Europe PMC Funders Group
Author Manuscript
J Immunol. Author manuscript; available in PMC 2016 March 16.

Published in final edited form as:
J Immunol. 2015 June 15; 194(12): 6068–6081. doi:10.4049/jimmunol.1401450.

The distal upstream promoter in Ly49 genes, Pro1, is active in mature NK cells and T cells, does not require TATA boxes, and displays enhancer activity[1]

Frances Gays[*], **Sally Taha**[*], and **Colin G. Brooks**[*,2]

[*]Institute of Cell and Molecular Biosciences, The University of Newcastle, Newcastle, U.K.

Abstract

Missing self recognition of MHC class I molecules is mediated in murine species through the stochastic expression of CD94/NKG2 and Ly49 receptors on NK cells. Previous studies have suggested that the stochastic expression of Ly49 receptors is achieved through the use of an alternate upstream promoter, designated Pro1, that is active only in immature NK cells, and operates via the mutually exclusive binding of transcription initiation complexes to closely opposed forward and reverse TATA boxes, forward transcription being transiently required to activate the downstream promoters, Pro2/Pro3, that are subsequently responsible for transcription in mature NK cells. Here we report that Pro1 transcripts are not restricted to immature NK cells but are also found in mature NK cells and T cells, and that Pro1-fragments display strong promoter activity in mature NK cell and T cell lines as well as in immature NK cells. However, the strength of promoter activity in vitro does not correlate well with Ly49 expression in vivo and forward promoter activity is generally weak or undetectable, suggesting that components outside of Pro1 are required for efficient forward transcription. Indeed, conserved sequences immediately upstream and downstream of the core Pro1 region were found to inhibit or enhance promoter activity. Most surprisingly, promoter activity does not require either the forward or reverse TATA boxes, but is instead dependent on residues in the largely invariant central region of Pro1. Importantly, Pro1 displays strong enhancer activity suggesting that this may be its principal function in vivo.

Keywords

Rodent; NK cells; Cell Surface Molecules

Introduction

Research over the last two decades has provided compelling evidence that one of the principal functions of NK cells is to destroy diseased cells via the recognition of stress associated molecules (1). Unlike effector T cells that require many days to develop from inactive precursors, mature NK cells are pre-armed. The advantage to animals of possessing

[1]This work was generously supported by a grant from the Medical Research Council, U.K.

[2]Address correspondence to Dr. Colin G. Brooks, Institute of Cell and Molecular Biosciences, The Medical School, Newcastle NE2 4HH, U.K. colin.brooks@ncl.ac.uk Tel. 44-191-222-7813 Fax. 44-191-222-7424 .

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

such “natural” killer cells is counterbalanced by the potential self harm caused by inappropriate triggering of these cells by low levels of stress molecules on healthy cells. To prevent this, NK cells are endowed with inhibitory receptors including ones that recognize ubiquitously expressed MHC class I (cI)[3] molecules (2, 3). Thus, triggering of effector function only occurs if the activating signals the NK cell receives from stress molecules are sufficient to exceed a threshold set by the normal levels of inhibitory signals it receives from cI recognition, or if the inhibitory signals themselves are weakened by loss of cI expression on diseased cells. This latter mode of triggering NK cell effector function is known as “missing self recognition” and allows NK cells to counteract the subversion of T cell immunity by parasites that downregulate cI expression (4).

Although some inhibitory receptors recognize monomorphic cI molecules, notably CD94/NKG2A recognition of Qa1 or HLA-E, others recognize polymorphic cI molecules and are able to distinguish polymorphic variations in these cI molecules, thereby potentially endowing NK cells with the capacity to detect the downregulation of individual cI molecules. The receptors that perform this function belong to the Ly49 family of C-type lectin receptors in rodents and to the KIR family of Ig-type receptors in primates (2). Some members of the Ly49 and KIR families have acquired activatory function, such as the Ly49H receptor in mice that recognizes virus-encoded cI-like molecules (5, 6). Unlike the cI receptors on T cells, Ly49s and KIRs are not the products of rearranging genes, and the capacity to recognize different cI molecules is achieved by polygenism and polymorphism. Thus, amongst the total of ~60 Ly49 genes that have been identified in the four mouse Ly49 gene complexes that have so far been sequenced (7) there are only two examples of alleles encoding identical proteins.

Because Ly49 genes and cI genes are located on different chromosomes and are therefore inherited independently, in order to maintain functional recognition the specificity of individual Ly49 molecules needs to be relatively broad, an expectation confirmed experimentally (8-10). Consequently, if all Ly49 receptors encoded in a heterozygous mouse were expressed on all NK cells there would be a high probability that all NK cells would recognize all self cI molecules, and thereby be insensitive to the down regulation of individual cI molecules. To avoid this, Ly49s are expressed in a stochastic manner such that each NK cell displays on its surface only a randomly selected subset of all available Ly49s from both homologous chromosomes (11). The same is true of KIRs (12).

The mechanism responsible for this unusual pattern of gene expression is unclear, except that it is achieved at the transcriptional level (13). Differential DNA methylation and histone acetylation clearly play a role in its maintenance (14-16), but cannot easily explain its initiation - for example it is hard to understand how ubiquitously expressed methylating and demethylating enzymes could by themselves achieve effective monoallelic expression of identical alleles. Another possibility is that the concentration of key transcription factors required for Ly49/KIR expression is extremely low, limiting the total number of genes that can be activated at the relevant stage of NK cell development (17). However, to date no transcription factor that affects the frequency of expression of all Ly49 genes without

[3]Abbreviations used in this paper: cI, MHC class I; fwd, forward; rev, reverse; HSVtk, herpes simplex virus thymidine kinase gene

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

affecting other aspects of NK cell development has been identified. Lastly, it has been suggested that Ly49/KIR gene expression is initiated by the differential action of multiple promoters, some of which are bidirectional (18, 19).

In humans it has been suggested that in immature NK cells the chance production of overlapping transcripts from forward (fwd) and reverse (rev) promoters in the upstream region of KIR genes generates small interfering double stranded RNA products that lead to methylation-induced repression of the cognate gene (19). In mice, Ly49 genes contain a distal upstream promoter, Pro1, located several kb upstream of the promoters, designated Pro2 and Pro3, from which Ly49 transcripts are initiated in mature NK cells (20). Transcripts initiating from Pro1 contain additional untranslated exons, designated E-1a and E-1b, located upstream of exon 1 (E1). Subsequent studies indicated that Pro1 is active only in immature NK cells and is a bidirectional promoter containing an upstream rev TATA box and a downstream fwd TATA box that are so close together (~100bp) that only a single transcriptional complex can bind (21). It was hypothesized that if by chance the transcriptional complex binds to the downstream TATA box, fwd transcription is initiated in the adjacent exon E-1a and traverses the Pro2 promoter. This transcription prevents the binding of repressors near to Pro2, so that when the cell matures transcription can switch from Pro1 to Pro2 and the gene is permanently switched on. By contrast, if by chance the transcriptional complex binds to the upstream Pro1 TATA box, rev transcription occurs, allowing repressors free access to Pro2, and causing the gene to be permanently switched off. The binding of C/EBP transcription factors to sites close to the 5′ and 3′ TATA boxes was postulated to modulate transcription by competitively inhibiting the binding of TBP to the proximal TATA box while enhancing the binding of TBP to the distal TATA box, these interactions collectively determining the relative strength of the competing promoters and the probability of transcription in a given direction (21).

In the present study a detailed analysis of the transcriptional activity of the Pro1-region revealed that Pro1 transcripts are present not only in immature NK cells but also in mature NK cells and mature CD4 and CD8 T cells. In line with this, Pro1 constructs displayed promoter activity in vitro in mature NK and T cell lines as well as in immature NK cell lines. Most surprisingly, this promoter activity did not require either of the two TATA boxes or the associated C/EBP binding sites but instead was dependent on residues in the central highly conserved region of Pro1. Coupled with the finding that Pro1 displays enhancer activity, the data leads to a new hypothesis for the function of Pro1.

Materials and Methods

Animals and cells

C57BL/6 mice were obtained from OLAC UK. $RAG^{-/-}$ mice were kindly provided by Dr. B. Seddon (National Institute for Medical Research, London), and $RAG^{-/-}$ $\gamma c^{-/-}$ mice by Dr. M. Batey, Newcastle. Bone marrow, thymus, and spleen cells were prepared using standard procedures. The immature fetal NK cell lines I2/22 and 1608b (C57BL/6 strain) and CBA (CBA strain) were derived as described previously (22, 23). Briefly, thymocytes from day 14 embryos were cultured for 1–3 days in medium containing 10 U/ml IL-4 and 10ng/ml PMA, then transferred to medium containing 10^4 U/ml recombinant human IL-2. After continuous

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

growth had been established, cells were cloned by limiting dilution and adapted to growth in 100 U/ml IL2. The BALB/c immature liver NK cell line LNK (20, 24) was obtained from Dr. S. Anderson (Frederick National Lab, MD) and was grown in the same medium. The derivation of the mature adult C57BL/6 NK cell line D- was described previously (25); it was maintained in medium containing 10^4 U/ml IL2. The tumor cell lines used are commonly used lines of C57 and BALB/c origin that have been maintained in our laboratory for many years; all are available from the ATCC. The RMA cell line has been shown to be a subline of C57 thymoma EL4 (26), and is here designated EL4/R.

Immunofluorescence staining, flow cytometry, and cell purification

Immunofluorescence staining was carried out using standard methods and reagents as described previously (25, 27), and analyzed using a FACS Canto cytometer (BD Biosciences). Purified spleen B cells and T cells were obtained by first separating white cells from red cells and dead cells by centrifugation at 1000g for 20min on a cushion of Lympholyte M (CL5031, Cedarlane). Cells were then stained with FITC anti-CD19 (Becton-Dickinson) and eF600 anti-CD3 (eBioscience) and sorted on a FACS AriaII (Becton-Dickinson). Purified NK cells were obtained by sorting red cell depleted spleen cells that had been stained with PECy7 anti-NK1.1 (Becton-Dickinson) and placing them in culture in medium containing 10^4 U/ml IL2. After 12 days they were harvested for RNA preparation. Staining with anti-NK1.1 and anti-CD3 mAbs showed them to contain >99% NK cells. To obtain pure CD4 and CD8 blast cells, spleen cells were incubated for 30min in ice with mAb SW3A4 anti-NK1.1 (28, 29) (kindly donated by Dr. V. Kumar, Chicago) together with either mAb 172.4 anti-CD4 (kindly provided by Dr. R. Macdonald, Lausanne) or mAb 3.168 anti-CD8 (kindly provided by Dr. F. Fitch, Chicago), then washed and incubated for 45min at 37^o with normal rabbit serum diluted 1/20 as a source of complement. Cells were then cultured for 3 days in DMEM containing 10% FBS, 200 U/ml IL2, and 2ug/ml ConA, then for a further 7 days in medium containing IL2 but not ConA. Staining showed them to comprise >99% CD4+ or CD8+ T cells with no detectable NK1.1+ cells.

RT-PCR analysis

RNA was prepared from aliquots of 5×10^6 freshly obtained cells or 2.5×10^6 cultured cells using RNA Bee (Biogenesis) according to the manufacturer's instructions, dissolved in water containing 1000 U/ml RNasin (N2511, Promega), and stored at −80°. cDNA was prepared by incubating RNA at 300μg/ml for 10min at 70^o, followed by incubation at 150μg/ml with 2500 U/ml MMV reverse transcriptase (M1701, Promega), 500 U/ml RNasin, 0.5mM dNTP, and 6μM random primers (S1330S, New England Biolabs) for 60min at 42^o, then 5min at 95^o before storage at −80^o. RNA samples for detection of Pro1 rev transcripts were treated with gDNA Wipeout reagent (205310, Qiagen) according to instructions so as to ensure absence of genomic DNA. To confirm the effectiveness of this, PCR reactions for detection of Pro1 rev transcripts were set up with cDNA samples prepared in parallel with or without inclusion of reverse transcriptase in the reaction: no PCR products were detected in reactions containing cDNA material prepared without reverse transcriptase. PCR reactions were set up in AB Gene Mastermix (PCR-300-600G, Fisher) with cDNA at 1/15 or 1/150 dilution and the primers (Eurofins) listed in Supplementary Table 1 at 0.2μM. Reactions were run with initial denaturation for 1min at 95^o, followed by 40 cycles of 1min at 95^o/1min at 58^o/1min

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

at 72°, followed by a final 2min at 72°. Products were analyzed on 1.0% agarose gels containing GelRed (730-2958, VWR) at 1/30,000 for 120min at 50V. In some cases the products were then purified using a Minelute Gel Extraction kit (28606, Qiagen) according to instructions, and ligated into pCR4-TOPOTA (K4575-01, Invitrogen) according to instructions. Plasmid minipreps were prepared from colonies using standard methods followed by sequencing of inserts. Oligo-capping 5'RACE was performed as described previously (25) using Generacer kits (Invitrogen).

Constructs and luciferase reporter assays

The desired fragments from Ly49 genes were amplified from appropriate BACs from the RP23 and RP24 C57BL/6 genomic libraries (BACPAC Resources Center, CHORI) using primers containing 5′ EcoRV and HindIII sites and using KOD HiFi polymerase (424771U, VWR). The HSV thymidine kinase (HSVtk) promoter was amplified from the pNS vector in the same manner; its sequence is identical to bases 24-276 of Genbank Accession J04327 preceded by the sequence GGATCAAGATC. Some fragments were prepared by gene synthesis (Eurofins). The products were digested with EcoRV and HindIII and cloned into the EcoRV and HindIII sites, or in the case of enhancer constructs into the SalI and BamHI sites, of the promoterless luciferase reporter vector pGL4 (Promega). Plasmids were purified using Genelute midiprep kits (NA0200, Sigma), and the correctness of the inserts was confirmed by sequencing. They were transfected into cells using DEAE-dextran. Briefly, cells were washed in Ca-free PBS and aliquots of 1.5 million cells were incubated with 1 μg/ml DNA and 100 μg/ml DEAE-dextran for 30 min at room temperature in a volume of 2 ml AIMV medium (31035-025, Invitrogen), then spun down, resuspended in DMEM/HEPES (D6171, Sigma) containing 10% FBS and 100 U/ml IL2, and duplicate 1 ml aliquots added to wells of 24 well plates. For D- NK cells IL2 was used at 10^4 U/ml. One day later, non-adherent cells were resuspended and transferred to microfuge tubes, spun down, and supernatant discarded. Meanwhile, 100μl of Glow Lysis Buffer (E2661, Promega) was added to the remaining adherent cells in each well and, after 10min incubation at room temperature with occasional mixing, the lysed cells were resuspended and transferred to the cell pellets in microfuge tubes. Following mixing, 30μl of lysate was mixed with 10μl of Bright-Glo reagent (E2610, Promega) in the wells of luminescence plates (7705, Thermo). These were read in a Thermo Varioskan luminometer after 10min dark adaptation. In all experiments empty pGL4 was used as a negative control. The luciferase activity of all samples was calculated as a percentage of that of the appropriate positive control plasmid. The data shown in diagrams represent the mean values from at least 2 experiments (and usually >3 experiments) in which at least 2 (and usually 3 or more) independent preparations of each plasmid were tested. Statistical analysis was done using the 2-tailed Student's t-test algorithm in Microsoft Excel.

Sequence comparisons

The sequences of mouse genes were obtained from the 2011 assembly. The sequences of rat genes were from the 2004 assembly with replacement of missing sections where possible from the 2012 assembly. Rat gene nomenclature was that of Hao & Nye (30). Sequences were aligned using the MAFFT program (31) after removing non-shared repeat sequences and large indels. Percentage identities were computed using the Genedoc program (http://

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

genedoc.software.informer.com/2.7/), and average identities were calculated as the arithmetic average of all pairwise comparisons between sets of genes.

Results

Comparative analysis of the upstream regions of mouse and rat Ly49 genes

Figure 1A shows a scale diagram of the ~7-12kb region upstream of exon 1 in each of the Ly49 genes expressed in C57 mouse NK cells, with exons E-1a, E-1b, and E1 shown in black and Pro1 as a protruding grey bar. In mature NK cells most transcripts are initiated from sites in the non-coding exon E1 or from sites upstream of and within the first coding exon E2 (not shown), under the influence of putative promoters designated Pro2 and Pro3. However, in a recent study no conventional promoter could be found at these locations (25).

The organization of all 9 genes is very similar throughout most of their length, an exception being the insertion of large LINE elements in the introns of the Ly49D and H genes. At the upstream ends of all 9 genes except Ly49J are large blocks of LINE or LTR repeat elements that extend beyond the left end of the diagram, and which potentially define the 5′ boundary of each gene. In Ly49A,C,F,G, and I these blocks of repeats begin ~1.2kb upstream of Pro1, but in Ly49D, E, and H they begin just ~400bp upstream of Pro1. In Ly49J there are no intergenic LINE/LTR blocks, the gene instead abutting directly against exon 7 of the Ly49M pseudogene. The disposition of other repeat elements is remarkably similar in the 9 genes. Of particular note is the consistent presence of an LTR-like repeat immediately upstream of, and in fact overlapping with, E-1b. Ly49 transcripts initiating from this repeat are found in mature NK cells (25).

Alignment of the non-repetitive sequences between E-1a and E1 revealed that they fall into 3 homology groups, two groups comprising the “inhibitory” genes Ly49A+G and Ly49C+E+F +I+J, and the third group comprising the “activatory” genes Ly49D+H (together with the activatory pseudogenes K, M, N, X). Within the three groups the average pairwise identity between sequences is remarkably high (89%, 88%, and 84% respectively), often exceeding that found in exons. However, between groups, homology is much lower, averaging 75% between the two groups of inhibitory genes, and 68% between the activatory genes and the two sets of inhibitory genes. Within the core Pro1 region (shown in detail in Fig 1C) the distinctiveness between the two sets of inhibitory genes is lost, the average pairwise identity between the Pro1 regions of all inhibitory genes being 92%. By contrast, although the Pro1 regions of the two activatory genes are also 92% identical to each other, they share only 76% average identity to the Pro1 regions of inhibitory genes.

Upstream of Pro1, the high levels of identity within the 3 groups of genes continue as far as the LINE/LTR blocks, but between the 3 groups there are major variations. As noted above, one variation is the extent to which the upstream LINE/LTR blocks encroach on the Pro1 region. More remarkably, ~100bp upstream of Pro1, the sequences of the Ly49A and G genes become completely unrelated to those of the other 2 groups of genes, due to the presence of a sequence that has ~65% identity to a ~750bp sequence that spans exons 7-8 of Ly49B (Supplementary Fig. 1). The striking divergence that begins at this point (marked by

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

an arrowhead in Fig. 1A) strongly suggests that sequences that reside more than ~100bp upstream of Pro1 are unlikely to be involved in controlling the expression of Ly49 genes.

To determine whether these features are conserved across species, we examined the BN rat genome which contains 34 identifiable Ly49 genes, of which 18 are potentially functional and have complete upstream sequence. In addition, genes 16 and 17 which have downstream defects in the BN strain but are expressed in other strains were included. Of these 20 genes, genes 1-4 lack a recognizable Pro1 element, and in gene 28 the first half of Pro1 has been replaced by a LINE repeat sequence. The upstream regions of the remaining 15 Pro1-containing potentially functional genes were organized in a similar manner to those in mouse Ly49 genes, with clearly recognizable E-1a and E-1b exons and large blocks of LINE or LTR-like repeats generally beginning ~1kb upstream of Pro1. Two representative inhibitory genes (11 and 25) and two representative activatory genes (16 and 17) are shown in Fig. 1B. However, in many of the rat genes, such as genes 16, 17, and 25, the size of the intron between E-1a and E-1b was considerably shorter than its ~3kb size in mouse genes due to the loss of an ~2kb segment that included most of the LTR-like sequence adjacent to E-1b.

Alignment of the sequences between the upstream LINE/LTR blocks and E1 revealed that most of these 15 rat Ly49 genes existed in 4 homology groups comprising genes 11+19, 16+17, 22+25+30+32, and 23+26+27. Within each group sequences were ~90% identical throughout the entire upstream region, whereas between groups sequences were no more identical than in introns (~67%). A striking exception, however, was the Pro1 region which was well conserved even between groups with an average identity of ~80%. Upstream of Pro1 there were, as in the mouse, some major sequence variations. In particular, a large proportion of the region >120bp upstream of Pro1 has been lost in genes 16 and 17. Assuming that genes 16 and 17 are expressed in the same manner as other Ly49s, then the sequences residing >120bp upstream of Pro1 in rat Ly49 genes are unlikely to be involved in controlling their expression.

When the upstream sequences of rat genes were compared to those of mouse genes, the average pairwise identity was similar to that between rat gene groups, namely ~65%. Once again, the Pro1 region was an exception, the average identity between rat and mouse Pro1 regions being ~80%, and in the immediate surrounding region (E-1a and the 100bp upstream of Pro1) it was 70-80%. However, the Pro1 regions of the mouse activatory genes Ly49D and H were clear outliers, having an average identity with rat Pro1 regions of only 73%. As is the case for coding sequences, there are no clear orthologous relationships between mouse and rat genes. In particular, the upstream regions of rat activatory genes, including the Pro1 regions, are no more similar to those of mouse activatory genes than to those of mouse inhibitory genes. Taken together this analysis of the upstream regions of mouse and rat Ly49 genes confirms the notion that the Ly49 genes in both species have evolved rapidly and independently from potentially a single ancestral gene, but with selective conservation of Pro1 and to a lesser extent of the ~100bp region upstream of Pro1 and of the ~400bp E-1a region downstream of Pro1.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Pro1-forward and reverse transcripts are widely expressed in NK cells and T cells

In the previous study of Saleh et al. (20), transcripts emanating from the Pro1 region were found in immature NK cells in various tissues and in the immature NK cell line, LNK. To examine whether they might also be present in other cells, we first performed RT-PCR analysis on a range of cells using primers designed to specifically detect fwd and rev transcripts in the Pro1-region of the Ly49A, C, E, and G genes. Amongst the cells tested were an immature fetal thymus NK cell line developed in our own laboratory, I2/22, that shows uniform surface expression of Ly49E but no other Ly49, sorted adult splenic NK cells grown for 12 days in IL2 and the mature adult splenic NK cell line D- that display surface expression of all four Ly49s, the RMA T cell line (which is a subline of EL4 (26), and is therefore designated here EL4/R) that expresses Ly49A and G but no other Ly49, the EL4 subline EL4/S and the NKT cell line C1498 that have no detectable surface Ly49 expression, and a wide range of other cell lines that lack detectable expression of Ly49s except for J774 that weakly expresses Ly49C.

Ly49E Pro1 fwd transcripts were readily detectable not only in the immature NK cell line I2/22, but also in highly purified adult splenic NK cells, the mature NK cell line D-, and the T cell lines EL4-R and C1498 **(**Fig. 2**)**. By contrast, Ly49E Pro1 fwd transcripts could not be detected in EL4/S or in a wide range of non-NK non-T cell lines, including the YAC-1 leukemia line, the J774 macrophage line, the P815 mastocytoma line, the A20 and NS0 B cell lines **(**Fig. 2**),** and also L1210, L1578, P388D1, FBL3, RAW264, MPC11, and WEHI3 (data not shown). The amplimers generated from all of the positive cell lines displayed a ladder of bands. Cloning and sequencing of the largest amplimer from each of these cell lines revealed a product containing all expected exons, with E-1a spliced directly to E1 exactly as previously reported (Genbank AF419251). The smaller amplimers also began with E-1a, but this was spliced directly to exons 2, 3, 4, 6, or 7. In addition, various products with missing downstream exons were found. Although transcripts lacking the non-coding exon 1 might be expected to be functional, most or all of the other alternatively spliced transcripts would be expected to be nonfunctional due to the loss of critical elements, such as the translation start codon in exon 2, the transmembrane region encoded by exon 3, the interchain disulfide bonds and stalk region encoded by exon 4, and the ligand binding site formed by the folding of the NK domain encoded by exons 5-7. A summary of all the Pro1fwd splice variants found in this study is provided in Supplementary Fig. 2.

Ly49E Pro1 rev transcripts were detected in immature and mature NK cells, and sequencing of the amplimers confirmed their identity. Very low levels were occasionally detected in EL4/R, C1498, P815 and NS0 lines but not in EL4/S, YAC, J774 or A20. Ly49E Pro2 transcripts were analyzed using a fwd primer specific for sequence in the distal part of E1, upstream of the splice point used in Pro1 transcripts (20), thereby enabling the specific detection of Pro2 transcripts independently of Pro1 transcripts. Although the frequency of Pro2 transcripts initiating so far upstream would be expected to be low (25), they could be clearly detected in immature and mature NK cells, but not in any of the other cells, including EL4/R and C1498.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Ly49G Pro1 fwd transcripts displayed the same pattern of expression as Ly49E Pro1 fwd transcripts, being clearly present in mature NK cells and in the T cell lines EL4/R and C1498. Cloning and sequencing of the amplimers, which had much less heterogeneity than the Ly49E Pro1 fwd amplimers, showed that most had a sequence corresponding to that previously reported (Genbank AF419249), with E-1a spliced to E-1b which in turn was spliced directly to E2; however, in 2 clones E-1b was absent (see Supplementary Fig. 2). The sequences of all the alternately spliced Ly49G Pro1 transcripts that we detected have been deposited in Genbank (KP687233-KP687241. Ly49G Pro1 fwd transcripts could also be detected, albeit at lower levels, in the immature NK line I2/22. Although these cells lack detectable surface expression of Ly49G, Pro2 transcripts, detected using a fwd primer specific for sequence in E1 which is absent from Ly49G Pro1 transcripts (20), were present at low levels. As expected, Ly49G Pro2 transcripts were abundant in the three cells that expressed Ly49G on the surface, and were also present at low levels in C1498 which lacked detectable surface Ly49G. Ly49G Pro1 rev transcripts were found in immature and mature NK cells, and also in the EL4/R line, but not in any other cells including C1498. The primers used in these Pro1 rev PCRs matched sequences in both the Ly49G and Ly49A genes, and cloning of the amplimers from D- cells showed the presence of both products.

Pro1 fwd transcripts from the Ly49A and Ly49C genes were detected in immature and mature NK cells and in the EL4/R T cell line, but not in C1498 or other cells with the exception that Ly49C Pro1 transcripts were present in the J774 macrophage line which curiously expresses Ly49C. Ly49A and C Pro1 transcripts could only be detected using reverse primers specific for sequences in exon 1, but not with reverse primers specific for sequences in exon 4 or exon 7, suggesting that they frequently terminate prematurely (unpublished observations). Even then, the signals were weak, suggesting low abundance compared to Ly49E and G Pro1 transcripts. Cloning of the Ly49A Pro1 fwd amplimers revealed not only products corresponding to those described previously (Genbank AY078436), but also novel variants with an enlarged E-1a exon (Genbank KP339879), or with a novel 182 base exon located between E-1a and E-1b (Genbank KP339880; see Supplementary Fig. 2). Cloning of Ly49C Pro1 fwd amplimers revealed a Pro1 transcript sequence different to any of those from other genes, that lacked E-1b and instead contained a novel 265bp exon located between E-1a and E-1b which was spliced directly to exon 1 via a non-canonical splice donor site (Genbank KP339881; see Supplementary Fig. 2).

Fresh cells were also tested. As reported previously, fresh adult bone marrow cells contained Ly49G Pro1 transcripts, but so too did adult thymus and spleen cells **(**Fig. 3**)**. Spleen cells from $RAG^{-/-}$ but not from $RAG^{-/-}$ $\gamma c^{-/-}$ mice contained Ly49G Pro1 transcripts, consistent with expression in NK cells, and sorting revealed that fresh splenic T cells that were >99% pure, but not B cells, expressed Ly49G Pro1 transcripts. Thymus, bone marrow, and spleen cells also showed low expression of Ly49E Pro1 transcripts, which was increased in $RAG^{-/-}$ spleen cells and absent in $RAG^{-/-}$ $\gamma c^{-/-}$ spleen cells. In addition, pure populations of T cell blasts that contained no detectable $NK1.1^{+}CD3^{-}$ cells expressed Ly49E Pro1 and Ly49G Pro1 transcripts. Surprisingly, Pro1 transcripts were more prominent in CD4 blasts than CD8 blasts but despite their presence (and also of Ly49G Pro2 transcripts) no expression of Ly49E or Ly49G at the cell surface could be detected (not shown).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Pro1-fragments display predominantly reverse promoter activity, and neither forward nor reverse promoter activity correlates with gene expression in NK cells

Having established the prevalence of Pro1 fwd and rev transcripts in various cells, the promoter activities of core Pro1 regions were systematically examined by preparing exactly homologous ~120bp fragments (shown in Fig. 1C) that began 8bp upstream of the 5′ TATA box, thereby including the potentially important C/EBP-1 site adjacent to the 5′ TATA box (21), and ended 13bp downstream of the 3′ TATA box so as to give fragments corresponding to the average length of the constructs used by Saleh et al (see Fig. 6 of their paper (21)) and which contained all of the C/EBP sites adjacent to the 3′ TATA box that were postulated to be important for Pro1 function. All Pro1 fragments were cloned in both fwd and rev orientations into the EcoRV-HindIII sites in pGL4, immediately in front of its promoterless luciferase gene. When transfected into the LNK line the Pro1 rev constructs from all of the genes tested except Ly49D and H displayed strong activity, often exceeding that of the control HSVtk promoter, whereas the corresponding fwd constructs had only low activity (Fig. 4A**)**. When transfected into various immature NK cell lines developed in our own laboratory, exemplified by the CBA line in Figure 4A, the bias towards rev activity was even more pronounced, with none of the fwd constructs displaying more than minimal activity.

These results differ from those reported by Saleh et al. (21), in that the promoter activity of the core Pro1 constructs is essentially unidirectional (in the reverse direction) rather than bidirectional, and the promoter strength of Pro1 fwd or Pro1 rev fragments does not correlate with the frequency of expression of the corresponding Ly49 products in NK cells. For example, the core Ly49G Pro1 rev fragment displayed similar activity to the Ly49A and Ly49J Pro1 rev fragments even though Ly49G is expressed much more frequently in NK cells than Ly49A or J, and the Ly49C, G, and I Pro1 fwd constructs did not display significantly greater activity than the Ly49A, E, and F fwd constructs even though Ly49C, G, and I are expressed much more frequently than Ly49A, E, and F. A possible explanation for the differences found in the present study could be the employment of the pGL4 luciferase vector rather than the pGL3 luciferase vector used in the Saleh et al. study. However, when we inserted the Ly49G and Ly49E Pro1 fwd and rev constructs into the EcoRV-HindIII sites of pGL3, their relative activity was essentially identical to that of the pGL4 constructs shown in Fig. 4A (data not shown). Another possibility could be that the 5′ and 3′ boundaries of the Pro1 constructs used in the present study differed slightly from those of Saleh et al. However, when we tested the exact same constructs shown in Figure 6 of the Saleh et al study the results were again very similar to those shown in Fig. 4A. A further possibility could be that the different results obtained by Saleh et al were due to the presence of pCR4-TOPO-TA vector sequences in their subcloned fragments or to the variable use of SacI-XhoI v. XhoI-HindIII sites in the pGL3 vector, but this was not tested.

We also wondered whether extension of the core Pro1 constructs to include the conserved regions upstream and downstream of Pro1 might influence activity. Inclusion of the ~180bp downstream E-1a region (ending 4bp upstream of the 3′ end of the E-1a exon so as to minimize the likelihood of creating a splice donor site) caused little general change in the pattern of activity in CBA cells but caused a marked increase in the activity of the Ly49C, I, and J fwd constructs in LNK cells (Fig. 4B). Despite this, the activity of Pro1-E-1a

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

constructs in LNK cells still failed to correlate with expression levels: for example, the Ly49I and Ly49J Pro1-E-1a fwd constructs had similar activity despite Ly49I being expressed much more frequently in NK cells than Ly49J, and the Ly49C and I Pro1-E-1a fwd constructs had markedly higher activity than the Ly49G Pro1-E-1a construct even though Ly49C, G, and I are all expressed at a similar high frequency in NK cells.

Subsequent experiments showed that deletion of just 65bp from the 3′ end of the E-1a region of the Ly49C and Ly49I Pro1-E-1a fragments significantly reduced fwd activity, and deletion of 149bp reduced it further (Fig. 4C**)**, indicating that the fwd promoter activity of the Ly49C and Ly49I Pro1-E-1a fragments in LNK cells is dependent on multiple sites in the E-1a region. Surprisingly, the E-1a regions from the Ly49C, and to a lesser extent from the Ly49I genes, by themselves displayed significant promoter activity in both fwd and rev orientations (Fig. 4C**)**. By contrast, the E-1a regions from the Ly49E and Ly49G genes had little or no promoter activity (data not shown).

When Pro1 constructs were extended at the 5′ end to include the highly conserved ~100bp region that lies immediately upstream of Pro1 there was no improvement in the low promoter activity of fwd constructs, but a marked reduction in the promoter activity of rev constructs (Fig. 4D**)**. This latter result might be because the upstream region introduced additional potential start codons ahead of the luciferase gene that disrupted its expression or function. However, as noted above, inclusion of the E-1a region which also contains multiple potential start codons had no obvious adverse effect on promoter activity. Therefore a more likely explanation is that the conserved upstream regions contain sites that bind repressor factors.

Taken together, the data indicate that the Pro1 region is a more complex regulatory region than previously realized with some genes having promoter activity in the E-1a region and many genes potentially having repressor sites in the conserved upstream region. The lack of promoter activity in the core Pro1 regions of the Ly49D and H genes could not be reversed by the inclusion of either the E-1a region (Fig. 4B**)** or the conserved upstream region (Fig. 4D**)**. Indeed, the core Pro1 fragment from these genes often inhibited the background luciferase activity of the pGL4 plasmid (not shown), indicating that they contain repressor sites.

Pro1 constructs display preferential activity in NK and T cell lines

The data presented above show that Pro1 transcripts are found in vivo in immature NK cells, mature NK cells, and T cells, but not generally in B cells and other cells. To examine whether the promoter activity of the Pro1-region displays a similar tissue specificity in vitro, the Ly49G Pro1-E-1a rev construct was transfected into a range of cultured cell lines. As noted above and confirmed in Figure 5, it had strong promoter activity not only in the LNK immature NK cell line, but also in other immature NK cell lines (CBA, I2-22, and 1608b) developed in our own laboratory. It also displayed strong activity in the mature NK cell line D- and in several T cell lines. By contrast, its activity in non-T cell lines was generally low, although moderate activity was observed in P815. Interestingly, it displayed high activity in the rat NK cell line RNK, but no activity in the rat B cell line YB2. An essentially identical cell-type specificity was found for Ly49A and Ly49E Pro1-E-1a rev constructs and for the

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

core Ly49G Pro1 rev construct (not shown). By contrast, the Ly49A, E, and G Pro1-E-1a fwd constructs displayed little or no activity in any of the cell lines, and the Ly49C Pro1-E-1a fwd construct that was active in LNK cells displayed minimal activity in all other immature NK, mature NK cell, and T cell lines that were tested.

The Pro1 sequences of rat and mouse genes are highly conserved, but core Pro1 promoter activity is not conserved

The finding that mouse Pro1 constructs had high activity in rat NK cells raised the question of whether rat Pro1 constructs would be active in mouse cells. As noted above, there is conservation between mouse and rat Pro1 sequences, and this is particularly striking in the boxed central region shown in Figure 1C. Of the 64-68 residues in this central region, 62 are invariant amongst the genes encoding mouse inhibitory Ly49s, 47 amongst all mouse Ly49 genes, and 31 amongst rat Ly49 genes. By contrast, homology is markedly lower at the 5′ and 3′ ends with poor conservation of the TATA boxes. Indeed, analysis of the entire set of Pro1-regions using the Jaspar algorithms (32) indicated that the only gene that contained a recognizable TBP binding site at the 3′ end of the forward strand was Ly49G, and that although most genes contained a high scoring TBP binding site at the 5′ end this was on the forward not reverse strand. Ly49C and I contained a second forward TATA box ~20bp downstream of the 5′ TATA box.

To test whether rat Pro1-regions displayed promoter activity we cloned the core Pro1 regions from two representative “inhibitory” genes (11 and 25) and from two representative “activatory” genes (16 and 17) into pGL4 (sequences shown in Fig. 1C). Unexpectedly, even the fragments from “inhibitory” genes displayed little activity in either fwd or rev orientation in mouse LNK or CBA cells or in rat RNK cells (Fig. 6A). However, when the E-1a regions were included in the constructs, the fwd and rev fragments from gene 25 displayed markedly enhanced promoter activity, and the rev fragment from gene 17 had higher promoter activity than that of the Ly49G control construct (Fig. 6B). These data indicate that in contrast to the situation for mouse Pro1 regions, promoter activity is not consistently found in the core Pro1 regions of rat “inhibitory” genes, but in parallel to the situation in the mouse, promoter activity may be present in the conserved downstream E-1a region, including in “activatory” genes.

Pro1 promoter activity does not require TATA boxes or C/EBP binding sites, and is determined predominantly by the conserved central region

To establish which parts of the core Pro1-region were essential for promoter activity a series of truncated constructs was tested. Deletion of 5 or 12bp from the 5′ end of the 122bp Ly49G Pro1 rev fragment caused no diminution in promoter activity **(**Fig. 7A**)**. Surprisingly, deletion of a further 9bp that included the upstream TATA box, giving a total deletion of 21bp, only partially reduced promoter activity. (Note that the upstream TATA box in the Ly49G Pro1 rev construct corresponds to the downstream TATA box in the native gene.) Similarly, deletion of 15bp from the 3′ end that included the downstream TATA box of the Ly49G Pro1 rev construct caused negligible reduction in promoter activity, and more substantial deletions of 20, 25, or 32bp from the 3′ end also had little effect **(**Fig. 7B**)**. However, fragments in which the deletions extended into the conserved core region shown in

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Figure 1C, namely by removing 43bp from the 5′ end **(**Fig. 7A**)** or 37bp from the 3′ end **(**Fig. 7B**)**, had virtually no activity.

Deletions were then made at the 5′ end of the 3'Del15 fragment that lacked a 3′ TATA box **(**Fig. 7C**)**. Deletions of 5 or 12 bases had no effect on promoter activity. Importantly, deletion of 21 bases created a fragment (5'Del21-3'Del15) that lacked both the 5′ and 3′ TATA boxes but which retained substantial promoter activity. In addition, although removal of 43bp from the 5′ end of the full Pro1 core fragment virtually eliminated promoter activity **(**Fig. 7A**)**, further removal of 15bp from the 3′ end resulted in a 64bp fragment (5'Del43-3'Del15) that had regained almost full promoter activity **(**Fig. 7C**)**. This fragment, which corresponds to alignment positions 16-89 in Figure 1C, and has the sequence GAACAGTGTAAATGCAAGAACAGGAAATCTCAAAATAGAGCTGTTTGTGGTTTTC CTAAGATCT, not only lacks the 5′ and 3′ TATA boxes but also all of the C/EBP binding sites postulated to be important for Pro1 function (21), demonstrating that neither the TBP nor C/EBP binding sites are required, and that Pro1 activity is primarily determined by residues in the central highly conserved core region. In addition, the high activity of the 5'Del43-3'Del15 fragment compared with the 5'Del43 fragment confirms that Pro1 is a complex regulatory region that includes potential repressor sites.

Pro1 promoter activity is dependent on residues in the predicted AML1 and NFkB binding sites

To examine in more detail which residues in the core region are important, and also to gain some understanding of why the Pro1-regions of Ly49D, Ly49H, and rat Ly49 genes are devoid of promoter activity, we first prepared an Ly49D Pro1 rev construct that corresponded to the 5'Del12-3'Del15 Ly49G Pro1 rev construct, together with four chimeric constructs, DG1-DG4, in which different segments of the 95bp 5'Del12-3'Del15 Ly49G Pro1 Ly49G fragment were replaced with the corresponding segments from Ly49D. The sequences of these fragments are shown in Figure 8A. As expected, the 5'Del12-3'Del15 Ly49G rev fragment had strong promoter activity, whereas the corresponding Ly49D rev fragment had no activity **(**Fig. 8B**)**. The DG2 chimera was also devoid of activity, whereas the DG3 chimera had about 50% activity, and the DG1 and DG4 chimeras had close to full activity despite their multiple substitutions compared to the parental Ly49G rev fragment. These data confirm the importance of the central region of Pro1, and indicate that the lack of promoter activity in the Ly49D Pro1-fragment is mainly due to one or more of the six Ly49D residues in the DG2 chimera, with perhaps a minor contribution from the three Ly49D residues in the DG3 chimera.

To determine which of these differences was important, mutant constructs were prepared in which individual residues in the 5'Del12-3'Del15 Ly49G rev fragment were changed to those in Ly49D. The sequences of these, designated DG2-1 to DG2-6 and DG3-1 to DG3-3, are again shown in Figure 8A. Luciferase assays performed with these mutants showed that none of the three Ly49D residues present in the DG3 chimera could by themselves significantly reduce the promoter activity of the Ly49G rev fragment **(**Fig. 8C**)**, nor did mutations corresponding to Ly49D residue 4 (mutant DG2-4) or Ly49D residue 6 (a single bp deletion, mutant DG2-6) from the DG2 chimera **(**Fig. 8D**)**. By contrast, a mutation

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

corresponding to Ly49D residue 2 from DG2 (mutant DG2-2) reduced activity by ~50%, and mutations corresponding to Ly49D residues 1 and 3 (mutants DG2-1 and DG2-3) were each by themselves sufficient to ablate the activity of the Ly49G rev fragment.

Variant residues 1, 2, and 3 of the DG2 chimera represent a trio of adjacent bases that in Ly49D disrupt the central NFkB and AML1 binding sites present in Ly49G and all other Pro1 regions that display significant promoter activity. Interestingly, variable residues 1 and 3 have undergone identical alterations in Ly49H, while variable residue 2 in Ly49H is G compared to T in Ly49G and A in Ly49D **(**Fig. 8A**)**. When a G residue was introduced at this position in the Ly49G rev fragment it greatly reduced promoter activity (mutant HG2-2, Fig. 8D). Interestingly, a G is also found at this position in the majority of rat Ly49 Pro1 regions (position 72 in the alignment in Fig. 1C). Another change consistently found in rat Ly49 Pro1 regions, and also in Ly49D and Ly49H, is a C to T transition 9 bases downstream at position 81 **(see** Fig. 1C**)**, near the 3′ end of the NFkB binding site (Fig. 8A**)**. Although this change by itself does not affect Pro1 activity (mutant DG2-4, Fig. 8D**)**, when combined with the T to G change at position 72 it virtually eliminated promoter activity (mutant HG2-24, Fig. 8D**)**. Taken together, these data demonstrate that the base changes that disrupt the NFkB and/or AML binding sites in the mouse Ly49D, Ly49H, and rat Pro1 regions are sufficient by themselves to account for the lack of promoter activity in the core Pro1 regions of these genes.

Pro1 displays enhancer activity

In our previous paper analysing transcription from the Pro2 and Pro3 regions, we speculated that Pro1 may act as an enhancer (25). To test this hypothesis, we inserted the Ly49G core Pro1 fragment in both orientations into the distal BamHI-SalI site in pGL4 such that it was located 2.2kb upstream of the luciferase gene. Such constructs had negligible luciferase activity, demonstrating that the Ly49G Pro1 fragment in this location was too distal from the luciferase gene to exert measurable promoter activity **(**Fig. 9A**)**. However, when present in the same position in a pGL4 plasmid that contained the HSVtk promoter in the proximal EcoRV-HindIII site immediately upstream of the luciferase gene, the distal Ly49G Pro1-fragment caused a substantial increase in luciferase expression in the immature NK cell lines LNK and CBA, but not in the T cell line EL4/R **(**Fig. 9A and 9B**)**. Importantly, and in marked contrast to the situation for promoter activity, the enhancer activity of the distal Ly49G Pro1 fragment was bidirectional with essentially identical activity in both orientations.

The enhancement of HSVtk promoter activity in LNK cells caused by the distal Pro1G fwd and Pro1G rev fragments was significant at the $P = 4 \times 10^{-12}$ and $P = 4 \times 10^{-8}$ levels respectively, with similar figures for the CBA line. By contrast, fragments from the coding sequences of Ly49G, Ly49E, and K^b MHC class I genes had no significant enhancer effect **(**Fig. 9A and 9B**)**. Neither did the Pro1 fragments from the mouse Ly49D and H genes nor from the rat Ly49-11, 16, and 17 genes **(**Fig. 9C**)**. By contrast, the Pro1 fragment from the Ly49C gene had strong enhancer activity. So too did the Pro1 fragment from the rat Ly49-25 gene **(**Fig. 9C**)** despite its very limited promoter activity **(**Fig. 6**)**, indicating that the promoter and enhancer activities of Pro1 fragments are distinct and separable functions.

Europe PMC Funders Author Manuscripts

Enhancer activity was generally unaffected by the presence of the E-1a region or the conserved upstream region **(**Fig. 9D**)**, with the striking exception of the rat Ly49-17 gene, where extended fragments containing the E-1a region had very high enhancer activity. Thus, rather surprisingly, the E-1a region of the rat Ly49-17 gene possesses or contributes to both strong promoter activity **(**Fig. 6**)** and strong enhancer activity **(**Fig. 9D**)**.

To determine which elements in the core Pro1 fragment were necessary for enhancer activity, truncated Ly49G Pro1 fragments were tested. As shown in Figure 9E, deletion of either the 5′ or 3′ TATA boxes, together with the associated C/EBP binding sites, had no effect on enhancer activity, the 64bp 5'Del43-3'Del15 fragment having the same high level of enhancer activity as the 122bp full length core Pro1 fragment. However, the same point mutations in the AML and NFkB binding sites that abolished promoter activity **(**Fig. 8**)** also greatly reduced enhancer activity **(**Fig. 9F**)**.

In conclusion, although examples were found in which fragments possessed strong enhancer activity but little of no promoter activity (the Ly49G fwd and rat Ly49-25 rev core Pro1 fragments), there was a striking correlation between enhancer and promoter activity, both being dependent on critical residues in the AML and NFkB binding sites, and with promoter activity only detectable in fragments that also possessed enhancer activity. Coupled with the lack of requirement for TATA boxes, the data suggest that the transcription factors responsible for enhancer activity, which presumably bind to the AML and NFkB sites in the center of the Pro1 element, are necessary and perhaps sufficient for the recruitment of transcription initiation complexes to Pro1, and it is this interaction that drives the promoter activity observed in Pro1 fragments.

Discussion

Understanding how the unusual stochastic expression of MHC cI receptors is achieved represents one of the greatest challenges in NK cell biology. Initial studies of the Ly49A gene suggested that the transcription of Ly49 genes in mature NK cells was controlled by a conventional TATA-box-containing promoter, subsequently designated Pro2, located ~130bp upstream of the exon1/intron1 boundary (33). Short sequences of ~250bp from this region displayed promoter activity when transfected into the EL4 T cell line, whereas slightly longer sequences containing an additional ~150bp had markedly reduced activity, suggesting the presence of an upstream repressor site (34, 35). This conventional model of Ly49 gene transcription has now been invalidated by studies showing that (a) in many Ly49 genes, including Ly49A, transcripts originate from random sites dispersed throughout exon 1, including upstream of the supposed site of the Pro2 promoter, (b) transcripts also frequently initiate from sites in intron 1 and exon 2, under the influence of the putative Pro3 promoter, including downstream of the translation start codon, (c) the TATA box in the putative Pro2 promoter is not present in many Ly49 genes, (d) although Pro2 fragments from the Ly49A gene have promoter activity in the EL4 T cell line, corresponding fragments from other Ly49 genes have little or no promoter activity in EL4 cells, (e) the presence of an inhibitory site upstream of the Ly49A Pro2 fragment cannot be consistently demonstrated, (f) Pro2 fragments from Ly49 genes, including Ly49A, have no significant promoter activity in mature NK cells (25, 36). Interestingly, in some Ly49 genes, namely Ly49B and Q that are

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

expressed predominantly in myeloid cells, Ly49F that is expressed predominantly in T cells, and Ly49D and H that encode activatory receptors in NK cells, transcript initiation is concentrated at an Inr element ~110bp upstream of the exon1/intron1 boundary (25). However, even in these genes, transcripts initiating from distal positions can be found.

Given that Ly49 transcription in mature NK cells does not initiate from a conventional proximal promoter, we examined whether the distal Pro1 promoter might be active in mature NK cells. Transcripts containing the E-1a exon of Ly49A, C, E, and G spliced to downstream exons were readily detected not only in immature NK lines but also in highly purified adult splenic NK cells and in the cloned adult mature NK cell line D-. Importantly, Pro1 fwd transcripts could be readily detected in vivo, not only in bone marrow where immature NK cells reside, but also in thymus and spleen. They were markedly enriched in the spleens of $RAG^{-/-}$ mice, but were undetectable in $RAG^{-/-}$ $\gamma c^{-/-}$ mice, indicating that they reside mainly in NK cells. 5'RACE experiments on D- cells showed that the E-1a-containing Ly49G transcripts initiated at the same point in E-1a as previously reported, and revealed that the Ly49A transcripts also initiated in E-1a but at a point 59bp further downstream (Gays and Brooks, unpublished). Pro1 fwd transcripts were also expressed in some T cell lines, in activated CD4 and CD8 T cells in vitro, and in splenic T cells in vivo, but not generally in other cell types including B cells and myeloid cells. Similarly, Pro1 rev transcripts were clearly expressed in mature NK cells and T cells, although whether these initiated close to Pro1 was not formally demonstrated. Taken together, these data, which confirm and extend our earlier preliminary observation of Ly49E Pro1 transcripts in cultured adult splenic NK cells (37) and a more recent one where Pro1 transcripts were found in the spleen, bone marrow, and liver of mice following bone marrow transplantation or injection of IL2 (38), demonstrate that the production of Pro1 transcripts is not switched off in mature NK cells as previously believed, and is not restricted to NK cells but can also occur in T cells.

Interestingly, we found two examples of cells that contained readily detectable Pro1 fwd transcripts but not Pro2 transcripts or surface protein from the corresponding gene (Ly49E in EL4/R and C1498 cells). This indicates that the production of Pro1 fwd transcripts does not automatically lead to the production of significant amounts of functional Pro2 transcripts, perhaps because this requires additional epigenetic changes. The failure of Pro1 transcripts themselves to generate measurable surface protein may be at least in part due to extensive alternate splicing (Ly49E), premature termination (Ly49A and C), and low abundance. We also found three examples of monoclonal cell lines that displayed uniform expression of a particular Ly49 at the cell surface and expressed not only the corresponding Pro2 transcript, but also the corresponding Pro1 fwd and rev transcripts (Ly49E in I2/22 and D- cells, and Ly49G in D- cells). Assuming that the cells contain no more than two functional copies of the relevant gene, the only explanation compatible with the original Pro1 hypothesis would be that one allele is producing Pro1 rev transcripts and is thereby switched off, and the other allele is producing both Pro1 fwd and Pro2 transcripts. An alternative explanation would be that in active Ly49 genes that are producing functional Pro2 transcripts, the Pro1-region is also active and generates sterile transcripts in both directions.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Consistent with the cellular distribution of Pro1 transcripts in vivo, DNA fragments from the Pro1 region of all “inhibitory” C57 Ly49 genes displayed clear promoter activity in vitro not only in immature NK cells but also in mature NK cells and in some T cell lines, with little or no activity in non-NK non-T cell lines. A striking feature of the results was the generally weak promoter activity of Pro1 fwd fragments, the only ones having clear activity being those of the Ly49C, I, and J genes, and even for these three genes fwd activity was restricted to the LNK cell line and was dependent on the inclusion of the E-1a region in the constructs. Remarkably, the E-1a regions from at least two of these genes displayed substantial promoter activity by themselves, in the complete absence of Pro1. No improvement in fwd promoter activity could be achieved by extending the constructs to include the highly conserved region upstream of Pro1. The lack of fwd promoter activity in core Pro1 constructs cannot be due to the lack of appropriate transcription factors in the test cells because the LNK and other test cells that we used express endogenous Pro1 fwd transcripts. Taken together, the data indicates that the efficient generation of Pro1 fwd transcripts is dependent on additional sequence elements in the E-1a region and more distally from Pro1.

Most surprisingly, the promoter activity of Pro1-fragments was not dependent on sequences at the 5′ and 3′ ends of the core Pro1-fragment that contain the the fwd and rev TATA boxes and C/EBP binding sites that were hypothesized to be important for the binding of the transcription complex (21). This is in line with the fact that the 5′ and 3′ ends of the core Pro1-fragments of mouse and rat Ly49 genes show substantial sequence variation and the TATA boxes and C/EBP sites located in these regions are not well conserved. In fact the TATA box at the 3′ end of the core Pro1 fragment is found only in the Ly49G gene. These observations are consistent with genome wide analyses showing that only a minority of promoters in eukaryotes contain recognizable TATA boxes (39, 40). Instead, promoter activity mapped to a central 64bp fragment that is highly conserved in mice and rats. Surprisingly, despite the high degree of homology between rat and mouse Pro1-regions, especially in the central 64bp fragment, none of the four rat Pro1-fragments tested displayed significant promoter activity even in rat NK cells that supported high promoter activity of appropriate mouse Pro1-fragments. Mutational analysis indicated that the lack of promoter activity in Ly49D, Ly49H, and rat Ly49 core Pro1 fragments is accounted for by base changes in the predicted AML and NFkB sites that are found near the center of all of the Pro1-regions that display promoter activity in vitro.

Compelling evidence that the Pro1-region plays an important role in the expression of Ly49 genes in vivo was provided by Tanamachi et al (41) who found that in NK cells but not in liver cells a DNAseI hypersensitivity site is present in the Pro1-E-1a region of Ly49 genes, and that deletion of an ~2kb fragment beginning ~800bp upstream of Pro1 prevented the expression of Ly49A from a 32kb genomic transgene. However, the results of the present study indicate that the function of Pro1 is different from that originally proposed. Firstly, it is now clear that the Pro1-region is active not only in immature NK cells but also in the other cell populations that have the potential to express Ly49 molecules, namely mature NK cells and T cells, whereas it is generally not active in other cell types. Despite this, its activity, as measured by the expression of Pro1 transcripts does not correlate with the expression of the associated Ly49 gene as exemplified by the presence of various Pro1 transcripts in the immature NK line LNK that lacks detectable expression of Ly49s (20) and of Ly49E Pro1

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

transcripts in the T cell lines EL4/R and C1498 that do not express Ly49E. This is consistent with the finding of Tanimachi that the DNAseI hypersensitivity site in the Pro1-region of Ly49A is open in both Ly49A+ and Ly49A− NK cells (41). Secondly, promoter activity of the Pro1-region can be largely attributed to an ~60bp fragment that lacks TATA boxes (and also Inr sequences). This does not invalidate the notion that fwd and rev transcription at Pro1 is initiated in a mutually exclusive competitive manner (20). However, since the publication of the original data on Pro1 it has become clear that bidirectional transcription from promoters is common, and may routinely be initiated at all nucleosome free (DNaseI hypersensitivity) sites (40, 42-44), with the predominance of fwd coding transcripts being determined by factors that prevent premature termination (45). Thirdly, Pro1 promoter activity in vitro is often weak or undetectable. Thus most mouse Pro1-regions display at best weak fwd activity in most relevant cells, Ly49D and H have no detectable fwd or rev activity in any relevant cell, and despite the high conservation of Pro1 sequences between mice and rats, the Pro1-regions of rat genes have little or no promoter activity in either mouse or rat cells. This suggests that regions distal from Pro1 are required to promote efficient transcription, or that the conservation of sequence within Pro1 is not driven by its promoter activity but by some other function. Fourthly, we demonstrate here that Pro1 indeed has a clearly demonstrable alternative function at least in vitro, namely acting as an enhancer. This does not conflict with the evidence that Pro1 can act as a promoter, because recent data indicates that enhancers are frequently sites of active transcription (46). Fifthly, the generation of coding transcripts from most Ly49 genes is initiated over a wide area that includes exon1, intron 1, exon 2, and sites further downstream, with no evidence of the existence of a classical localized promoter.

We therefore propose the following modification of the original Pro1 hypothesis. A necessary but not sufficient step in the activation of Ly49 genes is the creation of a nucleosome free region at Pro1. This event is cell lineage specific, occurring in NK cells and T cells, but not in most other cell types, and at least in NK cells is an irreversible step that occurs early in NK cell development. In conjunction with additional “remodeling” events elsewhere in the gene, two linked stochastic events occur during Ly49 gene activation: the probabilistic formation of a stable physical association of Pro1 with a downstream point in the gene, and variation in the exact location of that point. For most inhibitory Ly49 genes the site may be anywhere in the exon 1-intron 1-exon 2 region. In this model, transcription from Pro1 might either be a likely but non-essential consequence of the binding of transcription factors to the Pro1 nucleosome free region, or it may be an essential part of creating and maintaining enhancer function, perhaps by facilitating the interaction of the enhancer with downstream regions (47), or it may be a reflection of the fact that Pro1 acts not only as an enhancer but as the effective functional promoter in the gene, recruiting RNA polymerase and transferring it via chromatin looping to downstream sites (48-50). In the case of mouse “activatory” Ly49 genes, whose sequences, including in the Pro1 region, are distinct from all other “inhibitory” and “activatory” genes in mice and rats, the disruption of enhancer activity in the Pro1 region caused by the absence of AML and NFkB binding sites may be the critical factor responsible for the distinct pattern of expression of these receptors (51-54).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We are grateful to Dr. Stephen Anderson for the provision of cells and Ly49 constructs and for many helpful discussions, and to Ian Harvey and Ian Dimmick for expert advice and assistance with flow cytometry and cell sorting.

References

1. Raulet DH, Guerra N. Oncogenic stress sensed by the immune system: role of natural killer cell receptors. Nat Rev Immunol. 2009; 9:568–580. [PubMed: 19629084]
2. Lanier LL. NK cell recognition. Annu Rev Immunol. 2005; 23:225–274. [PubMed: 15771571]
3. Joncker NT, Raulet DH. Regulation of NK cell responsiveness to achieve self-tolerance and maximal responses to diseased target cells. Immunol Rev. 2008; 224:85–97. [PubMed: 18759922]
4. Ljunggren HG, Karre K. In search of the 'missing self': MHC molecules and NK cell recognition. Immunol Today. 1990; 11:237–244. [PubMed: 2201309]
5. Arase H, Mocarski ES, Campbell AE, Hill AB, Lanier LL. Direct recognition of cytomegalovirus by activating and inhibitory NK cell receptors. Science. 2002; 296:1323–1326. [PubMed: 11950999]
6. Smith HR, Heusel JW, Mehta IK, Kim S, Dorner BG, Naidenko OV, Iizuka K, Furukawa H, Beckman DL, Pingel JT, Scalzo AA, Fremont DH, Yokoyama WM. Recognition of a virus-encoded ligand by a natural killer cell activation receptor. Proc Natl Acad Sci U S A. 2002; 99:8826–8831. [PubMed: 12060703]
7. Belanger S, Tai LH, Anderson SK, Makrigiannis AP. Ly49 cluster sequence analysis in a mouse model of diabetes: an expanded repertoire of activating receptors in the NOD genome. Genes Immun. 2008; 9:509–521. [PubMed: 18528402]
8. Scarpellino L, Oeschger F, Guillaume P, Coudert JD, Levy F, Leclercq G, Held W. Interactions of Ly49 family receptors with MHC class I ligands in trans and cis. J Immunol. 2007; 178:1277–1284. [PubMed: 17237373]
9. McFall E, Tu MM, Al-Khattabi N, Tai LH, St-Laurent AS, Tzankova V, Hall CW, Belanger S, Troke AD, Wight A, Mahmoud AB, Zein HS, Rahim MM, Carlyle JR, Makrigiannis AP. Optimized tetramer analysis reveals Ly49 promiscuity for MHC ligands. J Immunol. 2013; 191:5722–5729. [PubMed: 24154624]
10. Mickiewicz KM, Gays F, Lewis RJ, Brooks CG. Mutagenesis of Ly49B reveals key structural elements required for promiscuous binding to MHC class I molecules and new insights into the molecular evolution of Ly49s. J Immunol. 2014; 192:1558–1569. [PubMed: 24403531]
11. Held W, Kunz B. An allele-specific, stochastic gene expression process controls the expression of multiple Ly49 family genes and generates a diverse, MHC-specific NK cell receptor repertoire. Eur J Immunol. 1998; 28:2407–2416. [PubMed: 9710218]
12. Valiante NM, Uhrberg M, Shilling HG, Lienert-Weidenbach K, Arnett KL, D'Andrea A, Phillips JH, Lanier LL, Parham P. Functionally and structurally distinct NK cell receptor repertoires in the peripheral blood of two human donors. Immunity. 1997; 7:739–751. [PubMed: 9430220]
13. Kubota A, Kubota S, Lohwasser S, Mager DL, Takei F. Diversity of NK cell receptor repertoire in adult and neonatal mice. J Immunol. 1999; 163:212–216. [PubMed: 10384118]
14. Rouhi A, Gagnier L, Takei F, Mager DL. Evidence for epigenetic maintenance of Ly49a monoallelic gene expression. J Immunol. 2006; 176:2991–2999. [PubMed: 16493057]
15. Santourlidis S, Trompeter HI, Weinhold S, Eisermann B, Meyer KL, Wernet P, Uhrberg M. Crucial role of DNA methylation in determination of clonally distributed killer cell Ig-like receptor expression patterns in NK cells. J Immunol. 2002; 169:4253–4261. [PubMed: 12370356]
16. Chan HW, Kurago ZB, Stewart CA, Wilson MJ, Martin MP, Mace BE, Carrington M, Trowsdale J, Lutz CT. DNA methylation maintains allele-specific KIR gene expression in human natural killer cells. J Exp Med. 2003; 197:245–255. [PubMed: 12538663]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

17. van Bergen J, Stewart CA, van den Elsen PJ, Trowsdale J. Structural and functional differences between the promoters of independently expressed killer cell Ig-like receptors. Eur J Immunol. 2005; 35:2191–2199. [PubMed: 15940669]
18. Anderson SK. Transcriptional regulation of NK cell receptors. Curr Top Microbiol Immunol. 2006; 298:59–75. [PubMed: 16329185]
19. Cichocki F, Miller JS, Anderson SK. Killer immunoglobulin-like receptor transcriptional regulation: a fascinating dance of multiple promoters. J Innate Immun. 2011; 3:242–248. [PubMed: 21411970]
20. Saleh A, Makrigiannis AP, Hodge DL, Anderson SK. Identification of a novel Ly49 promoter that is active in bone marrow and fetal thymus. J Immunol. 2002; 168:5163–5169. [PubMed: 11994471]
21. Saleh A, Davies GE, Pascal V, Wright PW, Hodge DL, Cho EH, Lockett SJ, Abshari M, Anderson SK. Identification of probabilistic transcriptional switches in the Ly49 gene cluster: a eukaryotic mechanism for selective gene activation. Immunity. 2004; 21:55–66. [PubMed: 15345220]
22. Brooks CG. Cloning and culturing of fetal mouse natural killer cells. Methods Mol Biol. 2000; 121:13–24. [PubMed: 10818713]
23. Brooks CG, Georgiou A, Jordan RK. The majority of immature fetal thymocytes can be induced to proliferate to IL-2 and differentiate into cells indistinguishable from mature natural killer cells. J Immunol. 1993; 151:6645–6656. [PubMed: 8258682]
24. Tsutsui H, Nakanishi K, Matsui K, Higashino K, Okamura H, Miyazawa Y, Kaneda K. IFN-gamma-inducing factor up-regulates Fas ligand-mediated cytotoxic activity of murine natural killer cell clones. J Immunol. 1996; 157:3967–3973. [PubMed: 8892629]
25. Gays F, Koh AS, Mickiewicz KM, Aust JG, Brooks CG. Comprehensive analysis of transcript start sites in ly49 genes reveals an unexpected relationship with gene function and a lack of upstream promoters. PLoS One. 2011; 6:e18475. [PubMed: 21483805]
26. Gays F, Unnikrishnan M, Shrestha S, Fraser KP, Brown AR, Tristram CM, Chrzanowska-Lightowlers ZM, Brooks CG. The mouse tumor cell lines EL4 and RMA display mosaic expression of NK-related and certain other surface molecules and appear to have a common origin. J Immunol. 2000; 164:5094–5102. [PubMed: 10799866]
27. Aust JG, Gays F, Mickiewicz KM, Buchanan E, Brooks CG. The expression and function of the NKRP1 receptor family in C57BL/6 mice. J Immunol. 2009; 183:106–116. [PubMed: 19535641]
28. Sentman CL, Hackett J Jr. Moore TA, Tutt MM, Bennett M, Kumar V. Pan natural killer cell monoclonal antibodies and their relationship to the NK1.1 antigen. Hybridoma. 1989; 8:605–614. [PubMed: 2613267]
29. Orange JS, Wang B, Terhorst C, Biron CA. Requirement for natural killer cell-produced interferon gamma in defense against murine cytomegalovirus infection and enhancement of this defense pathway by interleukin 12 administration. J Exp Med. 1995; 182:1045–1056. [PubMed: 7561678]
30. Hao L, Nei M. Genomic organization and evolutionary analysis of Ly49 genes encoding the rodent natural killer cell receptors: rapid evolution by repeated gene duplication. Immunogenetics. 2004; 56:343–354. [PubMed: 15316687]
31. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013; 30:772–780. [PubMed: 23329690]
32. Mathelier A, Zhao X, Zhang AW, Parcy F, Worsley-Hunt R, Arenillas DJ, Buchman S, Chen CY, Chou A, Ienasescu H, Lim J, Shyr C, Tan G, Zhou M, Lenhard B, Sandelin A, Wasserman WW. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res. 2014; 42:D142–147. [PubMed: 24194598]
33. Kubo S, Itoh Y, Ishikawa N, Nagasawa R, Mitarai T, Maruyama N. The gene encoding mouse lymphocyte antigen Ly-49: structural analysis and the 5′-flanking sequence. Gene. 1993; 136:329–331. [PubMed: 8294025]
34. Kubo S, Nagasawa R, Nishimura H, Shigemoto K, Maruyama N. ATF-2-binding regulatory element is responsible for the Ly49A expression in murine T lymphoid line, EL-4. Biochim Biophys Acta. 1999; 1444:191–200. [PubMed: 10023056]
35. Gosselin P, Makrigiannis AP, Nalewaik R, Anderson SK. Characterization of the Ly49I promoter. Immunogenetics. 2000; 51:326–331. [PubMed: 10803845]

Europe PMC Funders Author Manuscripts

36. Wilhelm BT, McQueen KL, Freeman JD, Takei F, Mager DL. Comparative analysis of the promoter regions and transcriptional start sites of mouse Ly49 genes. Immunogenetics. 2001; 53:215–224. [PubMed: 11398966]
37. Gays F, Martin K, Kenefeck R, Aust JG, Brooks CG. Multiple cytokines regulate the NK gene complex-encoded receptor repertoire of mature NK cells and T cells. J Immunol. 2005; 175:2938–2947. [PubMed: 16116180]
38. Barao I, Alvarez M, Ames E, Orr MT, Stefanski HE, Blazar BR, Lanier LL, Anderson SK, Redelman D, Murphy WJ. Mouse Ly49G2+ NK cells dominate early responses during both immune reconstitution and activation independently of MHC. Blood. 2011; 117:7032–7041. [PubMed: 21498673]
39. Dikstein R. The unexpected traits associated with core promoter elements. Transcription. 2011; 2:201–206. [PubMed: 22231114]
40. Lenhard B, Sandelin A, Carninci P. Metazoan promoters: emerging characteristics and insights into transcriptional regulation. Nat Rev Genet. 2012; 13:233–245. [PubMed: 22392219]
41. Tanamachi DM, Moniot DC, Cado D, Liu SD, Hsia JK, Raulet DH. Genomic Ly49A transgenes: basis of variegated Ly49A gene expression and identification of a critical regulatory element. J Immunol. 2004; 172:1074–1082. [PubMed: 14707081]
42. Buratowski S. Transcription. Gene expression--where to start? Science. 2008; 322:1804–1805. [PubMed: 19095933]
43. Rhee HS, Pugh BF. Genome-wide structure and organization of eukaryotic pre-initiation complexes. Nature. 2012; 483:295–301. [PubMed: 22258509]
44. Seila AC, Core LJ, Lis JT, Sharp PA. Divergent transcription: a new feature of active promoters. Cell Cycle. 2009; 8:2557–2564. [PubMed: 19597342]
45. Almada AE, Wu X, Kriz AJ, Burge CB, Sharp PA. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. Nature. 2013; 499:360–363. [PubMed: 23792564]
46. Natoli G, Andrau JC. Noncoding transcription at enhancers: general principles and functional models. Annu Rev Genet. 2012; 46:1–19. [PubMed: 22905871]
47. Lai F, Shiekhattar R. Enhancer RNAs: the new molecules of transcription. Curr Opin Genet Dev. 2014; 25C:38–42. [PubMed: 24480293]
48. Koch F, Jourquin F, Ferrier P, Andrau JC. Genome-wide RNA polymerase II: not genes only! Trends Biochem Sci. 2008; 33:265–273. [PubMed: 18467100]
49. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, Markenscoff-Papadimitriou E, Kuhl D, Bito H, Worley PF, Kreiman G, Greenberg ME. Widespread transcription at neuronal activity-regulated enhancers. Nature. 2010; 465:182–187. [PubMed: 20393465]
50. Mousavi K, Zare H, Dell'orso S, Grontved L, Gutierrez-Cruz G, Derfoul A, Hager GL, Sartorelli V. eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. Mol Cell. 2013; 51:606–617. [PubMed: 23993744]
51. Ortaldo JR, Winkler-Pickett R, Mason AT, Mason LH. The Ly-49 family: regulation of cytotoxicity and cytokine production in murine CD3+ cells. J Immunol. 1998; 160:1158–1165. [PubMed: 9570529]
52. Smith HR, Chuang HH, Wang LL, Salcedo M, Heusel JW, Yokoyama WM. Nonstochastic coexpression of activation receptors on murine natural killer cells. J Exp Med. 2000; 191:1341–1354. [PubMed: 10770801]
53. Lees RK, Ferrero I, MacDonald HR. Tissue-specific segregation of TCRgamma delta+ NKT cells according to phenotype TCR repertoire and activation status: parallels with TCR alphabeta+NKT cells. Eur J Immunol. 2001; 31:2901–2909. [PubMed: 11592065]
54. Rouhi A, Lai CB, Cheng TP, Takei F, Yokoyama WM, Mager DL. Evidence for high bi-allelic expression of activating Ly49 receptors. Nucleic Acids Res. 2009; 37:5331–5342. [PubMed: 19605564]
55. Nylenna O, Naper C, Vaage JT, Woon PY, Gauguier D, Dissen E, Ryan JC, Fossum S. The genes and gene organization of the Ly49 region of the rat natural killer cell gene complex. Eur J Immunol. 2005; 35:261–272. [PubMed: 15593300]

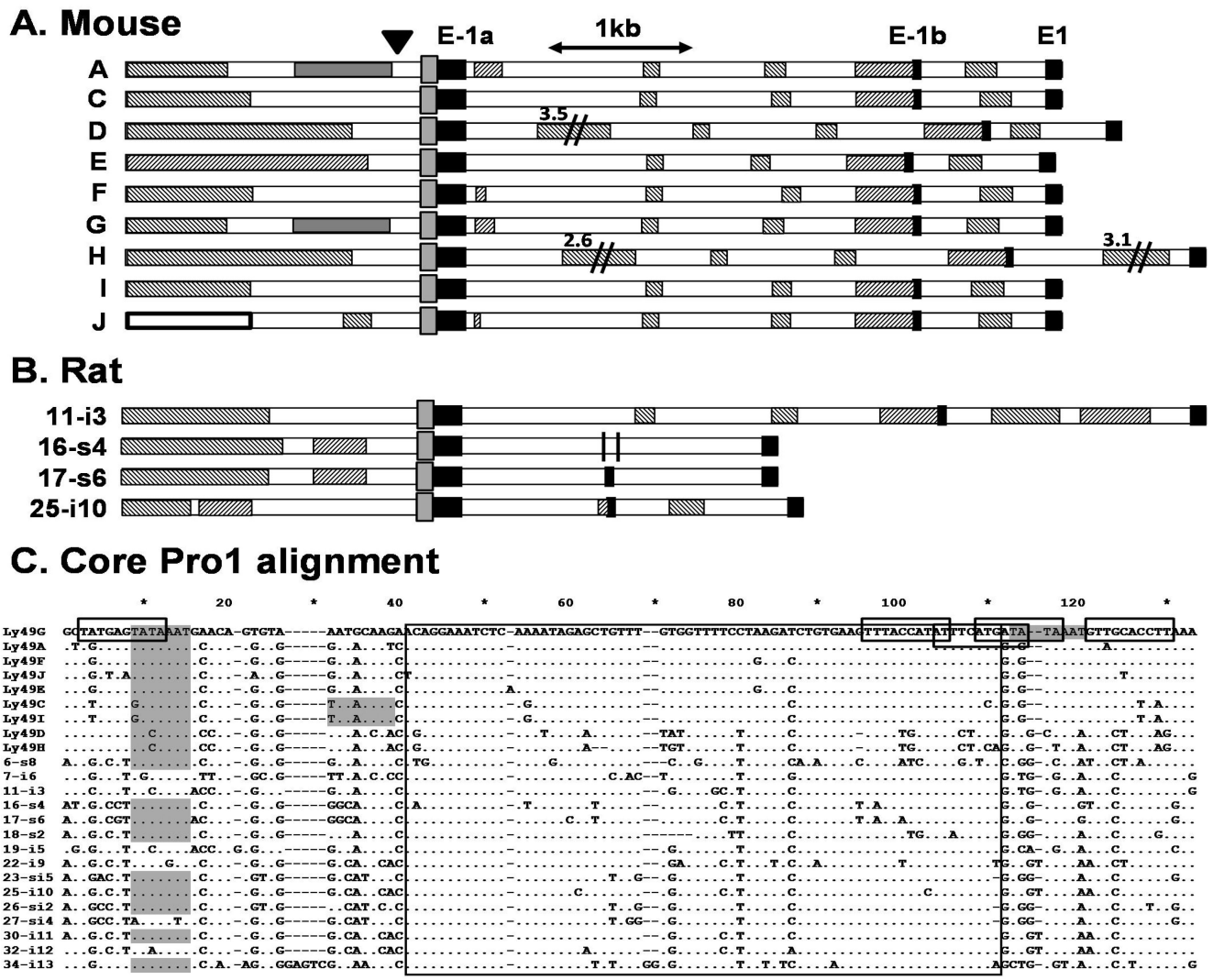

Figure 1. The upstream regions of Ly49 genes

A and B. Scale diagrams of the upstream regions (from exon 1 (E1) to a point ~2kb upstream of Pro1) of the 9 functional Ly49 genes in the C57BL/6 mouse strain, and of 4 representative putatively functional Ly49 genes in the BN rat strain. The rat Ly49 genes are identified by both their number (30) and by the functional potential of their protein products (55) (i, contains cytoplasmic ITIM; s, contains transmembrane charge; si, contains both). Exons E-1a, E-1b, and E1 are shown in black, LINE/SINE repeats as boxes with downward sloping hash lines, LTR repeats as boxes with upward sloping hash lines, and Pro1 elements as protruding grey bars. The 3′ end of the Ly49M gene immediately upstream of Ly49J is shown as an open box. The large LINE repeats in Ly49D and H have been truncated and their actual sizes are shown in kb. The arrowhead above the Ly49A gene shows the point upstream of which the Ly49A and G sequences become unrelated to those in other Ly49 genes. The grey boxes upstream of the Pro1 elements in the Ly49A and G genes denote Ly49B-like sequences. The vertical lines in the rat 16-s4 gene correspond to a gap in the sequence. **C.** Alignment of the core Pro1 regions of potentially functional mouse and rat

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Ly49 genes. Dots designate identity with the top Ly49G sequence, dashes gaps. TBP binding sites (TATA boxes) were analyzed with the JASPAR algorithms (32), and those with scores >8.0 are shaded grey. Boxes on the top line show the C/EBP binding sites identified by Saleh et al. (21). The large box shows the central highly conserved region.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

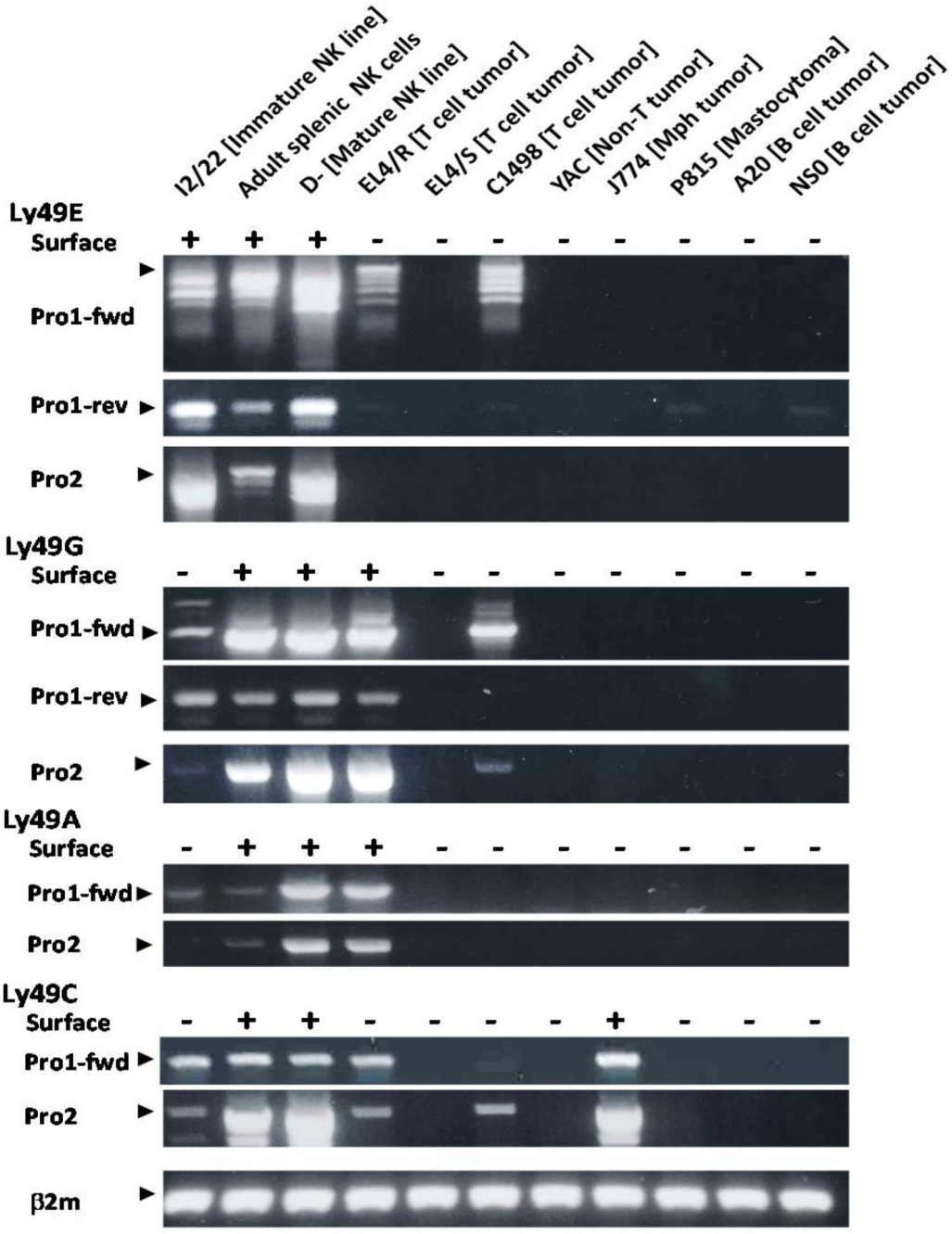

Figure 2. RT-PCR analysis of Pro1 and Pro2 transcripts in cell lines

cDNA prepared from the cell lines shown was analyzed for the presence of Pro1 fwd, Pro1 rev, Pro2, and β2m transcripts. Arrowheads show the expected position of products from RNA containing all exons (note that in heavily loaded lanes products run faster than expected). Surface expression of the relevant Ly49s, as determined by staining with Ly49-specific mAbs, is shown as + and −. The results shown are representative of multiple experiments performed with the primers described in Supplementary Table 1, and with additional primer combinations (data not shown).

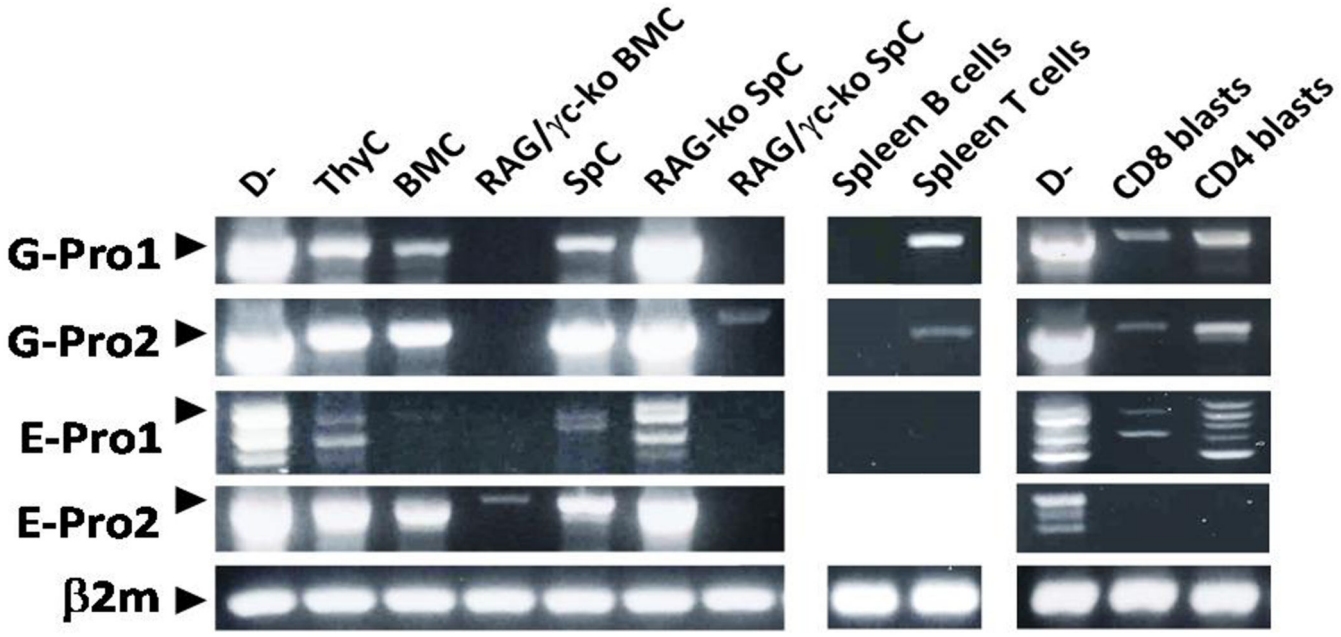

Figure 3. RT-PCR analysis of Pro1 and Pro2 transcripts in fresh cells

cDNA prepared from freshly obtained thymocytes (ThyC), bone marrow cells (BMC) from wild type and RAG$^{-/-}$ γc$^{-/-}$ mice, spleen cells (SpC) from wild type, RAG$^{-/-}$, and RAG$^{-/-}$ γc$^{-/-}$ mice, purified spleen B cells, purified spleen T cells, and CD8 or CD4 T cell blasts was analyzed for the presence of Pro1fwd and Pro2 transcripts from Ly49G (G) or Ly49E (E), and also β2m transcripts, using the primers described in Supplementary Table 1. cDNA from the D- adult NK cell line was included as a positive control. Arrowheads show the expected position of products amplified from RNA containing all exons (heavily loaded products run faster than expected). The results shown are representative of multiple experiments.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

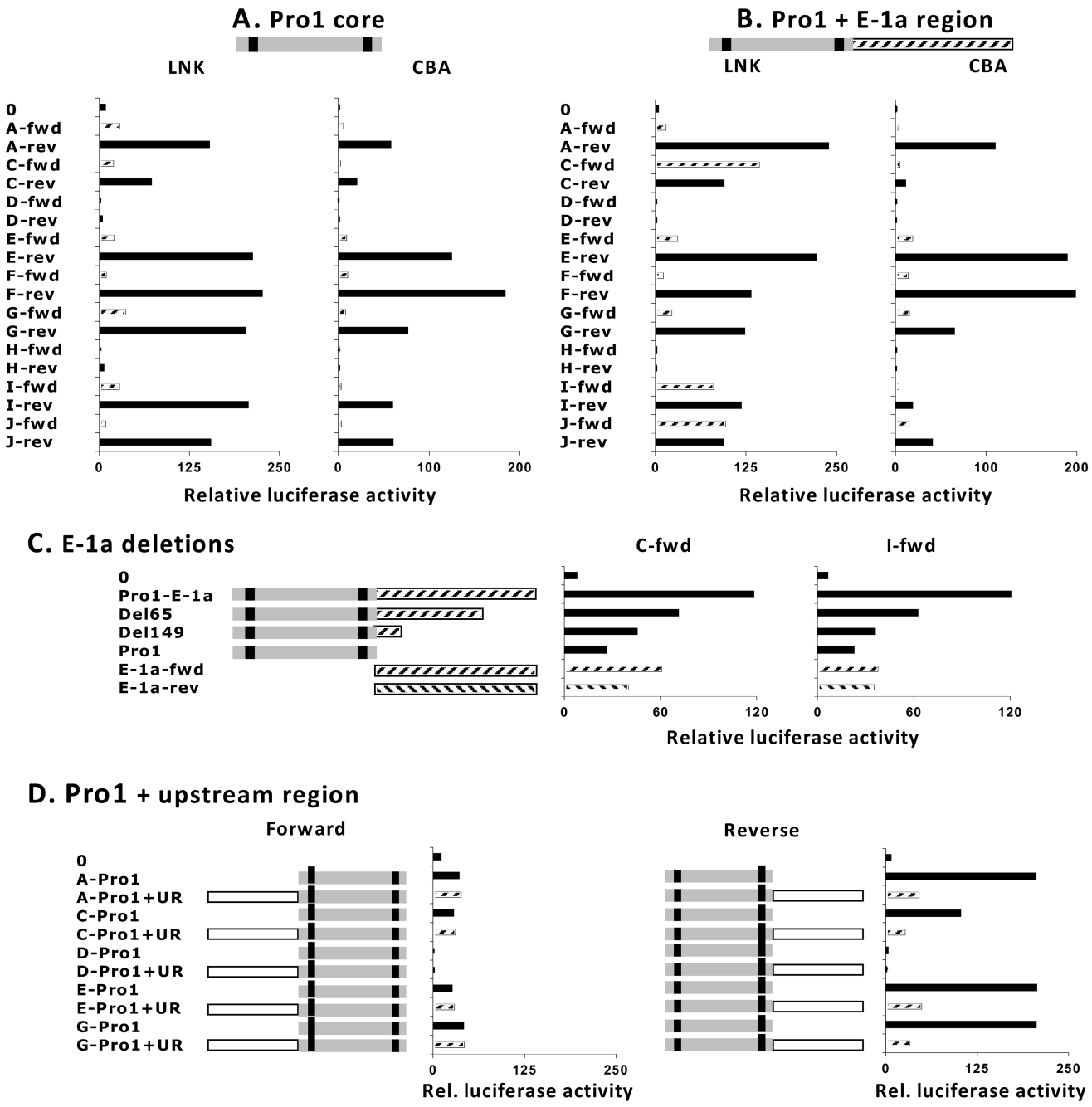

Figure 4. Promoter activity of mouse Pro1-fragments

The LNK and CBA immature NK cell lines were transfected with pGL4 plasmids containing the Pro1 fragments listed on the left side of each graph, and whose structure is illustrated in the accompanying schematics (grey bar, core Pro1 region; black line, TATA box; hatched bar, exon E-1a; open bar, the ~100bp conserved upstream region). **A.** Core Pro1 regions in fwd (hatched bars) or rev (solid bars) orientations. **B.** Pro1-E-1a regions in fwd (hatched bars) or rev (solid bars) orientations. C. LNK cells were transfected with pGL4 plasmids containing the Pro1-E-1a fwd fragments from the Ly49C (C-fwd) or Ly49I (I-fwd) genes,

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Pro1-E-1a fragments lacking 65bp (Del65) or 149bp (Del149) from the 3′ end of E-1a, core Pro1 fragments, or E-1a fragments. D. LNK cells were transfected with pGL4 plasmids containing the core Pro1 regions or core Pro1 regions + the ~100bp conserved upstream region (UR) of the genes shown in either forward or reverse orientation. In the schematic the 5′ TATA box is shown protruding. In all four panels, controls consisted of pGL4 with no insert (0) or with the HSVtk promoter, and relative luciferase activity was calculated as a percentage of that of the HSVtk promoter.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

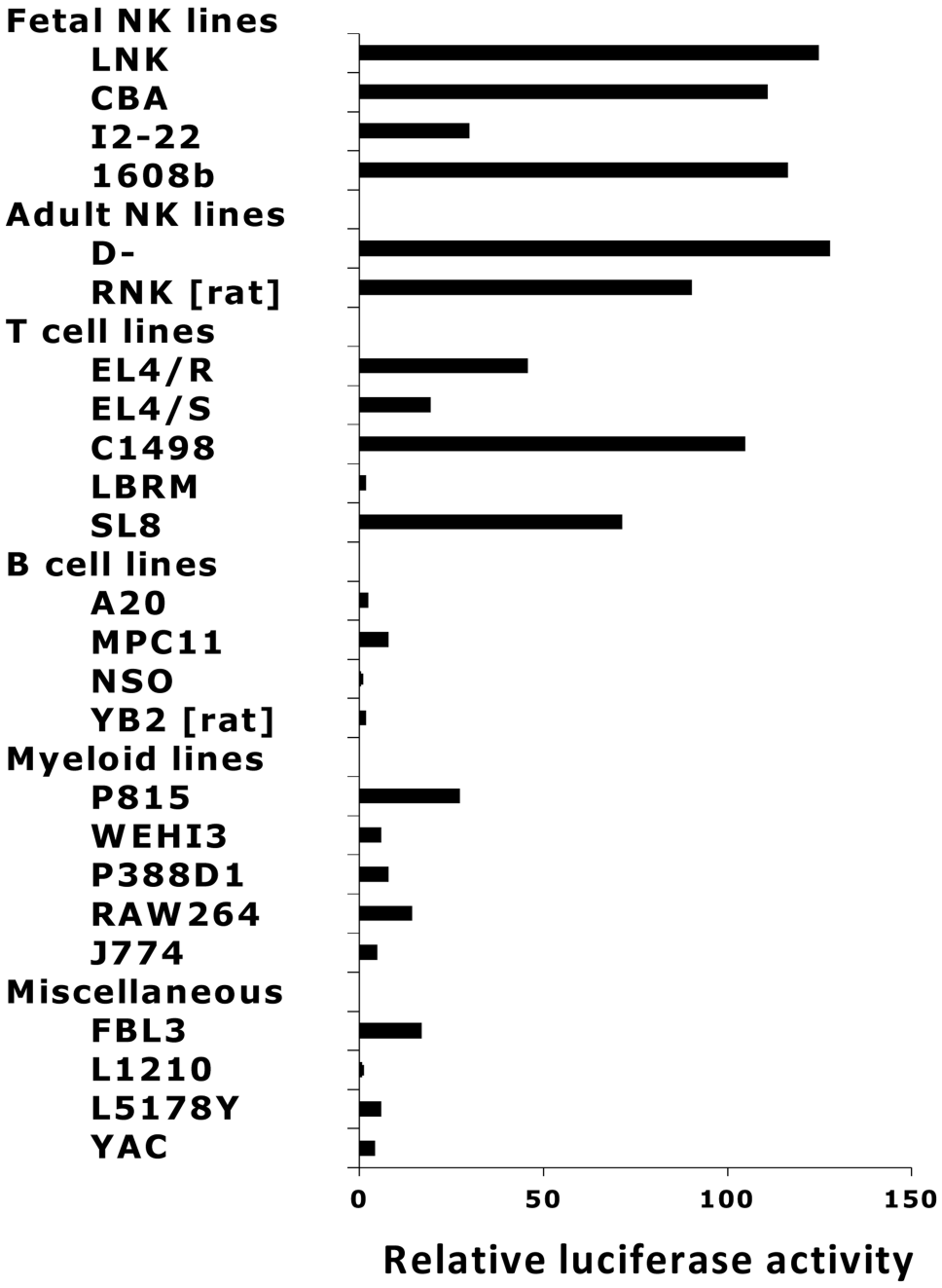

Figure 5. Pro1 displays selective activity in NK and T cell lines

The cell lines shown were transfected with pGL4 plasmids containing the Ly49G Pro1-E-1a fragment in rev orientation or the HSVtk promoter, and luciferase activity was calculated as a percentage of that of the HSVtk promoter.

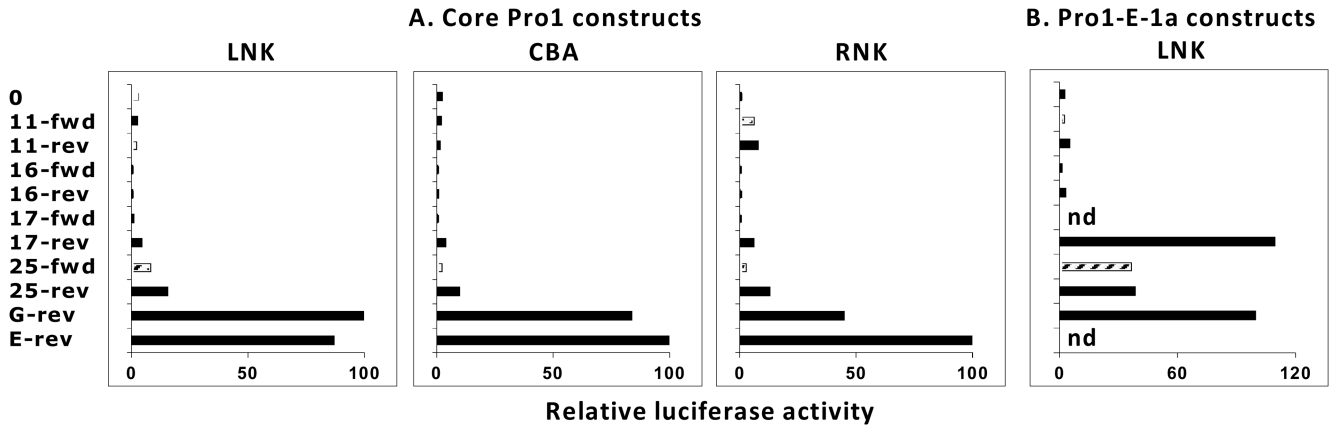

Figure 6. Promoter activity of rat Pro1-fragments

A. The cell lines shown were transfected with pGL4 plasmids containing the core Pro1 fragments from rat genes 11, 16, 17, and 25 in fwd (hatched bars) or rev (solid bars) orientation, together with the Pro1 rev fragments from mouse Ly49G and E as positive controls. Relative luciferase activity was calculated as a percentage of that of the Ly49G or E Pro1-fragments, whichever was greater. **B.** LNK cells were transfected with pGL4 plasmids containing the Pro1-E-1a fragments from the same rat genes in fwd (hatched bars) or rev (solid bars) orientation, together with the Pro1 rev fragment from mouse Ly49G as positive control. nd = not done.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

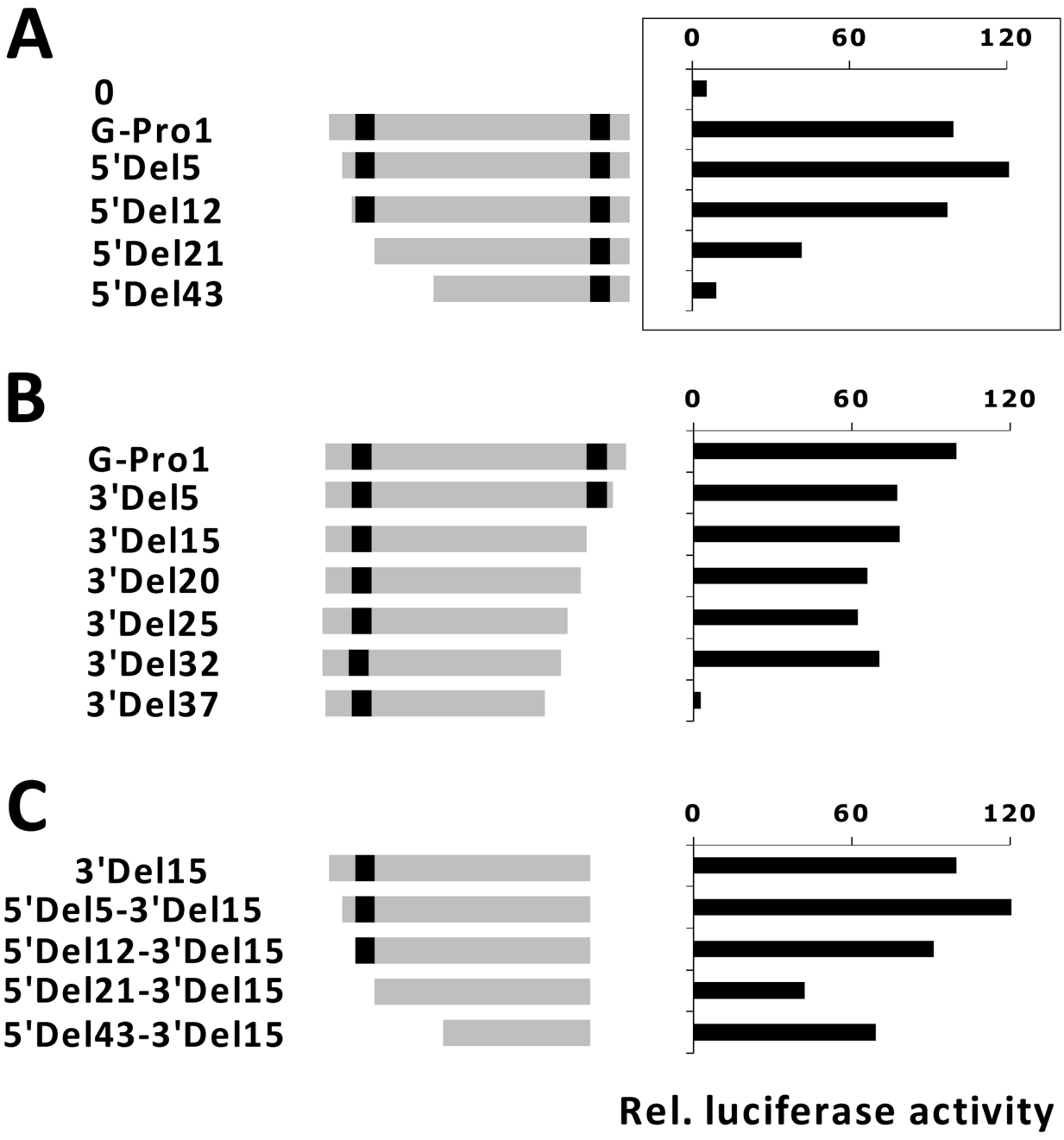

Figure 7. Pro1 promoter activity does not require TATA boxes

LNK cells were transfected with pGL4 plasmids containing the following: **A.** No insert (0), the 122bp core Ly49G Pro1 rev fragment (G-Pro1), or fragments of this lacking 5, 12, 21, or 43bp at the 5′ end; **B.** The 122bp core Ly49G Pro1 rev fragment (G-Pro1) or fragments of this lacking 5, 15, 20, 25, 32, or 37bp at the 3′ end; **C.** The 107bp Ly49G Pro1 rev 3'Del15 fragment or fragments of this lacking 5, 12, 21, or 43bp at the 5′ end. In all cases, luciferase activity was calculated as a percentage of that of the relevant parental fragment. Schematics illustrate the size of the tested fragments, with TATA boxes shown in black. Note that the

upstream TATA box in the constructs corresponds to the downstream TATA box in the native gene.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

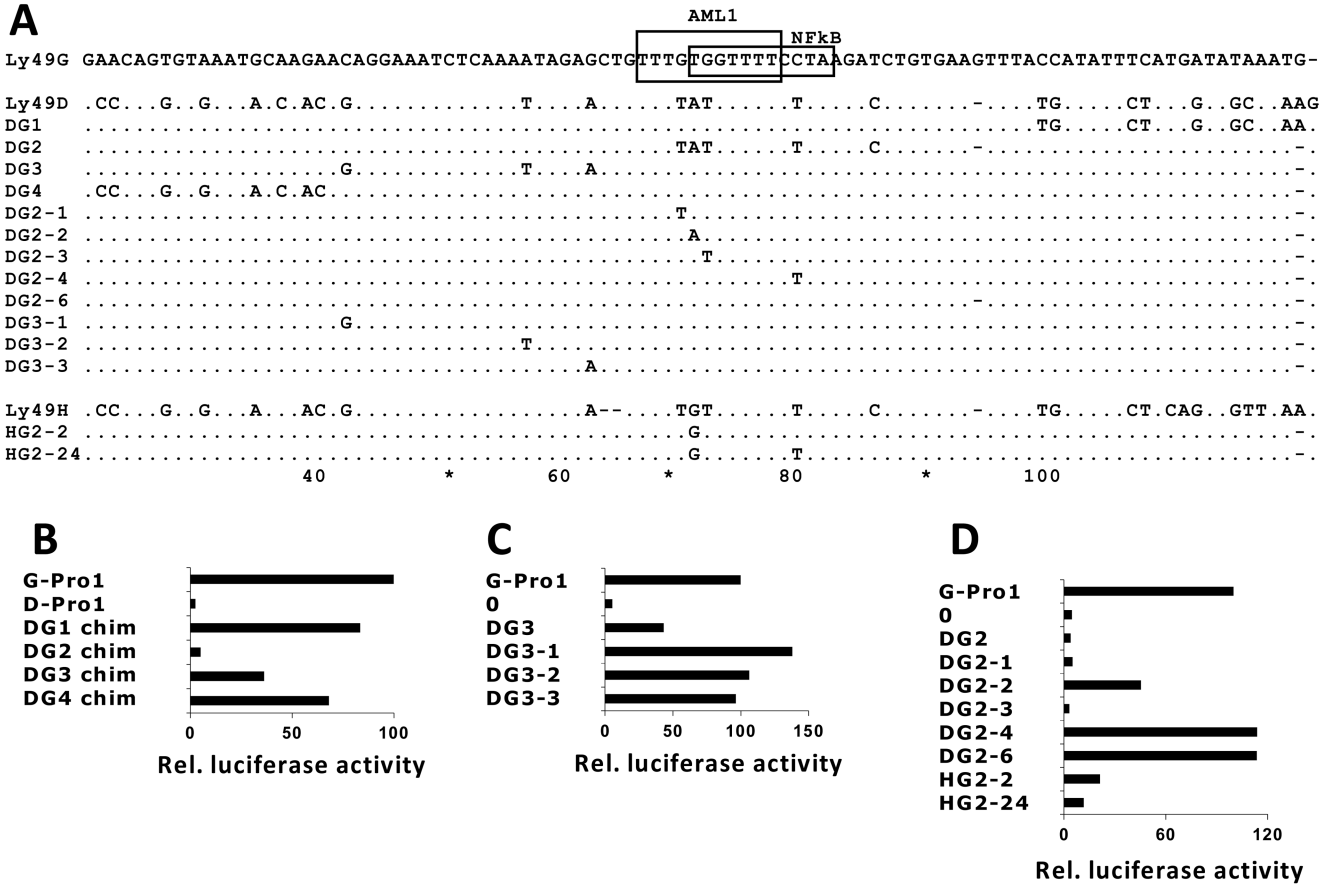

Figure 8. Mapping of key residues required for Pro1 promoter activity

A. An alignment showing the 5'Del12-3'Del15 Pro1 fragments of Ly49G, Ly49D, and Ly49H, together with the four chimeric fragments DG1-DG4, and various mutants of the Ly49G 5'Del12-3'Del15 Pro1 fragment containing individual residues from Ly49D and Ly49H. The boxes in the Ly49G fragment show the keyAML1 and NFkB binding sites identified in the previous study (21) which, according to the JASPAR algorithms, have high scores in the Ly49G fragment but are absent (score <80%) from the Ly49D and H fragments. The numbering at the bottom corresponds to that in Figure 1C with minor adjustment for gaps. **B-D.** LNK cells were transfected with pGL4 plasmids containing the 5'Del12-3'Del15 Pro1 fragments shown, all in reverse orientation. In all cases, luciferase activity was calculated as a percentage of that of the parental Ly49G-Pro15'Del12-3'Del15 Pro1 fragment.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

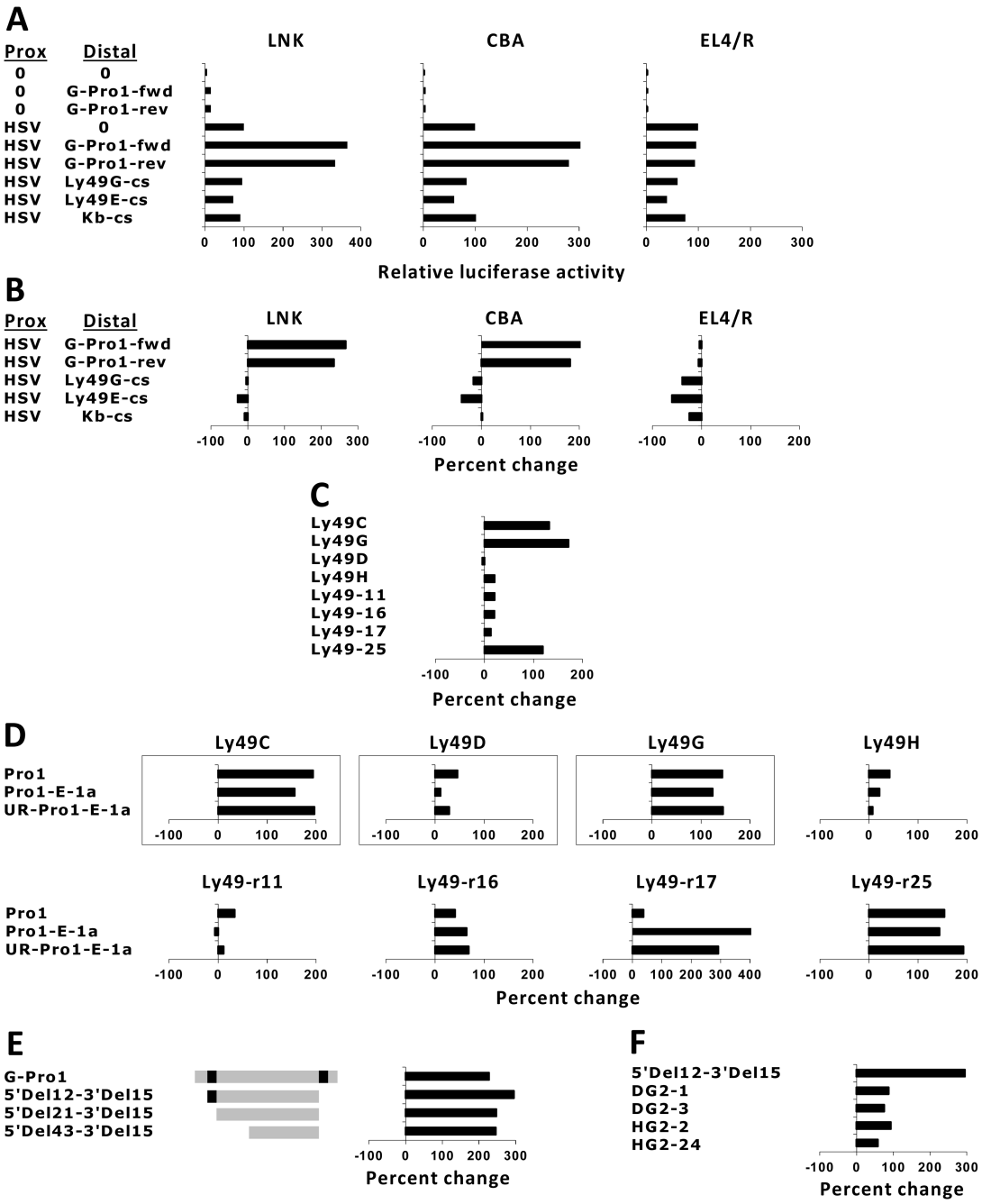

Figure 9. Pro1 displays enhancer activity

A. The cell lines shown were transfected with pGL4 plasmids that contained either no insert (0) or the HSVtk promoter in the proximal (Prox) EcoRVHindIII site, and no insert (0), the Ly49G core Pro1 fwd (G-Pro1fwd) or rev (G-Pro1rev) fragments, or various control fragments in the distal BamHI-SalI site. The control sequences correspond to bases 432-732 of the Ly49G mRNA RefSeq sequence NM014194 (Ly49G-cs), bases 109-906 of the Ly49E mRNA RefSeq sequence NM008463 (Ly49E-cs), and bases 886-1186 of the K^b MHC class I mRNA RefSeq sequence NM001001892 (Kb-cs). Luciferase activity was calculated as a

percentage of that of the pGL4-HSVtk construct that contained no insert in the distal site. **B.** The same data as in A but showing the percent change in luciferase activity of pGL4-HSVtk caused by the presence of fragments in the distal BamHI-SalI site, calculated as 100 × (HSVtk-X/HSVtk-0 - 1), where HSVtk-X is the luciferase activity in cells transfected with pGL4-HSVtk containing insert X in the distal site, and HSVtk-0 is the luciferase activity in cells transfected with pGL4-HSVtk containing no insert in the distal site. **C.** The enhancer activity of core Pro1 fragments from various Ly49 genes, inserted into the distal site of pGL4-HSVtk in reverse orientation, and calculated as in B. **D.** The enhancer activity of core Pro1 fragments (Pro1) alone, or fragments also containing the E-1a region (Pro1-E-1a), or both this and the conserved upstream region (UR-Pro1-E-1a), inserted into the distal site of pGL4-HSVtk in reverse orientation, and calculated as in B. **E.** The enhancer activity of the Ly49G core Pro1 fragment (G-Pro1) and fragments of this having 5′ and 3′ deletions, inserted into the distal site of pGL4-HSVtk in reverse orientation, and calculated as in B. Schematics illustrate the size of the tested fragments, with TATA boxes shown in black. Note that the upstream TATA box in these corresponds to the downstream TATA box in the native gene. **F.** The enhancer activity of the 5'Del12-3'Del15 Pro1 fragment of Ly49G and various mutant fragments of this (sequences shown in Fig. 8A) , inserted into the distal site of pGL4-HSVtk in reverse orientation, and calculated as in B.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts