Recognition of the Class Ib Molecule Qa-1\(^b\) by Putative Activating Receptors CD94/ NKG2C and CD94/ NKG2E on Mouse Natural Killer Cells

By Russell E. Vance, Amanda M. Jamieson, and David H. Raulet

From the Department of Molecular and Cell Biology and Cancer Research Laboratory, University of California, Berkeley, California 94720

Summary

The heterodimeric CD94/NKG2A receptor, expressed by mouse natural killer (NK) cells, transduces inhibitory signals upon recognition of its ligand, Qa-1\(^b\), a nonclassical major histocompatibility complex class Ib molecule. Here we clone and express two additional receptors, CD94/NKG2C and CD94/NKG2E, which we show also bind to Qa-1\(^b\). Within their extracellular carbohydrate recognition domains, NKG2C and NKG2E share extensive homology with NKG2A (93–95% amino acid similarity); however, NKG2C/E receptors differ from NKG2A in their cytoplasmic domains (only 33% similarity) and contain features that suggest that CD94/NKG2C and CD94/NKG2E may be activating receptors. We employ a novel blocking anti-NKG2 monoclonal antibody to provide the first direct evidence that CD94/NKG2 molecules are the only Qa-1\(^b\) receptors on NK cells. Molecular analysis reveals that NKG2C and NKG2E messages are extensively alternatively spliced and ~20-fold less abundant than NKG2A message in NK cells. The organization of the mouse CD94/Nkg2 gene cluster, presented here, shows striking similarity with that of the human, arguing that the entire CD94/NKG2 receptor system is relatively primitive in origin. Analysis of synonymous substitution frequencies suggests that within a species, NKG2 genes may maintain similarities with each other by concerted evolution, possibly involving gene conversion–like events. These findings have implications for understanding NK cells and also raise new possibilities for the role of Qa-1 in immune responses.

Key words: CD94 • NKG2 • Qa-1 • natural killer cell • MHC class I

Natural killer cells mediate innate immunity by secretion of inflammatory cytokines and by direct cytotoxicity against infected or transformed target cells (1, 2). NK cells are regulated in part by several families of cell surface inhibitory receptors specific for various alleles of MHC class I (3, 4). These receptors are organized into multiple families, including (a) the killer cell Ig-like receptors (KIRs), expressed in humans but not rodents; (b) the lectin-like Ly49 homodimers expressed functionally in rodents but not, apparently, in humans; and (c) the lectin-like heterodimers formed from CD94 and NKG2 subunits, which are the only class I–specific receptors found in both humans and rodents (for review see reference 5). By virtue of their expression of these class I–specific inhibitory receptors, NK cells preferentially lyse targets lacking appropriate class I MHC molecules. Thus, NK cells may defend against viruses and tumors that downregulate host cell MHC class I so as to evade detection by T cells (6–8).

An unusual feature of the CD94/NKG2 heterodimers is that they recognize a nonclassical MHC class I ligand, HLA-E in humans or Qa-1 in mice (9–12). Recognition of Qa-1 and HLA-E appears to require both the CD94/NKG2 receptor subunits (9, 10). Surprisingly, HLA-E and Qa-1 are not recognizably orthologs at the level of overall amino acid sequence (13). However, they are both able to bind specifically and predominantly to a nine–amino acid peptide derived from the signal sequences of some classical class I molecules (14–18). In the mouse, this peptide is referred to as Q dm (for 'Q a-1 determinant modifier') and has the sequence AMAPRTLLL (16). Although presumably cleaved from the nascent class I polypeptide by an endoplasmic reticulum–resident enzyme ('signal peptidase'), Q dm will not be presented by Qa-1 unless transported by TAP, the transporter associated with antigen presentation (16, 19).

Functional evidence has demonstrated that NK cells will not recognize cell surface Q a-1 unless Q dm (or possibly related peptides) is present (9, 20). Evidence also suggests that...
human CD94/NKG2 recognition of HLA-E may be exquisitely sensitive to the sequence of the particular HLA-E-bound peptide (21, 22). Taken together, these observations support an elegant mechanism by which NK cells can specifically monitor TAP function and the biosynthesis of highly polymorphic class I molecules by recognition of a relatively nonpolymorphic molecule.

Several NKG2 isoforms have been described. In humans, both the inhibitory CD94/NKG2A and the activating CD94/NKG2C receptors have been shown to recognize HLA-E (15, 22), but to date in mice, only the inhibitory CD94/NKG2A receptor has been shown to interact with Qa-1b (9). The capacity of human CD94/NKG2C to activate NK cells is due to its association with an immunoreceptor tyrosine-based activation motif–containing homodimer called DAP12 (originally called KARAP; 23–25). It is not known if mouse NKG2C, recently identified only at the message level (26, 27), associates with mouse DAP12. Both mouse and human NK cells also appear to express the NKG2D molecule. However, NKG2D exhibits little sequence similarity to other NKG2 molecules (see Fig. 1 B), does not appear to heterodimerize with CD94, and associates with the DAP10 signaling module instead of DAP12 (28). Moreover, the ligand for NKG2D in humans was recently shown to be a stress-induced class I–like molecule (29), arguing that NKG2D is functionally distinct from other NKG2 receptors.

Among mammals, the only known peptide ligands for CD94/NKG2 receptors are the Qd-m-like peptides found in class I signal sequences. However, several viruses encode proteins containing Qa-1 or HLA-E binding peptides. For example, the UL40 gene product of human CMV contains a Qa-1b homology in its signal sequence (available from SWISSPROT under accession number P16780). Similarly, the mouse hepatitis virus hemagglutinin (HA) esterase gene product contains the canonical Qa-1 binding peptide AMPRTLLL (13). It is tempting to speculate that these viruses may utilize these peptides as ‘decoys’ to evade NK-mediated lysis through engagement of the inhibitory CD94/NKG2A receptor. The presence of NK cells expressing activating CD94/NKG2 receptors specific for HLA-E or Qa-1 complexes might then be important in allowing the immune system to detect and lyse cells displaying such viral decoys.

T cells have also been proposed to recognize Qa-1 in various contexts (30–34). Although it seems reasonable to presume that recognition of Qa-1 by T cells is mediated by the TCR, it should be noted that ‘NK’ receptors (including CD94 and NKG2 receptors) can be expressed by T cells (35, 36), and it remains possible that some of the effects of Qa-1 on T cells might be due to T cell expression of CD94/NKG2 heterodimers.

Although previous studies demonstrated that NK cells can bind to soluble Qa-1 tetramers (9, 37), as can heterologous transfectants expressing CD94/NKG2A (9), it has never been formally demonstrated that Qa-1 tetramer binding by NK cells is due to NK expression of CD94/NKG2, and moreover, additional Qa-1 receptors might exist. Here we report the identification and characterization of two putative activating receptors on mouse NK cells, CD94/NKG2C and CD94/NKG2E, which, we demonstrate, bind Qa-1b. We also provide evidence that CD94/NKG2 receptors are the only Qa-1 receptors on NK cells. Our results have implications for our understanding of NK cells and the Qa-1 molecule and will permit dissection of the role of class I–specific activating receptors in the genetically and experimentally amenable mouse model.

Materials and Methods

Primers, Probes, PCR, and Sequencing. Table I contains a list of PCR primers used in this study. For PCR, an annealing temperature of 55°C was used throughout. Oligonucleotide probes were hybridized to Southern blots at 50°C and washed twice briefly in 2× SSC/1% SDS at room temperature and then for ~30 min at 50°C in 1× SSC/1% SDS before exposure to film. Sequencing reactions were performed using the Big Dye Terminator Ready Mix (PE Biosystems) and were resolved and analyzed using an ABI Prism 310 sequencer (PE Biosystems). Antibodies. To generate anti-CD94/NKG2 mAbs, male Lewis rats were immunized three times, at ~2-wk intervals, with stable Chinese hamster ovary (CHO) cell transfectants expressing the B6 alleles of CD94 and NKG2A. The rats were then boosted with the same transfectants, and 3 d later spleen cells were fused to P3X63-Ag8.653 (subclone of ATCC TIB-9) using PEG1500 (Boehringer Mannheim) according to the manufacturer’s instructions. HAT-resistant hybridoma supernatants were screened for staining of NKG2C3− spleen cells and for staining of CHO–CD94/NKG2A but not untransfected CHO cells. Hybridoma 20d5, specific for NKG2C, and hybridoma 18d3, specific for CD94, were cloned by serial dilution three times. Both mAbs are rat IgG2a, as assessed with a rat mAb isotyping kit (PharMingen) and were purified from hybridoma supernatants over protein G agarose (Boehringer Mannheim). Purified 20d5 was used to block NKG2 receptors at a concentration of 100 μg/ml. Anti-NK1.1 mAb was purified from PK136 supernatants and conjugated to FITC. PE-conjugated PK136 was purchased from PharMingen, and streptavidin–PE was from Molecular Probes, Inc.

Molecular Cloning of NKG2C and NKG2E. In the course of amplifying NKG2C cDNA ends (as described previously; reference 9), one clone similar to human NKG2C was obtained but was lacking its 5′ end and start codon. Based on this clone, a sequencing primer (NKG2C3′ex2) was designed and used to obtain the sequence of the 5′ coding and untranslated regions (UTR’s) directly from bacterial artificial chromosome (BAC) B6-3 (clone 91119). 3′-coding and untranslated sequence was obtained by sequencing the end of a genomic fragment, H3.5. Together, these 5′ and 3′ sequences were used to design two sets of primers to amplify full-length NKG2C and -E open reading frames (ORFs). The NKG2C 5′ and 3′ UTR primers recognized sites just outside of the predicted ORF, whereas the NKG2C5′ATG and NKG2C3′#3 primers bound just within the ORF. Both sets of primers were used to amplify NKG2C sequences from oligo dT–primed cDNA generated by standard methods from IL-2–cultured NK cell RNA. The products, generated with Tag polymerase (Promega Corp.), were cloned directly into the T-tailed T easy vector (Promega Corp.) and sequenced on both strands. Clones generated with the UTR primer pair were designated ‘C-UTR’ or ‘E-UTR’; whereas clones generated with the ATG primer pair were designated ‘C-ORF’ or ‘E-ORF.’ Two clones, 1802 R recognition of Qa-1b by CD94/NKG2C and CD94/NKG2E

Downloaded from on June 21, 2017
E-UTR 4 and C-ORF4, were found to encode full length NKG2E and NKG2C, respectively, and the sequences were judged free of PCR-induced coding errors based on comparisons of multiple independent clones. Clone C-ORF4 is identical to sequences previously reported to GenBank (26, 27). Clone E-UTR 4 contains a silent T → C mutation at nucleotide 414.

BACs and Genomic Analysis. BACs were identified from the Genome Systems C57BL/6 BAC library as previously described (9). Two HindIII fragments of 3.5 and 9 kb that hybridized to an NKG2C exon 5/6 probe were cloned into HindIII-digested pBluescript SKII (+) (Stratagene Inc.) to generate clones H3.5 and H3.6. These clones were sequenced using T3 and T7 primers, as well as the following primers, to determine the exon/intron boundaries for exons 4–7: NKG2Cex5, NKG2Cex6, NKG2Cex7, and NKG2C#3. Quantitation of NKG2A, NKG2C, and NKG2E mRNA. Sorted purified (>98% pure) NK1.1+CD3− IL-2–cultured NK cells or nylon wool–passed, IL-2–cultured NK cells were used as starting material to make total RNA using the Ultraspec II reagent. Approximately 2 μg of total RNA was reverse transcribed using an oligo-dT primer (GIBCO BRL). Once drug-resistant cells started to grow out, bulk transfections were sorted for bright surface expression of NKG2A, -C, or -E using the 20d5 anti-NKG2 mAb or Qa-1 tetramer as staining reagent. After sorting, cells were cloned by limiting dilution, and positive clones were expanded without G418. Qa-1 tetramers were used as templates for the identical PCR and digestion protocol. Stable CHO T transfectants and Qa-1 T tetramer staining. The expression vector for NKG2A has been described (9). To generate expression constructs, NKG2C and -E were amplified from clones C-ORF4 and E-UTR 4 by PCR using Pfu polymerase (Stratagene Inc.), primer NKG2C3 #3 (used for NKG2C) or NKG2AHANot3 (used for NKG2E). The products were digested with XhoI and NotI, and cloned into the pME18S expression vector (sequence available from EMBL/GenBank/DDBJ under accession number AB009864), and confirmed by sequencing. The latter primer added a nine–amino acid HA epitope tag (amino acid sequence YPYDVPDYA) to the COOH terminus (extracellular) of the NKG2E molecule, allowing its detection with an anti-HA mAb. Previous results with NKG2A-HA (9) showed the tag is unlikely to affect ligand binding. For our studies, the tag proved unnecessary, as we were subsequently successful in developing anti-NKG2 mAbs. The expression construct for NKG2C did not encode an HA–tagged molecule. The pME18S expression vectors were stably transfected into CHO cells along with a pIR ES vector (Clontech) encoding mouse CD94 and the neo gene. NKG2E (but not NKG2C) was also cotransfected with pME18S-DAP12, but resulting clones were not assessed for DAP12 expression. Transfection was with the Lipofectamine reagent (GIBCO BRL) as previously described (38). 48 h after transfection, CHO cells were passaged into RPMI supplemented with 1 mg/ml G418 (GIBCO BRL). O nce drug-resistant cells started to grow out, bulk transfectants were sorted for bright surface expression of NKG2A, -C, or -E using the 20d5 anti-NKG2 mAb or Qa-1 tetramer as staining reagent. After sorting, cells were cloned by limiting dilution, and positive clones were expanded without G418. Qa-1 tetramers, complexed with the Qdm peptide (AMAPRTLLL),
were generated and used as described previously (9, 39), except that a Superdex200 gel filtration column (Pharmacia) was used in all purification steps.

Sequence Alignment and Analysis. Nucleotide sequences were aligned with CLUSTAL W (40), using the portion of each sequence corresponding to the carbohydrate recognition domain (CRD) of NKG2A (taken here as the last 363 nucleotides of the mNKG2A ORF). For each pair of aligned sequences, the program DnaSP 3.0 (41) was used to calculate the number of synonymous substitutions per synonymous site (Ks) and the number of nonsynonymous substitutions per nonsynonymous site (Ka). This program uses the unweighted algorithm of Nei and Gojobori (42), which can be explained in brief. First, the total number of nonsynonymous sites (N) at which nucleotide changes give rise to amino acid changes and synonymous sites (S, at which nucleotide changes are silent) are determined for each sequence, and the average values of S and N for each pair of sequences are used. Then, for each pair of sequences, the number of synonymous (S_d) and nonsynonymous (N_d) nucleotide differences observed is determined. For a pair of codons that differ by more than one nucleotide, the different possible evolutionary paths that could give rise to the observed changes need to be taken into account (different evolutionary paths can involve different numbers of synonymous and nonsynonymous substitutions). In the unweighted algorithm, all of the different possible pathways are considered equally likely. S_d/S in is the proportion of synonymous differences, which is then used to estimate Ks. See Nei and Gojobori (42) for more details.

Results and Discussion

Molecular Cloning of NKG2C and NKG2E. In the course of amplifying mouse NKG2A cDNA ends (see Materials and Methods), we obtained the 3' end of a cDNA that exhibited similarity to human NKG2C. Based on genomic sequence obtained from a BAC clone (clone 91i19), we were able to design PCR primers to amplify the entire coding region of an NKG2C cDNA from C57BL/6 mice. The consensus sequence obtained was identical to the sequence also reported recently by others (26, 27). In addition, we also obtained a novel NKG2C-like cDNA that we termed mouse NKG2E. It must be emphasized that mouse NKG2C is no more related to human NKG2E than to human NKG2E, nor does mouse NKG2E exhibit more similarity to human NKG2E than to human NKG2C (Fig. 1 B). Thus, the mouse NKG2C and NKG2E genes are named in the order of their discovery, rather than on the basis of homology to their human counterparts.

Mouse NKG2C and NKG2E are 91% identical to each other at the amino acid level and are new members of the rapidly expanding family of lectin-like receptors expressed by NK cells. Mouse NKG2C and -E exhibit greatest homology with the extracellular CRD of mouse NKG2A (~95% similar and ~93% identical at the amino acid level), but this similarity disappears almost entirely outside of the CRD (Fig. 1). The previously identified mouse NKG2D molecule (38, 43), on the other hand, differs markedly from the other NKG2 molecules and is in fact no more related to these molecules than it is to CD94 (Fig. 1 B). These comparisons confirm that NKG2D is likely not a bona fide NKG2 family member.

Unlike mouse NKG2A, mouse NKG2C and -E do not contain immunoreceptor tyrosine-based inhibitory motif-like sequences in their cytoplasmic tails. Instead, like human NKG2C and -E, mouse NKG2C and -E appear to contain the typical immunoreceptor tyrosine-based activation motif (ITAM) residues (Fig. 1 B).
a positively charged residue (arginine) in their transmembrane domains. Notably, the position of the charged residue within the transmembrane domain is slightly different between the mouse and human molecules (Fig. 1A). Similar charges have been shown to play an important role in other receptors by mediating associations with activating signaling subunits such as DAP12 (44) or FcRγ (45), but there are no data on the relevant signaling partner for mouse NKG2C or -E.

The cytoplasmic tails of mouse NKG2C and -E do not appear to contain any obvious signaling motifs themselves, although their length (64 amino acids) allows for the possibility that at least some component of the signal transduction by NKG2C/E might be through their own receptor tails. The tails do not share any regions of substantial homology with the human NKG2C/E molecules, with the exception of an invariant four–amino acid motif (PPEK) that is presumed to directly abut the cytoplasmic face of the lipid bilayer. The role of this motif is not known.

NKG2 and CD94 mAbs establish CD94/NKG2 as the predominant Qa-1 receptor on NK cells. To study CD94/NKG2 expression by NK cells, we have generated mAbs against the mouse NKG2 and CD94 molecules. The specificity of these antibodies was confirmed by staining COS cells that had been transiently transfected with CD94, NKG2, or control (Ly49A) cDNAs (Fig. 2A). The 18d3 mAb recognized COS cells transfected with CD94 alone as well as CD94/NKG2 cotransfectants, but it did not recognize COS cells transfected with NKG2A alone. In contrast, the 20d5 mAb recognized COS cells transfected with NKG2A alone as well as CD94/NKG2A cotransfectants. The 20d5 mAb also appears to recognize NKG2C and NKG2E, as 20d5 did not stain COS cells transfected with CD94 alone but did stain CD94/NKG2C/DAP12 and CD94/NKG2E/DAP12 transfectants (Fig. 2A). NKG2C and NKG2E reach the cell surface inefficiently when transfected alone into COS cells (data not shown). Cotransfection with CD94 improved surface expression, and the further addition of DAP12 had a small additional effect (data not shown). As was previously documented (9), CD94 and NKG2A did not require each other to reach the cell surfaces of COS transfectants.

As shown previously (9, 37) and in Fig. 2B, tetrameric complexes of Qa-1β2 microglobulin/Qdm stain roughly half of adult mouse NK cells. Although the Qa-1 tetramers presumably bind CD94/NKG2 receptors on NK cells, this has not been formally demonstrated, and the possibility remains that NK cells express additional Qa-1 receptors. We observed staining of NKG2C and -E receptors on NK cells. As expected, the frequency of tetramer-reactive NKG2C receptors (48% in Fig. 2B) was similar to the frequency of NKG2E receptors (53% in Fig. 2B). The frequency of tetramer-reactive NK cells was somewhat variable and sometimes lower than (but never higher than) the frequency of 20d5+ cells, particularly as the tetramer preparation aged. Similar variability has been previously observed with class I tetramer recognition of Ly49 receptors and its possible explanations discussed (46). To confirm that CD94/NKG2 receptors are the predominant Qa-1 receptors on NK cells, we attempted to block Qa-1 binding with the 20d5 (anti-NKG2) mAb (Fig. 2B). We found that 20d5 completely blocked all binding of Qa-1 tetramers to NKG2C and -E receptors, strongly arguing that CD94/NKG2 receptors are indeed the predominant, if not the only, Qa-1 receptors on mouse NK cells. Similarly, it has been observed that human CD94 blocks all binding of HLA-E tetramers to the human NK cell line (10).

NKG2C and -E recognize Qa-1β2. Based on the high degree of similarity between the extracellular ligand binding domains of NKG2A, -C, and -E, it is reasonable to hypothesize that CD94/NKG2C and CD94/NKG2E heterodimers, like CD94/NKG2A heterodimers, would recognize Qa-1β2. We generated stable transfectants of CHO cells expressing high levels of CD94/NKG2A, CD94/NKG2C, and CD94/NKG2E (Fig. 3). All three stable cell lines, but not untransfected CHO cells, stained brightly with soluble, tetrameric Qa-1β2 microglobulin/Qdm complexes. We previously showed that CD94 alone is incapable of binding Qa-1β2 (9). The results therefore provide direct evidence that both CD94/NKG2C and CD94/NKG2E recognize Qa-1β2.

Figure 2. (A) The 18d3 mAb recognizes mouse CD94, and the 20d5 mAb recognizes mouse NKG2A, -C, and -E. COS cells were transiently transfected with expression vectors encoding the indicated cDNAs (as described in Materials and Methods). The same batch of transiently transfected cells was stained with either 18d3 or 20d5 hybridoma supernatants, followed by FITC-conjugated anti-rat IgG. Ly49A-transfected COS cells were used as a negative control. (B) CD94/NKG2 receptors are the only detectable Qa-1 receptors on NK cells. Nyl...
1806 Recognition of Qa-1b by CD94/NKG2C and CD94/NKG2E

The degree of staining appeared to correlate roughly with the levels of NKG2 and CD94, as independently assessed by staining with anti-NKG2 and anti-CD94 mAbs. In particular, the CD94/NKG2A transfectant stained most brightly with tetramer and with anti-CD94 and anti-NKG2 antibodies, whereas the CD94/NKG2C transfectant stained least brightly with all three staining reagents. Thus, our data do not reveal a gross difference in the avidity of the various CD94/NKG2 subunits for Qa-1b. Evidence in humans using quantitative surface plasmon resonance techniques suggests that CD94/NKG2A binds HLA-E with higher affinity than does CD94/NKG2C (21). Such differences may also exist for the mouse CD94/NKG2 receptors, as our data is not quantitative. Because the Qa-1 protein was produced in Escherichia coli and is therefore unglycosylated, our results also demonstrate that CD94/NKG2C and CD94/NKG2E, like CD94/NKG2A (9), can recognize carbohydrate-independent epitopes on their ligands.

Genomic Mapping and Structure of NKG2C and NKG2E. Previous Southern blot data were consistent with a small cluster of NKG2 genes located within the NKG2 complex near mouse CD94. To determine if the NKG2C and NKG2E cDNAs we obtained might map to the CD94/NKG2 locus, we designed an oligonucleotide probe specific for NKG2E, as well as a probe that recognized NKG2C and NKG2A but not NKG2E. These probes were tested for specificity on their corresponding cDNAs (not shown) and were then used on Southern blots shown in Fig. 4. We found that a 3.5-kb HindIII fragment, present in all the BACs, hybridized specifically to the NKG2E probe. In contrast, the NKG2A/C probe hybridized specifically to a 9-kb HindIII fragment, present only in BACs B6-1 and B6-4. The NKG2A/C probe also recognized a ~20-kb fragment present only in BAC B6-4, the only BAC known to contain NKG2A (9). The 3.5- and 9-kb HindIII fragments were subcloned and partially sequenced and were found to contain sequence that perfectly matched exons 4–7 of NKG2E and NKG2C, respectively (exon numbering based on homology to the human NKG2A gene; reference 47). The boundaries for the mouse exons are presented in Table II. As described below, the boundaries of exons 4 and 5 appear to be somewhat flexible and can give rise to alternatively spliced transcripts.

Together with previously presented data (9, 48), these results allow us to infer the order of mouse CD94 and NKG2 genes within the NKG2 complex on mouse chromosome 6 (Fig. 4 A). Interestingly, the overall structure of the locus is well conserved with that of the human CD94/NKG2 locus. In particular, the gene order CD94–NKG2D–NKG2E/C–NKG2A appears to be conserved (49). The human locus differs by the presence of NKG2F, a truncated
Table II. Splice Acceptors and Donors in NKG2C and NKG2E

<table>
<thead>
<tr>
<th>Gene</th>
<th>Exon</th>
<th>Splice</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Acceptor</td>
<td>agagagactttaacttcag ATGCTGAAGAAGAAGAAG...</td>
</tr>
<tr>
<td>E</td>
<td>4</td>
<td>Donor</td>
<td>CAGATCCACCCCTTCTCGattaaccgcaa...</td>
</tr>
<tr>
<td>C</td>
<td>4</td>
<td>Acceptor</td>
<td>agagagactttaacttcag ATGCTGAAGAAGAAGAAGAAG...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Donor</td>
<td>TTTGGCTCTCAAGAGCAAGAGGAGGAGT...</td>
</tr>
<tr>
<td>E</td>
<td>5</td>
<td>Acceptor</td>
<td>attttgattttaacttcag ATGATCCACCCCTTCTCG...</td>
</tr>
<tr>
<td>C</td>
<td>5</td>
<td>Donor</td>
<td>TTCACATAGATAG...</td>
</tr>
<tr>
<td>E</td>
<td>6</td>
<td>Acceptor</td>
<td>gttctctttgatg ACCTTTTCTCATTGCTATG...</td>
</tr>
<tr>
<td>C</td>
<td>6</td>
<td>Donor</td>
<td>AACACTCAATTTTCAAACAA...</td>
</tr>
<tr>
<td>E</td>
<td>7</td>
<td>Acceptor</td>
<td>atgacccattgaGATGACAGGAGGATG...</td>
</tr>
<tr>
<td>C</td>
<td>7</td>
<td>Donor</td>
<td>AGACTCAATTTTCAAACCAA...</td>
</tr>
</tbody>
</table>

Intronic sequence (based on clones C-ORF4 and E-UTR4) is shown in lowercase. The invariant dinucleotide donors (gt) or acceptors (ag) at the edges of introns are underlined, as are alternative donors or acceptors used in alternatively spliced transcripts.

NKG2 gene that lies between NKG2D and NKG2E (49, 50). We have not seen evidence of a mouse NKG2F gene, although its existence cannot be ruled out.

NKG2C and NKG2E Are Extensively Alternatively Spliced. In addition to the cDNA clones encoding full length NKG2C and NKG2E, we obtained multiple clones encoding NKG2C- or NKG2E-like sequences that differed by various insertions and deletions (Fig. 5). The positions of the insertions and/or deletions corresponded precisely to canonical splice acceptors and donors determined from the

Figure 5. Alternative splicing of Nkg2c and Nkg2e genes. Exons found to be involved in alternative splices are shown as boxes at the top, and various cDNA clones sequenced are depicted below. The portions of the cDNA sequences depicted with a solid black line are predicted to be translatable, whereas the portions of the cDNA depicted with a stippled line are predicted to follow an in-frame stop codon. Only clones encoding NKG2C1, -C1, or -C2 are predicted to generate full length protein, and the nucleotide positions of in-frame stop codons are indicated in parentheses. Exon-intron boundaries were confirmed by direct sequencing of genomic clones, but the distances between exons were not determined. Dashed vertical lines within exons represent the presence of cryptic splice acceptors or donors. Based on sequence polymorphisms between NKG2C and -E, clone E-UTR1 appears to be a fusion of transcripts from both genes. This might represent an artifact of PCR, or it may have been generated in vivo. The exons are numbered according to homology with the human NKG2A gene. The C2 sequence was not obtained in this study but was found previously (26). Additional splice variants appear to exist (Fig. 6) but have not yet been cloned and sequenced.
cDNA samples were derived from RNA obtained from: 1, highly purified, IL-2–cultured NK1.1+CD3+ NK cells; 2, highly purified, IL-2–cultured NK1.1–CD3− NK cells that had also been sorted for expression of NKG2 using the 2D5 mAb; 3, unsorted IL-2–cultured LAK (lymphokine-activated killer) cells; preparation A; 4, unsorted IL-2–cultured LAK cells, preparation B; and 5, unsorted IL-2–cultured LAK cells, preparation C. A total of eight distinct bands were visible, although some were very weak. Based on Southern blot experiments (B and C), these bands could be assigned to either NKG2E or NKG2C, as indicated on the right side of the gel. (B) The agarose gel shown in A was analyzed by Southern blot and probed with an oligonucleotide probe (E-probe), which recognizes NKG2E but not NKG2C. (C) The blot was then stripped and reprobed with an oligonucleotide probe (C-probe), which recognizes NKG2C but not NKG2E. The right portion of the blot was exposed longer than the left to enhance detection of weaker bands and prevent overexposure of the brighter control bands. It should be noted that not all splice variants detected by this method have been cloned. The NKG2C2 splice variant was only very weakly visible.

Figure 6. NKG2C and NKG2E are extensively alternatively spliced. (A) PCR primers (5‘ex2 and 3‘575) flanking the region that is alternatively spliced were used to amplify product from control templates and from various cDNA samples. The PCR products were resolved on a 3% agarose gel and stained with ethidium bromide. The control templates were −, distilled water; A, NKG2A cDNA; 1, NKG2E clone C-UTR1; 2, NKG2E clone E-UTR2; EI, NKG2E clone E-UTR4, which encodes a full length protein, NKG2E1; 3, NKG2E clone E-ORF3; and C1, NKG2C clone C-ORF4, which encodes a full length protein, NKG2C1. Alternative splicing of other exons was not observed. Some of the variants produce reading frame shifts that result in premature stop codons and truncated receptors that would not be predicted to bind ligand (Fig. 5). A previously identified NKG2C cDNA called NKG2C2 (26) likely arose from alternative use of the splice acceptors and donors identified here. One particularly interesting clone was E-UTR1, which was found to contain a chimeric NKG2C/NKG2E cDNA (Fig. 5). It is possible that this transcript was generated as an artifact of PCR. Alternatively, if NKG2C and -E share the same transcriptional orientation in the genome as they do in humans, it is possible that a single primary message encoding both genes was transcribed and spliced to form the chimeric receptor. Alternative splicing, particularly involving exons encoding the transmembrane and stalk domains, has also been observed for other C-type lectin-like receptors, including mouse Ly49 C, D, G, H, and J, human CD94, and human and mouse NKG2B, a splice variant of NKG2A (27, 47, 51–54). Despite their apparent prevalence, the function of these alternative splices is not clear.

To estimate the relative abundance of the splice variants in NK cells, we amplified the alternatively spliced region from various NK cell cDNA samples using flanking primers that recognize NKG2C and NKG2E but not NKG2A. The heterogeneous PCR product was resolved by gel electrophoresis and then analyzed by Southern blot and probed with NKG2C- and NKG2E-specific oligonucleotide probes (Fig. 6). The results confirm that NKG2C and -E are indeed extensively alternatively spliced, and a total of eight different cDNA variants could be detected. The pattern of splicing detected was remarkably consistent in five different RNA samples (Fig. 6), including RNA from highly purified (>98% pure) NK1.1+CD3− NK cells. Control PCR reactions using an equimolar mix of templates suggested that shorter products were not preferentially amplified (data not shown). Interestingly, although NKG2C and NKG2E are closely related genes, they exhibit quite different patterns of mRNA splicing. The majority of NKG2C cDNAs are identical in size to the NKG2C1 cDNA and would therefore be capable of producing a full length protein. On
the other hand, the majority of NKG2E cDNAs are considerably shorter and identical in size to the E-UTR2 cDNA. This cDNA contains a frame shift as compared with the full length ORF. Consequently, most NKG2E cDNAs are predicted to encode a truncated polypeptide that is unable to bind ligand. We conclude that a substantial fraction of the NKG2C and -E message in NK cells is of an alternatively spliced form.

NKG2A transcripts are substantially more abundant than NKG2C and -E transcripts. Previously, we observed that Qa-1 tetramer-reactive NK cells (presumably expressing CD94/NKG2A, -C, and/or -E) were inhibited, not activated, by the presence of Qa-1-Qdm complexes on target cells. These data raised the possibility that the activating NKG2C and NKG2E receptors might be expressed less frequently than the inhibitory NKG2A receptor. Alternatively, mouse NKG2C and -E might recognize Qa-1 relatively poorly or transmit weaker signals. To distinguish these nonexclusive possibilities, we developed a PCR-based assay to quantify the relative levels of NKG2A, -C, and -E message in purified NK cells. We took advantage of the high degree of nucleotide identity among the three cDNAs to design a primer pair that would match perfectly to all three cDNAs. As the primer binding sites and the size of the amplified product is identical for all three sequences, PCR should faithfully amplify each cDNA in proportion to its representation in the original sample. The various amplified cDNA products could be distinguished by PCR by differential digestion with gene-specific restriction enzymes. Radiolabeled nucleotides were included in the amplification reaction to permit quantification of the various digested PCR products by PhosphorImager analysis. The results (Fig. 7) show that the majority of NKG2C PCR product could be digested with an NKG2A-specific restriction enzyme, Mbol, whereas only a small fraction of the product could be digested with NKG2C- and NKG2E-specific restriction enzymes (PvuII and StuI). Using this method, we estimated that NKG2A transcripts are at least 20-fold more abundant than NKG2C and -E transcripts combined. It is unclear whether this difference is best accounted for by the presence of relatively few NKG2C/E+ NK cells or by relatively low NKG2C/E transcript levels per cell. Nevertheless, our molecular analysis suggests that the net inhibition of NK cells by Qa-1-expressing target cells may be explained, at least in part, by the low abundance of activating NKG2C/E receptors on NK cells. It should be emphasized that activating NKG2 genes have been evolutionarily conserved, and it thus seems likely that NKG2C/E genes play an important role that does not depend on abundant mRNA transcripts.

Evolution of Class I Receptors by NK Cells. The above findings document the striking conservation of the function and genomic organization of CD94/NKG2 receptors between mice and humans. In contrast, other class I-specific receptors, such as Ly49 receptors in mice and KIRs in humans, do not appear to be functionally conserved. Given the availability of NKG2 sequences from three species, it is interesting to analyze the evolutionary history of these genes. We therefore aligned the extracellular CRDs for each possible pair of NKG2 genes from mice, rats, and humans. Using the DnaSP algorithm (41), each aligned nucleotide position was designated as a synonymous site (at which nu-

### Table III. Calculation of Number of Substitutions per Synonymous Site (Ks) and per Nonsynonymous Site (Ka) for the CRDs of Various NKG2 Sequences

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Synonymous sites</th>
<th>Nonsynonymous sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence 1</td>
<td>Sequence 2</td>
<td>Ns</td>
</tr>
<tr>
<td>mNKG2A mNKG2E</td>
<td>77.75</td>
<td>0.1257</td>
</tr>
<tr>
<td>mNKG2A mNKG2C</td>
<td>76.33</td>
<td>0.0000</td>
</tr>
<tr>
<td>mNKG2E mNKG2C</td>
<td>77.42</td>
<td>0.1263</td>
</tr>
<tr>
<td>mNKG2A mNKG2A</td>
<td>77.67</td>
<td>0.6369</td>
</tr>
<tr>
<td>mNKG2A hNKG2A</td>
<td>77.67</td>
<td>0.6026</td>
</tr>
<tr>
<td>mNKG2A hNKG2E</td>
<td>77.50</td>
<td>0.6886</td>
</tr>
<tr>
<td>mNKG2E hNKG2A</td>
<td>78.75</td>
<td>0.6330</td>
</tr>
<tr>
<td>mNKG2E hNKG2C</td>
<td>78.75</td>
<td>0.5993</td>
</tr>
<tr>
<td>mNKG2C hNKG2E</td>
<td>77.33</td>
<td>0.6262</td>
</tr>
<tr>
<td>mNKG2C hNKG2C</td>
<td>77.33</td>
<td>0.5922</td>
</tr>
<tr>
<td>mNKG2C hNKG2A</td>
<td>77.17</td>
<td>0.6773</td>
</tr>
<tr>
<td>hNKG2A hNKG2A</td>
<td>78.67</td>
<td>0.0259</td>
</tr>
<tr>
<td>hNKG2A hNKG2E</td>
<td>78.50</td>
<td>0.0925</td>
</tr>
<tr>
<td>hNKG2C hNKG2C</td>
<td>78.50</td>
<td>0.0998</td>
</tr>
<tr>
<td>mNKG2A rNKG2A</td>
<td>77.58</td>
<td>0.2064</td>
</tr>
<tr>
<td>mNKG2A rNKG2C</td>
<td>77.25</td>
<td>0.1906</td>
</tr>
<tr>
<td>mNKG2E rNKG2A</td>
<td>78.67</td>
<td>0.2730</td>
</tr>
<tr>
<td>mNKG2E rNKG2C</td>
<td>78.33</td>
<td>0.2384</td>
</tr>
<tr>
<td>mNKG2C rNKG2A</td>
<td>77.25</td>
<td>0.2103</td>
</tr>
<tr>
<td>mNKG2C rNKG2C</td>
<td>76.92</td>
<td>0.1943</td>
</tr>
<tr>
<td>rNKG2A rNKG2C</td>
<td>78.17</td>
<td>0.0669</td>
</tr>
<tr>
<td>rNKG2A hNKG2A</td>
<td>78.58</td>
<td>0.7132</td>
</tr>
<tr>
<td>rNKG2A hNKG2C</td>
<td>78.58</td>
<td>0.6706</td>
</tr>
<tr>
<td>rNKG2C hNKG2A</td>
<td>78.42</td>
<td>0.7354</td>
</tr>
<tr>
<td>rNKG2C hNKG2C</td>
<td>78.25</td>
<td>0.7073</td>
</tr>
<tr>
<td>rNKG2C hNKG2A</td>
<td>78.25</td>
<td>0.6648</td>
</tr>
<tr>
<td>rNKG2C hNKG2E</td>
<td>78.08</td>
<td>0.6491</td>
</tr>
</tbody>
</table>

m, mouse; h, human; r, rat.

* All calculations performed with DnaSP version 3.0.
† Defined here as the last 363 nucleotides of mNKG2A. Alignment was performed using CLUSTAL W.
cleotide substitutions do not generate amino acid substitutions) or as a nonsynonymous site (at which nucleotide substitutions generate amino acid substitutions; see Materials and Methods). For each pair of aligned sequences we could then estimate \( K_s \), a commonly used parameter that represents the number of synonymous substitutions that have occurred per synonymous site (Table III). The number of nonsynonymous substitutions per nonsynonymous site, \( K_a \), could be similarly calculated. Synonymous (or 'silent') substitutions are usually considered selectively neutral, or nearly so. Thus, a high \( K_s \) value reflects the substantial selection-independent accumulation of nucleotide differences (55). The analysis demonstrates that within a species, NKG2 sequences are closely related, whereas between species they are not. The intraspecies similarity is most dramatically exemplified by the finding that there are, for example, no synonymous substitutions within the CRDs of mouse NKG2A and mouse NKG2C. Other intraspecies pairs of NKG2 genes give similarly low frequencies of synonymous substitution, i.e., 0.03–0.13 (Table III). In contrast, the frequency of synonymous substitutions for interspecies pairs of NKG2 genes is substantially higher, i.e., 0.60–0.68 (Table III).

These data are consistent with two broad evolutionary histories: (a) humans and rodents share a single ancestral NKG2 gene that underwent duplications separately in each lineage to generate 'independent' and relatively young multigene families; or (b) the common ancestor of humans and rodents contained all three genes, but the genes have since been undergoing 'concerted evolution,' a process by which homologous loci within a species maintain similarity to each other. Concerted evolution has been inferred in several multigene families (55). The mechanisms by which it occurs remain unclear but are thought to involve homogenizing events, such as unequal crossing over and/or gene conversion-like events (55).

Although it is difficult to adjudicate between the two above evolutionary histories, the second model may be more likely for the following reasons. First, all three species examined contain multiple NKG2 genes, including at least one inhibitory and one activating type, consistent with a primitive duplication event that was transmitted to all lineages. Second, the 5' (cytoplasmic) portions of the inhibitory and activating NKG2 genes tend to vary substantially from each other, even within species, suggesting that the genes have been diverging for some time. Unless gene conversion-like events—which can affect discrete parts of genes—are invoked, it is difficult to reconcile the evolutionary divergence apparent in the 5' portions of mouse NKG2A and -C with the finding that they do not exhibit any synonymous substitutions in their CRDs. Lastly, the remarkable overall similarity seen in the genomic structures of the mouse and human NKG2 loci (Fig. 4 and reference 49) suggests that the gene families did not result from independent duplication events. It is interesting that gene conversion events have also been proposed to play a role in the evolution of MHC class I genes (56).

Our previous work on mouse CD94/NKG2A provided evidence that recognition of class I was a primitive function of NK cells that predated the divergence of mouse and human ancestors. Assuming that NKG2 genes underwent a primitive duplication, as argued above, our work seems to suggest that both activating and inhibitory class I recognition may have been primitive functions of NK cells. This is interesting, as it suggests some fundamental role for activating receptors in the recognition of class I. This notion is further supported by the observation that families of class I-specific receptors, including Ly49, KIR, and leukocyte Ig-like receptors/Ig-like transcript families, invariably consist of both activating and inhibitory paralogs. Indeed, the pairing of inhibitory and activating isoforms extends beyond receptors specific for class I, and includes the paired Ig-like receptors on B cells and myeloid cells (57) and the CD28/CTLA-4 coreceptors on T cells. Although well understood in the case of T cells, the molecular logic behind the existence of paralogous activating and inhibitory receptors on NK cells is still a matter of conjecture. We anticipate that the identification of mouse CD94/NKG2C and CD94/NKG2E as activating receptors will permit the appropriate in vivo experimental manipulations required to dissect their function.
References


