Characterization of the Chicken C-Type Lectin-Like Receptors B-NK and B-lec Suggests That the NK Complex and the MHC Share a Common Ancestral Region

Sally L. Rogers, Thomas W. Göbel, Birgit C. Viertlboeck, Sarah Milne, Stephan Beck and Jim Kaufman


http://www.jimmunol.org/content/174/6/3475
Characterization of the Chicken C-Type Lectin-Like Receptors B-NK and B-lec Suggests That the NK Complex and the MHC Share a Common Ancestral Region

Sally L. Rogers,* Thomas W. Göbel,† Birgit C. Viertlboeck,† Sarah Milne,‡ Stephan Beck,‡ and Jim Kaufman²*

The sequencing of the chicken MHC led to the identification of two open reading frames, designated B-NK and B-lec, that were predicted to encode C-type lectin domains. C-type lectin domains are not encoded in the MHC of any animal described to date; therefore, this observation was completely unexpected, particularly given that the chicken has a “minimal essential MHC.” In this study, we describe the initial characterization of the B-NK and B-lec genes, and show that they share greatest homology with C-type lectin-like receptors encoded in the human NK complex (NKC), in particular NKR-P1 and lectin-like transcript 1 (LLT1), respectively. In common with NKR-P1 and LLT1, B-NK and B-lec are located next to each other and transcribed in opposite orientation. Like human NKR-P1, B-NK has a functional inhibitory signaling motif in the cytoplasmic tail and is expressed in NK cells. In contrast, B-lec contains an endocytosis motif in the cytoplasmic tail, and like LLT1, is an early activation Ag. Further analysis leads us to propose that there are four subgroups of C-type lectin-like receptors in the NKC, which arose as a result of duplication events. Moreover, this analysis suggests that the NKC may be considered a fifth paralogue region, and therefore shares an ancient common origin with the MHC. This provides evidence that C-type lectin-like receptors were present in the preduplication, primordial MHC region, and suggests that an original function of MHC molecules was for recognition by NK cell receptors encoded nearby. The Journal of Immunology, 2005, 174: 3475–3483.

The sequencing of the chicken MHC has shown that the chicken has a small, compact MHC, containing only 19 genes in 92 kb (1, 2). Many of these genes were expected from their presence in the mammalian MHC, for example MHC class I and class II, and the TAP genes. However, a number of genes commonly found in mammalian MHC regions are absent from the chicken MHC. The chicken MHC was also found to be organized differently, with the class I and class II regions located next to each other, and at least part of the class III region extending away from the class I region. In addition, intergenic distances and introns are generally much smaller than had been observed in mammalian MHC. This led to the hypothesis that the chicken has a “minimal essential MHC,” stripped of pseudogenes and multi-gene families, and containing only genes that are essential for immune function. The identification by bioinformatic analysis of two sets of exons predicted to encode C-type lectin domains, designated B-NK and B-lec, in the chicken MHC was completely unexpected, because no C-type lectin genes have been identified in the MHC of any mammal studied to date.

C-type lectins are a large superfamily of proteins with a variety of functions in the immune response (3). They were originally identified and characterized as calcium-dependent carbohydrate-binding proteins, and were later shown to contain a conserved structural fold, termed the carbohydrate recognition domain (CRD).³ More recently, a number of proteins have been identified that contain the conserved structural residues for correct folding of the C-type lectin domain but that lack the critical residues for calcium (and therefore carbohydrate) binding. This structure has been termed the C-type lectin-like domain (CTLD) and has been identified in a large number of proteins, including a group of mammalian NK receptors.

The entire C-type lectin superfamily has also been divided into seven structural groups, based on the overall architecture of the proteins, the positions of the CRD/CTLD relative to other domains, and the degree of similarity between the amino acid sequence of the CRD/CTLD (4). Briefly, these groups are as follows: I, proteoglycans; II, hepatic lectins, the Kupffer cell receptor, and CD23; III, the collectins including mannose binding protein; IV, the selectins; V, the type II membrane-bound CTLD; VI, the mannose receptors; and VII, the free CRD. Initial analysis of B-NK and B-lec suggested that the putative amino acid sequences shared greatest homology with receptors belonging to group V of the C-type lectin superfamily, which are those that contain a CTLD, including the C-type lectin-like NK receptors. Many of the C-type lectin-like NK receptors have been shown to bind MHC class I molecules (5, 6), an observation that formed the basis for one of our hypotheses as to why B-NK and B-lec had maintained their place in the chicken “minimal essential MHC.” We predicted that

³ Abbreviations used in this paper: CRD, carbohydrate recognition domain; CTLD, C-type lectin-like domain; NKC, NK complex; CHX, cycloheximide; C₉₅, threshold cycle value; UTR, untranslated region; AICL, activation-induced C-type lectin; LLT1, lectin-like transcript 1; POD, peroxidase; SHPTP, Src homology domain 2-containing protein tyrosine phosphatase; SHP, Src homology region 2 domain-containing phosphatase.

Copyright © 2005 by The American Association of Immunologists, Inc.

Received for publication April 15, 2004. Accepted for publication January 3, 2005.

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

¹ Institute for Animal Health, Compton, Berkshire, United Kingdom; ² Institute for Animal Physiology, University of Munich, Munich, Germany; and ³ Sanger Center, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, United Kingdom

Address correspondence and reprint requests to Dr. Jim Kaufman, Institute for Animal Health, Compton, Newbury, Berkshire RG20 7NN, U.K. E-mail address: jim.kaufman@bbsrc.ac.uk

* Address correspondence and reprint requests to Dr. Jim Kaufman, Institute for Animal Health, Compton, Newbury, Berkshire RG20 7NN, U.K. E-mail address: jim.kaufman@bbsrc.ac.uk

© 2005 by The American Association of Immunologists, Inc.
both B-NK and B-lec would be NK receptors that bind the MHC class I molecules encoded nearby, and that the close physical proximity of B-NK and B-lec to their putative ligands would allow coevolution to occur, such that receptors from one haplotype would specifically bind MHC class I molecules from the same haplotype. In support of this hypothesis, we have previously shown that B-NK cDNA is present in an activated NK cell line, but not in T cell, B cell, or macrophage cell lines (1). Strikingly, all of the human group V C-type lectin-like receptors are encoded in region on chromosome 12, termed the NK complex (NKC) (7, 8). This region contains a number of genes encoding receptors involved in the development and regulation of NK cells, and is entirely separate from the human MHC (HLA) on chromosome 6. However, the NKC has been proposed as a fifth paralogous region of the MHC (9). The MHC-paralogous regions were identified following studies on the number and location of genes encoding subunits of the proteasome (10). It was noted that regions on at least three other chromosomes had similarities to the MHC, both in the gene families that they contain and their organization. It has been proposed that these four regions share a common origin, and arose from successive duplications of an ancestral region. Genes within a single species that arose by such duplication events are termed paralogs, and the chromosomal segments that contain these genes are called paralogous regions. At least six members of the MHC-paralogous gene families have copies near the NKC on chromosome 12. These include α2-macroglobulin (a member of the complement C3/C4/C5 family) and a tapasin-related gene called TAPBP-R (11). Two C-type lectin genes (CD23 and DC-SIGN) have been mapped to the paralogous region on chromosome 19 (12), despite the absence of any C-type lectin genes in the human MHC. Together with the abundance of C-type lectin genes encoding subunits of the proteasome C3/C4/C5 family) and a tapasin-related gene.

For further codification and identification of transcriptional start site for B-lec was obtained by performing 5′-RACE using a 5′/3′-RACE kit (Roche). LPS-stimulated, monocye-derived macrophage RNA (kindly provided by L. Rothwell, Institute for Animal Health), which was shown to contain B-lec cDNA, was PCR amplified. The PCR products were designed against the predicted CTLD-encoding exons (data not shown). One microliter of this RNA sample was used in the 5′-RACE reaction according to the manufacturer’s instructions, using gene-specific reverse primers as follows: SP1, 5′-TGATGCAGAAGTCCCTCTCG-3′; SP2, 5′-TTGTTGGCAAAACCTATCCCTG-3′; SP3, 5′-CTCTTGTGTCATGAGG-3′. The products obtained after three rounds of amplification were cloned and sequenced.

Sequence alignment and phylogenetic analysis
C-type lectin-like sequences were obtained from public databases, and a multiple sequence alignment was produced with Pileup (GCG10), and optimized using GeneDoc. Phylogenetic trees of CTLDs were constructed using the Neighbor Joining method in the PHYLIP software package.

Quantitative real-time RT-PCR (TaqMan)
Tissues were collected from a 12-wk-old male line C-B12 (White Leghorn) chicken, cut into 0.5-cm chunks, and immediately immersed in RNAlater (Ambion) and stored at −20°C. PBL were isolated by density gradient centrifugation (Ficoll-Paque; density = 1.077), and stimulated using 50 ng/ml PMA (Sigma-Aldrich) with or without 50 μM cycloheximide (CHX; Sigma-Aldrich). Total RNA was isolated from stimulated and unstimulated cells using RNeasy spin columns (Qiagen). B-NK and B-lec mRNA levels were detected using gene-specific primer and probe sets that were designed using the Primer Express software program (Applied Biosystems). B-NK and B-lec probes were labeled with the fluorescent reporter dye FAM at the 5′ end and with the quencher TAMRA at the 3′ end. Primer and probe sequences were designed to cross exon boundaries to prevent amplification of genomic DNA, and are as follows: B-NK probe, 5′-FAM)-(CCTACCAACAGCAGACTGGCTAGT-3′)-(TAMRA); B-NK forward primer, 5′-GGAAATGGTGAGCAACT-3′; B-NK reverse primer, 5′-ATTTCTTACACAGGAGG-3′; B-lec probe, 5′-FAM)-(CCATCGGCTCGGACCTGCAACAGCAGT-3′)-(TAMRA); B-lec forward primer, 5′-GAACCCCTCCTGCTGCTG-3′; B-lec reverse primer, 5′-TCCCCAGGATGACGACA-3′. RT-PCR was performed using the TaqMan EZ RT-PCR kit (Applied Biosystems) for the tissue samples, and using the Reverse Transcriptase qPCR MasterMix (Eurogentec) for the PBL samples, according to the manufacturers’ instructions. Amplification and detection of specific products was performed using the ABI PRISM 7700 Sequence Detection System (Applied Biosystems) with the following cycle profile: one cycle of 50°C for 2 min, 96°C for 3 min, 60°C for 30 min, and 95°C for 5 min, followed by 40 cycles of 94°C for 20 s, 59°C for 1 min. Results were detected as threshold cycle value (Ct), which is the cycle at which the change in the reporter dye passes a significance threshold set at 0.05 for all the primer/probe sets described in this work. To normalize the Ct values of B-NK and B-lec to control for variation in sampling and RNA preparation, 28s specific primers and probe (kindly provided by P. Kaiser and L. Rothwell (Institute for Animal Health)) were used to detect the sequences of which are as follows: 28s probe, 5′-VIC-AGGAC CGCAGCGAATTCCCAACCA-(TAMRA)-3′; 28s forward primer, 5′-GCC GAAGGCAAGAGGAAAACT-3′; and 28s reverse primer, 5′-GACCAC CGATTTGCCAGCTC-3′. The mean Ct value for 28s RNA-specific

Materials and Methods
RNA isolation, cDNA production, and isolation of B-NK and B-lec cDNA sequences
Spleen and bone marrow mRNA was isolated using oligo(dT) magnetic Dynabeads according to the manufacturers instructions (Dynal). First-strand cDNA was produced by RT-PCR using the reverse transcriptase Superscript II (Invitrogen Life Technologies) following manufacturer’s instructions, and an oligo(dT) primer (Invitrogen Life Technologies). B-NK-specific cDNA products were amplified using primers designed against the genomic sequence (accession no. AL023516), as follows: B-NK forward primer c591 5′-ATGATGAGGAAATGACTATGCT-3′ (predicted start codon underlined). B-NK reverse primer c588 5′-CATTAGCAGAGGACGCTAAAGG-3′ (reverse complement predicted stop codon underlined). Three potential cytoplasmic exons containing the start codon (underlined) were identified for B-NK, and forward primers were designed for each of these as follows: primer c592 in exon 1a (5′-ATGGTTGTTTTTCTTAAA TACAC-3′), c587 in exon 1b (5′-ATGATCACCATATTTCGCATTCCG-3′), and c588 in exon 1c (5′-TGTTTCCTTTATTATTTGAGGAG-3′). A reverse primer underlined). B-NK cDNA was amplified using Pwo DNA polymerase (Hybaid) for B-NK in a 30-μl reaction. The cycling conditions were as follows: 96°C for 2 min, followed by 30 cycles of 96°C for 30 s, a primer-specific annealing step of 55°C for primers c590/c591, c587/c588, and c588/c589, or 49°C for primers c592/c589 for 30 s and 1 min at 72°C. Two microliters of the PCR product was used in a second round of amplification using the same protocol. Thirty microliters of the second PCR product was analyzed by agarose gel electrophoresis, gel extracted, and cloned. Since the first isolations that revealed the genomic organization for both genes, numerous clones from different tissues and chicken MHC haplotypes have been isolated using this method. The B-lec start-stop codon sequence shown in Fig. 1 was amplified using Pwo DNA polymerase (Hybaid) and primers c592 (5′-AAGCGGTTACCATGGAGAAGTGAGGAAATGACT-3′) and c584 (5′-CAAGTCTCCTGGCTGTAGTGCAAGAATGGCTTCT-3′) from 5 μl of cDNA prepared from the bone marrow cell line C-B12 (White Leghorn) chicken. This bursal cDNA was prepared using RNeasy spin columns (Qiagen) and Superscript II (Invitrogen Life Technologies) following manufacturers’ instructions. The cycling conditions were as follows: 94°C for 2 min, followed by 35 cycles of 94°C for 30 s, 55°C for 30 s, and 1 min at 68°C. All clones obtained were sequenced on both strands using an ABI 377 automated sequencer (Applied Biosystems). Comparison of all sequences has revealed identical genomic organization for all clones of each particular gene.
product was calculated by pooling values from all tissue or all cell experiments. Tube-to-tube variations in 28S rRNA C_T values about the experimental mean were calculated, and used to normalize the B-NK and B-lec C_T values.

**Pervanadate treatment of B-NK-transfected cells and Western blot**

A B-NK cDNA amplified from NK cells derived from B19 chickens (1) was cloned into the plasmid vector pcDNA3.1/V5/His-Topo vector (Invitrogen Life Technologies) with a C-terminal V5 epitope. Ten million cells of the UG9 T cell line were transfected with 20 μg of construct DNA using electroporation at 1.0 kV, selected using 1 mg/ml G418 (Applichem), and screened using a mAb to the V5 epitope (Invitrogen Life Technologies). A clone with bright surface fluorescence was designated UG9-BNK.

Fifty million UG9-BNK cells in 5 ml of medium were stimulated for 5 or 15 min with 0.1 mM Na_3VO_4 containing 0.05% H_2O_2 at 37°C, washed twice in cold PBS containing 0.4 mM EDTA and 0.4 mM Na_3VO_4, and then treated for 45 min on ice with lysis buffer (1% Triton X-100, 150 mM NaCl, 50 mM Tris-Cl (pH 7.4), 1 mM EDTA) supplemented with protease inhibitors (Complete; Roche) and phosphatase inhibitors (10 mM NaF, 2 mM EDTA, 1 mM Na_3VO_4), and then centrifuged at 16,000 × g for 30 min at 4°C. Immunoprecipitates with amab to V5 (Invitrogen Life Technologies) were analyzed by SDS gel electrophoresis and Western blot using a nylon membrane and mouse mAbs (anti-V5-peroxidase (POD) conjugate (1:5000; Sigma-Aldrich), anti-phosphotyrosine RC20H-POD conjugate (1:5000; Transduction Laboratories)) and rabbit antiserum (1 μg/ml rabbit anti-mouse Src homology domain 2-containing protein tyrosine phosphatase (SHPTP)-1/Src homology region 2 domain-containing phosphatase (SHP)-1 antisemur (Upstate) and 1 μg/ml rabbit anti-mouse SHPTP-2 (C18) antisemur (Santa Cruz Biotechnology), both detected with goat anti-rabbit Ig-POD (1:5000; Amersham Biosciences), and then visualized using chemiluminescent detection (Super Signal; Pierce).

**Results**

**B-NK and B-lec encode type II transmembrane proteins containing an extracellular CTLD**

The sequence of the 678-bp fragment amplified from spleen with the B-NK-specific primers designed around the start and stop codons of B-NK is shown in Fig. 1 (accession no. AJ634338), together with the deduced amino acid. Comparison with the genomic sequence of the chicken MHC revealed that the cDNA sequence obtained is interrupted by five introns into six exons (Fig. 2). The first three exons of B-NK correspond to separate functional domains of the protein. Exon 1 encodes the cytoplasmic domain (aa

**FIGURE 1.** Coding nucleotide and deduced amino acid sequences for B-NK and B-lec. Identity between the two amino acid sequences is indicated with a vertical line. Dashed lines in the sequences indicate gaps introduced for optimal alignment. The predicted transmembrane region is underlined. The endocytosis motif of B-lec (YxxV) and the ITIM of B-NK (IxYxxL) are double-underlined. Diamonds indicate the last nucleotide of an exon, and an asterisk indicates a stop codon.
FIGURE 2. Genomic organization of B-NK and B-lec in the chicken MHC. The top cartoon shows the organization of genes (represented as filled boxes) in the B-F/B-L region. The enlarged section shows the genomic organization of B-NK and B-lec. Coding sequences are represented by filled boxes, numbered with roman numerals above and are drawn to scale. The functional properties of exons are indicated (TM, transmembrane; CT, cytoplasmic tail). Unfilled boxes represent UTRs, and the dotted-line boxes at the start of B-lec represent two further initiation sites as described in the text. Arrows show the direction of transcription. The scale along the bottom is numbered to reflect the genomic organization, with numbers indicating the first and last nucleotide of each exon. The A of the start codon is designated 1, and numbering finishes at the end of the polyadenylation site. The expression of the 3′-UTRs have not been experimentally determined.

1–24) and contains a canonical ITIM motif (IYXxL). This is shorter than the previously published sequence of B-NK (1) (accession no. GGA245903) isolated from an NK cell line, which contains an extra 5 aa at the 3′ end of exon I. These extra amino acids were not detected in any B-NK clones from seven more independent PCR from normal tissues, presumably indicating alternative splicing between the lymphokine-activated splenic NK cell line and unactivated tissue preparations. Exon II (aa 25–52) encodes the transmembrane region, and exon III (aa 53–97) encodes a stalk domain and the start of the CTLD. Exons IV, V, and VI (aa 98–225) encode the remainder of the CTLD. A single polyadenylation site (AATAAA) was identified in the genomic sequence 199 bp downstream of the stop codon.

B-lec was originally identified by the Sanger Sequencing Centre as four putative exons encoding a predicted transmembrane region and a CTLD. Analysis of the genomic sequence by eye revealed a further three putative exons predicted to encode cytoplasmic domains (on the basis that they encode a methionine followed by tyrosine-based signaling motif and a splice site). The ATG start codons of these predicted exons are located 507, 1356, and 1613 bp away from the start ATG of B-NK, but are transcribed in the opposite orientation. PCR containing forward primers designed against the first two predicted cytoplasmic exons were used in pairs with a reverse primer designed against the predicted stop codon region, but amplification gave nonspecific products of an unexpectedly large size from cDNA preparations. However, PCR containing a forward primer designed against the third predicted cytoplasmic exon used with the reverse primer from the stop yielded identical products of 846 bp from bone marrow and spleen cDNA preparations (accession no. AJ634334). This cDNA sequence contains two potential start ATG codons, the first of which is at the 5′ end of the sequence (the third predicted cytoplasmic exon). However, there is a stop codon 171 bp downstream and in frame with this first ATG, as a result of which there is no open reading frame into the CTLD. Moreover, this ATG is not part of a good Kozak consensus sequence. The second ATG is part of a good Kozac consensus sequence located 264 bp downstream from the first ATG, and is followed by a 567-bp open reading frame that contains the CTLD as predicted. We believe therefore that the second ATG is the correct translation start point (1877 bp away from the B-NK start ATG), which means that the sequences isolated contain a 5′-untranslated region (UTR) of at least 279 bp. We cannot rule out the possibility that alternative splicing may occur, but no cDNA sequences containing the two predicted upstream cytoplasmic exons have been isolated from bone marrow, spleen, or LPS-stimulated macrophage cDNA samples.

The cytoplasmic protein-coding exon that was identified in B-lec was not one that had originally been predicted, and to confirm that the sequence was correct, 5′-RACE was also performed. LPS-stimulated, monocyte-derived macrophage RNA was used as a template for this experiment, because we had found it to contain a significant amount of B-lec RNA using primers designed against the CTLD encoding exons (data not shown). 5′-RACE yielded three different size products, one of which was nonspecific. Sequencing of the other two products revealed identical exon organization as the B-lec cDNA sequences isolated by PCR, confirming that the cytoplasmic exon identified using this method was correct. The two differently sized specific 5′-RACE products varied only in the length of the 5′-UTR, having either a 22-bp (accession no. AJ634336) or 182-bp 5′-UTR (accession no. AJ634335). Together with the results from the earlier PCR experiments, we have so far identified a total of three transcriptional start sites for B-lec. This result is consistent with the fact that there is no TATA box in the B-lec promoter, because multiple transcriptional start sites are a general feature of TATA-less promoters (13). In contrast, a single band was obtained by 5′-RACE giving a 5′-UTR of 21 bp for B-NK, which does contain a TATA box in the promoter region (data not shown). Finally, a cDNA fragment coding for the B-lec protein (from the start-to-stop codon) was amplified, which contained the same exon organization as has been described here so far (accession no. AJ634337).

Comparison of the B-lec cDNA sequence with the genomic sequence of the chicken MHC revealed that the cDNA sequence obtained is interrupted by four introns into five exons (Fig. 1). Exon I (amino acid residues 1–30) encodes the cytoplasmic domain and, in contrast to the signaling motif identified in B-NK, contains a canonical endocytosis motif (YxxV). Exon II (residues 31–49) encodes a hydrophobic stretch characteristic of a transmembrane region. Unlike B-NK, no separate stalk exon was identified either in the PCR or in the RACE products. Exons III, IV, and V (residues 49–188) encode a short 9-aa stalk region, followed by a CTLD. A single polyadenylation site (AATAAA) was identified in the genomic sequence 394 bp downstream of the stop codon.

All the introns identified in both B-NK and B-lec start with GT and end with AG, and correspond to the phases and sites of introns that interrupt the genes of many mammalian C-type lectin receptors. Finally, no charged residues (which would suggest association with signaling adaptor molecules such as DAP12) were identified in the transmembrane region of either B-NK or B-lec.
B-NK is most closely related to the human receptor NKR-P1, whereas B-lec is most closely related to activation-induced receptors.

The amino acid sequences of the CTLDs of B-NK and B-lec were compared with a number of CRDs and CTLDs, and a multiple sequence alignment was made (Fig. 3). All of the motif residues that are essential for correct folding of the CTLD are present in B-NK and B-lec, including the tryptophan/hydrophobic/glycine/hydrophobic (WIGL) motif that is characteristic of the family. Furthermore, a pair of cysteines (one in the predicted b0 strand and the other in the predicted b1 strand) is observed in the B-NK and B-lec sequences, which characterize both receptors as long-form CTLDs. In common with other group V CTLD sequences, B-NK and B-lec lack the critical residues for calcium-dependent binding of carbohydrate by mannose binding protein.

The CTLD amino acid sequences were then analyzed using the PHYLIP program, and a neighbor-joining phylogenetic tree was constructed (Fig. 4). The results obtained show that B-NK clusters with the NKR-P1 family of receptors, and the closely related molecules KLRF1 and MAFA-L in the tree shown in Fig. 4. A number of trees were constructed during the analysis of B-NK and B-lec, and the tree shown is representative of the results obtained. The majority of trees placed B-NK in a clade with KLRF1 and the NKR-P1 family; occasionally B-NK formed an outgroup by itself.

Pairwise comparisons showed that the CTLD of B-NK has the highest amino acid sequence identity to human MAFA-L (28%), followed by human NKR-P1A and murine NKR-P1B at 27 and 25%, respectively. Although B-NK has a relatively high level of amino acid identity with human CD69 (27%), it does not cluster with this protein in the phylogenetic tree, and shares lower levels of identity with other members of the group; 21% amino acid identity with human activation-induced C-type lectin (AICL) and 20% with human lectin-like transcript 1 (LLT1). The most closely related chicken gene is B-lec, which shares 24% amino acid identity with B-NK, followed by 17.5 and Y-lec2, at 21 and 19%, respectively.

B-lec is most closely related to a group of activation-induced lymphocyte receptors, which includes CD69, AICL, LLT1, and the murine Crr. Also clustering in this group are two predicted amino acid sequences, Y-lec2 and 17.5, which are encoded by genes situated in the chicken Rfp-Y region. Pairwise comparisons showed that the CTLD of B-lec is most similar to human LLT1, and the murine homologs, the Crr family of receptors, with 43% amino acid identity and 38, 37, and 41%, respectively. Moreover, the CTLD of B-lec has significant identity to other C-type lectin-like receptors encoded near LLT1 in the human NKC, for example, 37% with AICL and 36% with CD69. The most closely related

FIGURE 3. Amino acid alignment of the CTLDs of B-NK and B-lec with other CTLD sequences. Sequences start 3 aa before the conserved cysteine found in most of the long-form C-type lectins. Dashes in the sequence indicate gaps introduced for optimal alignment. Residues that are characteristic of the domain are indicated in the motif line above the alignment, and are denoted C for cysteine (C* indicates the pair of cysteines that form a disulfide bond in the long-form CTLDs), O for aromatic, H for hydrophobic, and G for glycine. Elements of secondary structure are shown above the alignment and are determined from a consensus of the crystal structures of Ly49A (1QO3), CD94 (1B6E), NKG2D (1HQ8), and CD69 (1FM5). Arrows represent /H-strands, and slashes indicate /H-helices. The critical residues for calcium binding are denoted by /H above the alignment. Accession nos. of the sequences used in this alignment are as follows: mLy49A (P20937), mNKG2A (Q9Z202), mNKG2D (AF054819), mCD94 (AF057714), hNKR-P1A (U11276), hKLRF1 (AF15206), hMAFA-L (O75613), hDECTIN1 (AY009090), mCrr (AAK70358), hLLT1 (Q9UHP7), hAICL (Q92478), chB1 (Q90644), chHepLec (P02707), mMBP (P11226), cMBP (AF022226), cY-Lec2 (Q08258). “h” denotes human sequence, “m” denotes murine sequence, and “ch” denotes chicken sequence.
other subgroups are discussed in the text. Subgroup 3 has been described previously (16), and the right mDCIR (MMU1335330), mDC-SIGN (AF373408), and hCLEC1 (AF200949), hCLEC2 (XP006932), hLOX1 mNKR-P1A (P27811), mNKR-P1C (P27814), mDECTIN1 hNKG2A (P26715), hNKG2D (P26718), hCD94 (Q13241), additions per site. Accession nos. are as in Fig. 4, with the distance given in the scale represents 0.143-aa substitutions per site. Accession nos. are as shown in Fig. 4, with the addition of mLy49D (Q60651), mLy49B (Q60660), hNKG2A (P26715), hNKG2D (P26718), hCD94 (Q13241), mNKR-P1A (P27811), mNKR-P1C (P27814), mDECTIN1 (AF262985), mNKG2A (AAK70357), mNKG2G (AAK70359), hCLEC1 (AF200949), hCLEC2 (XP006932), hLOX1 (NP002534), mCD69 (P37217), hDCIR (HA133532), hDCIR (MMU1335330), mDC-SIGN (AF373408), and hCD23(P06734). The group assignments are shown to the right of the tree, and as described by Drickamer (4). Subgroup 3 has been described previously (16), and the other subgroups are discussed in the text.

Chicken amino acid sequences are encoded by the C-type lectin-like genes in the Rfp-Y region, Y-lec2 and 17.5, at 40 and 36% identity, respectively. These associations are supported by the statistical significance of the bootstrap values.

**B-NK and B-lec are both expressed predominantly in lymphoid tissues**

Quantitative real-time RT-PCR (TaqMan) was used to characterize expression patterns of B-NK and B-lec in different tissues. The $C_T$ values for 28S rRNA did not alter significantly from sample to sample, but B-NK and B-lec values varied from sample to sample and between target probes. Thus, the $C_T$ values for 28s rRNA appeared to be independent of B-NK and B-lec production, as expected, and were taken to be representative of the level of RNA extracted from all samples for use in normalizing B-NK and B-lec $C_T$ values. Fold changes away from the sample containing the lowest amount of target mRNA for each primer/probe set were calculated, and are shown in Fig. 5A. The sample containing the lowest amount of B-lec mRNA was isolated from the kidney, and was set at an arbitrary value of 1. Expression was highest in the bursa and cecal tonsils, which contain 165- and 3272-fold more B-lec mRNA than kidney, respectively. The sample containing the lowest amount of B-NK mRNA was isolated from the proventriculus, which was set at a value of 1. Expression of B-NK was highest in the spleen, which contained 2466-fold more mRNA than the proventriculus; followed by the bursa, thymus, and bone marrow, which contain 768-, 574-, and 540-fold more mRNA, respectively.

**B-NK can function as an inhibitory receptor in chicken cells**

To assess the ability of B-NK to recruit the cytoplasmic phosphatase SHP-1, UG9-BNK cells (a clone of chicken UG9 T cells stably expressing B-NK) were treated with the phosphatase inhibitor pervanadate to increase the basal level of tyrosine phosphorylation. As shown in Fig. 5B, analysis before and after pervanadate treatment revealed an increase in phosphorylation of B-NK, with a corresponding increase in recruitment of both SHP-1 and SHP-2, but no increase in the amount of B-NK glycoprotein. Three isoforms of B-NK with different apparent m.w. were detected using the anti-V5 Ab, which have also been seen in transfected CHO cells. We have found that the three isoforms are due to differential glycosylation (14).

**B-lec is an early activation Ag**

Analysis of the amino acid sequence from the CTLD of B-lec showed close similarity to the early activation Ags CD69, LLT1, and AICL. mRNAs for these receptors are inducible in PBL with PMA, and superinduction was detected when PBL were stimulated with both PMA and the translational inhibitor CHX (15). To determine whether B-lec is also an activation Ag, chicken PBL were stimulated with PMA in the presence or absence of CHX. B-lec mRNA levels were detected using quantitative RT-PCR, normalized using $C_T$ values for 28s rRNA, and fold changes away from the $C_T$ value from unstimulated cells at the same time point (set at 1) were calculated (Fig. 5C). The results reveal a 2.76-fold increase in B-lec mRNA levels within 2 h in PMA-stimulated PBL compared with unstimulated cells. Moreover, a 331-fold increase of B-lec mRNA was detected in cells stimulated with PMA and CHX together within 2 h. Likewise, up-regulation of B-lec mRNA was detected at 6 and 20 h poststimulation, with 11.3- and 20.7-fold increase in levels, respectively, and superinduction of B-lec mRNA in response to PMA was seen in the presence of cycloheximide at 6 and 20 h poststimulation with 1813- and 2985-fold increases observed, respectively. These results demonstrate that B-lec is an early activation Ag. Moreover, the expression kinetics of B-lec was similar to that observed for CD69 and LLT1, regardless of whether the results were analyzed as described here (fold change away from unstimulated cells at the same time) or as described by Eichler et al. (15) (fold changes away from cells stimulated with PMA for 2 h). In contrast to B-lec, smaller (and slower) changes were found in levels of B-NK mRNA, with a decrease in message detected in cells stimulated for 2 h with PMA compared with unstimulated cells. B-NK mRNA levels were consistently low in cells stimulated with PMA compared with unstimulated cells, which was only slightly enhanced in the presence of CHX, with 5-fold increase at 2 h posttreatment, rising to 30-fold increase at 6 and 20 h poststimulation compared with unstimulated cells. This is similar to the pattern observed for NKR-P1 (15).

**Discussion**

The phylogenetic distance of birds and mammals can make identification of direct orthologous challenging, particularly when taking into account the diversity of the C-type lectin family as a whole...
and the NK receptors in particular. Comparisons even within mammals demonstrate this, because the functional analogs of the C-type lectin-like Ly-49 NK receptor family in rodents are the killer Ig receptors in primates. However, based on genomic organization, phylogenetic sequence analysis, locus organization, and functional characteristics, we have shown that the two C-type lectin-like genes B-NK and B-lec in the chicken MHC are most closely related to the mammalian receptors NKR-P1 and LLT1, respectively. We have shown that B-NK contains a functional ITIM in the cytoplasmic domain, and together with our previous observation that B-NK is expressed in an NK cell line, but not T, B, or macrophage cell lines (1), these data suggest that B-NK is the first nonmammalian NK receptor to be reported. We also show that, like LLT1, B-lec is an early activation Ag, which is up-regulated in response to PMA/CHX. Interestingly, B-lec shows the closest sequence identity of any nonmammalian C-type lectin to a specific mammalian NKC-encoded receptor described to date. The observation that B-NK and B-lec are more closely related to particular mammalian NKC-encoded receptors than to other chicken C-type lectins (Fig. 4) suggests that subgroups of CTLD-containing receptors appeared before the divergence of the avian and mammalian lineages some 330 million years ago.

A subgroup of C-type lectin-like receptors has been described previously (16), and consists of the Lox-1, Dectin-1, and CLEC receptors, which are expressed on monocytes, DCs, and endothelial cells. The results presented here have allowed us to expand this observation, using comparative information to organize the group V lectins into four distinct subgroups (Fig. 4). The first subgroup contains the NKR-P1 receptors, KLRF1 and B-NK, all of which are signaling molecules encoded by a gene containing a separate stalk exon expressed predominantly on NK cells. Moreover, their amino acid sequences are phylogenetically clustered. The second subgroup consists of early activation Ags including CD69, LLT1, and B-lec expressed on a wider range of lymphocytes. The third subgroup is the previously described subgroup found on DCs and endothelial cells. The fourth and final subgroup includes the Ly49 and NKG2D/CD94 signaling receptors, which are expressed predominantly on NK and T cells. Moreover, most of the receptors in this subgroup are known to bind MHC class I Ags or related molecules.

The organization of the group V CTLDs into subgroups based on comparison of the chicken and mammalian sequences is particularly useful in creating a hypothesis for the evolution of the NKC. The first and second subgroup receptor genes are found

![FIGURE 5](http://www.jimmunol.org/)

**FIGURE 5.** Expression and function of B-NK and B-lec. A, B-lec and B-NK RNA levels were determined using real-time quantitative RT-PCR in different tissues isolated from a 12-wk-old male chicken. Ct values for B-lec and B-NK were normalized using 28S Ct values. Fold changes were calculated in relative units away from the normalized sample containing the lowest amount of target mRNA, which is kidney for B-lec and proventriculus for B-NK. B, rB-NK was expressed in UG9 cells, which were then stimulated with pervanadate for the indicated time. Following stimulation, B-NK was immunoprecipitated using an anti-V5 mAb and Western blotted. Blots were screened with mAbs to anti-V5, anti-phosphotyrosine, anti-SHP-1, or anti-SHP-2. Molecular mass is indicated on the left-hand side. C, B-lec (left-hand side) and B-NK (right-hand side) RNA levels in cells stimulated with either with PMA alone (square; scale on the left axis), or PMA and CHX together (circle; scale on the right axis) measured using real-time quantitative RT-PCR. Ct values for B-NK and B-lec were normalized using 28S Ct values, and fold changes were calculated away from the unstimulated sample at each time point.

![FIGURE 6](http://www.jimmunol.org/)

**FIGURE 6.** Organization of B-NK and B-lec in the chicken MHC is the same as the organization of the most closely related genes in the mammalian NKC. Genes are shown as arrows representing orientation of transcription, but are not drawn to scale. Organization for human and mouse NKC determined from Ensembl Genome Browser (http://www.ensembl.org/). Unfilled arrows indicate B-NK and other subgroup 1 genes, black filled arrows represent B-lec and other subgroup 2 genes.
intermingled with each other, often arranged in pairs consisting of a subgroup 1 and a subgroup 2 receptor transcribed in opposite orientation (Fig. 6). This is seen in its most simple form in the chicken MHC, with B-NK and B-leu located next to each other in opposite orientation. This pattern can also be observed in the human NKC, where there are two subgroup 1–2 gene pairs next to each other, suggesting that one pair arose from the other in a relatively recent duplication event. In the mouse, the evolutionary history is more difficult to follow, due to the expansion in gene number that is such a key feature of the murine NKC, but the pattern can still be detected in at least two instances (Fig. 6). The fact that this organization is maintained from chicken through mammals would suggest that the subgroup 1–2 gene pair arrangement predates the divergence of the avian and mammalian lineages.

In contrast to the well-maintained organization of the subgroup 1–2 gene pair, the location of the pairs is strikingly different in mammals and chickens. In chickens, the subgroup 1–2 gene pair is found in the MHC, whereas in mammals, no C-type lectin-like genes are found in the MHC, but instead are in an entirely separate region, the NKC. There are two obvious possibilities for the location of C-type lectin-like genes in the chicken but not mammalian MHC. One possibility is that the MHC and NKC genes were originally independent loci, and that some C-type lectin-like genes were translocated to the MHC in the lineage leading to birds but not in the lineage leading to mammals.

The more likely possibility is that C-type lectin genes were originally part of the MHC, but were lost from the MHC in the lineage that led to mammals (Fig. 7). The simplest explanation is that the ancestral MHC contained MHC genes and C-type lectin genes. This ancestral MHC region is thought to have undergone genomewide duplication events during the emergence of the jawed vertebrates, leading eventually to the four MHC-paralogous regions on human chromosomes 1, 6, 9, and 19 (9). Thus, in the immediate descendents after the genome duplications, the C-type lectin genes would have been present in all the MHC-paralogous regions. In the different species that followed, the paralogous regions would have been free to evolve independently, with subsequent gene loss, multigene family expansion and contraction, translocation, and other large-scale events. We envisage the presence of subtype 1–2 gene pairs in at least two of the paralogous regions of some common ancestor of birds and mammals, with these genes deleted from the MHC but maintained in the NKC in lineage leading to mammals, but maintained in both the MHC and at a reduced level in the NKC of birds.

There are two independent pieces of evidence to support this view. First, there are group II C-type lectin genes located in the paralogous region of human chromosome 19, consistent with the idea that C-type lectin genes were present in the ancestral MHC (12). Second, the human NKC on chromosome 12p12–13 contains several genes that are homologs of genes in the MHC, including a tapasin-related gene and the complement component C3/C4/C5-related α1-macroglobulin gene (9). Taken together, this suggests that the NKC originated as part of a paralogous region, as proposed by Casahara (9), and is consistent with the idea that C-type lectin genes, including homologs of the subtype 1 and subtype 2 genes, were present in an ancestral MHC.

A region syntenic with the mammalian NKC has been identified in the chicken (17). However, the sequencing of the chicken genome has identified only two C-type lectin-like genes in this region (http://www.ensembl.org/). S. Rogers and J. Kaufman, unpublished data). The absence of an extended cluster of C-type lectin-like genes in this region suggests that the NKC has a very different function in the chicken immune response. For mammals, it is believed that the large number of C-type lectin-like receptors encoded by the NKC has arisen to provide a finely tuned detection system for abnormal expression levels of cell surface markers. The ligands for some of the NKC-encoded receptors are known to be MHC class I Ags, of which as many as six different classical molecules may be expressed by mammalian cells. One explanation for the lower number of C-type lectin-like genes in the chicken NKC may be the lower number of different MHC class I molecules expressed by chicken cells, with potentially as few as one class I molecule expressed (2). Moreover, the close linkage of two C-type lectin-like receptors to their potential ligands in the “minimal essential MHC” allows for receptors and ligands to coevolve as haplotypes in the chicken. This would remove the selective pressure on the NKC that is observed in mammals, in which many polymorphic ligands drive rapid expansion and diversity in the unlinked NKC-encoded receptors.

One method of evaluating our hypothesis that C-type lectin-like receptors should be considered members of the MHC-paralogous gene families is to examine their location in other species. To date, at least three C-type lectin genes with some homology to NK receptors have been reported in teleost fish (18–20), but neither the orthology to particular receptors nor location with respect to the MHC is clear. Searching the pufferfish (Fugu rubripes) and zebrafish (Danio rerio) genomes has not revealed the presence of C-type lectin genes near the MHC regions (S. Rogers and J. Kaufman, unpublished data), but the fragmented nature of the published

![FIGURE 7.](http://www.jimmunol.org/) Hypothetical model of evolution of the chicken MHC and the mammalian NKC from a common ancestral MHC-like region. C-type lectin genes are shown as triangles, with up-pointing triangles representing ancestral CRD, down-pointing triangles representing group II CRD, and side-pointing triangles representing group V CTLDs. Group V CTLDs are filled to represent the subgroups discussed in the text, with no fill representing subgroup 1, filled representing subgroup 2, vertical hatch representing subgroup 3, and horizontal hatch representing subgroup 4. One gene is shown in each subgroup for clarity, although there may be multiple genes in each subgroup. Only three other MHC-paralogous gene families are shown in the figure, and are denoted C* for the α2M/C3/C4/C5 family, TAPBP for the TAPBP (tapasin) family, and MHC for MHC class I family.
sequences makes it impossible to say whether C-type lectin-like receptors are present in these regions. Moreover, the presence of several unlinked regions containing MHC-related genes in teleost fish (which is believed to be a derived characteristic) may make it more likely that the C-type lectin genes will have lost their physical association with MHC regions. Finally, a receptor with reported homology to CD94/NKR-P1 has been identified in the urochordate Botryllus schlosseri (21), but again there is no clear homology to any particular receptor, and there is nothing known about the genetic location with respect to a putative ancestral MHC region.

In this study, we have reported the characterization of the first nonmammalian receptors with clear homology to specific receptors encoded in the mammalian NKC. Moreover, the location of chicken B-NK and B-lec in the MHC region has led us to suggest that this location has been maintained during the evolution of the chicken B-F/B-L region as a “minimal essential MHC” to allow coevolution of NK receptors with their predicted ligands, the MHC class I molecules. However, in the absence of experimental evidence for this hypothesis, alternative ligands for B-NK and B-lec have to be considered. The fact that some murine NKR-P1 receptors have recently been shown to bind to the Clr receptors (22, 23) may suggest that their closest chicken homologs, B-NK and Blec, may also bind to each other. Work is currently ongoing in our laboratory to elucidate the ligands for B-NK and B-lec.

In conclusion, we believe that these observations show that the primordial MHC likely contained both ancestral NK and ancestral MHC (class I) genes. This may suggest that an (or even the) original function of MHC molecules was to be recognized by NK cells in an innate role rather than by T cells in an adaptive role.

Acknowledgments

We thank Drs. Pete Kaiser and Adrian Smith for critical review of the manuscript, Dr. Colin Butter for graphical assistance, Dr. Pete Kaiser and Lisa Rothwell for technical assistance and for supplying RNA and the reagents used in the TaqMan analysis, and a reviewer for suggested additional experiments.

Disclosures

The authors have no financial conflict of interest.

References


