Immunogenetics (1995) 42: 101–111


ORIGINAL PAPER

**Mark J. Smyth · Mark D. Hulett · Kevin Y. T. Thia**
**Howard A. Young · Thomas J. Sayers**
**Clive R. D. Carter · Joseph A. Trapani**

# Cloning and characterization of a novel *NK* cell-specific serine protease gene and its functional 5′-flanking sequences

Received: 19 January 1995 / Revised: 17 March 1995

**Abstract** Rat natural killer cell Met-ase-1 (*RNK-Met-1*) is a 30000 $M_r$ serine protease (granzyme) found in the cytolytic granules of CD3⁻ large granular lymphocytes (LGL) with natural killer (NK) activity. To characterize the genomic sequences responsible for the CD3⁻ LGL-restricted expression of this gene, we screened a rat genomic library with *RNK-Met-1* cDNA, and obtained bacteriophage clones that contained the *RNK-Met-1* gene. The *RNK-Met-1* gene comprises 5 exons and spans approximately 5.2 kilobases (kb), exhibiting a similar structural organization to a class of CTL-serine proteases with protease catalytic residues encoded near the borders of exons 2, 3, and 5. The 5′-flanking region of the *RNK-Met-1* gene contains a number of putative promoter and enhancer regulatory elements and shares several regions of homology with the 5′-flanking region of the mouse perforin gene. We have prepared nested deletions from approximately 3.3 kb of the 5′-flanking region of the *RNK-Met-1* gene, and inserted these upstream of the chloramphenicol acetyltransferase (*CAT*) reporter gene. These 5′-flanking *RNK-Met-1-CAT* constructs were transiently transfected into rat LGL leukemia, T-lymphoma, and basophilic leukemia cell lines. The transcriptional activity of the *RNK-Met-1* 5′-flanking region was strong, restricted to the RNK-16 LGL leukemia and controlled by several positive *cis*-acting regions spread over at least 3.3 kb. The longest and most active 5′-flanking region (–3341 to –33) was also used to drive specific expression of β-galactosidase in RNK-16. These data are consistent with the *NK* cell-specific expression of *RNK-Met-1* and suggest the potential utility of this gene promoter in the development of transgene models of *NK* cell biology in vivo.

The nucleotide sequence data reported in this paper have been submitted to the EMBL/GenBank nucleotide sequence database and have been assigned the accession number L38482

M. J. Smyth[1] (✉) · M. D. Hulett · K. Y. T. Thia · J. A. Trapani
Cellular Cytotoxicity Laboratory, Austin Research Institute, Austin Hospital, Heidelberg, 3084 Victoria, Australia

H. A. Young
Laboratory of Experimental Immunology, Biological Response Modifiers Program, NCI-Frederick Cancer Research and Development Center, Frederick, MD 21702-1201, USA

T. J. Sayers · C. R. D. Carter
Biological Carcinogenesis and Development Program, PRI/DynCorp, NCI-Frederick Cancer Research and Development Center, Frederick, MD 21702–1201, USA

*Present address:*
[1] Cellular Cytotoxicity Laboratory, Austin Research Institute, Austin Hospital, Heidelberg, 3084 Victoria, Australia

## Introduction

A variety of serine proteases have been found to reside within the cytoplasmic granules of cytotoxic lymphocytes (Jenne and Tschopp 1988). These granules are believed to be involved in at least one lytic mechanism employed by both CD3⁻ large granular lymphocytes (LGL) with *NK* activity and cytotoxic T lymphocytes (CTL) to kill target cells (Tschopp and Nabholz 1990). The serine proteases have been localized to the same lytic granules as the pore-forming protein, perforin, which causes transmembrane lesions in target cells (Podack et al. 1991). Unlike perforin, isolated serine proteases are not intrinsically lytic. However, a role for serine proteases in cellular cytotoxicity has been supported by: the ability of protease inhibitors to completely abrogate lymphocyte cytotoxicity (Hudig et al. 1991); the demonstration that effector cells cotransfected with serine proteases and perforin can initiate DNA fragmentation in target cells (Shiver et al. 1992); and the DNA fragmentation observed in target cells treated with a sub-lytic concentration of perforin and granzyme *B* (Shi et al. 1992 a, b). These studies have been further supported by the generation and characterization of independent gene knock-out mutations in mice of the perforin (Kagi et al. 1994) and granzyme *B* (Heusel et al. 1994) genes.

Until recently, across all species, granzyme genes were known to encode just three enzyme activities (tryptase, Asp-ase, and chymase specificities). With the development

of model synthetic peptide substrates, we purified a novel serine protease, that cleaves after methionine [(Met-ase) (designated *RNK-Met-1*)], from the cytotoxic granules of the rat LGL cell leukemia, RNK-16. Isolation and cloning of the cDNA encoding *RNK-Met-1* revealed *RNK-Met-1* to be a unique granzyme with a predicted amino acid sequence less than 45% identical to any other member of the serine protease family (Smyth et al. 1992). A cDNA clone encoding a human LGL-specific Met-ase (*Hu-Met-1*) was obtained using the *RNK-Met-1* cDNA clone (Smyth et al. 1993).

Unlike other members of the granzyme family which are highly expressed in activated peripheral T cells, *RNK-Met-1* and *Hu-Met-1* have a restricted expression in cells of CD3$^-$ LGL phenotype. Thus far, *RNK-Met-1* and *Hu-Met-1* gene transcripts have not been detected in other resting or activated primary cell populations or cell lines including thymocytes, CD4$^+$ and CD8$^+$ T cells, B cells, myelomonocytic cells or granulocytes, or in a variety of non-lymphoid tissues (Smyth et al. 1992, 1993, 1995). The evolution of a subfamily of granzymes distinct from those previously described (Jenne and Tschopp 1988) has been suggested by chromosomal gene mapping studies. The *Hu-Met-1* gene (*GZMM*) is located on chromosome 19p13 (Baker et al. 1994; Pilat et al. 1994) and the equivalent mouse *LMet1* gene is located on a syntenic region of chromosome 10 (Thia et al. 1995; Pilat et al. 1994). Interestingly, both genes are closely linked to a cluster of neutrophil elastase-like serine proteases (Zimmer et al. 1992). Regulatory elements that confer LGL-specificity have yet to be defined in any gene. Therefore, to address the molecular basis for the LGL-specific expression of the *RNK-Met-1* gene, the present report defines the genomic organization and characterizes the 5′-flanking regions involved in regulation of the *RNK-Met-1* gene.

## Materials and methods

*Isolation of rat genomic RNK-Met-1 clones*

A full-length cDNA transcript encoding the rat RNK-Met-1 protein (Smyth et al. 1992), was radiolabeled by random priming (Sambrook et al. 1989) and used to screen a rat genomic library constructed in Charon 4A (kindly provided by A. Albiston, Baker Institute, Melbourne, Australia) using standard protocols (Sambrook et al. 1989). A single reactive phage clone was plaque purified, the phage DNA was isolated, and a number of overlapping rat genomic DNA fragments were subcloned into the pBluescript KS$^+$ (KS$^+$) plasmid vector (Stratagene, La Jolla, CA).

*Nucleotide sequencing*

Sequencing of the gene was performed in both orientations using synthetic oligonucleotides, the dideoxy method (Sanger et al. 1977) and single-stranded and double-stranded templates (Chen and Seeburg 1985). Areas of apparent sequence compression were resequenced, substituting dGTP with dITP in the extension reactions. The entire final sequence was determined from both strands. The approximate size of intron 1 was estimated by restriction enzyme mapping and sequencing of the 6*Taq* I-A (see Figure 1) subclone DNA. DNA sequence alignments were performed by M. Gunnell and G. Smythers at the Biomedical Supercomputer Center of the Frederick Cancer Research and Development Center using BLAST, BESTFIT and GAP programs following comparison of the *RNK-Met-1* nucleotide sequence against GenBank, EMBL, and accessed individual granzyme sequences. An analysis of potential transcriptional regulatory elements in the *RNK-Met-1* 5′-flanking region was performed by M. Sandrin using SIGNAL SCAN (Prestridge 1991).

*Genomic Southern analysis*

High molecular weight genomic DNA was extracted using guanidinium isothiocyanate, as reported previously (Pang et al. 1992). Ten micrograms of DNA was digested with 50 units of the designated restriction enzyme (Boehringer Mannheim, Indianapolis, IN or Life Technologies Inc., Gaithersburg, MD) overnight at 37 °C in the manufacturer's restriction enzyme buffer. DNA was then subjected to electrophoresis in a 0.8% agarose gel, transferred to Magna nylon membrane (MSI, Westboro, MA), and hybridized to the *RNK-Met-1* cDNA in Nylohybe hybridization buffer (Digene Inc., Silver Spring, MD) at 42 °C according to the manufacturer's recommendations. After 24 h of hybridization, the blot was washed for 10 min at 42 °C in 0.3 M sodium chloride, 0.03 M sodium citrate, 0.1% sodium dodecyl sulfate (SDS) and 10 min at 65 °C in 0.03 M sodium chloride, 0.003 M sodium citrate, 0.1% SDS. The membrane was then exposed to Kodak X-OMAT X-ray film for 6 h at –70 °C.

*Northern analysis*

Total cellular RNA was obtained from cell lines and tissues. The rat basophilic leukemia (RBL) was obtained from the American Type Culture Collection, Rockville, MD. A tumor cell line derived from the rat LGL leukemia [(RNK-16)(CRC$^-$)] and the rat T-cell lymphoma, NTD, were kindly provided by C. Reynolds (BRMP, Frederick Cancer Research and Development Center, Frederick, MD). RNK-16 and NTD

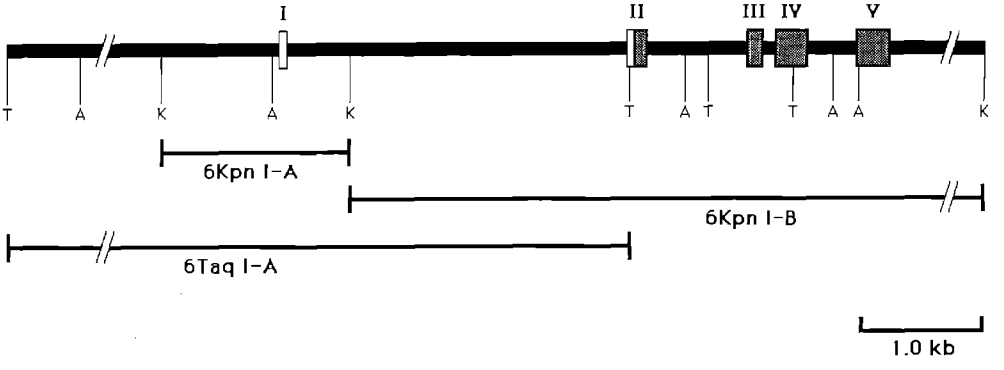



**Fig. 1** Genomic organization of the rat *RNK-Met-1* gene. The exon-intron structure of the *RNK-Met-1* gene is indicated (*top*). The entire gene is contained within one genomic clone λ6rii (not shown) and two overlapping *Kpn* I subclones of this genomic clone. A simple restriction map of these is shown including: A = *Apa* I, K = *Kpn* I, and T = *Taq* I. Sequencing strategy involved subcloning *Kpn* I or *Taq* I fragments into pBluescript KS$^+$ (Materials and methods)

**Exon I**

```
ATG GAG GTC CGC TGT TCC CTG CTG CTC CTG CTG GCC CTG AAA ACA CTG TGG GCA G
Met glu val arg trp ser leu leu leu leu leu ala leu lys thr leu trp ala val                    .. -7

GTGAGCATGAGGGAGAAACAGGCCTAGA.....(~2.7 kb)....CTGAGAAGGGAGGGAGGGCTGTTTGGAGGTGCTGTACCCCGGCTTTCCGCTAGGGCCTGGAAGCCCCTGAATCTGGC
```

**Exon II**

```
 TA GGC AAC AGA TTT GAG GCC CAG ATC ATT GGG GGT CGA GAG GCA GTC CCG CAC TCC CGC CCA TAC ATG GTC TCG CTA CAG
    gly asn arg phe glu ala gln ile ile gly gly arg glu ala val pro his ser arg pro tyr met val ser leu gln .. 19

AAT ACC AAG TCC CAC ATG TGT GGG GGA GTC CTC GTG CAT CAG AAG TGG GTG TTG ACC GCT GCC [CAC] TGC CTG TCT GAA CC
asn thr lys ser his met cys gly gly val leu val his gln lys trp val leu thr ala ala [his] cys leu ser glu pro .. 46

GTGAGTACCCTCTTGATGTCCCACCATGGCTTCAGGCACTCAAGGACTTAGGCCAACCCTCTCCAGGCTGTGTAGGCTGTGGGTTCAGGTTGGGTGGGTCACATGGTACTGGAAAGGACACCTCCT
CCCAACCCTTCCCGTGACACCTGGTAGTGTGTGAAAGTCAGCCTGTCAGTAAGACAGAGCCAGCGTAGGCAGAACCGCCATCAAGAAACCACTGACTACATCCCTCCGGTAGGGATGATAGAGTGA
CTGCTCGTGCCGGGTCTCATCCTCAACGCTAAAGAAAACATGGGCCCTGAACAAAAGGAGGCAGCTCTTGGGTGTGACAGCCTTTAGTCCTAGCCCTCAAGGAGAGCTGAGGCATGTAGATCTCTA
CCTGACCTGCATGACTCCAGACTACGCAGTGAGATCTTATCTTAAAGCCAAAGGGGATTGTTGGCTCATGTGTTCAGAAGCCTGGGCTTCAGGAACGGCTCGATCCAGATATTTACCTAAGCTTCT
CCTTCACGGCTTGTAAGTGTGAGGTAGGAGTCCTGTAGAAGGAGGTTATGTCTCTGCTATAGTTTGGATGAGTGCAGGACGTCAACACAAGCTGACTGACCTCAAGCCTAATCACAGAAGAGAGGC
TGCAGTTGAGGGGAGAGGCATTAAATCCCCCAATCTGACGCTAGGTGGGAACCAGAGCAGAAGTCTGGTTGGGAGAGGAGGACTGAGAAGATGGGAGATGACCTTCCAGGAGTGGGCTGGGGAGGC
TGCCAAGGATGTGTGGAAAGGGAGCTCTCGGGCTGTGGGCGGCTGTGGCGGCTGTGGGTGAGCAAGGTTTGCATTGCAG
```

**Exon III**

```
  G CTA CAG CAG CTG AAG CTG GTG TTC GGC CTG CAC AGC CTT CAT GAT CCC CAA GAT CCT GGC CTT ACC TTC TAC ATC AAG
    leu gln gln leu lys leu val phe gly leu his ser leu his asp pro gln asp pro gly leu thr phe tyr ile lys .. 72

CAA GCC ATT AAA CAC CCT GGT TAC AAC CTC AAA TAC GAG AAC [GAC] CTG GCC CTG CTT AAG
gln ala ile lys his pro gly tyr asn leu lys tyr glu asn [asp] leu ala leu leu lys                    .. 92

GTGCGGAATGGGGCGGGATAGCTGAGGGGGTGGGCACTGAAGGTGGATGAGATCCTGCTTCTGTCTGCCACTCACTGGTCCTCCTGTCCCCACAG
```

**Exon IV**

```
CTG GAT GGA CGG GTG AAG CCC AGC AAG AAT GTC AAA CCA CTG GCT CTG CCA AGA AAG CCC CGA GAC AAG CCT GCA GAA GGC
leu asp gly arg val lys pro ser lys asn val lys pro leu ala leu pro arg lys pro arg asp lys pro ala glu gly ..119

TGG CGG TGT AGC ACG GCT GGA TGG GGT ATA ACC CAC CAG AGG GGA CAG CTA GCC AAC TCC CTG CAG GAG CTC GAC CTG CGT
ser arg cys ser thr ala gly trp gly ile thr his gln arg gly gln leu ala lys ser leu gln glu leu asp leu arg ..146

CTT CTG GAC ACC CGG ATG TGT AAC AAC AGC CGC TTC TGG AAC GGT GTC CTC ACG GAC AGC ATG CTG TGC TTA AAG GCT GGG
leu leu asp thr arg met cys asn asn ser arg phe trp asn gly val leu thr asp ser met leu cys leu lys ala gly ..173

GCC AAG GGC CAA GCT CCT TGC AAG
ala lys gly gln ala pro cys lys                                                                    ..181

GTGACAGGGGTGGCCTGGAACCAGCCTCTAACACCCTGGGGAGTGGAGGAGGGCGGCCTAACACCCACCTCATTTCTCAGATACGGAAACCAAGGCCCCAGAGGGTTTAAAAAAAAAAAAAAAAAAA
AAAAAGGAGTCTAGGTTGGAAGTCAGGAAGCCAGAAAATTGTGGTTTGAACTCCAGTGCCAGGAATTAGTCATCTGGGCCCAGAGCAAAAGTCCTGGGGTAGAAAGGAGGAGAGAAGGGAAAGAGA
CAGGAAGGTGAGCAAGGAGGAAGGCTGCGGGCAAGGGTGCACTAGGACGCTTGAAAGGCTCCAGCACCTCCCAGGTCCCTGGTGCCCAGCTGCAAGCACAAAGGCTGAACCTCTGTGTCCCCCAA
CTCTTTCAG
```

**Exon V**

```
GGT GAC [TCT] GGA GGG CCC CTG GTG TGT GGC AAA GGC AAG GTG GAT GGG ATC CTG TCT TTC AGC TCC AAA AAC TGC ACA GAC
gly asp [ser] gly gly pro leu val cys gly lys gly lys val asp gly ile leu ser phe ser ser lys asn cys thr asp ..208

ATC TTC AAG CCC ACC GTG GCC ACT GCT GTA GCC CCC TAC AGC TCC TGG ATC AGG AAG GTC ATT GGT CGC TGG TCA CCC CAG
ile phe lys pro thr val ala thr ala val ala pro thr ser ser trp ile arg lys val ile gly arg trp ser pro gln ..235

CCT CTG ACC
pro leu thr                                                                                        ..238

TGATGTCCCAAACTATCTGGGACATCATTCTTGATGTCTGGGGCTGGGAAGGGACTAGGTGTGCCTCTGGGGATCAATAAATCCTGATATATCTGTTGAGTCTACCCTTGCCTGCCTCTGGTCTCT
GCAAACCTCCATTCCCTCTATAATGATGATGACCTTTCCGAATCGCTGTCAGGAGAGTCCTGGGACACACGCAG
```

**Fig. 2** A schematic representation of the approximate 5.2 kb encompassing RNK-Met-1. The locations of the active site residues His-, Asp-, and Ser- are *boxed*. DNA sequences of the *RNK-Met-1* gene that includes all of the exons, and some of the intron sequences. Intron 1 is not completely sequenced, its approximate size (in *brackets*) was determined by restriction enzyme mapping (Materials and methods). The inferred amino acid sequences are indicated *below* their corresponding DNA sequences. The stop codon (TGA) and the polyadenylation signal (AATAAA) are *underlined*

cells were cultured in RPMI 1640 medium supplemented with 10% heat-inactivated fetal calf serum, 2 mM glutamine, $5\times10^{-5}$ M 2-mercaptoethanol (Sigma, St. Louis, MO), 100 units/ml penicillin, and 100 µg/ml streptomycin (Gibco Laboratories, Life Technologies, Grand Island, NY). RBL were maintained in DME with the same additives. Highly purified rat LGL were isolated from rat peripheral blood leukocytes by passage over nylon wool (Reynolds et al. 1981) followed by discontinuous density gradient fractionation on Percoll as previously described (Timonen et al. 1982). Highly purified rat T cells were prepared from splenocytes by passage of total splenic leukocytes over nylon wool for enrichment of T cells and LGL. The LGL were depleted by surface labeling of the LGL with the *NK*-specific monoclonal antibody (mAb) 3.2.3 (10 µg/ml) at 4 °C for 1 h followed by passage over R&D T-cell enrichment columns (R&D Systems Inc., Minneapolis, MN) to remove Ab-labeled LGL. Cells were labeled with mAbs 3.2.3 (NK-specific), anti-CD5 (pan T cell), anti-CD4 (helper T cell), MQ-1 (macrophage-specific), and anti-surface Ig (B cell-specific) and then analyzed for immunofluorescence by using a

FACScan analyzer (Becton Dickinson, San Jose, CA) and a Hewlett-Packard Consort 30 microcomputer. Both T-cell and LGL populations were >85% pure on FACS analysis. Total cellular RNA was extracted from normal rat thymus, spleen, and bone marrow by homogenizing the tissues in 4M guanidine thiocyanate. Total cytoplasmic RNA was purified from all cells by the method of Chomczynski and Sacchi (1987). For northern analysis 25 μg RNA was subjected to electrophoresis on a 1.0% agarose formaldehyde gel, then transferred to Nytran (Schleicher and Schuell, Keene, NH). All blots were hybridized to $^{32}$P-labeled *RNK-Met-1* cDNA and human γ-actin (Gunning et al. 1983) cDNA as described previously (Thomas 1983). The blots were then exposed to Kodak X-OMAT AR film for 1 to 7 days at –70 °C.

*Transient transfection*

Exponentially growing cells were washed and resuspended at $5\times10^7$ cells/ml in DMEM with 10–20 μg plasmid DNA. Cells (250 μl suspension) were electroporated at 960 μF and 260 V (RNK-16) or 240 V [(NTD, RBL) (Gene-pulser; BioRad, Richmond, CA)]. After incubation for 10 min on ice, the cells were diluted with warm medium and cultured in a humidified $CO_2$ incubator.

*CAT, β-galactosidase, and control vectors*

Due to the lack of convenient restriction enzyme sites in the *RNK-Met-1* gene, nested promoter fragments were derived by polymerase chain reaction (PCR) and cloned upstream of a promoterless and enhancerless *CAT* gene (pCATbasic; Promega Corporation, Madison, WI). Various 5′ primers were designed to incorporate a *Hin* d III site at the 5′ end for ease of subcloning. The 3′ primer, identical for all but the largest construct, was designed to incorporate a *Sal* I site. The numbering of the constructs refers to the first and last genuine genomic *RNK-Met-1* nucleotide of each construct (+1 corresponds to the ATG codon) (Fig. 1B). The PCR-derived fragments were digested with *Hin* d III and *Sal* I and force-cloned into pCATbasic. All constructs were verified by restriction endonuclease mapping and nucleotide sequencing. The PCR/sequencing primers were as follows:
5′-TTTGTCGACATCCTGGCACAGGAC-3′
[+3 to –13 (antisense 3′ primer)];
5′-TTTGTCGACTTGGTGCAGGCCTTTG-3′
[–284 to –299 (antisense)];
5′-TTTGTCGACGCCTGCCCCTGCTGTG-3′
[–627 to –642 (antisense)];
5′-TTTAAGCTTCTAGCAGAGGACCCACTTG-3′
(–283 to –265);
5′-TTTAAGCTTGAGTGAATTCTCTGACC-3′
(–625 to –609);
5′-TTTAAGCTTAGGGCTACCTGGGTAC-3′
(–954 to –939);
5′-TTTAAGCTTAGACTACCTCACCGAGG-3′
(–1264 to –1248); and
5′-TTTAAGCTTGGACTGGGTGACAGACCG-3′
(–1700 to –1683).
The largest fragment was generated by PCR using 5′-GCTAAGCTTGGGCCCCCAGAGCACGG-3′ (–3341 to –3325) and 5′-GCTGTCGACGGGCCCCACACTTTCTCT-3′ [–33 to –50 antisense 3′ primer)]. The RSVCAT plasmid was obtained from H. Young, NCI-Frederick Cancer Research and Development Center, Frederick, MD. The largest fragment of the *RNK-Met-1* gene 5′-flanking region (–3341 to –33) was also cloned upstream of a promoterless and enhancerless β-galactosidase gene containing a nuclear localization signal (Tan 1991). A 5′ primer was designed to incorporate a *Not* I site at the 5′ end for ease of subcloning and the 3′ primer was designed to incorporate a *Xba* I site. An identical β-galactosidase plasmid driven by an upstream SV40 promoter, rather than the *RNK-Met-1* gene 5′-flanking region, was provided by I. van Driel, Monash University, Prahran, Australia. All plasmid DNA for electroporation was purified by two rounds of banding on CsCl gradients.

*Quantitation of CAT enzyme activity*

CAT enzyme produced by transfected cells was assayed quantitatively using a sandwich enzyme linked immunoadsorbent assay (ELISA) technique (CAT-ELISA kit; Boehringer Mannheim, Sydney, Australia) 48 h after transfection, according to the manufacturer's instructions. The absorbance at 410 nm was read at 40 min intervals on a microtiter ELISA plate reader. Values within the linear range of the standard curve were used for calculation of CAT enzyme activity. The protein concentration of each sample was assessed using the Bradford assay (Bradford 1976). Levels of CAT enzyme were expressed as pg/100 μg of cell extract. Reported results are from a minimum of three independent experiments and were verified with at least two independently produced plasmid preparations.

*Staining procedure for β-galactosidase*

Transfected cells ($5\times10^5$/ml) were centrifuged onto glass slides at 700 rpm for 5 min (Shandon Inc., Pittsburgh, PA). Cells were fixed on the slide for 15 min at room temperature with 4% paraformaldehyde in phosphate buffered saline (PBS), pH 7.3. The slides were washed three times with PBS and cells were covered with 500 μl of staining buffer [1 mg/ml, 5-bromo-4-chloro-3-indoyl-b-D-galactopyranoside (X-gal)], 5 mM potassium ferrocyanide, and 5 mM potassium ferricyanide). After overnight incubation, staining was quantitated under an inverted microscope. Three fields of 400 cells were counted for the number of stained blue nuclei indicating the presence of β-galactosidase. The results were calculated as mean ± standard error of at least four different transfections for each cell line. *RNK-Met-1* promoter activity (–3341 to –33) in each cell line was expressed as a percentage of SV40 promoter activity (from the % β-galactosidase positive cells). The control (Con.) was the enhancerless and promoterless parental β-galactosidase construct.

## Results and discussion

*Isolation of the RNK-Met-1 gene*

In order to isolate genomic clones encoding the *RNK-Met-1* gene, a genomic library was screened with a full length *RNK-Met-1* cDNA (Smyth et al. 1992). DNA derived from a single reactive phage clone, designated λ 6rii, was restricted with *Taq* I, or *Kpn* I (Fig. 1). The 5′ *Taq* I fragment containing 5′ flanking sequences, exon 1, intron 1, and a portion of exon 2 was subcloned into KS+. Adjoining 5′ *Kpn* I (including exon 1) and 3′ *Kpn* I (including intron 1 to polyadenylation signal) fragments were also subcloned into KS+.

*Genomic organization and comparison with other granzyme genes*

We sequenced 5.8 kb of genomic DNA encompassing the gene, including 3.8 kb of 5′-flanking sequence, the 5′ untranslated region, the complete coding region, and a portion of the 3′ untranslated region. Nucleotide sequence analysis of clone 6rii confirmed the presence of the entire *RNK-Met-1* gene, which was approximately 5.2 kb (ATG start site to polyadenylation signal) in length (Fig. 2). By comparison with the cDNA sequence, the gene was found to encode the 238 amino acid protein on five exons, with

four intervening introns of approximately 2.9 kb, 833 base pairs (bp), 95 bp, and 386 bp (Fig. 2). There are distinctions between the granzymes with regard to their genomic organization. The mouse granzyme *A* gene (trypsin-like) is composed of six exons (Hershberger et al. 1992) compared with the genes for mouse granzymes *B*, *C* (Lobe et al. 1988), and *F* (Jenne et al. 1991), *Hu-Met-1* [(*GZMM*) (Pilat et al. 1994)], human granzymes *B* (Klein et al. 1989), and *H* (Haddad et al. 1990), and the structurally and functionally related genes, rat mast cell chymases I and II (Benfey et al. 1987), human adipsin (Min and Spiegelman 1986), human neutrophil elastase (Takahashi et al. 1988), and human neutrophil cathepsin G (Hohn et al. 1989), which share a four intron/five exon gene organization. Seven of eight exon/intron transitions of the *RNK-Met-1* gene fulfill the GT-AG rule (Mount 1982). The introns 1 and 2 occur in phase 1 and 2, respectively, and introns 3 and 4 occur in phase 0. Intron 1 falls at the codon for Gly (–7), which is the first residue of the putative RNK-Met-1 propeptide Gly (–7) – Gln (–1). The second intron falls at the codon for residue 46 and the third and fourth intron occur between the codons for residues 92 and 93 and 181 and 182, respectively. In keeping with the more distant members of the serine protease superfamily, however, each of the residues of the charge relay system is encoded by a separate exon (Benfey et al. 1987; Bleackley et al. 1988; Phillips et al. 1986). The substrate specificity site (in the case of the *RNK-Met-1* gene, an alanine [(residue 178) (Fig. 2)] is encoded within the fourth exon, as it is in *Hu-Met-1* (*GZMM*) and other chymotrypsin-like serine proteases (Bell et al. 1984).

Except for the position of intron 1 and the slightly different length of exons due to the addition or deletion of single amino acids, the positions and phase type of the other introns are completely conserved between the *RNK-Met-1* gene and the following serine protease genes that are expressed in different hematopoietic cell lineages: human lymphocyte granzymes *B* and *H*, mouse lymphocyte granzymes *B* and *C*, neutrophil-associated serine protease cathepsin G, mast cell-associated chymases I and II of the rat, and complement factor D (adipsin) that is expressed in macrophages, adipocytes and hepatocytes. On the basis of the number and position of introns, the *RNK-Met-1* serine protease most closely resembles the recently described novel sixth class of serine proteases (Irwin et al. 1988). In particular, the characteristic peculiar position of the third intron (between a region that encodes a highly conserved peptide segment) in *RNK-Met-1* and this sixth class of genes contrasts with that found in all five other classes of serine protease genes which are not expressed in cells of the immune system. In addition, the unusual position of intron 1 at residue (–7) in *RNK-Met-1* is shared with the human equivalent, *Hu-Met-1* [(*GZMM*) (Pilat et al. 1994)] and the neutrophil elastase-like genes, azurocidin (*AZU1*), proteinase-3 (*PRTN3*), and neutrophil elastase [(*ELA2*) (Zimmer et al. 1992)]. This configuration is therefore a distinguishing feature of this group of serine proteases, indicating their close evolutionary relationship. In support of this relationship, it is intriguing that both the human equivalent of *RNK-Met-1* and three neutrophil elastase-like genes map to human chromosome 19p13 (Smyth et al. 1993; Zimmer et al. 1992; Pilat et al. 1994), suggesting that they may have arisen by duplication and divergence of a common ancestral precursor gene.

### *Southern blot analysis*

To ascertain that our cloned genomic DNA comprised entirely genuine sequences, we compared its restriction pattern with that of the endogenous *RNK-Met-1* gene locus in genomic DNA. In Southern blotting experiments (Fig. 2 and data not shown), each hybridizing band in rat genomic DNA corresponded to fragments of the λ 6rii *RNK-Met-1* clone. No polymorphism was detected when

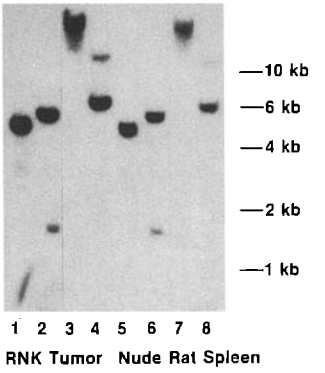
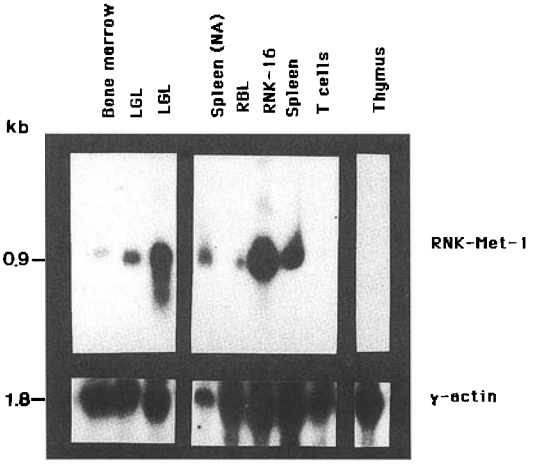
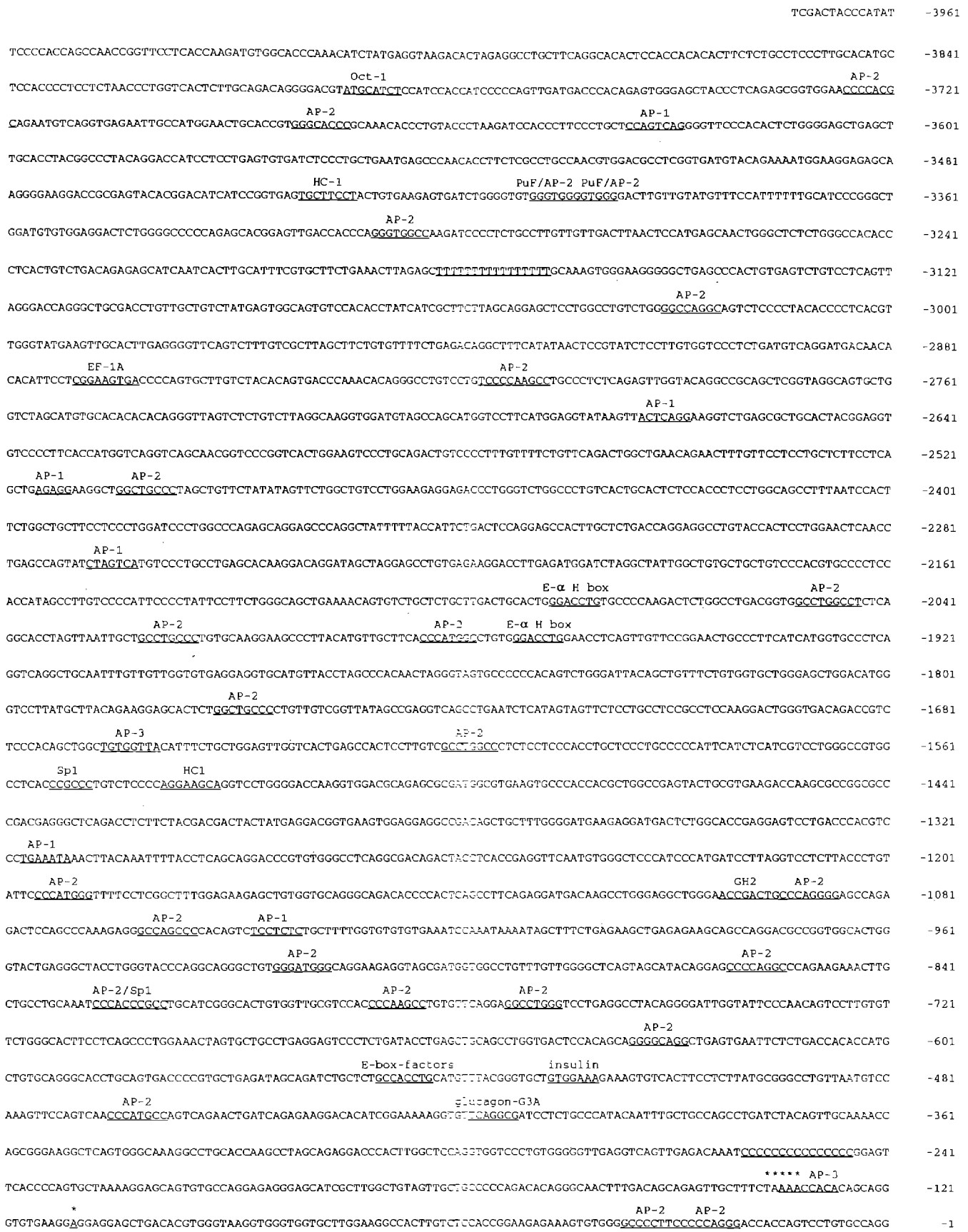



**Fig. 3** Southern blot analysis of *RNK-Met-1* genomic sequences. Twenty μg of DNA derived from either rat LGL leukemia, RNK-16, or nude rat spleen was digested to completion with *Bam* HI (lanes 1, 5, *Bgl* II (lanes 2, 6), *Eco* RV (lanes 3, 7), and *Hin* dIII (lanes 4, 8). The samples were electrophoresed on a 0.8% agarose gel and transferred to a nylon membrane. The blot was probed with the full-length *RNK-Met-1* cDNA and stringently washed. The sizes (in kb pairs) of *Hin* dIII cleaved λ phage markers are shown on the *right*

**Fig. 4** Northern blot analysis of RNK-Met-1 mRNA expression in various rat cell lines, primary cells, and tissues. Total cellular RNA was isolated from the following tissues and cell lines (as described in the Materials and methods): lane (1) normal rat bone marrow, (2&3) two different populations of normal rat peripheral blood CD3⁻ LGL, (4) normal rat spleen non-adherent cells, (5) rat RBL basophilic leukemia, (6) rat RNK-16 NK leukemia, (7) normal rat spleen (unseparated), (8) normal rat spleen CD3⁺ T cells, and (9) normal rat thymus. Northern blot analysis was performed on 25 μg of RNA and the filters were sequentially hybridized with $^{32}$P-labeled cDNA probes for RNK-Met-1 (exposure, 2 days) and human γ-actin (exposure, 1 day)




TCGACTACCCATAT -3961
TCCCCACCAGCCAACCGGTTCCTCACCAAGATGTGGCACCCAAACATCTATGAGGTAAGACACTAGAGGCCTGCTTCAGGCACACTCCACCACACACTTCTCTGCCTCCCTTGCACATGC -3841
Oct-1 AP-2
TCCACCCCTCCTCTAACCCTGGTCACTCTTGCAGACAGGGGACGTATGCATCTCCATCCACCATCCCCCAGTTGATGACCCACAGAGTGGGAGCTACCCTCAGAGCGGTGGAACCCCACG -3721
AP-2 AP-1
CAGAATGTCAGGTGAGAATTGCCATGGAACTGCACCGTGGGCACCCGCAAACACCCTGTACCCTAAGATCCACCCTTCCCTGCTCCAGTCAGGGGTTCCCACACTCTGGGGAGCTGAGCT -3601
TGCACCTACGGCCCTACAGGACCATCCTCCTGAGTGTGATCTCCCTGCTGAATGAGCCCAACACCTTCTCGCCTGCCAACGTGGACGCCTCGGTGATGTACAGAAAATGGAAGGAGAGCA -3481
HC-1 PuF/AP-2 PuF/AP-2
AGGGGAAGGACCGCGAGTACACGGACATCATCCGGTGAGTGCTTCCTACTGTGAAGAGTGATCTGGGGTGTGGGTGGGGTGGGGACTTGTTGTATGTTTCCATTTTTTGCATCCCGGGCT -3361
AP-2
GGATGTGTGGAGGACTCTGGGGCCCCCAGAGCACGGAGTTGACCACCCAGGGTGGCCAAGATCCCCTCTGCCTTGTTGTTGACTTAACTCCATGAGCAACTGGGCTCTCTGGGCCACACC -3241
CTCACTGTCTGACAGAGAGCATCAATCACTTGCATTTCGTGCTTCTGAAACTTAGAGCTTTTTTTTTTTTTTTTGCAAAGTGGGAAGGGGGCTGAGCCCACTGTGAGTCTGTCCTCAGTT -3121
AP-2
AGGGACCAGGGCTGCGACCTGTTGCTGTCTATGAGTGGCAGTGTCCACACCTATCATCGCTTCTTAGCAGGAGCTCCTGGCCTGTCTGGGGCCAGGCAGTCTCCCCTACACCCCTCACGT -3001
TGGGTATGAAGTTGCACTTGAGGGGTTCAGTCTTTGTCGCTTAGCTTCTGTGTTTTCTGAGACAGGCTTTCATATAACTCCGTATCTCCTTGTGGTCCCTCTGATGTCAGGATGACAACA -2881
EF-1A AP-2
CACATTCCTCGGAAGTGACCCCAGTGCTTGTCTACACAGTGACCCAAACACAGGGCCTGTCCTGTCCCCAAGCCTGCCCTCTCAGAGTTGGTACAGGCCGCAGCTCGGTAGGCAGTGCTG -2761
AP-1
GTCTAGCATGTGCACACACACAGGGTTAGTCTCTGTCTTAGGCAAGGTGGATGTAGCCAGCATGGTCCTTCATGGAGGTATAAGTTACTCAGGAAGGTCTGAGCGCTGCACTACGGAGGT -2641
GTCCCCTTCACCATGGTCAGGTCAGCAACGGTCCCGGTCACTGGAAGTCCCTGCAGACTGTCCCCTTTGTTTTCTGTTCAGACTGGCTGAACAGAACTTTGTTCCTCCTGCTCTTCCTCA -2521
AP-1 AP-2
GCTGAGAGGAAGGCTGGCTGCCCTAGCTGTTCTATATAGTTCTGGCTGTCCTGGAAGAGGAGACCCTGGGTCTGGCCCTGTCACTGCACTCTCCACCCTCCTGGCAGCCTTTAATCCACT -2401
TCTGGCTGCTTCCTCCCTGGATCCCTGGCCCAGAGCAGGAGCCCAGGCTATTTTTACCATTCTGACTCCAGGAGCCACTTGCTCTGACCAGGAGGCCTGTACCACTCCTGGAACTCAACC -2281
AP-1
TGAGCCAGTATCTAGTCATGTCCCTGCCTGAGCACAAGGACAGGATAGCTAGGAGCCTGTGAGAAGGACCTTGAGATGGATCTAGGCTATTGGCTGTGCTGCTGTCCCACGTGCCCCTCC -2161
E-α H box AP-2
ACCATAGCCTTGTCCCCATTCCCCTATTCCTTCTGGGCAGCTGAAAACAGTGTCTGCTCTGCTTGACTGCACTGGGACCTGTGCCCCAAGACTCTGGCCTGACGGTGGCCTGGCCTCTCA -2041
AP-2 AP-2 E-α H box
GGCACCTAGTTAATTGCTGCCTGCCCTGTGCAAGGAAGCCCTTACATGTTGCTTCACCCATGGCCTGTGGGACCTGGAACCTCAGTTGTTCCGGAACTGCCCTTCATCATGGTGCCCTCA -1921
GGTCAGGCTGCAATTTGTTGTTGGTGTGAGGAGGTGCATGTTACCTAGCCCACAACTAGGGTAGTGCCCCCCACAGTCTGGGATTACAGCTGTTTCTGTGGTGCTGGGAGCTGGACATGG -1801
AP-2
GTCCTTATGCTTACAGAAGGAGCACTCTGGCTGCCCCTGTTGTCGGTTATAGCCGAGGTCAGCCTGAATCTCATAGTAGTTCTCCTGCCTCCGCCTCCAAGGACTGGGTGACAGACCGTC -1681
AP-3 AP-2
TCCCACAGCTGGCTGTGGTTACATTTCTGCTGGAGTTGGTCACTGAGCCACTCCTTGTCGCCTGGCCCTCTCCTCCCACCTGCTCCCTGCCCCCATTCATCTCATCGTCCTGGGCCGTGG -1561
Sp1 HC1
CCTCACCCGCCCTGTCTCCCCAGGAAGCAGGTCCTGGGGACCAAGGTGGACGCAGAGCGCGATGGCGTGAAGTGCCCACCACGCTGGCCGAGTACTGCGTGAAGACCAAGCGCCGGCGCC -1441
CGACGAGGGCTCAGACCTCTTCTACGACGACTACTATGAGGACGGTGAAGTGGAGGAGGCCGACAGCTGCTTTGGGGATGAAGAGGATGACTCTGGCACCGAGGAGTCCTGACCCACGTC -1321
AP-1
CCTGAAATAAACTTACAAATTTTACCTCAGCAGGACCCGTGTGGGCCTCAGGCGACAGACTACCTCACCGAGGTTCAATGTGGGCTCCCATCCCATGATCCTTAGGTCCTCTTACCCTGT -1201
AP-2 GH2 AP-2
ATTCCCCATGGGTTTTCCTCGGCTTTGGAGAAGAGCTGTGGTGCAGGGCAGACACCCCACTCAGCCTTCAGAGGATGACAAGCCTGGGAGGCTGGGAACCGACTGCCCAGGGGAGCCAGA -1081
AP-2 AP-1
GACTCCAGCCCAAAGAGGGCCAGCCCCACAGTCTCCTCTCTGCTTTTGGTGTGTGTGAAATCCAAATAAAATAGCTTTCTGAGAAGCTGAGAGAAGCAGCCAGGACGCCGGTGGCACTGG -961
AP-2 AP-2
GTACTGAGGGCTACCTGGGTACCCAGGCAGGGCTGTGGGATGGGCAGGAAGAGGTAGCGATGGTGGCCTGTTTGTTGGGGCTCAGTAGCATACAGGAGCCCCAGGCCCAGAAGAAACTTG -841
AP-2/Sp1 AP-2 AP-2
CTGCCTGCAAATCCCACCCGCCTGCATCGGGCACTGTGGTTGCGTCCACCCCAAGCCTGTGTTCAGGAGGCCTGGGTCCTGAGGCCTACAGGGGATTGGTATTCCCAACAGTCCTTGTGT -721
AP-2
TCTGGGCACTTCCTCAGCCCTGGAAACTAGTGCTGCCTGAGGAGTCCCTCTGATACCTGAGCTGCAGCCTGGTGACTCCACAGCAGGGGCAGGCTGAGTGAATTCTCTGACCACACCATG -601
E-box-factors insulin
CTGTGCAGGGCACCTGCAGTGACCCCGTGCTGAGATAGCAGATCTGCTCTGCCACCTGCATGTTTACGGGTGCTGTGGAAAGAAAGTGTCACTTCCTCTTATGCGGGCCTGTTAATGTCC -481
AP-2 glucagon-G3A
AAAGTTCCAGTCAACCCATGCCAGTCAGAACTGATCAGAGAAGGACACATCGGAAAAAGGTGTTCAGGCGATCCTCTGCCCATACAATTTGCTGCCAGCCTGATCTACAGTTGCAAAACC -361
AGCGGGAAGGCTCAGTGGGCAAAGGCCTGCACCAAGCCTAGCAGAGGACCCACTTGGCTCCAGGTGGTCCCTGTGGGGGTTGAGGTCAGTTGAGACAAATCCCCCCCCCCCCCCCCGGAGT -241
***** AP-3
TCACCCCAGTGCTAAAAGGAGCAGTGTGCCAGGAGAGGGAGCATCGCTTGGCTGTAGTTGCTGCCCCCAGACACAGGGCAACTTTGACAGCAGAGTTGCTTTCTAAAACCACACAGCAGG -121
* AP-2 AP-2
GTGTGAAGGAGGAGGAGCTGACACGTGGGTAAGGTGGGTGGTGCTTGGAAGGCCACTTGTCTCCACCGGAAGAGAAAGTGTGGGGCCCCTTCCCCCAGGGACCACCAGTCCTGTGCCAGG -1

restriction fragments generated from genomic DNA of RNK-16 tumor and nude rat spleen were compared (Fig. 3). These data also suggested that *RNK-Met-1* is a single-copy gene.

### *Expression of RNK-Met-1 in CD3⁻ LGL*

Northern analysis using a full-length *RNK-Met-1* cDNA detected a single 0.9 kb RNK-Met-1 mRNA in the total cellular RNA from the purified CD3⁻ spleen LGL of two different rats (Fig. 4). Similar levels of *RNK-Met-1* mRNA were also detected in total cellular RNA from the rat LGL leukemia cell line, RNK-16 and whole and non-adherent spleen (devoid of B cells and monocytes, but containing CD3⁻ LGL). Some weak expression of *RNK-Met-1* mRNA was observed in total cellular RNA from rat bone marrow and this signal may be due to a small number of mature large granular lymphocytes or their precursors. By comparison, *RNK-Met-1* mRNA could not be detected in total cellular RNA from rat thymus and purified CD3⁺ spleen T cells. Furthermore, the rat basophilic leukemia cell line, RBL, did not express *RNK-Met-1* mRNA (Fig. 4) nor does the rat T-cell lymphoma cell line, NTD (data not shown). This data is consistent with previous reports that have suggested that both *RNK-Met-1* and its human equivalent, *Hu-Met-1*, can be detected in CD3⁻ LGL and not T or other leukocyte and non-leukocyte cell populations (Smyth et al. 1992, 1993, 1995).

### *5′-flanking DNA sequence motifs*

The 5′ upstream region of the *RNK-Met-1* gene is characterized by the absence of a typical TATAAA consensus sequence, but by the presence of a similar motif, TAAAA [(positions –137, –228, –1014) (Fig. 5)]. Although departures from the TATA rule are unusual, this variation exists in eukaryotic genes that are transcribed normally, including the other granzymes (Klein et al. 1989; Corden et al. 1980). Therefore, the predicted likely transcriptional start site is at position –111 relative to the ATG translational start codon. The following experiments were attempted to determine the site for initiation of transcription: 1) Primer extension was performed with several different antisense oligonucleotides (30 to 45mers from exons 1 and 2) annealed to RNK-16 and RBL (total cellular or poly A⁺ RNA); 2) RNase protection mapping, using several different antisense transcripts covering the region from –1264 to +60. Both of these approaches did not yield conclusive results. Therefore, in order to assess the possibility of additional intron(s) existing further 5′, we attempted to clone longer *RNK-Met-1* cDNAs from a λgt11 RNK-16 cDNA library. However, no clones with 5′ sequence beyond the +1 ATG codon were isolated. Indeed, the majority of isolated clones terminated within a few bp of each other in this region and anchored PCR of RNK-16 RNA also failed to yield clones of greater length (data not shown). Based on these results we therefore conclude that the 5′ region of *RNK-Met-1* RNA is highly complex and precludes analysis of the transcriptional start site by conventional means. It should be noted that our prediction of a start site at position –111 relative to the ATG codon is based on a general conservation of start sites within 200 bp of the ATG in other granzymes sharing a similar organization to the *RNK-Met-1* gene. Transcription of the human Met-ase gene (*GZMM*) appears to be initiated at two major sites (positions 800 and 826; Pilat et al. 1994); however, neither of these sites has been characterized in the 5′ region of the *RNK-Met-1* gene.

A large number of potential regulatory motifs occurred throughout the first 3.9 kb of 5′-flanking region, and it is likely that some of these are involved in transcriptional control of the *RNK-Met-1* gene (Fig. 5). These potential regulatory elements included motifs that are involved in activation-induced expression such as AP-1 (Angel et al. 1988), AP-2 (Imagawa et al. 1987), and AP-3 (Mitchell et al. 1987). Also identified were several motifs that mediate tissue-specific expression including OCT-1 (Singh et al. 1986) and glucagon-G3A (Cockell et al. 1989). Other elements of interest included: E-box factors (Ephrussi et al. 1985), SP-1 (Briggs et al. 1986), insulin (Ohlsson and Edlund 1986), and atypical repetitive 15–16 bp DNA sequences $(T)_{16}$ and $(C)_{15}$ (Fig. 5). Examination of the 5′-flanking sequence of the *RNK-Met-1* gene revealed some limited sequence similarity but no significantly long homologous regions (>50 bp) when compared with the putative 5′-promoters of other serine protease genes including: mouse granzyme *A* (Hershberger et al. 1992); mouse granzyme *D* (Accession X56990); mouse granzyme *E* (Accession X56988); mouse granzyme *F* (Jenne et al. 1991); human *Met-ase-1* [(*GZMM*) (Pilat

**Fig. 5** Analysis of the *RNK-Met-1* 5′-flanking region. The nucleotide sequence of the 5′ upstream region of the *RNK-Met-1* gene. Numbers on the *left* and *right* indicate the nucleotide positions starting from the translation initiation site. The putative TAAAA box and transcriptional start site (A) are asterisked. Several of many potential transcription control elements are shown in *bold* above the respective underlined sequence. These sequence data have been submitted to the EMBL/GenBank Data Libraries.

**Fig. 6** Matching 5′-flanking sequences of RNK-Met-1 and mouse and human perforin genes. Nucleotide numbers refer to positions described as in Figure 4A and Trapani and co-workers (1990) and Lichtenheld and Podack (1989)

| | |
|---|---|
| human perforin | 5'-TGGGTGCTCGTGG-3' (1310 to 1322) |
| mouse perforin | 5'-TGGGTGCTGGTGG-3' (-92 to -80) |
| RNK-Met-1 | 5'-TGGGTACTGAGGG-3' (-963 to-951) |
| human perforin | 5'-ACCTGTG- ACCACA-3' (1341 to 1353) |
| mouse perforin | 5'-ACCTGTG- ACCACA-3' (-77 to -65) |
| RNK-Met-1 | 5'-ACCTGGGTACC-CA-3' (-948 to -936) |
| human perforin | 5'-CTGGGGGCGGGGCAGGAAGTAGAAGTGAT-3' (1355 to 1383) |
| mouse perforin | 5'-CTGGGGGCAGGGCAGGAAGTAGTAATGAT-3' (-62 to -34) |
| RNK-Met-1 | 5'-CTGTGGGATGGGCAGGAAGAGGTAGCGAT-3' (-928 to -900) |

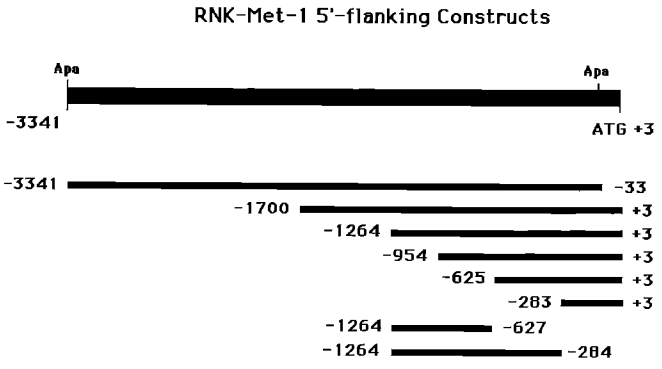
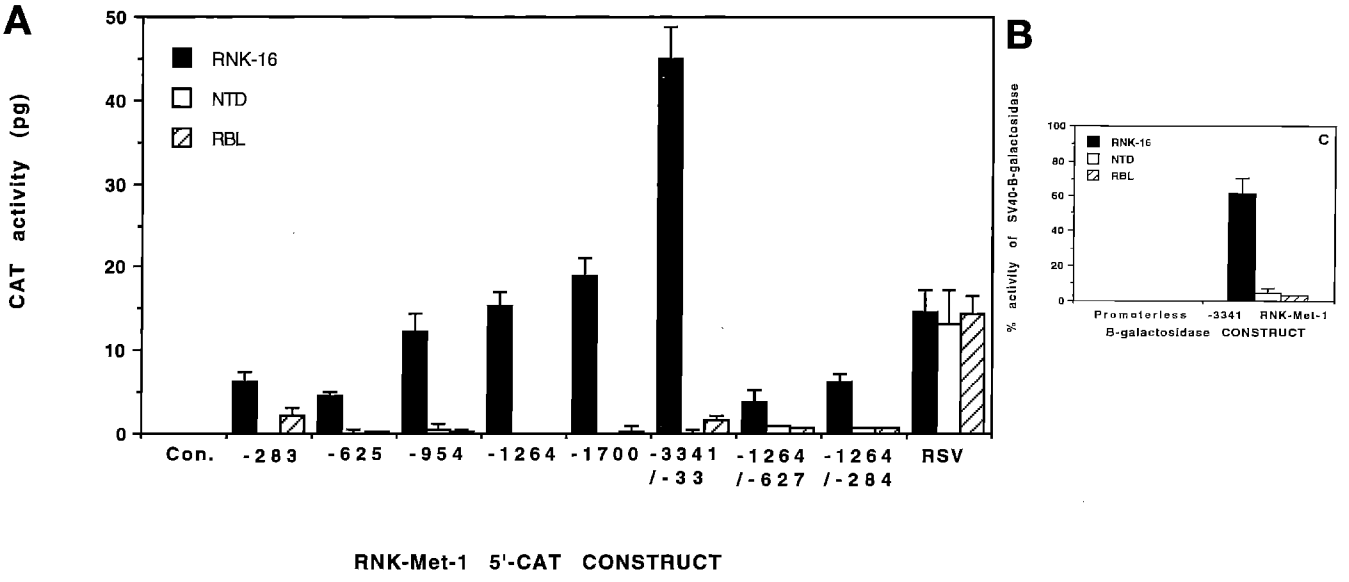



**Fig. 7** Diagrammatic representation of the *RNK-Met-1* 5′-flanking region and nested *RNK-Met-1 5′* fragments used to drive CAT expression. Residue positions are taken from Figure 2 and indicate distances from the ATG translational start site. Nested *RNK-Met-1* 5′-flanking fragments were derived by PCR and cloned upstream of a promoterless and enhancerless *CAT* gene. The numbering of the constructs refers to the first nucleotide of each construct.

*Function of the RNK-Met-1 5′-flanking region*

The *RNK-Met-1* 5′-flanking and promoter region was analyzed for its ability to control a promoterless *CAT* reporter gene to determine whether, and how, this DNA could confer CD3⁻ LGL-restricted expression of *RNK-Met-1*. Nested portions of the promoter were incorporated into CAT constructs to facilitate examination of six regions in the *RNK-Met-1* 5′-flanking region (Fig. 7). These DNA constructs were transfected into a *RNK-Met-1*-expressing rat large granular lymphocyte leukemia cell line (RNK-16), into a *RNK-Met-1*-negative rat T lymphoma (NTD), and into a *RNK-Met-1*-negative rat basophilic leukemia cell line (RBL). CAT activity was determined by ELISA assay 48 h after transfection. Figure 8 A depicts the mean expression of CAT protein (pg) ± SE calculated from three to fifteen independent transfection experiments using a CAT standard curve reference. All three cell lines tested were able to produce CAT in approximately equal quantity in response




et al. 1994)]; human granzyme *B* [(*CGL1*) (Klein et al. 1989)]; human granzyme *H* [(*CTSGL*) (Haddad et al. 1990)]; human *AZU1* and *PRTN3* [(Zimmer et al. 1992) (data not shown)]. However, a 5′-GCCCCTTCC-3′ (AP-2) motif in the *RNK-Met-1* 5′-flanking region (–36 to –28) has also been described in the 5′-flanking sequences of cathepesin G and myeloperoxidase (both granular enzymes) genes (Hohn et al. 1989). More interestingly, the region –963 to –900 of the 5′-flanking region of *RNK-Met-1* shares three successive sequence elements that are highly homolgous with the immediate 5′-flanking sequences of both mouse (Trapani and Dupont 1990) and human (Lichtenheld and Podack 1989) perforin (Fig. 6). These three elements of the *RNK-Met-1* 5′-flanking region share approximately 87% (–963 to –951), 73% (–948 to –936, with two gaps inserted), and 79% (–928 to –900) identity with corresponding regions of the mouse and human perforin genes.

**Fig. 8** **A** Functional analysis of the *RNK-Met-1* 5′-flanking region. RNK-16, NTD, and RBL were transfected with deletion constructs of the *RNK-Met-1* 5′-flanking region driving the *CAT* gene (as described in Materials and methods). The CAT activities shown are expressed as pg/100 μg of protein. Their mean ± SE from two to twelve independent transfections are shown. –283 is the construct containing the *RNK-Met-1* 5′-flanking sequence –283 to +3. –625, –954, –1264, and –1700 have similar 3′ ends (+3). –3341/–33, –1264/–284, and –1264/–627 indicate the other deletion constructs of the *RNK-Met-1* 5′-flanking region. RSVCAT is a positive control plasmid for equivalent transfection into all three cell lines. The control (Con.) was the enhancerless and promoterless pCAT basic parental construct. **B** Functional analysis of the *RNK-Met-1* 5′-flanking region. RNK-16, NTD, and RBL were transfected with deletion constructs of the *RNK-Met-1* 5′-flanking region driving the β-galactosidase gene (see Materials and methods). The percentage of cells expressing staining following transient transfection of the SV40-β-galactosidase was RNK-16 (5.7 ± 0.9%), NTD (12.1 ± 0.1%), and RBL (9.9 ± 1.1%)

to the RSV promoter (CAT activity ~13 to 15 pg/100 μg protein). The two most immediate proximal 5′-flanking regions, –283 to +3 and –625 to +3, yielded only minor levels of CAT expression in RNK-16 (<6 pg), although less than 2 pg in *RNK-Met-1*-negative cells (Fig. 8A). The inclusion of more distal promoter sequences, –954 to +3, was associated with enhanced CAT expression compared with the shorter *RNK-Met-1* promoters (–283 to +3 and –625 to +3) in *RNK-Met-1*-expressing RNK-16 (12±2 pg) with no effect observed in NTD and RBL cells. Longer constructs, –1264 to +3 and –1700 to +3, further enhanced CAT expression in *RNK-Met-1*-positive RNK-16. Despite the fact that the two most immediate proximal 5′-flanking regions, –283 to +3 and –625 to +3, yielded only minor levels of CAT expression in RNK-16, it appeared that the region (–283 to +3) was important for the activity of longer constructs (–1264 to +3), as constructs (–1264 to –283) and (–1264 to –625) displayed <6 pg of CAT activity (Fig. 8A). The inclusion of sequences up to –3341 increased CAT expression in RNK-16 further (up to 45±4 pg CAT/100 μg protein) with no effect in *RNK-Met-1*-negative NTD and RBL cells. This level of CAT activity in RNK-16 was approximately three-fold higher than even that observed following transfection with the strong, ubiquitously expressed RSV-CAT construct. These data clearly demonstrate that the region –3341 to –33 contains sequence elements conferring the observed high levels of LGL-specific expression.

The most specifically active *RNK-Met-1* 5′-flanking sequence tested, 3341 to –33, was also examined for its ability to drive expression of a β-galactosidase gene in all three rat cell lines (Fig. 8B). SV40-β-galactosidase was a positive control plasmid for transfection into all three cell lines. The negative control (Con.) used was the enhancer-less and promoterless basic parental β-galactosidase construct. *RNK-Met-1* (–3341 to –33) 5′-promoter activity in each cell line was expressed as a percentage of SV40 promoter activity (from the % β-galactosidase positive cells). In these experiments, the percentage of cells with nuclear staining following transient transfection of the SV40-β-galactosidase was RNK-16 (5.7±0.9%), NTD (12.1±0.1%), and RBL (9.9±1.1%). Strikingly, the RNK-Met-1 5′-β-galactosidase construct was 60%–70% as active as the SV40 driven construct in RNK-16 (3.5±0.6%), but less than 5% as active in the NTD (0.5±0.3%) and RBL (0.3±0.0%) cell lines. By contrast, all cells transfected with the control promoterless β-galactosidase construct failed to exhibit staining. These data further illustrate the potential strength and specificity of the *RNK-Met-1* 5′-flanking region to target exogenous gene expression in rat LGL.

The functional reporter gene analysis demonstrates that the first 3.3 kb of the *RNK-Met-1* 5′ promoter is sufficient to enable high levels of transcription specifically in a rat CD3⁻ LGL leukemia cell line. Transcriptional activity of this promoter was restricted to *RNK-Met-1*-positive LGL among the different cell types analyzed. The mechanisms by which this specific expression pattern is achieved is interesting because the immediate proximal *RNK-Met-1* 5′-flanking region, –625 to +1, functions weakly. The *RNK-Met-1* gene appears to be controlled by *cis*-acting regions further upstream as determined by the comparative CAT activities obtained from the 5′ extending *RNK-Met-1* promoter constructs. The sequences between –625 and –954, –954 and –1264 and –1264 and –1700 progressively enhance the proximal promoter activity in RNK-16, but have no activity in *RNK-Met-1*-negative cell lines NTD and RBL. Overall, however, the greatest difference in CAT activity between *RNK-Met-1*-positive and *RNK-Met-1*-negative cell lines (approximately a 30-fold difference) was observed with all of these regions acting in concert (–3341 to –33). Currently, there is no evidence for regulation of the *RNK-Met-1* mRNA or protein levels in rat LGL, but the effect of physiological extracellular signals that activate LGL (such as IL-2, IL-12) on these constructs is also relevant for future studies. Further extension of these analyses and fine mapping of the identified regulatory regions by DNase footprinting and gel retardation should define the sequences of the critical regulatory elements and aid our isolation of the transcription factors involved.

The regulatory elements responsible for the exclusive expression of granule proteins by cytolytic lymphocytes have yet to be accurately defined and relatively little is understood about granzyme gene regulation. One exception is the regulation of granzyme *B* expression as a model for the induction of "late" genes during T-cell activation. Several regulatory regions in the 5′-flanking sequence that specifically promote granzyme *B* expression in activated T lymphocytes (Hanson and Ley 1990, 1993) and target cytotoxic lymphocyte-specific gene expression in transgenic mice (Hanson et al. 1991) have been identified. Clearly, however, several *cis*-acting sequences involved in constitutive or inducible T-lymphocyte-specific expression have already been identified, including upstream activating sequences 5′ to the IL-2 (Durand et al. 1987) and IL-2 receptor α chain (Cross et al. 1987) genes. By contrast, there has been no description of 5′ promoter elements that target genes specifically to CD3⁻ LGL with NK activity. The gene elements that control the expression of other NK-cell-associated genes such as mouse *NK 1.1*, rat *NKRP1*, human *NKG* family members, and p75 IL-2 receptor are poorly characterized and many of these molecules are also expressed in some T-cell subsets (Giorda et al. 1992; Houchins et al. 1993; Koyashu 1994). Therefore, these studies of the *RNK-Met-1* 5′-flanking region represent an important step towards understanding the control of gene expression specifically in CD3⁻ LGL.

Furthermore, elucidating a role for NK cells in immune responses has been hindered by the lack of a satisfactory model of NK cell deficiency in animals. Depletions of NK cells by various cytotoxic antisera (anti-asialoGM1, anti-NK1.1, or anti-p75 IL-2 receptor) are flawed by their transience, their inconvenience, and the likelihood of non-specific effects on other leukocytes (Seaman et al. 1987; Suttles et al. 1986; Tanaka et al. 1993). The use of beige (bg/bg) mice is complicated by the retention of many NK cell functions and the loss of some neutrophil and T-cell functions (Takeuchi et al. 1986; Halle-Pannenko and Bru-

ley-Rosset 1985). Similarly, the recently described CD3ε transgenic mice appear to have a complete loss of T lymphocytes in addition to their lack of NK cells (Wang et al. 1994). Our studies suggest that the *RNK-Met-1* gene 5′-flanking promoter may be a useful tool for targeting exogenous gene products (such as suicide genes) to $CD3^-$ NK cells in transgenic rats or mice, thereby creating better in vivo models of NK cell function.

**Acknowledgments** We express our appeciation to Michael Kershaw for his helpful assistance. M. J. S. and J. A. T. were supported by Wellcome Trust Australasian Senior Research Fellowships in Medical Science. M. D. H. was supported by an Australian Post-Doctoral Award from the National Health and Medical Research Council of Australia. This project was supported by a National Health and Medical Research Council of Australia Project Grant.

## References

Angel, P., Allegretto, E. A., Okino, S. T., Hattori, K., Boyle, W. J., Hunter, T., and Karin, M. Oncogene jun encodes a sequence-specific transactivator similar to AP-1. *Nature 332:* 166–171, 1988

Baker, E., Sutherland, G. R., and Smyth, M. J. The gene encoding a human natural killer cell granule serine protease, *Met-ase 1*, maps to chromosome 19p13.3. *Immunogenetics 39:* 294–295, 1994

Bell, G. I., Quinto, M., Quiroga, P., Valenzuela, C. J., Craik, C. S., and Rutter, W. J. Isolation and sequence of a rat chymotrypsin B gene. *J Biol Chem 259:* 14 265–14 270, 1984

Benfey, P. N., Yin, F. H., and Leder, P. Cloning of the mast cell protease, RMCPII. Evidence for cell-specific expression and a multi-gene family. *J Biol Chem 262:* 5377–5384, 1987

Bleackley, R. C., Lobe, C. G., Duggan, B., Ehrman, N., Fregeau, C., Meier, M., Letellier, M., Havele, C., Shaw, J., and Paetkau, V. The isolation and characterization of a family of serine protease genes expressed in activated cytotoxic T lymphocytes. *Immunol Rev 103:* 5–19, 1988

Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem 72:* 248–253, 1976

Briggs, M. R., Kadonga, J. T., Bell, S. P., and Tijan, R. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. *Science 234:* 47–52, 1986

Chen, E. Y. and Seeburg, P. H. Supercoil sequencing: fast and simple method for sequencing plasmid DNA. *DNA 4:* 165–170, 1985

Chomczynski, P. and Sacchi, N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem 162:* 156–159, 1987

Cockell, M., Stevenson, B. J., Strubin, M., Hagenbuchle, O., and Wellauer, P. K. Identification of a cell-specific DNA-binding activity that interacts with a transcriptional activator of genes expressed in acinar pancreas. *Mol Cell Biol 9:* 2464–2476, 1989

Corden, J., Wasylyk, A., Sassone-Corsi, P., Kedinger, C., and Chambon, P. Promoter sequences of eukaryotic protein-coding genes. *Science 209:* 1406–1414, 1980

Cross, S. L., Feinberg, M. B., Wolf, J. B., Holbrook, N. J., Wong-Staal, F., and Leonard, W. J. Regulation of the human interleukin-2 receptor alpha chain promoter: activation of a nonfunctional promoter by the transactivator gene of HTLV-I. *Cell 49:* 47–56, 1987

Durand, D. B., Bush, M. R., Morgan, J. G., Weiss, A., and Crabtree, G. R. A 275 basepair fragment at the 5′ end of the interleukin 2 gene enhances expression from a heterologous promoter in response to signals from the T cell antigen receptor. *J Exp Med 165:* 395–407, 1987

Ephrussi, A., Church, G. M., Tonegawa, S., and Gilbert, W. B lineage-specific interactions of an immunoglobulin enhancer with cellular factors in vivo. *Science 227:* 134–140, 1985

Giorda, R., Weisberg, E. P., Ip, T.-K., and Trucco, M. Genomic structure and strain-specific expression of the natural killer cell receptor NKR-P1. *J Immunol 149:* 1957–1963, 1992

Gunning, P., Ponte, P., Okayama, H., Engel, J., Blau, H., and Kedes, L. Isolation and characterization of full-length cDNA clones for human alpha-, beta-, and gamma-actin mRNAs: skeletal but not cytoplasmic actins have an amino-terminal cysteine that is subsequently removed. *Mol Cell Biol 3:* 787–795, 1983

Haddad, P., Jenne, D., Tschopp, J., Clement, M. V., Mathieu-Mahul, D., and Sasportes, M. Structure and evolutionary origin of the human granzyme H gene. *Int Immunol 3:* 57–66, 1990

Halle-Pannenko, O., and Bruley-Rosset, M. Decreased graft-versus-host reaction and T cell cytolytic potential of beige mice. *Transplantation 39:* 85–87, 1985

Hanson, R. D. and Ley, T. J. Transcriptional activation of the human cytotoxic serine protease gene CSP-B in T lymphocytes. *Mol Cell Biol 10:* 5655–5662, 1990

Hanson, R. D., Sclar, G., Kanagawa, O., and Ley, T. J. The 5′-flanking region of the human CGL-1/granzyme B gene targets expression of a reporter gene to activated T-lymphocytes in transgenic mice. *J Biol Chem 266:* 24 433–24 438, 1991

Hanson, R. D., Grisolano, J. L., and Ley, T. J. Consensus AP-1 and CRE motifs upstream from the human cytotoxic serine protease B (CSP-B/CGL-1) gene synergize to activate transcription. *Blood 82:* 2749–2757, 1993

Hershberger, R. J., Gershenfeld, H. K., Weissman, I. L., and Su, L. Genomic organization of the mouse granzyme A gene. Two mRNAs encode the same mature granzyme A with different leader peptides. *J Biol Chem 267:* 25 488–25 493, 1992

Heusel, J. W., Wesselschmidt, R. L., Shresta, S., Russell, J. H., and Ley, T. J. Cytotoxic lymphocytes require granzyme B for the rapid induction of DNA fragmentation and apoptosis in allogeneic target cells. *Cell 76:* 977–987, 1994

Hohn, P. A., Popescu, N. C., Hanson, R. D., Salvesen, G., and Ley, T. J. Genomic organization and chromosomal localization of the human cathepsin G gene. *J Biol Chem 264:* 13 412–13 419, 1989

Houchins, J. P., Kricek, F., Chujor, S. N., Heise, C. P., Yabe, T., McSherry, C., and Bach, F. H. Genomic structure of *NKG5*, a human NK and T cell-specific activation gene. *Immunogenetics 37:* 102–107, 1993

Hudig, D., Allison, N. J., Pickett, T. M., Winkler, U., Kam, C.-M., and Powers, J. C. The function of lymphocyte proteases. Inhibition and restoration of granule-mediated lysis with isocoumarin serine protease inhibitors. *J Immunol 147:* 1360–1368, 1991

Imagawa, M., Chiu, R., and Karin, M. Transcription factor AP-2 mediates induction by two different signal transduction pathways: protein kinase C and cAMP. *Cell 51:* 251–260, 1987

Irwin, D. M., Robertson, K. A., and MacGillivray, R. T. Structure and evolution of the bovine prothrombin gene. *J Mol Biol 200:* 31–45, 1988

Jenne, D. E. and Tschopp, J. Granzymes, a family of serine proteases released from granules of cytolytic T lymphocytes upon T cell receptor stimulation. *Immunol Rev 103:* 53–71, 1988

Jenne, D. E., Zimmer, M., Garcia-Sanz, J. A., Tschopp, J., and Lichter, P. Genomic organization and subchromosomal in situ localization of the murine granzyme F, a serine protease expressed in $CD8^+$ T cells. *J Immunol 147:* 1045–1052, 1991

Kagi, D., Ledermann, B., Burki, K., Seiler, P., Odermatt, B., Olsen, K. J., Podack, E. R., Zinkernagel, R. M., and Hengartner, H. Cytotoxicity mediated by T cells and natural killer cells is greatly impaired in perforin-deficient mice. *Nature 369:* 31–37, 1994

Klein, J. L., Shows, T. B., Dupont, B., and Trapani, J. A. Genomic organization and chromosomal assignment for a serine protease gene (CSPB) expressed by human cytotoxic lymphocytes. *Genomics 5:* 110–117, 1989

Koyashu, S. $CD3^+$ $CD16^+$ $NK1.1^+$ $B220^+$ large granular lymphocytes arise from both $\alpha$–$\beta$ $TCR^+$ $CD4^-$ $CD8^-$ and $\gamma$–$\delta$ $TCR^+$ $CD4^-$ $CD8^-$ cells. *J Exp Med 179:* 1957–1972, 1994

Lichtenheld, M. G. and Podack, E. R. Structure of the human perforin gene. A simple gene organization with interesting potential regulatory sequences. *J Immunol 143:* 4267–4274, 1989

Lobe, C. G, Upton, C., Duggan, B., Ehrman, N., Letellier, M., Bell, J., and Bleackley, R. C. Organization of two genes encoding murine cytotoxic T-lymphocyte-specific serine proteases CCPI and CCPII. *Biochemistry 27:* 6941–6946, 1988

Min, H. Y. and Spiegelman, B. M. Adipsin, the adipocyte serine protease: gene structure and control of expression by tumor necrosis factor. *Nucleic Acids Res 14:* 8879–8892, 1986

Mitchell, P. J., Wang, C., and Tijan, R. Positive and negative regulation of transcription in vitro: enhancer-binding protein AP-2 is inhibited by SV40 T antigen. *Cell 50:* 847–861, 1987

Mount, S. M. A catalogue of splice junction sequences. *Nucleic Acids Res 10:* 459–472, 1982

Ohlsson, H. and Edlund, T. Sequence-specific interactions of nuclear factors with the insulin gene enhancer. *Cell 45:* 35–44, 1986

Pang, Y., Norihisa, Y., Benjamin, D., Kantor, R. R. S., and Young, H. A. Interferon-gamma gene expression in human B cell lines: induction by interleukin-2, protein kinase C activators, and possible effect of hypomethylation on gene regulation. *Blood 80:* 724–732, 1992

Phillips, M., Dijan, P., and Green, H. The nucleotide sequence of three genes participating in the adipose differentiation of 3T3 cells. *J Biol Chem 261:* 10821–10827, 1986

Pilat, D., Fink, T., Obermaier-Skrobanek, B., Zimmer, M., Wekerle, H., Lichter, P., and Jenne, D. E. The human Met-ase gene (*GZMM*): Structure, sequence, and close physical linkage to the serine protease gene cluster on 19p13.3. *Genomics 24:* 445–480, 1994

Podack, E. R., Hengartner, H., and Lichtenheld, M. G. A central role of perforin in cytolysis? *Annu Rev Immunol 9:* 129–157, 1991

Prestridge, D. S. SIGNAL SCAN: a computer program that scans DNA sequences for eukaryotic transcriptional elements. *CABIOS 7:* 203–206, 1991

Reynolds, C. W., Timonen, T., and Herberman, R. B. Natural killer (NK) activity in the rat. I. Isolation and characterization of the effector cells. *J Immunol 127:* 282–287, 1981

Sambrook, J., Fritsch, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory, Cold Spring Harbor, 1989

Sanger, F., Nicklen, S., and Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA 74:* 5463–5467, 1977

Seaman, W. E., Sleisenger, M., Eriksson, E., and Koo, G. C. Depletion of natural killer cells in mice by monoclonal antibody to NK1.1. Reduction in host defense against malignancy without loss of cellular or humoral immunity. *J Immunol 138:* 4539–4544, 1987

Shi, L., Kraut, R. P., Aebersold, R., and Greenberg, A. H. A natural killer cell granule protein that induces DNA fragmentation and apoptosis. *J Exp Med 175:* 553–566, 1992a

Shi, L., Kam, C. M., Powers, J. C., Aebersold, R., and Greenberg, A. H. Purification of three cytotoxic lymphocyte granule serine proteases that induce apoptosis through distinct substrate and target cell interactions. *J Exp Med 176:* 1521–1529, 1992b

Shiver, J. W., Su, L., and Henkart, P. A. Cytotoxicity with target DNA breakdown by rat basophilic leukemia cells expressing both cytolysin and granzyme A. *Cell 71:* 315–322, 1992

Shingh, H., Sen, R., Baltimore, D., and Sharp, P. A. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. *Nature 319:* 154–158, 1986

Smyth, M. J., Wiltrout, T., Trapani, J. A., Ottaway, K. S., Sowder, R., Henderson, L., Powers, J., and Sayers, T. J. Purification and cloning of a novel serine protease, RNK-Met-1, from the granules of a rat natural killer cell leukemia. *J Biol Chem 267:* 24418–24425, 1992

Smyth, M. J., Sayers, T. J., Wiltrout, T., Powers, J. C., and Trapani, J. A. Met-ase: cloning and distinct chromosomal location of a serine protease preferentially expressed in natural killer cells. *J Immunol 151:* 6195–6205, 1993

Smyth, M. J., Browne, K., Kinnear, B. F., Trapani, J. A., and Warren, H. Distinct granzyme expression in human $CD3^-$ $CD56^+$ large granular- and $CD3^-$ $CD56^+$ small high density-lymphocytes displaying non-MHC-restricted cytolytic activity. *J Leukoc Biol 57:* 88–93, 1995

Suttles, J., Schwarting, G. A., and Stout, R. D. Flow cytometric analysis reveals the presence of asialo GM1 on the surface membrane of alloimmune cytotoxic T lymphocytes. *J Immunol 136:* 1586–1591, 1986

Takahashi, H., Nukiwa, T., Yoshimura, K., Quick, C. D., States, D. J., Whang-Peng, J., Knutsen, T., and Crystal, R. G. *J Biol Chem 263:* 14739–14747, 1988

Takeuchi, K., Wood, H., and Swank, R. T. Lysosomal elastase and cathepsin G in beige mice. Neutrophils of beige (Chediak-Higashi) mice selectively lack lysosomal elastase and cathepsin G. *J Exp Med 163:* 665–677, 1986

Tan, S. S. Liver-specific and position-effect expression of a retinol-binding protein-lacZ fusion protein (RBP-lacZ) in transgenic mice. *Dev Biol 146:* 24–37, 1991

Tanaka, T., Kitamura, F., Nagasaka, Y., Kuida, K., Suwa, H., and Miyasaka, M. Selective long-term elimination of natural killer cells in vivo by an anti-interleukin 2 receptor beta chain monoclonal antibody in mice. *J Exp Med 178:* 1103–1107, 1993

Thia, K. Y. T., Jenkins, N. A., Gilbert, D. J., Copeland, N. G., and Smyth, M. J. The natural killer cell serine protease gene *Lmet1* maps to mouse chromosome 10. *Immunogenetics 41:* 47–49, 1995

Thomas, P. S. Hybridization of denatured RNA transferred or dotted nitrocellulose paper. *Methods Enzymol 100:* 255–266, 1983

Timonen, T., Reynolds, C. W., Ortaldo, J. R., and Herberman, R. B. Isolation of human and rat natural killer cells. *J Immunol Methods 51:* 269–277, 1982

Trapani, J. A. and Dupont, B. Novel putative promoter/enhancer sequences are shared by the mouse and human perforin (Pfp) genes. *Tissue Antigens 36:* 228–234, 1990

Tschopp, J. and Nabholz, M. Perforin-mediated target cell lysis by cytolytic T lymphocytes. *Annu Rev Immunol 8:* 279–302, 1990

Wang, B. P., Biron, C., She, J., Higgins, K., Sunshine, M. J., Lacy, E., Lonberg, N., and Terhorst, C. A block in both early T lymphocyte and natural killer cell development in transgenic mice with high-copy numbers of the human CD3ε gene. *Proc Natl Acad Sci USA 91:* 20–24, 1994

Zimmer, M., Medcalf, R. L., Fink, T. M., Mattmann, C., Lichter, P., and Jenne, D. E. Three human elastase-like genes coordinately expressed in the myelomonocytic lineage are organized as a single genetic locus on 19pter. *Proc Natl Acad Sci USA 89:* 8215–8219, 1992