THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 10, Issue of March 9, pp. 7312–7319, 2001

Printed in U.S.A.

Crystal Structure of the C-type Lectin-like Domain from the Human Hematopoietic Cell Receptor CD69*

Received for publication, September 19, 2000, and in revised form, October 17, 2000
Published, JBC Papers in Press, October 17, 2000, DOI 10.1074/jbc.M008573200

Andrea S. Llera‡, Fernando Viedma§, Francisco Sánchez-Madrid§, and José Tormo¶

From the Department of Macromolecular Structure, Centro Nacional de Biotecnología, Universidad Autónoma de Madrid, 28049 Madrid, Spain, ‡Instituto de Estudios de la Inmunidad Humoral, Consejo Nacional de Investigaciones Científicas y Técnicas-Universidad de Buenos Aires, Junín 956, 1113 Buenos Aires, Argentina, and §Servicio de Inmunología, Hospital de la Princesa, Universidad Autónoma de Madrid, 28006 Madrid, Spain

CD69, one of the earliest specific antigens acquired during lymphoid activation, acts as a signal-transducing receptor involved in cellular activation events, including proliferation and induction of specific genes. CD69 belongs to a family of receptors that modulate the immune response and whose genes are clustered in the natural killer (NK) gene complex. The extracellular portion of these receptors represent a subfamily of C-type lectin-like domains (CTLDs), which are divergent from true C-type lectins and are referred to as NK-cell domains (NKDs). We have determined the three-dimensional structure of human CD69 NKD in two different crystal forms. CD69 NKD adopts the canonical CTLD fold but lacks the features involved in Ca^{2+} and carbohydrate binding by C-type lectins. CD69 NKD dimerizes noncovalently, both in solution and in crystalline state. The dimer interface consists of a hydrophobic, loosely packed core, surrounded by polar interactions, including an interdomain β sheet. The intersubunit core shows certain structural plasticity that may facilitate conformational rearrangements for binding to ligands. The surface equivalent to the binding site of other members of the CTLD superfamily reveals a hydrophobic patch surrounded by conserved charged residues that probably constitutes the CD69 ligand-binding site.

The C-type lectin-like domain (CTLD)[1] is a conserved protein module that has been found in numerous proteins with a wide range of functions (1). This fold was initially identified in a group of C-type (Ca^{2+}-dependent) animal lectins (sugar-binding proteins) that mediate both pathogen recognition and cell-cell communication by means of protein-carbohydrate interactions (2). Based on amino acid sequence comparisons, multidomain arrangement, and functional criteria the extensive superfamily of CTLDs has been subdivided in a variety of groups (3, 4). Among them, those that are predicted to bind sugars through coordination to a conserved Ca^{2+} ion are also known as carbohydrate-recognition domains (CRDs). The structure of several CRDs has been solved, and their binding to sugars has been thoroughly characterized (1, 5). Increasing evidence shows, however, that many of the modules that adopt a CTLD fold lack the Ca^{2+}-coordinating residues that mediate the classical C-type lectin sugar binding properties, suggesting that they may serve functions other than saccharide recognition (1).

The CTLD fold is widely represented among proteins that mediate the innate immune response (6). In particular, a conserved genomic region known as natural killer (NK) gene cluster (NKC) encodes for a group of receptors with CTLD-containing sequences that are involved in modulation of NK-cell activity and natural host defense (6, 7). Most of these proteins are type II transmembrane receptors, usually expressed as disulfide-linked homo- or heterodimers. Each subunit comprises a single extracellular CTLD, specifically named NK domain (NKD), connected by a neck region of variable length to a single membrane-spanning region and a short intracellular N-terminal portion (reviewed in Refs. 8 and 9). The cytoplasmic regions are often involved in signaling by recruitment, through specific sequence motifs, of kinases and phosphatases. Alternatively, some activating receptors, bearing short cytoplasmic tails devoid of any specific signaling sequence, associate noncovalently with membrane-anchored signaling molecules (10). Some receptors encoded in the NKC, such as the CD94/NKG2 heterodimers and the mouse Ly49 family, have proved to bind molecules of the major histocompatibility complex (MHC) (11, 12), but for many other receptors ligands have not yet been identified.

CD69, one of the first described members of the NKC family of receptors (13–15), is present at the cell surface as a disulfide-linked homodimer, with subunits of 28 and 32 kDa resulting from the differential glycosylation at a single extracellular *N*-linked glycosylation site (see Ref. 16 and reviewed in Ref. 17). Contrary to other NKC gene products, whose expression is restricted to NK cells, CD69 has been found in the surface of most hematopoietic lineages (reviewed in Ref. 18). It is one of the earliest markers induced upon activation in T and B lymphocytes, NK cells, macrophages, neutrophils, and eosinophils. In addition, it is constitutively expressed on monocytes, platelets, Langerhans cells, and a small percentage of resident lymphocytes in thymus and secondary lymphoid tissues (19). CD69 is also present on B cell precursors in the bone-marrow, and recent studies with CD69-deficient mice revealed its modulatory role on B cell development and antibody synthesis (20).

* This work was supported by the Spanish Dirección General de Enseñanza Superior (PB96-0271) and Fundación Antorchas, Argentina. 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.

The atomic coordinates and structure factors (code 1E87 and 1E8I) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

¶ To whom correspondence should be addressed: Centro Nacional de Biotecnología, Campus de la Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain. Tel.: 34-915854917; Fax: 34-915854506; E-mail: jtormo@cnb.uam.es.

[1] The abbreviations used are: CTLD(s), C-type lectin-like domain(s); NK, natural killer; NKD(s), NK-cell receptor domain(s); NKC, NK gene cluster; CRD(s), carbohydrate-recognition domain(s); MHC, major histocompatibility complex; MBP, mannose-binding protein; r.m.s., root mean square; DTT, dithiothreitol; MES, 4-morpholinoethanesulfonic acid.

It has been demonstrated that CD69 acts as a signal-transmitting receptor. Its cytoplasmic portion is constitutively phosphorylated on serine residues (21, 22). Even when the actual ligand that triggers this receptor is not known, cross-linking of CD69 by specific antibodies activates the extracellular signal-regulated kinase signaling pathway (23) and has been shown to induce rise in intracellular calcium concentration, synthesis of different cytokines and/or proliferation (24–28), and target lysis in interleukin-2-activated NK cells (22). In summary, CD69 wide distribution, along with its activating signal-transducing properties, suggest an important role of this receptor in the physiology of leukocyte activation.

As a first step toward better understanding the structural basis for CD69 function, we have undertaken the production and analysis of soluble constructs of the CD69 extracellular region. We report here the three-dimensional structure of human CD69 NKD determined in two crystal forms. We compare it with other known CTLD structures, specifically those present in NKC receptors, describe its dimeric oligomerization, and suggest a putative ligand-binding site.

EXPERIMENTAL PROCEDURES

Protein Expression, Refolding, and Purification—The complete NKD of human CD69 (residues 82–199) was amplified from cDNA (14) by polymerase chain reaction and subcloned into *Nde*I-*Bam*HI restriction sites of pET 26b plasmid (Novagen) (29) by double digestion and subsequent ligation. Restriction sites and a TAA termination codon were added to the insert using the following primers: CD695s, 5′GCGCGC-GCGCATATGGTTTCTTCATGCTCTG; CD693, 5′GCGCGCGGATCCT-TATTTGTAAGGTTTG. Automatic DNA sequencing (ABI PRISM, PerkinElmer Life Sciences) of the cloned 1insert using T7 promoter primer rendered the correct sequence.

The resulting plasmid was transformed into *Escherichia coli* strain BL21(DE3), and these cells were grown in LB medium at 37 °C until the A_{600} reached 0.7 cm^{-1} and then induced by addition of isopropyl-1-thio-β-D-galactopyranoside to 0.5 mM. After 4 h, cells were harvested and resuspended in Tris-HCl, pH 8.0, 0.2 M NaCl, 5 mM EDTA, 5 mM DTT. They were lysed by adding lysozyme to a final concentration of 1 mg/ml, and the viscosity was reduced by sonication. The protein was obtained as insoluble aggregates forming inclusion bodies, which were extensively washed three times in 50 mM Tris-HCl, pH 8.0, 0.1 M NaCl, 1 mM EDTA, 1 mM DTT, 0.5% (v/v) Triton X-100 and once in 50 mM Tris-HCl, pH 8.0, 2 M NaCl, 1 M urea, 1 mM EDTA, 1 mM DTT. The protein was solubilized in 25 mM MES, pH 6.5, 8 M urea, 10 mM EDTA, 1 mM DTT, and insoluble material was discarded by centrifugation. The CD69 NKD was refolded by the method of dilution of denaturing conditions following a modification of the protocol originally described for MHC class I molecules (30). Urea-solubilized CD69 NKD was added by slow dilution to 1 liter of 0.1 M Tris-HCl, pH 8.5, 400 mM L-arginine, 2 mM EDTA, 6.3 mM cysteamine, 3.7 mM cysteamine, 0.1 mM phenylmethylsulfonyl fluoride. Repeated pulses were added every 12 h. After 36 h, the refolding mixture was concentrated under nitrogen to a volume of 5 ml and purified by gel filtration chromatography in 25 mM Tris-HCl, pH 7.5, 100 mM NaCl, 0.1 mM EDTA on a Superdex 200 column (Amersham Pharmacia Biotech). In these conditions, CD69 NKD elutes as a non-covalent dimer, although the peak shows a slight asymmetry, consistent with an equilibrium between the monomeric and dimeric forms. Correctly sized fractions were further purified by cation exchange chromatography using a Mono S column (Amersham Pharmacia Biotech). Protein was loaded in 25 mM Tris-HCl, pH 7.0 and eluted with a linear gradient to 500 mM NaCl in the same buffer.

Crystallization and Data Collection—For crystallization, the protein was buffer-exchanged into 15 mM Hepes, pH 7.2, 50 mM NaCl and concentrated to 5 mg/ml. Crystals were obtained by mixing aliquots of the protein solution with an equal volume of the reservoir solution containing 0.1 M sodium acetate buffer, pH 4.8, 150 mM zinc sulfate or sodium sulfate, and 15% polyethylene glycol 6000. Two different crystal forms grew in these conditions. Long and thin prisms of square section, which belong to the tetragonal system, predominated, whereas larger irregular crystals, occasionally with a triangular shape and that belong to the trigonal system, appeared infrequently. The tetragonal crystals belong to the space group $P4_32_12$ with unit cell dimensions a = b = 69.6 Å and c = 118.6 Å and contain one dimer in the asymmetric unit. The trigonal crystals belong to the space group $P3_121$ with unit cell dimensions a = b = 48.4 Å, c = 119.9 Å, a = β = 90°, and γ = 120° and contain a monomer in the asymmetric unit. For cryogenic data collection, crystals were harvested in a modified reservoir solution, transferred to harvest buffer containing 15% glycerol or ethylene glycol, and flash-cooled by plunging into liquid propane. The high resolution data sets used for the structure refinement of both crystal forms were collected at beam line ID14-3 of the European Synchrotron Radiation Facility in Grenoble, France, using a MarResearch charge-coupled device, whereas the initial data set in the tetragonal form used for the structure determination was collected at beam line ID14–2 using an ADSC Quantum4 charge-coupled device. Data were integrated, scaled, and merged with the HKL package (31) (see Table I).

Structure Determination and Refinement—The CD69 structure was determined in the tetragonal crystal form, which became available first. The structure was solved by molecular replacement using the AMoRe package (32) with truncated coordinates of the CD94 dimer (33) as the search model. Structure factors from 15.0 to 5.0 Å were used for the rotation and translation functions. Model phases were improved and extended from 5.0 to 2.6 Å by iterative cycles of density modification in DM (34), which consisted of solvent flattening and 2-fold averaging. The resulting electron density maps allowed unambiguous building of the molecule, including extensive portions, like helix α2, not present in the initial truncated model.

Crystallographic refinement was carried out in CNS (35) using standard procedures that included a bulk solvent correction and overall anisotropic scaling. Automatic refinement, employing the maximum likelihood amplitude target, was alternated with manual rebuilding in the graphics program O (36) using both averaged and unaveraged σ_A-weighted (37) $2F_o - F_c$, $F_o - F_c$, and omit electron density maps. Tight noncrystallographic symmetry restrains were initially applied to all regions except the flexible loops involved in lattice contacts. When the high resolution data set for this crystal form became available, the noncrystallographic symmetry restrains were gradually relaxed based on the behavior of the R_{free}. All regions of CD69 NKD are well ordered, with the exception of the tip of the β2-β2′ hairpin, one residue at the N terminus, and the last two residues at the C terminus, which show poor density and high B factors. The model contains residues from 83 to 199 for both subunits, and 69 solvent atoms. The present R_{crys} is 24.8, and R_{free} is 27.0 for all data (F > 0) between 25.0 and 1.95 Å.

The trigonal crystal form was readily solved using a partially refined model from the tetragonal crystal form and was similarly refined in CNS. Electron density maps clearly showed two possible conformations for residues at the carboxy end of helix α2. Therefore, residues 133 to 136 were modeled in two alternate conformations with half-occupancy. All residues, including the β2-β2′ hairpin, showed in the electron density maps. The model contains residues from 83 to 199 for both subunits, and 115 solvent atoms. The present R_{crys} is 22.9%, and R_{free} is 24.4% for all data (F>0) between 20.0 and 1.50 Å. Refinement statistics for both crystal forms are given in Table I.

Structure Analysis—Structure superpositions were done with SHP (38). Solvent-accessible surface areas were determined with NACCESS (39) using a probe radius of 1.4 Å and default atom radii (40), and cavity volumes were determined with SURFNET (41). Calculation of hydrogen bonds was carried out with HBPLUS (42) using the program's default values. Figures were produced with GRASP (43) and BOBSCRIPT (44) and rendered with RASTER3D (45).

Alignments—The human CD69-NKD sequence was aligned to representatives of the NKD family using ClustalW at ExPASy on the Internet and were subsequently edited manually based on the known structures of rat mannose-binding protein (MBP)-A (PDB accession code 1ytt), mouse Ly49A (1qo3), and human CD94 (1b6e). Sequences were retrieved from GenBank™ (hCD69, NP_001772.1; hMAFA-L, AAC32200; hAICL, NP_005118; hLLT1, NP_037401; hNKRP1A, NP_002249; hKLRF1, AAF37804; hCLEC2, AAF36777.1; hCD94, NP_002253; hNKG2A, NP_002250.1; mLy49A, I49361) and Swiss-Prot (mCD69, P37217; rMBP-A, P19999). The alignment figure was drawn using ESPript (46).

RESULTS AND DISCUSSION

Structure Determination and Quality of the Models—Human CD69 consists of a 40-residue intracellular domain, a 21-residue transmembrane region, and an extracellular portion that comprises a 20-residue neck and an NKD of 118 amino acids. Soluble forms of CD69 NKD, comprising residues 82 to 199, were prepared by *in vitro* refolding from material expressed in *E. coli*. Recombinant CD69 NKD retains binding to a panel of

TABLE I
Data collection and refinement statistics for human CD69

Data collection			
Data sets	$P4_32_12$	$P4_32_12$	$P3_121$
Resolution range[a] (Å)	50.0–2.30 (2.35–2.30)	50.0–1.95 (2.00–1.95)	50.0–1.50 (1.53–1.50)
Measurements	80597	98431	151066
Unique reflections	13929	21836	26542
Completeness (%)	98.7 (98.4)	99.1 (99.2)	98.6 (100.0)
I/σ(I)	20.2 (4.0)	27.5 (5.4)	33.5 (15.6)
R_{merge} (%)	6.8 (31.1)	6.1 (20.2)	6.2 (10.6)
Refinement			
Resolution range		25.0–1.95	20.0–1.50
R_{cryst}/R_{free}		24.8/27.0	22.9/24.4
Reflections			
working set		19826 (90.0%)	24561 (91.3%)
test set		2010 (9.1%)	1961 (7.3%)
Number of copies in asymmetric units		2	1
Number of nonhydrogen protein atoms[b]		1926	956 [54]
Number of solvent atoms		69	115
r.m.s. deviations from ideality			
bond length (Å)		0.007	0.009
bond angle (°)		1.41	1.57
bonded B factors (Å^2)		2.27	1.82
Ramachandran plot[c]			
most favored (%)		90	92.4
allowed (%)		10	7.6
generously allowed (%)		0	0
disallowed (%)		0	0

[a] Numbers in parentheses correspond to the highest resolution shell.
[b] Numbers in brackets represent atoms in dual conformation.
[c] As calculated by PROCHECK (60).

specific monoclonal antibodies recognizing four distinct epitopes (16), and it behaves as a noncovalent dimer during gel filtration (results not shown).

The structure of CD69 NKD was determined by molecular replacement in two different crystal forms, tetragonal and trigonal, and refined to 1.95 and 1.50 Å, respectively. The quality of the diffraction data and refinement statistics are given in Table I. The tetragonal crystal form contains two molecules in the asymmetric unit, related by a molecular 2-fold axis. The electron density is continuous in the final $2F_o - F_c$ map from residues 83 to 199, except for the exposed β2-β2′ hairpin that shows weak density and appears to be disordered (see below for a description of secondary structure elements). In the trigonal crystal, the CD69 NKD dimer is crystallographic, and there is a single molecule in the asymmetric unit. In this crystal form, all residues (83 to 199), including the β2-β2′ hairpin, are in good, continuous density in the final electron density maps. The C-terminal end of helix α2 shows static disorder and has been modeled in two alternate conformations in the trigonal form.

Pairwise superpositions of the three independent copies for CD69 NKD give r.m.s. deviations from 0.27 to 0.40 Å for main chain atoms between residues 89 and 198. The largest differences are focused in the domain N terminus, around the β turn (residues 86–89), which precedes strand β0, and the β2-β2′ hairpin. In the tetragonal form, this flexible hairpin is exposed to the solvent, whereas in the trigonal crystal it is better ordered because of crystal packing interactions. Description of the structure is based on the high resolution trigonal model unless specifically stated.

The CD69 NKD Fold—The α-carbon trace of CD69 NKD is shown in Fig. 1*A*. As predicted by its amino acid sequence, the overall structure of CD69 NKD displays the salient features of the CTLD fold. The domain, with overall dimensions of 44 × 32 × 30 Å, consists of two connected antiparallel β sheets and two α helices, like in the CRD of C-type animal lectins (1) and mouse Ly49A NKD (47) (Fig. 1*B*). Strand β2 acts as a connection between the two β sheets formed by strands β0, β1, β5, β2, and β1′ in the lower part of the molecule (following the standard view for CTLD folds as shown in Fig. 1) and strands β2′, β2, β3, and β4 in the upper part. This portion of the molecule is also characterized by a long stretch, connecting strands β2′ and β3, lacking regular secondary structure. This region corresponds to the Ca^{2+}-binding site of true C-type lectins. The two helices, α1 and α2, are located one on each side of the extended β structure.

There are three intrachain disulfide bonds in CD69 NKD (Fig. 1), two of which (Cys^{113}–Cys^{194} and Cys^{173}–Cys^{186}) correspond to the characteristic invariant disulfides found in all CTLDs. The third disulfide bond, Cys^{85}–Cys^{96}, connects a loop at the N terminus, which precedes the first β strand (β0), with strand β1 by linking two cysteines separated by 10 residues in this segment. This disulfide is only found in long-form C-type lectins (including lithostathine, tetranectin, and factors IX/X-binding protein) (4) and appears to be present in most NKD domains (Fig. 2). In members of the rodent Ly49 gene family, however, these two bonded cysteines are separated by only four residues and are located on contiguous β strands (47) (Fig. 1*B*).

Despite the low sequence identity between CD69 NKD and other CTLDs, which ranges from around 20% with CRDs of animal lectins to almost 30% for other NKDs, the overall structure of the domain is highly conserved. Superposition of CD69 with other CTLDs gives rise to r.m.s. deviations of between 1.2 Å for 100 equivalent Cα atoms with Ly49A and 1.4 Å for 95 equivalences with rat MBP (48). The major differences among them occur at the N terminus, the position of helix α2, hairpin β2-β2′, the loop connecting it to strand β3, and hairpin β3-β4 (Fig. 1*B*). These regions coincide with segments displaying higher variability in length and amino acid sequence among members of the NKD family (Fig. 2). Overall, CD69 NKD is more similar to long-form CTLDs, like tetranectin or lithostathine, and to other NKDs, like Ly49A. Although based on sequence identity the NKD structure closer to CD69 is that of the CD94 subunit of the NK-cell receptor CD94/NKG2 (around 28% identical), their superposition gives an r.m.s. deviation of 1.4 Å for 97 equivalent Cα positions. This relatively poorer match is mostly due to differences between helix α2 in CD69 and the equivalent region in CD94. In the crystal structure of

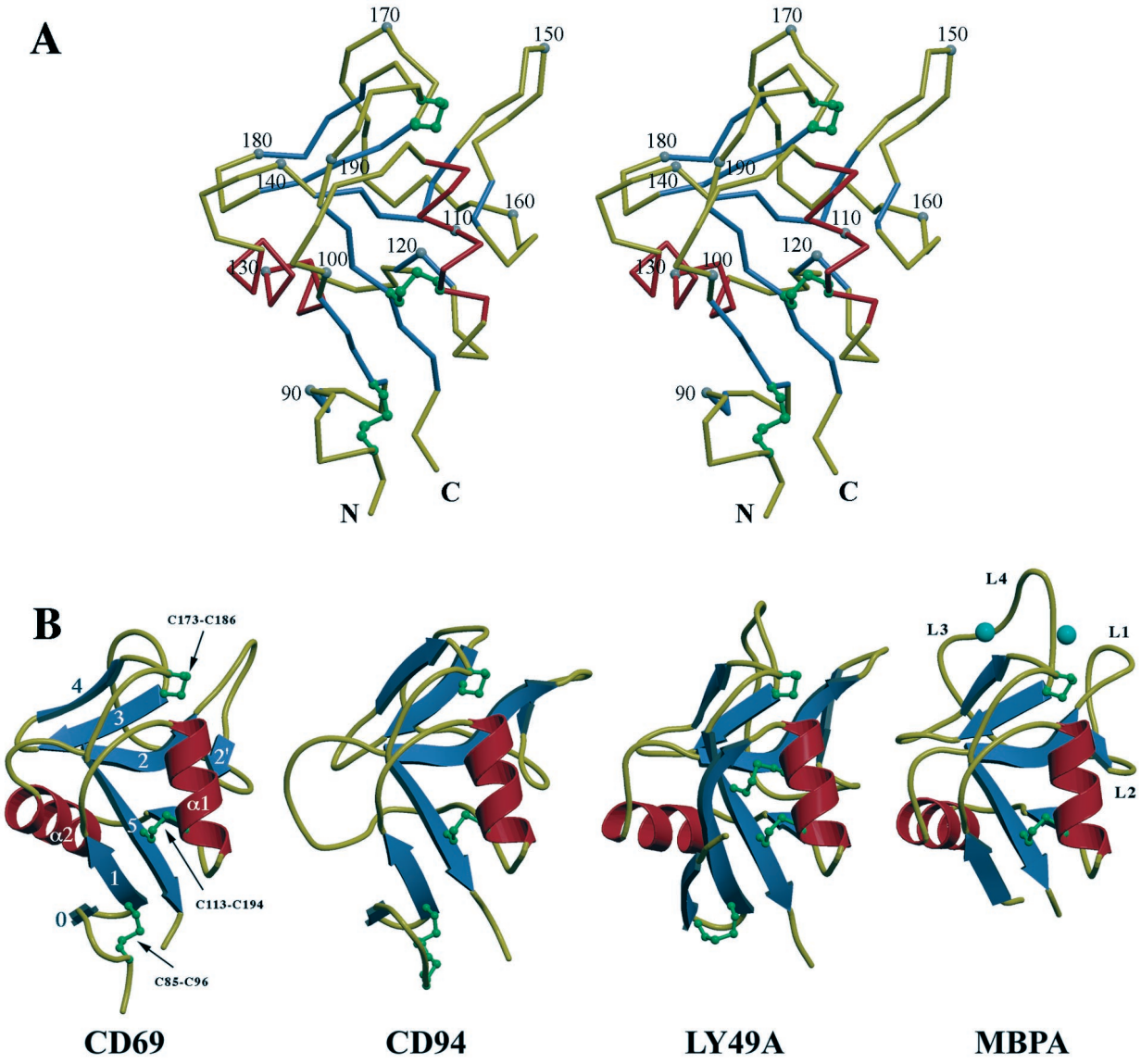

FIG. 1. **Structure of human CD69 and comparison with other members of the NK-cell receptor family and the CRD of rat MBP.** *A*, stereoview of the α-carbon trace of CD69 NKD. Every tenth residue is represented as a *gray sphere* and numbered according to its order in the full-length protein. Segments are colored based on secondary structure assignments, as calculated with program DSSP (59). The α helices are shown in *red*, β strands are shown in *blue*, and the rest of the residues are shown in *yellow*. Disulfide bonds are represented by *green ball-and-stick* models. The N- and C termini of the domain, which are less than 7 Å apart, are labeled. *B*, fold of CD69 NKD. Comparison with other members of the NK-cell receptor family and the CRD of rat MBP-A as a prototype of animal lectins is shown. *Ribbon diagrams* of the NKDs from human CD69, human CD94 (PDB entry 1b6e), mouse Ly49A (1qo3), and the CRD of rat MBP-A (1ytt) are shown in a common orientation obtained by pairwise superpositions. Ca^{2+} ions bound to MBP-A are shown as *light blue spheres*, whereas its loop regions without regular secondary structure, three of them involved in Ca^{2+} coordination, are labeled *L1* to *L4*. Secondary structure elements in CD69 have been labeled following the numbering for MBP-A, prototype of the family. Therefore, the first β strand, which is absent in MBP-A and is characteristic of the long-form CTLDs, has been labeled as β0, whereas the strand that forms a β-hairpin with strand β2 has been named β2′. Differences in the orientation of helix α2, which is absent in CD94, are evident.

the CD94 NKD homodimer, this helix is replaced by a loop that is involved in the dimerization interface (33). CD94 forms, with distinct members of the NKG2 family, heterodimers that are involved in the recognition of the nonclassical MHC class I molecule HLA-E (49–51). The unraveling of this α helix in the crystal structure of CD94 NKD could be because of the formation of the homodimer, whose physiological role is uncertain. The amino acid sequence at this region retains a distribution of hydrophobic residues that appears suitable for the formation of an α helix, but analysis of sequence alignments reveals that, in CD94 and NKG2 proteins, this segment is two residues shorter than in the majority of NKD sequences. Whether the lack in CD94 of this α helix is due to the formation of the homodimer or consequence of the deletion at the C-terminal end of this segment would have to await the determination of the structure of a CD94/NKG2 heterodimer.

CD69 and other NKDs differ substantially from canonical C-type lectins in that they lack the Ca^{2+}-binding site that is critical for carbohydrate recognition in CRDs. The departure of the equivalent regions in different NKDs from the canonical Ca^{2+}-binding site appears gradual. In Ly49A, three side chains are in a position similar to that of Ca^{2+}-coordinating residues, whereas in CD94 two appear in an equivalent position. CD69 shows a deletion of five residues that constitute loops 3 and 4 in the CRDs (Fig. 1*B*), both structures involved in Ca^{2+} and carbohydrate binding, and only one residue (Asp^{171}) is located

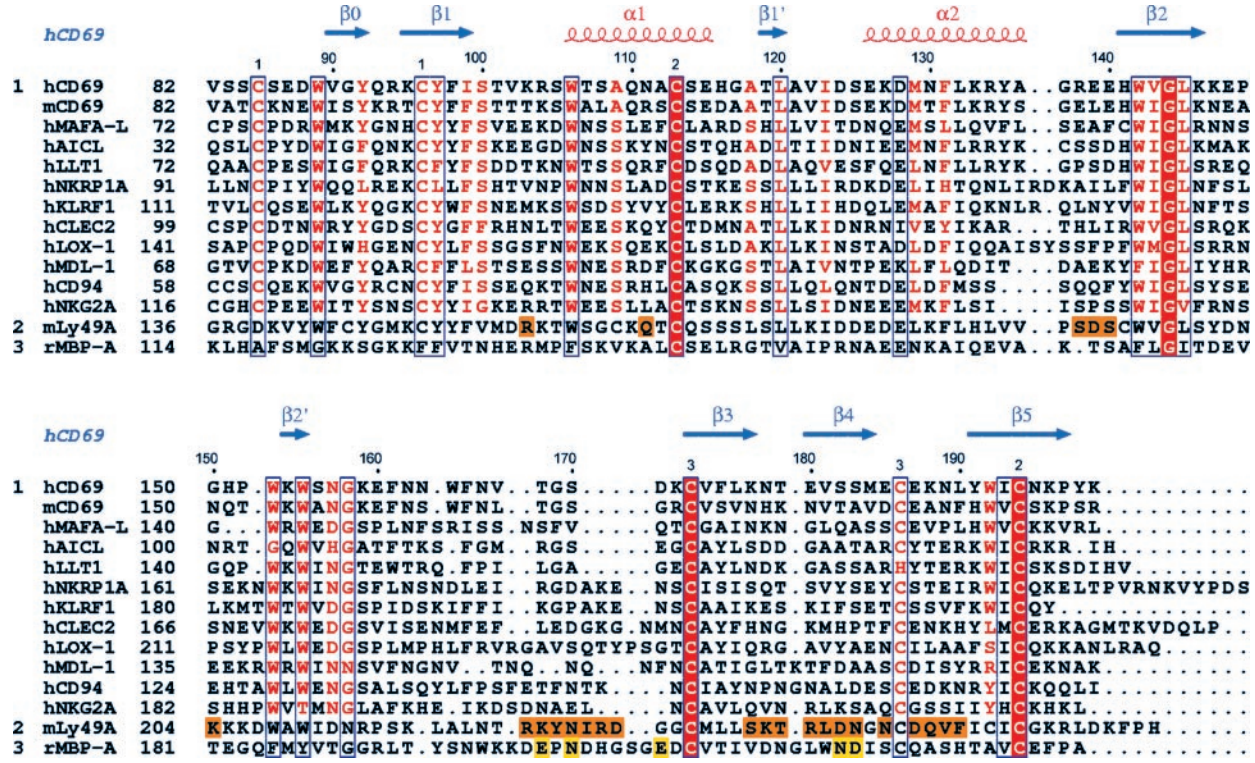

FIG. 2. **Structure-based sequence alignment of NKDs for CD69, selected NK-cell receptors, and the rat MBP-A CRD.** NK-cell receptors have been arranged in two groups. The first group contains human CD69, its mouse homologue, and other human C-type lectin-like receptors encoded in the NK complex, which, based on their amino acid sequences, do not appear to use carbohydrates as ligands. The second group contains the mouse Ly49A receptor, also encoded in the NK complex, whereas the third group includes the prototype animal lectin, rat MBP-A. Residues strictly conserved are shown by *white characters on a red background*. They include the four cysteines involved in the two disulfide bridges conserved in the C-type lectin-like superfamily and the glycine residue where strand $\beta 2$ switches from one β sheet to another. Residues well conserved within the first NKD group are depicted by *red letters*, and the rest are in *black*. A *blue frame* denotes similarity across groups. The secondary structural elements and every tenth residue for CD69 are displayed over the sequences. Helices are represented by *red squiggles*, and β strands are represented by *blue arrows*. The paired numbers (1 to 3) correspond to the bonded cysteines in CD69. Ly49A residues labeled as *orange boxes* correspond to those involved in molecular contacts with its ligand, the mouse MHC class I molecule H-2D^d, as observed in the crystal structure of their complex (47). Rat MBP-A residues enclosed in *yellow boxes* indicate those whose side chains coordinate the Ca^{2+} ion involved in carbohydrate binding (48).

in a position roughly equivalent to any of the side chains involved in the coordination of the Ca^{2+} ion in C-type lectins. Furthermore, this aspartate is not conserved and is replaced by a glycine in mouse CD69. The extensive alteration of the putative binding site for carbohydrates in CD69 argues against its binding to sugars, at least using the same mechanism as C-type lectins. This is consistent with a recent report showing binding neither to monosaccharides nor to various polysaccharides but only a weak binding signal to fucoidan (52).

The CD69 NKD Dimer—CD69, along with other receptors encoded in the NKC, is a dimeric type II membrane protein. Most receptors in this group appear to be covalent dimers through the formation of one or a few interchain disulfide bridges in the neck, a region that connects the NKD to the single membrane-spanning segment and presents high variability among members of the family. The extracellular part of CD69 encompasses a neck of around 20 amino acids (residues 62–84), which contains the cysteine (Cys^{68}) involved in the interchain disulfide bridge, and the globular NKD. The recombinant CD69 NKD used in this work forms noncovalent dimers in both crystal forms analyzed. Furthermore, it behaves as a dimer during size-exclusion chromatography. In the trigonal crystal, the dimer is crystallographic, whereas in the tetragonal form there are two CD69 copies in the asymmetric unit related by a molecular 2-fold axis. The two dimers are almost identical, with an r.m.s. deviation of 0.35 Å for their main chain atoms. This difference is in the same range as that between the two independent copies in the asymmetric unit of the tetragonal crystal form (0.4 Å).

The dimer has overall dimensions of $30 \times 35 \times 70$ Å and is stabilized by both intermolecular hydrogen bonds and hydrophobic interactions (Figs. 3 and 4). An intermolecular β sheet extending through both subunits is formed through antiparallel pairing of their N-terminal β strands ($\beta 0$). At the ends of the paired β strands and over them, additional polar interactions are established. The main-chain carbonyl of Glu^{87}, located in the β turn preceding chain $\beta 0$, forms two hydrogen bonds with main-chain and side-chain nitrogen atoms of Gln^{83} at the β turn connecting strands $\beta 0$ and $\beta 1$. Asp^{88}, in the same β turn as Glu^{87}, establishes a salt bridge with Lys^{127}, located at the N terminus of helix $\alpha 2$. At the domain N terminus, and over the intermolecular β sheet, Ser^{84} from both subunits associate with each other through several hydrogen bonds. Although these interactions, between Ser^{84} at the N termini of both subunits, are seen in the two crystal forms, it cannot be excluded that they are a consequence of the truncated domain used for these crystallographic studies. Interestingly, in a group of recently described receptors expressed on the surface of human macrophage and dendritic cells, which appear to have CTLDs with a functional Ca^{2+}-binding site, Ser^{84} is replaced by a cysteine residue (53–55). Based on the structure of CD69, where the side chains of Ser^{84} are hydrogen bonded through their hydroxyl groups, it seems probable that the equivalent cysteine in those receptors could well be involved in an interchain disulfide bridge.

Beneath the intersubunit β sheet, the two subunits associate through the juxtaposition of the C-terminal half of helix $\alpha 2$ (Fig. 4). A hydrophobic core is formed by side chains projecting

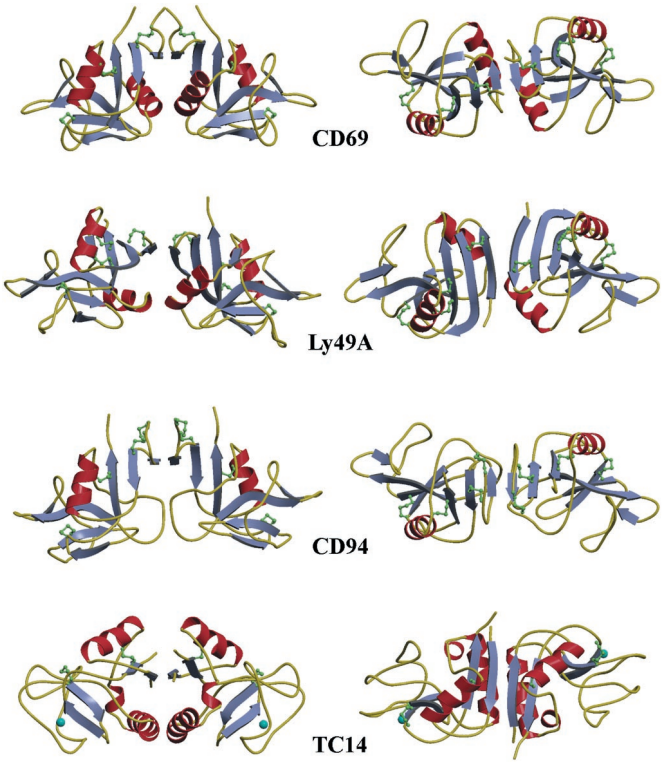

FIG. 3. **The CD69 dimer and its comparison to other dimeric CTLDs.** *Ribbon diagrams* showing two views of dimers for CD69, CD94, Ly49A, and the tunicate lectin TC14 are displayed. In the *left* view, the molecular 2-fold axes run along the vertical. The *right* view is rotated 90° from the left one around the horizontal axis, so the molecular 2-fold is perpendicular to the plane of the figure. Secondary structure elements, as defined by the program DSSP (59), are shown in *blue* for β strands and in *red* for α helices. The disulfide bridges are represented by *green ball-and-stick* models, and the Ca^{2+} ions in TC14 are represented as *blue spheres*. The first three dimers (CD69, CD94, and Ly49A) are all NKDs encoded in the NK gene cluster. For CD94 and Ly49A, their left view has been aligned to that of CD69 by superimposing only the subunit on the right, so differences in dimerization are shown as different orientations of their left subunits. The dimerization arrangements in NKDs are very similar but with some significant differences. The CD94 dimer is the closest to CD69, although the segment of CD94 corresponding to helix α2 in other CTLDs does not form a helix in the homodimer. The Ly49A dimer, which was determined in complex with the mouse MHC class I molecule $H\text{-}2D^d$, does not follow strict 2-fold symmetry.

from helix α2 (Phe^{131}, Tyr^{135}) and from the bottom of the intermolecular β sheet (Val^{90}, Tyr^{92}). The side chain of Arg^{134}, also on helix α2, extends across the dimer interface and forms a hydrogen bond with the carbonyl of Tyr^{135} at the carboxy end of the same helix on the neighboring subunit; this interaction probably stabilizes the helix dipole. Underneath the hydrophobic core there are contacts between Tyr^{138}, in the loop connecting helix α2 to strand β2, and Asn^{178} and Thr^{179}, both at the β turn between strands β3 and β4. The packing at the hydrophobic core is loose, and there is an interdomain cavity of 75.5 $Å^3$ between the aromatic side chains of Phe^{131} and Tyr^{135} (Fig. 4*B*). In the high resolution trigonal form, there is static disorder at the C terminus of helix α2, and residues 133 to 136 show two alternate conformations (Fig. 5). This disorder appears to be triggered by a displacement of the side chain of Tyr^{135} to fill the empty cavity at the dimer interface. In one conformation, Tyr^{135} fills a cavity formed by residues Val^{90}, Tyr^{97}, Ile^{99}, Phe^{131}, Ile^{132}, and Ile^{193} from its own subunit. In the second conformation, Tyr^{135}, together with the backbone of neighboring residues, has been displaced to pack against residues Phe^{131}, Tyr^{135}, and Arg^{134} of the symmetry-related subunit. The r.m.s. difference for Tyr^{135} in the two conformations is 2.8 Å. This alternate conformation cannot occur simultaneously in both subunits, because it would give rise to steric clashes around the 2-fold axis. Although these dual conformations are not observed in the tetragonal crystal, perhaps because of its more limited resolution, disorder in this region is manifested as a diffuse electron density and higher thermal factors for the side chain of Tyr^{135}, which are also consistent with certain structural variability at the interface. The residues involved in the formation of this interdomain cavity, especially the two aromatic side chains, are conserved in other members of this family (Figs. 2 and 3), suggesting that this feature may also be present in other receptors. Plasticity at the dimer interface was also observed in Ly49A when bound to its MHC class I ligand (Fig. 3). In the complex, the Ly49A subunits are not related by a strict 2-fold axis, probably because of packing restrictions in the simultaneous binding to different sites on two MHC molecules (47). There is a shift in the position of the 2-fold axis with respect to the N-terminal β strand that results in a slightly different arrangement of the subunits so that juxtaposed α helices do not run roughly parallel but are approximately aligned. This departure from dyad symmetry is accompanied by structural differences between several residues at the interface.

These packing defects and flexibility at the dimer interface of NKDs can provide a structural mechanism to allow conformational rearrangements to facilitate the simultaneous binding of the two subunits of the dimer to oligomeric or different ligands. Protein cavities at domain interfaces have been suggested to facilitate such movements (56). Furthermore, the flexible neck region would allow various orientations of the NKD domains with respect to the cell membrane, whereas the intermolecular disulfide bridge would ensure that subunits would not dissociate.

The dimer arrangement in CD69 is very similar to that observed for the CD94 crystallographic dimer, even though helix α2 has been replaced by a loop in the latter. After superposition of one subunit of both dimers, the other two can be superimposed after a rotation of only 6°. In CD69, the total buried surface area between the subunits is 1673 $Å^2$. Atoms buried at the interface are 48% nonpolar, and there are twelve hydrogen bonds and two salt bridges. The CD94 dimer buries a similar overall surface (1224 $Å^2$), but the percentage of nonpolar atoms involved is higher (61%), and there are only 2 hydrogen bonds. Ly49A, partly because of the absence of an N-terminal loop preceding strand β0, buries a smaller surface (930 Å), with only 3 hydrogen bonds. The dimer from the tunicate lectin TC14, an example of a naturally dimeric C-type lectin, buries 1687 $Å^2$ and forms 10 hydrogen bonds, very similar values to those of CD69, but the atoms buried at the interface are more hydrophobic (73%). The formation of dimers in TC14 also occurs between helix α2 and the N-terminal β strand. However, the arrangement of the subunits is different, because it involves much more extensive contacts between the sides of the helices, which are arranged side by side, and uses strand β1 instead of β0, because TC14 is a short-form lectin. Therefore, although the overall arrangement of the subunits in these dimeric CTLDs is similar, the detailed interactions involved in oligomerization are different among the various receptors. These differences could affect the structural plasticity and stability of the dimers.

The Putative Ligand-binding Site—In all NKD structures known to date, the apposition of the two monomers generates a relatively flat, continuous surface that is located opposite to the N and C termini, and that includes the regions equivalent to the carbohydrate-binding site of CRDs at opposite ends (Figs. 3 and 6). This is one of the most variable regions, both in struc-

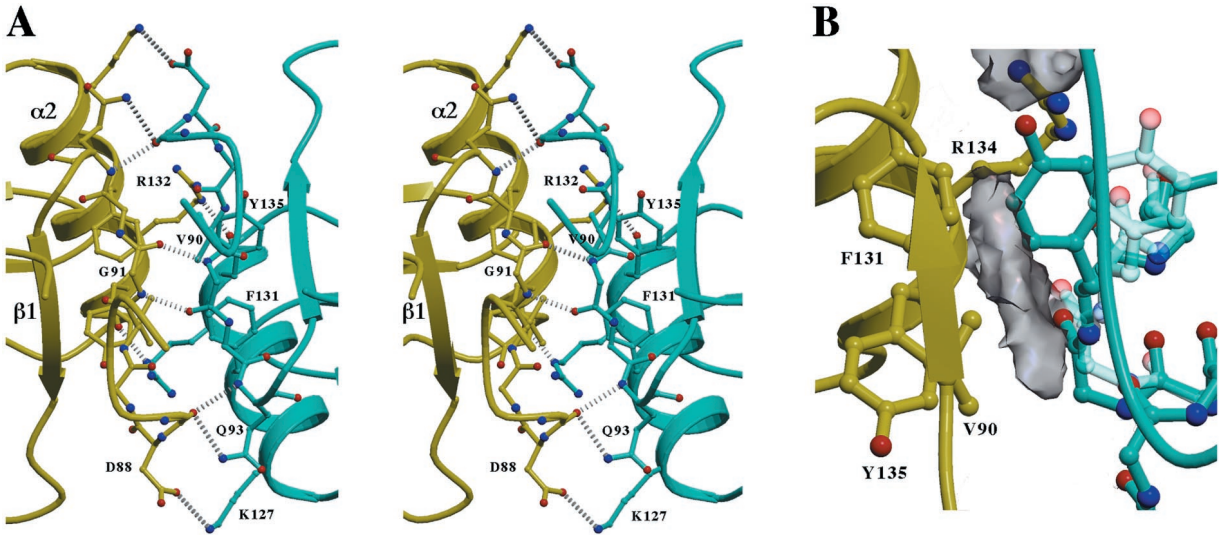

FIG. 4. **The dimerization interface in CD69.** *A*, stereoview of the dimerization interface in CD69. The two subunits of the dimer are shown as *ribbon diagrams* in different colors, *yellow* and *green*. The backbone of strand β0 and side chains involved in hydrogen bonds or hydrophobic interactions are represented by *ball-and-stick models*. Hydrogen bonds are shown as *white broken lines*. *B*, cavity and static disorder at the dimer interface. A close-up of the hydrophobic cluster at the dimer interface is shown. The solvent-accessible surface and the intersubunit cavity, as calculated with program SURFNET (41), are shown as a *semitransparent* surface. For the *right* subunit, the two alternate conformations for Tyr^{135} side chain and the backbone atoms of surrounding residues are shown in different colors, *solid green* and *semitransparent green*. In the *green* subunit, strand β0 is shown as a *coil* instead of an *arrow* for clarity.

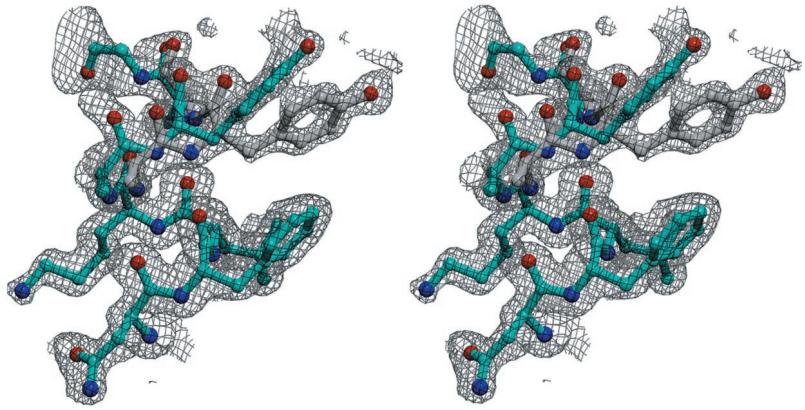

FIG. 5. **Electron density map for helix α2.** A 1.5-Å resolution electron density map calculated with σ_A-weighted (37) $2F_o - F_c$ amplitudes and model phases for residues Asn^{130} to Gly^{137} in helix α2 is shown together with the final refined model. Static disorder in the C-terminal half of the helix, because of rearrangements at the dimer interface, is evident, especially for the side chain of Tyr^{135}. The two alternate conformations modeled for residues 133 to 136 are shown in *cyan* and *white*, respectively.

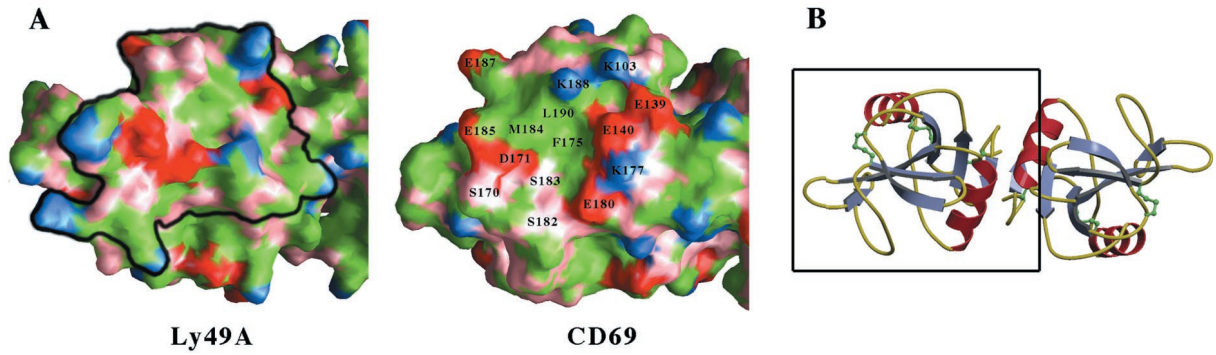

FIG. 6. **Surface analysis of the hypothetical ligand-binding site in CD69.** *A*, representation of the solvent-accessible surfaces of the ligand-binding site in Ly49A (*left*) and the equivalent region in CD69 (*right*). Surfaces have been colored based on the nature of the underlying atoms (carbons and sulfurs in *green*, polar nitrogens and oxygens in *pink*, charged nitrogens in *dark blue*, and charged oxygens in *red*). For Ly49A, the footprinting of its ligand, the mouse MHC class I molecule $H\text{-}2D^d$, has been contoured with a *black line*. In CD69, relevant residues in the putative ligand-binding site have been labeled. *B*, overview of the CD69 dimer with the region highlighted in *panel A* enclosed in a *black box*.

ture and amino acid sequence, in the NKDs (Fig. 2), and this variability is consistent with this part of the molecule being involved in selective binding to ligands by receptors bearing these domains. For Ly49A, this region has indeed been shown to contain the binding site for its ligand, MHC class I molecules (47). This same region has been postulated to constitute the

ligand-binding site of the CD94/NKG2 heterodimers (33).

Analysis of the amino acid conservations between human and mouse CD69 that, at the same time, differ significantly from other CTLDs, specially those of the NKD family, also pinpoints this region as functionally relevant, supporting the suggestion that this region is the potential ligand-binding site (57). In CD69, this surface is characterized by a central hydrophobic patch (side chains of Phe^{175}, Met^{184}, and Leu^{190} and the aliphatic portions of Glu^{185} and Lys^{188}) surrounded by polar side chains, among which acidic residues (Glu^{140}, Asp^{171}, Glu^{180}, Glu^{185}, and Glu^{187}) predominate (Fig. 6). The topology and chemical nature of this putative binding region appear more similar to a protein-protein recognition site than to a carbohydrate binding surface (5, 58). This observation is consistent with recent data for recombinant CD69 that showed no binding signals to the various monosaccharides tested (52).

In summary, this work provides the first structural insights on human CD69 and constitutes a firm basis for future analysis of the interactions between this receptor and its ligand(s), once they are identified. The structure reveals that, whereas the molecule presents the overall characteristics of the CTLD fold, it completely lacks the structural features responsible for carbohydrate binding. Rather, the putative binding site is more consistent with this receptor having a proteinaceous ligand. The identification of a buried cavity at the dimer interface and the conservation of several key residues involved in its formation suggest the possibility that this feature, which may confer flexibility at the dimer, could be also present in other receptors of the family.

Acknowledgments—We thank F. Torrents for computing facilities, J. Navaza for valuable advice, J. Colom for assistance during purification, and the laboratories of M. Coll and I. Fita for help and support. We are grateful to the staff at European Synchrotron Radiation Facility (ESRF) beam lines ID14-2 and 14-3. We also thank the European Molecular Biology Laboratory Grenoble Outstation for providing support for measurements at ESRF under the European Union 'Access to Research Infrastructures' action of the Improving Human Potential Program.

REFERENCES

1. Drickamer, K. (1999) *Curr. Opin. Struct. Biol.* **9,** 585–590
2. Drickamer, K., and Taylor, M. E. (1993) *Annu. Rev. Cell Biol.* **9,** 237–264
3. Drickamer, K. (1993) *Curr. Opin. Struct. Biol.* **3,** 393–400
4. Day, A. J. (1994) *Biochem. Soc. Trans.* **22,** 83–88
5. Weis, W. I., and Drickamer, K. (1996) *Annu. Rev. Biochem.* **65,** 441–473
6. Weis, W. I., Taylor, M. E., and Drickamer, K. (1998) *Immunol. Rev.* **163,** 19–34
7. Brown, M. G., Scalzo, A. A., Matsumoto, K., and Yokoyama, W. M. (1997) *Immunol. Rev.* **155,** 53–65
8. Lanier, L. L. (1998) *Annu. Rev. Immunol.* **16,** 359–393
9. López-Botet, M., and Bellón, T. (1999) *Curr. Opin. Immunol.* **11,** 301–307
10. Bléry, M., Olcese, L., and Vivier, E. (2000) *Hum. Immunol.* **61,** 51–64
11. Ugolini, S., and Vivier, E. (2000) *Curr. Opin. Immunol.* **12,** 295–300
12. López-Botet, M., Llano, M., Navarro, F., and Bellón, T. (2000) *Semin. Immunol.* **12,** 109–119
13. Hamann, J., Fiebig, H., and Strauss, M. (1993) *J. Immunol.* **150,** 4920–4927
14. López-Cabrera, M., Santis, A. G., Fernández-Ruiz, E., Blacher, R., Esch, F., Sánchez-Mateos, P., and Sánchez-Madrid, F. (1993) *J. Exp. Med.* **178,** 537–547
15. Ziegler, S. F., Levin, S. D., Johnson, L., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Baker, E., Sutherland, G. R., Feldhaus, A. L., and Ramsdell, F. (1994) *J. Immunol.* **152,** 1228–1236
16. Sánchez-Mateos, P., and Sánchez-Madrid, F. (1991) *Eur. J. Immunol.* **21,** 2317–2325
17. Testi, R., D'Ambrosio, D., De Maria, R., and Santoni, A. (1994) *Immunol. Today* **15,** 479–483
18. Marzio, R., Mauel, J., and Betz-Corradin, S. (1999) *Immunopharmacol. Immunotoxicol.* **21,** 565–582
19. Sánchez-Mateos, P., Cebrián, M., Acevedo, A., López-Botet, M., De Landázuri, M. O., and Sánchez-Madrid, F. (1989) *Immunology* **68,** 72–79
20. Lauzurica, P., Sancho, D., Torres, M., Albella, B., Marazuela, M., Merino, T., Bueren, J. A., Martínez,-A, C., and Sánchez-Madrid, F. (2000) *Blood* **95,** 2312–2320
21. Testi, R., Phillips, J. H., and Lanier, L. L. (1988) *J. Immunol.* **141,** 2557–2563
22. Lanier, L. L., Buck, D. W., Rhodes, L., Ding, A., Evans, E., Barney, C., and Phillips, J. H. (1988) *J. Exp. Med.* **167,** 1572–1585
23. Zingoni, A., Palmieri, G., Morrone, S., Carretero, M., López-Botet, M., Piccoli, M., Frati, L., and Santoni, A. (2000) *Eur. J. Immunol.* **30,** 644–651
24. Cebrián, M., Yague, E., Rincón, M., López-Botet, M., De Landázuri, M. O., and Sánchez-Madrid, F. (1988) *J. Exp. Med.* **168,** 1621–1637
25. Nakamura, S., Sung, S. S., Bjorndahl, J. M., and Fu, S. M. (1989) *J. Exp. Med.* **169,** 677–689
26. Testi, R., Phillips, J. H., and Lanier, L. L. (1989) *J. Immunol.* **143,** 1123–1128
27. Santis, A. G., Campanero, M. R., Alonso, J. L., Tugores, A., Alonso, M. A., Yague, E., Pivel, J. P., and Sánchez-Madrid, F. (1992) *Eur. J. Immunol.* **22,** 1253–1259
28. De Maria, R., Cifone, M. G., Trotta, R., Rippo, M. R., Festuccia, C., Santoni, A., and Testi, R. (1994) *J. Exp. Med.* **180,** 1999–2004
29. Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990) *Methods Enzymol.* **185,** 60–89
30. Garboczi, D. N., Hung, D. T., and Wiley, D. C. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 3429–3433
31. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
32. Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50,** 157–163
33. Boyington, J. C., Riaz, A. N., Patamawenu, A., Coligan, J. E., Brooks, A. G., and Sun, P. D. (1999) *Immunity* **10,** 75–82
34. Cowtan, K. D. (1994) *Joint CCP4 ESF-EACMB Newslett.* **31,** 34
35. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 905–921
36. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
37. Read, R. J. (1986) *Acta Crystallogr. Sect. A* **42,** 140–149
38. Stuart, D. I., Levine, M., Muirhead, H., and Stammers, D. K. (1979) *J. Mol. Biol.* **134,** 109–142
39. Hubbard, S. J., and Thornton, J. M. (1993) *NACCESS*, Department of Biochemistry and Molecular Biology, University College, London
40. Chothia, C. (1976) *J. Mol. Biol.* **105,** 1–12
41. Laskowski, R. A. (1995) *J. Mol. Graph.* **13,** 323–330
42. McDonald, I. K., and Thornton, J. M. (1994) *J. Mol. Biol.* **238,** 777–793
43. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins* **11,** 281–296
44. Esnouf, R. M. (1997) *J. Mol. Graph.* **15,** 133–138
45. Merrit, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 869–873
46. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) *Bioinformatics* (*Oxf.*) **15,** 305–308
47. Tormo, J., Natarajan, K., Margulies, D. H., and Mariuzza, R. A. (1999) *Nature* **402,** 623–631
48. Weis, W. I., Kahn, R., Fourme, R., Drickamer, K., and Hendrickson, W. A. (1991) *Science* **254,** 1608–1615
49. Braud, V. M., Allan, D. S., O'Callaghan, C. A., Soderstrom, K., D'Andrea, A., Ogg, G. S., Lazetic, S., Young, N. T., Bell, J. I., Phillips, J. H., Lanier, L. L., and McMichael, A. J. (1998) *Nature* **391,** 795–799
50. Borrego, F., Ulbrecht, M., Weiss, E. H., Coligan, J. E., and Brooks, A. G. (1998) *J. Exp. Med.* **187,** 813–818
51. Lee, N., Llano, M., Carretero, M., Ishitani, A., Navarro, F., López-Botet, M., and Geraghty, D. E. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 5199–5204
52. Childs, R. A., Galustian, C., Lawson, A. M., Dougan, G., Benwell, K., Frankel, G., and Feizi, T. (1999) *Biochem. Biophys. Res. Commun.* **266,** 19–23
53. Suzuki, N., Yamamoto, K., Toyoshima, S., Osawa, T., and Irimura, T. (1996) *J. Immunol.* **156,** 128–135
54. Matsumoto, M., Tanaka, T., Kaisho, T., Sanjo, H., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., and Akira, S. (1999) *J. Immunol.* **163,** 5039–5048
55. Bates, E. E., Fournier, N., García, E., Valladeau, J., Durand, I., Pin, J. J., Zurawski, S. M., Patel, S., Abrams, J. S., Lebecque, S., Garrone, P., and Saeland, S. (1999) *J. Immunol.* **163,** 1973–1983
56. Hubbard, S. J., and Argos, P. (1996) *J. Mol. Biol.* **261,** 289–300
57. Bajorath, J., and Aruffo, A. (1994) *J. Biol. Chem.* **269,** 32457–32463
58. Lo Conte, L., Chothia, C., and Janin, J. (1999) *J. Mol. Biol.* **285,** 2177–2198
59. Kabsch, W., and Sander, C. (1983) *Biopolymers* **22,** 2577–2637
60. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291