of murine perforin 1 cDNA, a component of cytolytic T-cell granules with homology to complement component C9. *Proc. Natl. Acad. Sci. USA.* 86:247.
16. Kwon, B. S., M. Wakulchick, C.-C. Liu, P. M. Persechini, J. A. Trapani, A. K. Haq, Y. Kim, and J. D.-E Young. 1989. The structure of the lymphocyte pore-forming protein perforin. *Biochem. Biophys. Res. Commun.* 158:1.
17. Brunet, J. F., M. Dosseto, F. Denizot, M-G. Mattei, W. R. Clark, T. M. Haggi, P. Ferrier, M. Nabholtz, A.-M. Schmitt-Verhulst, M.-F. Luciani, and P. Golstein. 1986. The inducible cytotoxic T-lymphocyte-associated gene transcript, CTLA-1: sequence and localization to mouse chromosome 14. *Nature (Lond.).* 322:268.
18. Gershenfeld, H. K., and I. L. Weissman. 1986. Cloning of a cDNA for a T-cell specific serine protease from a cytotoxic T-lymphocyte. *Science (Wash. DC).* 232:854.
19. Gershenfeld, H. K., R. J. Hershberger, T. B. Shows, and I. L. Weissman. 1986. Cloning and chromosomal assignment of a human cDNA encoding a T-cell and natural killer cell-specific trypsin-like serine protease. *Proc. Natl. Acad. Sci. USA.* 85:1184.
20. Jenne, D., C. Ray, D. Masson, K. K. Stanley, J. Herz, G. Plaetinck, and J. Tschopp. 1988. cDNA cloning of granzyme C, a granule-associated serine protease of cytotoxic T-lymphocytes. *J. Immunol.* 140:318.
21. Jenne, D., C. Ray, D. Masson, K. K. Stanley, J. Herz, G. Plaetinck, and J. Tschopp. 1988. Identification and sequencing of cDNA clones encoding the granule-associated serine protease granzymes D, E, and F of cytotoxic T-lymphocytes. *Proc. Natl. Acad. Sci. USA.* 83:150.
22. Lobe, C. G., B. B. Finlay, W. Paranchych, V. A. Paetkau, and R. C. Bleakley. 1986. Novel serine esterases encoded by two cytotoxic T-lymphocyte-specific genes. *Science (Wash. DC).* 232:858.
23. Kwon, B. S., D. Kestler, E. Lee, M. Wakulchik, and J. D.-E Young. 1988. Isolation and sequence analysis of serine protease cDNAs from mouse cytotoxic T lymphocytes. *J. Exp. Med.* 168:1839.
24. Schmid, J., and C. Weissman. 1987. Induction of mRNA for serine protease and a beta-thromboglobulin-like protein in mitogen stimulated leukocytes. *J. Immunol.* 139:250.
25. Trapani, J. A., J. L. Klein, T. B. Shows, and B. Dupont. 1988. Molecular cloning of an inducible serine esterase gene from human cytotoxic lymphocytes. *Proc. Natl. Acad. Sci. USA.* 85:6924.

26. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463.
27. Chen, E. Y., and P. H. Seeburg. 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. *DNA (NY).* 4:165.
28. Levinson, A., D. Silver, and B. Seed. 1984. Minimal size plasmids containing an M13 origin for production of single stranded transducing particles. *J. Mol. Appl. Genet.* 2:507.
29. Young, J. D.-E, E. R. Podack, and Z. A. Cohn. 1986. Properties of a purified pore-forming protein (perforin 1) isolated from H-2-restricted cytotoxic T cell granules. *J. Exp. Med.* 164:144.
30. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science (Wash. DC).* 239:487.
31. Hoggan, M. D., N. F. Halden, C. E. Buckler, and C. A. Kozak. 1988. Genetic mapping of the mouse c-*fms* proto-oncogene to chromosome 18. *J. Virol.* 62:1055.
32. Killary, A. M., and R. E. K. Fournier. 1984. A genetic analysis of extinction: trans dominant loci regulate expression of liver-specific traits in hepatoma hybrid cells. *Cell.* 38:523.
33. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
34. Marazziti, D., G. Eggertsen, G. H. Fey, and K. K. Stanley. 1988. Relationships between the gene and protein structure in human complement component C9. *Biochemistry.* 27:6529.
35. Loh, E. Y., J. F. Elliott, S. Cwirla, L. L. Lanier, and M. M. Davis. 1989. Polymerase chain reaction with single-sided specificity: analysis of T cell receptor delta chain. *Science (Wash. DC).* 243:217.
36. Mitchell, M., D. Simon, N. Affara, M. Ferguson-Smith, P. Avner, and C. Bishop. 1989. Localization of murine X and autosomal sequences homologous to the human Y located testis-determining region. *Genetics.* 121:803.
37. Bode, V. C., J. D. McDonald, J.-L. Guenet, and D. Simon. 1988. *hph-1*: a mouse mutant with hereditary hyperphenylalanemia induced by ethylnitrosourea mutagenesis. *Genetics.* 118:299.
38. Lichtenfeld, M. G., M. Birnbaum, K. J. Olsen, P. Lu, and E. R. Podack. 1989. Structure of the human perforin gene. *In* Mechanisms of Cell-mediated Cytotoxicity. Abstracts from the Seventh International Congress of Immunology, Berlin, FRG. 333.
39. DiScipio, R. G., M. H. Gehring, E. R. Podack, C. C. Kan, T. E. Hugli, and G. H. Fey. 1984. Nucleotide sequence of cDNA and derived amino acid sequence of human complement component C9. *Proc. Natl. Acad. Sci. USA.* 81:7298.
40. Stanley, K. K., h. P. Kocher, J. P. Luzio, P. Jackson, and J. Tschopp. 1985. The sequence and topology of human complement component C9. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:375.
41. Stanley, K. K. 1989. A family of killer proteins. *Nature (Lond.).* 334:475.
42. Howard, O. M. Z., A. G. Rao, and J. M. Sodetz. 1987. Complementary DNA and derived amino acid sequence of the beta subunit of human complement component C8: identification of a close structural and ancestral relationship to the alpha subunit and C9. *Biochemistry.* 26:3565.
43. Patthy, L. 1988. Detecting distant homologies of mosaic proteins. Analysis of the sequences of thrombomodulin, thrombospondin, complement components C9, C8alpha and C8beta, vitronectin and plasma cell membrane glycoprotein, PC-1. *J. Mol. Biol.* 202:689.
44. Stanley, K. K., and J. Herz. 1987. Topological mapping of complement component C9 by recombinant DNA techniques suggests a novel mechanism for its insertion into target membranes. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:1951.
45. DiScipio, R. G., D. N. Chakravarti, H. J. Muller-Eberhard, and G. H. Fey. 1988. Struc-

ture of the human complement component C7 and the C5b-C7 complex. *J. Biol. Chem.* 263:549.

46. Chakravarti, D. N., B. Chakravarti, C. A. Parra, and H. J. Muller-Eberhard. 1989. Structural homology of complement component C6 with other channel-forming proteins of complement. *Proc. Natl. Acad. Sci. USA.* 86:2799.
47. Williams, A. F., and A. N. Barclay. 1988. The immunoglobulin superfamily-domains for cell surface recognition. *Annu. Rev. Immunol.* 6:381.
48. Nonaka, M., K. Nakayama, Y. D. Yeul, A. Shimuzu, and M. Takahashi. 1985. Molecular cloning and characterization of complementary and genomic clones for mouse C4 and Slp. *Immunol. Rev.* 87:81.
49. Campbell, R. D., and D. R. Bentley. 1985. The structure and genetics of the C2 and factor B genes. *Immunol. Rev.* 87:19.
50. Liu, C.-C., S. Rafii, A. Granelli-Piperno, J. A. Trapani, and J. D.-E Young. 1989. Perforin and serine esterase gene expression in stimulated human T cells. Kinetics, mitogen requirements, and effects of cyclosporin A. *J. Exp. Med.* 170:2105.