Conservation of a CD1 Multigene Family in the Guinea Pig

Christopher C. Dascher, Kenji Hiromatsu, Jerome W. Naylor, Pamela P. Brauer, Kara A. Brown, James R. Storey, Samuel M. Behar, Ernest S. Kawasaki, Steven A. Porcelli, Michael B. Brenner and Kenneth P. LeClair

*J Immunol* 1999; 163:5478-5488; ;
http://www.jimmunol.org/content/163/10/5478

---

**References**
This article **cites 51 articles**, 23 of which you can access for free at:
http://www.jimmunol.org/content/163/10/5478.full#ref-list-1

**Subscription**
Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

**Permissions**
Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**
Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
Conservation of a CD1 Multigene Family in the Guinea Pig

Christopher C. Dascher,* Kenji Hiromatsu,* Jerome W. Naylor,† Pamela P. Brauer,* Steven A. Porcelli,* Michael B. Brenner,* and Kenneth P. LeClair†‡

CD1 is a family of cell-surface molecules capable of presenting microbial lipid Ags to specific T cells. Here we describe the CD1 gene family of the guinea pig (Cavia porcellus). Eight distinct cDNA clones corresponding to CD1 transcripts were isolated from a guinea pig thymocyte cDNA library and completely sequenced. The guinea pig CD1 proteins predicted by translation of the cDNAs included four that can be classified as homologues of human CD1b, three that were homologues of human CD1c, and a single CD1d homologue. These guinea pig CD1 protein sequences contain conserved amino acid residues and hydrophobic domains within the putative Ag binding pocket. A mAb specific for human CD1b cross-reacted with multiple guinea pig CD1 isoforms, thus allowing direct analysis of the structure and expression of at least a subset of guinea pig CD1 proteins. Cell-surface expression of CD1 was detected on cortical thymocytes, dermal dendritic cells in the skin, follicular dendritic cells of lymph nodes, and in the B cell regions within the lymph nodes and spleen. CD1 proteins were also detected on a subset of PBMCs consistent with expression on circulating B cells. This distribution of CD1 staining in guinea pig tissues was thus similar to that seen in other mammals. These data provide the foundation for the development of the guinea pig as an animal model to study the in vivo function of CD1. *The Journal of Immunology, 1999, 163: 5478–5488.

R

cent structural and functional data have led to the identification of the CD1 family as a third lineage of Ag-presenting molecules distinct from those encoded by the MHC (reviewed in Ref. 1). CD1 proteins are expressed as heterodimeric cell-surface glycoproteins consisting of an ~45-kDa mature glycosylated H chain that is noncovalently associated with β2-microglobulin (β2m).7 CD1 H chains are type I integral membrane proteins that consist of three extracellular domains (α1, α2, and α3) similar in size to those of MHC class I, followed by a transmembrane segment and short cytoplasmic tail at the carboxyl terminus. Analysis of protein sequence data demonstrate that the CD1 family shares approximately equal homology with both MHC class I and class II. This suggests that CD1 diverged from a common ancestral gene for MHC proteins at a distant point in vertebrate evolution, possibly close to the time that MHC class I and class II diverged from each other. Consistent with this proposed distant evolutionary relationship to the MHC, genetic mapping has shown the CD1 locus in humans and mice to be unlinkd to the MHC in both species.

Human CD1 represents a small multigene family composed of five nonpolymorphic members, designated CD1a, -b, -c, -d, and -e. The five CD1 protein sequences are divided into two groups based on homology, with group 1 comprised of CD1a, -b, and -c and group 2 comprised of only CD1d (1, 2). The CD1e protein sequence has an intermediate homology between group 1 and 2 (2). Group 1 CD1 proteins are expressed at the surface of a wide range of APCs, including Langerhans cells, dendritic cells, B cells, and cytokine-activated monocytes (reviewed in Ref. 1). While the tissue distribution and gene sequence data for CD1 proteins have been known for some time, only recently has the function of CD1 been appreciated. Initial studies established a role for the human CD1b protein in the recognition of an Ag derived from Mycobacterium tuberculosis by a CD4+ CD8– T cell line (3). These studies have been extended by the characterization of additional CD1-restricted T cell lines that now include examples belonging to most of the major phenotypic T cell subsets, including CD4+ CD8–, CD4+ CD8+, and CD4–CD8– TCR-αβ+ T cells (3–6). In addition, γδ TCR+ T cells recognizing CD1 proteins have also been described (3, 7).

The first isolation of an Ag recognized by a CD1b-restricted human T cell line by Beckman et al. yielded the surprising finding that the M. tuberculosis Ag presented to these T cells was mycolic acid, an abundant mycobacterial cell wall lipid (8). A greater appreciation of the structural features of CD1 Ags is now emerging as the number of defined CD1-presented lipid and glycolipid Ags increases (9–12). All of the CD1-presented Ags identified to date have two hydrophobic lipid tails coupled to a charged or hydrophilic head group (e.g., a carboxylate, simple sugar, or oligosaccharide) (13, 14). Studies of T cell recognition of these Ags have shown exquisite specificity for the hydrophilic head group structure, but not for the hydrophobic lipid tails (10). This supports the prediction that the hydrophilic head groups of Ags bound to CD1...
proteins are exposed to the aqueous environment and thus accessible for direct recognition by TCRs. Recognition of the CD1-lipid complex by the TCR-αβ has recently been confirmed by TCR transfection experiments (15).

Substantial insight into the mechanism by which CD1 proteins can bind and present lipid Ags was recently provided by the elucidation of the crystal structure of the mouse CD1d1 protein (16). This revealed a remarkable resemblance in the overall three-dimensional structure of CD1 to that of MHC class I, with two anti-parallel α1 and α2 helices forming the sides of a putative Ag binding pocket and a β-sheeted sheet forming the floor of the pocket. However, compared with a typical MHC class I molecule, the α1 and α2 helices are closer together and more elevated above the β sheet platform. This creates an Ag binding pocket that is deeper and larger in volume than the MHC class I groove. Importantly, the CD1 Ag binding groove is lined primarily with nonpolar or hydrophobic amino acids, making its molecular surface electrostatically neutral with essentially no capacity for hydrogen bonding. These characteristics are consistent with its potential to bind hydrophobic lipids rather than peptides.

Whereas humans possess both group 1 (CD1a, -b, -c) and group 2 CD1 (CD1d) proteins, the genomes of muroid rodents (mice and rats) appear to contain only genes encoding homologues of the group 2 CD1d proteins (17, 18). Studies from our laboratory and others demonstrate that in vitro-derived CD1-restricted T cell lines specific for mycobacterial lipid Ags can be restricted by one of the group 1 CD1 proteins, either CD1a, CD1b, or CD1c (5, 8–10, 19). In contrast, no CD1d-restricted T cells have been demonstrated against bacterially derived Ags, although CD1d has been shown to present GPI-linked proteins derived from protozoal pathogens (11). Moreover, while most group 1 CD1-reactive T cells secrete Th1 cytokines, murine CD1d-reactive T cells have been shown to rapidly secrete high levels of IL-4 following activation and, therefore, may function more as immunoregulatory T cells than as effectors in the control of invading pathogens (4). Thus it is possible that group 1 and group 2 CD1 proteins have distinct roles in the host immune response, either with respect to the effector functions of the CD1-restricted T cells or with respect to the types of Ags presented by each group.

Investigation of the role of group 1 CD1 proteins in vivo requires an animal model in which group 1 CD1 proteins are expressed. The mouse, having only group 2 CD1, cannot be used as a model for the study of group 1 CD1 in vivo. Therefore, characterization of other potential animal models is needed. Here we report the cloning and analysis of the guinea pig CD1 gene family and demonstrate the expression of CD1 proteins by cells in a variety of lymphoid and nonlymphoid tissues. Our findings demonstrate that, unlike the muroid rodents, guinea pigs have an extended family of CD1 genes that includes genes encoding clear evolutionary homologues of the human CD1b, CD1c, and CD1e proteins. We propose that the guinea pig provides a uniquely relevant small animal model for in vivo studies of the functions of group 1 CD1 proteins in infectious disease, autoimmunity, and cancer.

Materials and Methods

Animals

Hartley guinea pigs were obtained from Charles River Breeder (Wilmington, MA). Strain 2 guinea pigs were obtained from the National Cancer Institute (Frederick, MD). Animals were housed under specific pathogen-free conditions at the Brigham and Women’s Hospital Animal Facility. All appropriate animal use protocols were obtained and followed.

Preparation of cells and tissues

Guinea pig whole blood was obtained by cardiac puncture of anesthetized animals. PBMC were isolated by Ficoll gradient as described previously (20). Thymus tissue was obtained from 3-wk-old guinea pigs. Other tissues were obtained from adult animals. Single-cell suspensions of spleen, thymus, and lymph nodes were generated by mincing tissue and passing through a no. 60 wire mesh screen. Cell suspensions were washed three times in RPMI 1640 medium (Life Technologies, Gaithersburg, MD) containing 10% FCS. The guinea pig cell lines used here were obtained from the American Type Culture Collection (Manassas, VA) and included 104C1 (CRL-1405), JH4 (CCL-158), and GP16 (CCL-242). These cells were grown in DMEM (11995-065; Life Technologies) with 10% FCS (HyClone, Logan, UT), 100 U/ml penicillin/streptomycin (15140-122; Life Technologies), 60 μM nonessential amino acids (11040-050; Life Technologies), and 0.8 mM t-glutamine (25030-081; Life Technologies). Cells were cultured in a humidified 10% CO2 incubator at 37°C.

Southern blot and hybridization analysis

Genomic DNA was isolated from the livers of Hartley or strain 2 guinea pigs using standard extraction techniques (21). Mouse and human genomic DNA were gifts from Drs. Christina Parker and Hamid Band, respectively. Southern blotting and hybridization were performed as described elsewhere (21). Briefly, genomic DNA was digested with restriction enzymes and electrophoresed on a 0.7% agarose gel, blotted onto Hybond-N membrane (Amersham Life Sciences, Little Chalfont, U.K.), and cross-linked by a 3-min exposure to UV light. The blot was prehybridized at 42°C for 2 h with prehybridization solution (5x SSC, 5x Denhardt, 50% formamide, 50 mM HEPES, pH 7.0, 0.5% SDS, and 160 μg/ml salmon sperm DNA). The probe was generated by PCR with 5′ primer (5′-AGTGAACATGGCCTTCAGGGCCACCC-3′) and 3′ primer (5′-GGGTGCGACAGGGATGATGTCCTGGCC-3′) to generate a product containing the nucleotides encoding the α1, α2, and α3 domains of the human CD1b cDNA. The CD1b probe was radioactively labeled with [α-32P]dCTP using the Rediprime random primer system (Amersham Life Sciences). Southern blots were hybridized with the CD1b probe overnight at 42°C. Blots were washed sequentially as described (22) and then exposed to X-Omat AR film (Eastman Kodak, Rochester, NY).

Cloning and sequencing of guinea pig CD1 genomic fragments and cDNA clones

Genomic DNA was used as the template for PCR using combinations of consensus oligonucleotide primers corresponding to the beginning and end of the CD1 α3-encoding genomic DNA exon. The four forward primers were 5′-CCCHGARGCTGTCGTTGGCC-3′, 5′-CCCHGARGCTGTCGTTGGCC-3′, 5′-CCANYCTGGGTYCTG GCC-3′, and the four reverse primers were 5′-CCTAGACTGCTGTYCTG GCC-3′, 5′-CCTAGACTGCTGTYCTG GCC-3′, 5′-CCTAGACTGCTGTYCTG GCC-3′, and 5′-CCTAGACTGCTGTYCTG GCC-3′. Ambiguity codes for these primers are B = C, T, or G; H = A, C, or T; M = A or C; N = C, A, or T; R = A or G; Y = C or T. Typical PCRs were performed with 266 nM concentrations of each primer and 1 μg genomic DNA template in a reaction buffer containing 1.5 mM MgCl2, 50 μM of each dNTP, and 1.25 U Taq polymerase for 30 cycles on a Perkin-Elmer thermal cycler (Norwalk, CT) with annealing temperatures between 45 and 55°C.

For isolation of CD1 cDNAs, total thymocyte RNA was prepared from Hartley or strain 2 guinea pig thymus tissue by homogenization and extraction with TRIzol as recommended by the manufacturer (Life Technologies). First-strand cDNA was reverse-transcribed from oligo(dT) or random-primed RNA templates using Superscript II reverse transcriptase (Life Technologies). Alternatively, mRNA was prepared by oligo(dT) priming of total RNA isolated from thymus tissue using a FastTrack 2.0 mRNA isolation kit (Invitrogen, Carlsbad, CA). Double-stranded cDNA libraries were prepared from thymus RNA template using a Marathon cDNA amplification kit (Clontech Laboratories, Palo Alto, CA). Marathon cDNA adapter ends containing AP1 and AP2 sequences were ligated to the double-stranded cDNAs to facilitate obtaining full-length CD1 cDNAs using rapid amplification of cDNA ends (RACE) methodologies. In this process, PCRs from thymus cDNA template were performed using a ppCD1 gene-specific internal primer paired with either the AP1- or the AP2-specific end primers.

PCR-derived genomic DNA and cDNA insert fragments were ligated into the pCR2.1 TA cloning vector and used to transform Escherichia coli strain INVaF competent cells (Invitrogen). Miniprep DNA was prepared using the Wizard DNA purification systems (Promega, Madison, WI).
Larger quantities of DNA were produced using endo-free plasmid maxi-prep kits (Qiagen, Valencia, CA).

DNA sequencing and data analyses
Sequencing of DNA inserts was performed enzymatically by the dideoxy method using the Sequenase version of T7 DNA polymerase (United States Biochemicals, Cleveland, OH) with 32P- or 35S-labeled deoxyribonucleotide incorporation. DNA sequencing reactions were resolved by electrophoresis on denaturing polyacrylamide gels followed by XAR-5 film (Eastman Kodak) autoradiography. Alternatively, automated DNA sequencing was performed on Applied Biosystems model 373 or 377 instruments at the Molecular Biology Core Facility of the Dana-Farber Cancer Institute (Boston, MA). Nucleotide and protein sequences obtained were compared with currently available sequences in the GenBank database using the BLAST program (23, 24). Alignment and dendrogram of CD1 protein sequences were made using the PILEUP program from the Wisconsin GCG software package (25).

Transfection of cells with a guinea pig CD1b3 gene
A cDNA for the gpCD1b3 gene was subcloned into the pcDNA3.1 eukaryotic expression vector (Invitrogen) and transfected into recipient cells using Lipofectamine Plus lipofection reagents and methodologies (Life Technologies). Briefly, subconfluent target cells in OptiMEM (Life Technologies) were treated with plasmid DNA and Lipofectamine Plus reagent for 3–5 h. Cells were then rinsed with and cultured in DMEM with 10% FCS (HyClone) overnight. To establish a stable gpCD1b3-expressing cell line, transfected cells were transferred into fresh DMEM with 10% FCS containing 0.5–0.75 mg/ml G418 (Life Technologies) for drug selection. Drug-resistant cells were cloned by limiting dilution and sterile sorted by flow cytometry using the human CD1b-specific mAb ABCDb3.1.

mAbs and FACS
The mAbs used in these studies include CT6 and CT7 (specific for guinea pig CD8 and CD4, respectively) (Serotec, Raleigh, NC); BCD1b3.1 (specific for human CD1b and cross-reactive with guinea pig CD1 proteins) has been described previously (26); PCA/188A was used as anti-CD3 (27). P3 was used as the nonspecific mouse IgG1 isotype control Ab (28). FACS analysis was performed as previously described (29). Briefly, primary Abs were added to single-cell suspensions at saturating concentrations for 1 h, washed with staining buffer (PBS containing 2% FCS and 0.01% azide), and then incubated for 1 h with 30 μg/ml FITC-conjugated donkey antimouse IgG (Jackson Immunologicals, West Grove, PA). After staining, cells were washed with staining buffer and analyzed with a FACS Sort flow cytometer (Becton Dickinson, Mountain View, CA). The forward and side scatter profiles of the analyzed cells (PBMC, splenocytes, and lymph node lymphocytes) were used to gate on the lymphocyte subpopulation. Dead cells were excluded using propidium iodide (Sigma, St. Louis, MO).

Cell-surface labeling and immunoprecipitation
Freshly isolated thymocytes (2 × 107) were cell-surface labeled with Na235S (DuPont-New England Nuclear, Boston, MA) using lactoperoxidase and hydrogen peroxide as described previously (30). The cells were solubilized in lysis buffer (50 mM Tris, pH 7.6, 140 mM NaCl (TBS) with 0.5% Nonidet P-40, 1% iodoacetamide, and 1 mM PMSF) for 1 h. After centrifugation to remove insoluble debris, the lysates were preclarified with 50 μl of a 10% suspension of Staphylococcus aureus Cowen strain I (Pansorbin, Calbiochem, La Jolla, CA). Preclarified lysates containing 106 cell equivalents were immunoprecipitated with ~1 μg of purified BCD1b3.1 mAb or 1 μl of P3, CT6, or CT7 ascites followed by incubation with 40 μl of a 10% suspension of Protein A-Sepharose (Pharmacia LKB Biotechnology, Uppsala, Sweden). The immunoprecipitates were washed five times with TBS containing 0.1% Triton X-100, then eluted with sample buffer and analyzed by SDS-PAGE using a 12% polyacrylamide gel under reducing conditions as described (31).

Immunohistochemistry
Tissue samples were mounted in OCT compound (Tissue-Tek, Torrance, CA), frozen in liquid nitrogen, and stored at −80°C. Frozen tissue sections (5 μm thick) were fixed in acetone for 10 min, air dried, and stained by an indirect immunoperoxidase method using avidin-biotin-peroxidase complex (Vector Laboratories, Burlingame, CA) and 3-aminobenzyltetrazolium (Sigma) as the chromogen.

Results
Presence of multiple CD1 genes in the guinea pig
The absence of group 1 CD1 proteins (CD1α, -b, and -c) in the mouse prompted us to investigate alternative small laboratory animal model systems for functional studies of group 1 CD1 proteins. Given its more distant relationship to muroid rodents (32, 33), we chose to investigate whether the guinea pig CD1 repertoire included group 1 CD1 homologues. Preliminary experiments to address this question were conducted using Southern blot analysis. Genomic DNA obtained from an outbred Hartley guinea pig was analyzed by Southern blot using a 32P-labeled probe generated from human CD1b cDNA containing the nucleotides encoding the α1–α3 exons. The α3 exon of CD1 is the most highly conserved among all known CD1 protein sequences and was included to maximize cross-species hybridization of the probe. The resulting autoradiogram revealed a minimum of 11 bands, suggesting that a relatively large CD1 family exists in the guinea pig (Fig. 1, lanes 1–3). Samples of mouse and human DNA were analyzed simultaneously and exhibited the expected banding pattern as described in published reports (34, 35). Five bands corresponding to the five CD1 genes were detected in the lanes containing human genomic DNA (Fig. 1, lanes 7–9). Two weakly cross-reactive bands, corresponding to the CD1D1 and CD1D2 genes, were observed with mouse genomic DNA (Fig. 1, lanes 4–6). The absence of non-specific bands in the human and mouse samples indicates that the hybridization and wash conditions employed were stringent enough to detect guinea pig CD1 genes, yet excluded other more distantly related sequences within the genome (e.g., MHC class I and class II). Subsequent Southern blot analysis of genomic DNA from another individual Hartley guinea pig and from the inbred strain 2 guinea pig resulted in an identical pattern of bands to that shown in Fig. 1 (data not shown). These data suggested that the
guinea pig genome contains an extended nonpolymorphic CD1 gene family with ~11–14 genes.

Cloning and sequencing of guinea pig CD1 α3 exon genomic DNA fragments

Molecular cloning and sequencing of multiple CD1 genes was undertaken to determine the diversity within the guinea pig CD1 family. To clone guinea pig CD1 genes, we took advantage of the high degree of nucleotide sequence conservation observed among CD1 genes, especially in the α3 exon. We aligned all of the published α3-encoding nucleotides and were able to identify multiple consensus sequences at the beginning and end of the α3 exon. Multiple PCRs were performed with pairs of these consensus, or slightly degenerate, synthetic primers using liver genomic DNA as a template to obtain guinea pig CD1 α3 DNA fragments (Fig. 2A). Sequencing of the individual genomic α3 fragments yielded 13 distinct CD1 α3 nucleotide sequences. The sequences of these three fragments revealed the presence of frame shifts, stop codons, or both within the α3 exon (data not shown). This suggests that the guinea pig genome contains CD1 genes that are incapable of encoding functional CD1 protein and are referred to here as pseudogenes. We have submitted these pseudogene sequences to GenBank (accession numbers AF178942, AF178943, and AF178944). The cloning of full-length cDNAs for these three CD1 pseudogenes was not performed. In contrast, the nucleotide and deduced amino acid sequences of the remaining 10 CD1 gene fragments were consistent with intact genes. Determination of the precise nucleotide sequences of these 10 CD1 gene fragments allowed us to design a strategy to isolate the full-length cDNAs.

Cloning and analysis of cDNAs encoding guinea pig CD1 proteins

It has been shown previously that all of the CD1 proteins are coexpressed on cortical thymocytes during thymic development in humans (36). We prepared adapter-ligated, doubled-stranded cDNAs from guinea pig thymocyte mRNA to serve as template for the isolation of guinea pig CD1 cDNAs. Using the α3 nucleotide sequences, we were able to specify gene-specific PCR primers for the isolation of corresponding full-length CD1 cDNA sequences using the RACE technique (Fig. 2, B and C). This strategy allowed us to isolate full-length cDNAs for the 10 guinea pig CD1 genes. Of these 10 genes, we identified two additional apparent CD1 pseudogenes (see below).

We then investigated the relationship between the guinea pig CD1 protein sequences and CD1 proteins from other species. Because the CD1 α1 and α2 domains form the putative Ag binding cleft, the CD1 protein isoforms are most divergent in these domains. Therefore, we focused on the amino acids comprising these two domains to evaluate the similarity of the guinea pig CD1 proteins to each other and to CD1 proteins from other species. Fig. 4
shows a dendrogram that compares the sequence similarity between the guinea pig CD1 proteins and those of several different animal species. These data revealed the clustering of the individual guinea pig CD1 isoforms with corresponding CD1 proteins from human and other species.

The percent amino acid identity between the guinea pig CD1b isoforms and the human CD1b protein sequence ranged from 57 to 68%, whereas the identity between the guinea pig CD1b group and guinea pig CD1c group members ranged from 41 to 54%. Therefore, members of specific isoform groups (CD1b or CD1c) from guinea pigs were more homologous to the human isoform counterpart than to other guinea pig isoform groups. Based on these criteria, we have assigned the individual guinea pig CD1 genes to various CD1 isoform groups. Henceforth, we refer to the guinea pig CD1 proteins as gpCD1 followed by the specific isoform (e.g., gpCD1b3). Using these criteria, the guinea pig thymocytes express...
mRNAs encoding four different CD1b proteins (gpCD1b1, -b2, -b3, -b4), three CD1c proteins (gpCD1c1, -c2, -c3), and one CD1e (gpCD1e) protein.

Isoforms corresponding to human CD1a and CD1d have not been isolated with the techniques described here even though several guinea pig tissues, including thymus, skin, and bone marrow-derived cell lines, were used as sources for mRNA. Evidence from Southern blotting analyses suggested that the guinea pig genome may lack both the CD1a and the CD1d isoforms. Identical blots of restriction enzyme-digested guinea pig genomic DNA were hybridized separately with probes corresponding to the CD1a, CD1b, or CD1d cDNAs, the gpCD1e cDNA has an additional 33 nucleotides encoding 11 amino acids precisely at the leader exon-exon 1 junction (Fig. 3). The mouse CD1d1 crystal structure indicates two disulfide bridges within the H chain of the CD1 molecule (16). One bridge is found between paired cysteine residues within the α2 domain (Fig. 3, consensus positions 131 and 195) and another between paired cysteines located within the α3 domain (Fig. 3, consensus positions 235 and 290). All of the guinea pig CD1 isoforms have conserved these paired cysteine residues, suggesting an importance in the stability of CD1 tertiary structure. In fact, the presence of these cysteine residues is conserved in all known CD1 protein sequences with the exception of mouse CD1d2, which substitutes tryptophan for cysteine at position 167 of the α2 domain (2).

The primary amino acid sequences of the guinea pig CD1 α and α2 domains contain a high proportion of hydrophobic amino acids, a common characteristic among CD1 proteins. The positions of hydrophobic regions within the α1 and α2 domains of the gpCD1b and gpCD1c groups is similar to that of human CD1b and CD1c, respectively. Moreover, structural modeling of the guinea pig CD1 proteins based on the mouse CD1d1 crystal structure reveals that the putative Ag binding pocket of the guinea pig CD1 proteins is lined primarily by nonpolar residues.

Human CD1b and CD1c have been shown to contain a tyrosine-based or YXXZ motif (where Y = tyrosine, X = any residue, Z = bulky hydrophobic residue) in their cytoplasmic tails that is important for their endosomal targeting. Human CD1a lacks this motif and remains primarily surface localized (M. Sugita and M. B. Brenner, unpublished observations), whereas CD1b possesses the YXXZ motif and shows strong steady-state localization to endosomes (37, 38). Interestingly, the gpCD1b3 cytoplasmic domain, like human CD1a, lacks the YXXZ motif (Fig. 3, consensus position 341–344). In addition, the gpCD1b3 protein sequence, while still grouping with CD1b, has a lower degree of homology to the other gpCD1b isoforms than those isoforms have to each other (Fig. 4). All other guinea pig CD1 proteins, with the exception of gpCD1e, possess the endosomal targeting motif.

**A guinea pig homologue of human CD1e**. The gpCD1e sequence provides the first description of a CD1e homologue in a species other than human. Both the human and guinea pig CD1e proteins fall outside the group 1 and group 2 classification scheme (1, 39). Northern analysis using the α1- and α2-encoding fragment of the gpCD1e cDNA as a probe revealed high levels of transcript expression in the guinea pig thymus (data not shown). The human CD1E gene is also transcribed and can be detected by Northern analysis, but no expression of a protein product has been described (39).

Two unusual features were observed in the guinea pig CD1e cDNA. First, compared with all of the other guinea pig CD1 cDNAs, the gpCD1e cDNA has an additional 33 nucleotides encoding 11 amino acids precisely at the leader exon-exon 1 junction (Fig. 3, consensus positions 19–29). Second, the gpCD1e protein has an extended cytoplasmic tail that lacks the YXXZ endosomal targeting motif. Both of these features are conserved.

FIGURE 4. Dendrogram of CD1 protein sequences. Multiple protein sequence alignment of the CD1 α1 and α2 domains only was performed with the program PILEUP, and a dendrogram representation of these data was generated. Horizontal distance along the tree is proportional to the degree of divergence. The guinea pig CD1 protein sequences described in Fig. 3 are indicated by asterisks. The division of the CD1 protein sequences into group 1 and group 2 is indicated. Other CD1 protein sequences were obtained from GenBank and the species are indicated. The non-guinea pig CD1 proteins and corresponding GenBank accession numbers are as follow: muCD1d1 (mouse CD1d1; I49581 (17)); mucd1d2 (mouse CD1d2; P11610 (17)); ratCD1d (rat CD1d; I56235 (18)); huCD1d (human CD1d; P15813 (2)); Crb1 (rabbit CD1d; B45887 (22)); scd1 (sheep CD1d; CAA07200 (45)); scd1a25 (sheep CD1b; CAA85359 (41)); scd1b42 (sheep CD1b; S47246 (41)); huCD1b (human CD1b; P29061 (55)); hucd1c (human CD1c; P29017 (55)); hucd1a (human CD1a; P06126 (55)); hucd1e (human CD1e; P15812 (2)).

Conservation of structural features in guinea pig CD1 proteins. The striking homology observed between guinea pig and human CD1 proteins is reflected in the conservation of several structural features common to all CD1 members. The overall domain structure predicted for the guinea pig CD1 proteins is conserved when compared with that of human CD1 protein sequences. There is an additional leader peptide at the amino terminus of each protein sequence followed by the α1, α2, and α3 domains encoding the extracellular portion of the CD1 protein. As with other species, the α1 and α2 domains exhibit a greater degree of amino acid sequence variability than the Ig-like α3 domain (Fig. 3). The α3 domain is followed by an ~20-aa transmembrane domain and a short cytoplasmic tail. Another feature the guinea pig CD1 proteins share with CD1 proteins of other species is the presence of multiple sites for Asn-linked carbohydrate attachment at canonical (Asn-X-Ser/Thr) motifs in the α1 and α2, but not in the α3 domains.

The mouse CD1d1 crystal structure indicates two disulfide bridges within the H chain of the CD1 molecule (16). One bridge is found between paired cysteine residues within the α2 domain (Fig. 3, consensus positions 131 and 195) and another between paired cysteines located within the α3 domain (Fig. 3, consensus positions 235 and 290). All of the guinea pig CD1 isoforms have conserved these paired cysteine residues, suggesting an importance in the formation of CD1 tertiary structure. In fact, the presence of these cysteine residues is conserved in all known CD1 protein sequences with the exception of mouse CD1d2, which substitutes tryptophan for cysteine at position 167 of the α2 domain (2).
between the predicted human and guinea pig CD1e protein sequences.

Identification of several CD1 pseudogenes in the guinea pig genome. Our initial cloning of the guinea pig α3 exon identified three genomic DNA fragments of CD1 genes that had frame-shift mutations. In addition to these three pseudogenes, a CD1b-like cDNA was cloned that was also defective. We have designated this pseudogene gpCD1Bφ1 and deposited the nucleotide sequence into the GenBank database (accession number AF147503). The DNA sequence of this cDNA revealed a single nucleotide deletion at a position ~30 bases into the α2 domain exon resulting in a frame shift. This deletion was also seen in genomic DNA from both Hartley and strain 2 guinea pigs and, therefore, was not merely a strain-specific polymorphism. The second pseudogene contained a CD1c-related sequence from the middle of the α1 domain through to the stop codon, but lacked a corresponding CD1-like 5’ end. We have designated this pseudogene as gpCD1Cφ1 and have also deposited it into GenBank (accession number AF159261). These results indicate that the guinea pig genome contains at least five CD1 pseudogenes.

A mAb to human CD1b cross-reacts with guinea pig CD1

Given the apparent conservation of CD1 among mammals, we tested mAbs specific for the CD1 of other species for cross-reactivity to guinea pig CD1. A panel of CD1 reactive mAbs was assembled and screened by FACS analysis using guinea pig thymocytes as target cells. These cells were chosen because cortical thymocytes in humans express high levels of all CD1 proteins. Surprisingly, a significant number of anti-human CD1 Abs were positive in this preliminary screen (data not shown). One of these Abs, BCD1b3.1, an anti-human CD1b-specific mAb, was chosen for more extensive analysis.

To confirm that the Ag recognized by the BCD1b3.1 mAb on guinea pig cells was in fact CD1, we characterized the anti-human CD1b cross-reacting Ag by immunoprecipitation of surface-iodinated guinea pig thymocytes. Immunoprecipitations were then subjected to SDS-PAGE and autoradiography. Radiolabeled species of 45 and 14 kDa were specifically visualized (Fig. 5, lane 4). These sizes are consistent with the relative mobilities of a glycosylated CD1 H chain and the βm m L chain, respectively. To further confirm the identity of the cross-reacting Ag as a CD1 gene product, we expressed the gpCD1b3 cDNA sequence in the guinea pig thymocytes. Immunoprecipitations were then subjected to SDS-PAGE and autoradiography. Radiolabeled species of 45 and 14 kDa were specifically visualized (Fig. 5, lane 4). Strong staining of the transfecant was observed, confirming that the BCD1b3.1 mAb recognized at least one of the guinea pig CD1 isoforms when expressed on the cell surface.

Guinea pig thymocytes showed high staining levels with BCD1b3.1 when used in FACS (Fig. 7A). The profile indicated that about 90% of thymocytes were highly positive. The pattern and intensity was similar to that seen when human thymocytes were stained with this Ab (data not shown). Single-cell suspensions were generated from normal spleen and lymph nodes and then subjected to flow cytometry after staining with the BCD1b3.1 Ab as described above. A representative experiment using BCD1b3.1 demonstrated a subpopulation of CD1-positive cells in the lymphocyte population of lymph node and spleen (Fig. 7, C and D). A subpopulation of lymphocytes gated from whole PBMC (11.8%) also stained positively with the Ab (Fig. 7B). These data indicated the presence of CD1 molecules cross-reactive with human CD1b on lymphocytes in the blood and lymphoid organs of guinea pigs.

We also investigated the presence of CD1 on mature guinea pig T cells. Previous data has shown that CD1 is present only on immature T cell precursors in the thymic cortex, but not on more mature medullary thymocytes or circulating T cells. A guinea pig T cell line propagated in vitro by repeated stimulation with PHA and recombinant human IL-2 revealed no staining with BCD1b3.1 (data not shown). This was consistent with the observation in humans that mature T cells lack expression of CD1a, -b, or -c. In addition, various guinea pig cell lines of nonhemopoietic origin including GP16, JH4, and the strain 2 guinea pig tumor line 104C1 were also unreactive with this Ab (Fig. 6 and data not shown).

Expression of CD1 proteins in normal guinea pig tissues

The studies described above showed that the mAb BCD1b3.1 recognized at least one guinea pig CD1 isoform, indicating that this would be a useful reagent for initial characterization of CD1 protein expression in normal guinea pig tissues. Immunoperoxidase staining of frozen sections of guinea pig thymus with mAb BCD1b3.1 revealed dense staining of cells in the cortical region consistent with CD1 expression on immature thymocytes (Fig. 8A). The spleen was also examined for CD1 expression. Positively stained cells in the white pulp of the spleen were observed with the BCD1b3.1 mAb (Fig. 8B). To identify the stained cells, two-color staining was conducted on spleen sections using the CD3-specific Ab PCA/188A together with BCD1b3.1. Spleen sections simultaneously stained with these two Abs demonstrated that CD3+ T cells and CD1-positive lymphoid cells represented closely associated but distinct populations in the white pulp of the spleen (data
This staining pattern was consistent with CD1 being present in B cell follicles adjacent to the T cells in periarteriolar lymphoid sheaths. In addition to the spleen, staining within lymph nodes with the BCD1b3.1 mAb was also consistent with CD1 expression on lymphocyte corona B cells of secondary lymphoid follicles, while staining of the associated germinal centers was not observed (Fig. 8). To confirm the identity of the CD1+ cells in these tissues as B cells, spleen and lymph node serial sections were stained with anti-guinea pig IgG and BCD1b3.1. This revealed colocalization of CD1 and IgG on the same cell population, further supporting the expression of CD1 on B cells in the guinea pig (data not shown).

The skin was also analyzed for expression of CD1. The elongated morphology and dermal location of stained cells suggested CD1 staining of dermal dendritic cells in the guinea pig skin (Fig. 8, C and D). Serial sections of skin were also stained with the anti-guinea pig MHC class II Ab CI13.1 and exhibited a pattern of staining similar to sections stained with BCD1b3.1, although the frequency of MHC class II+ to CD1+ cells in the dermis was ~3:1 (data not shown). These results indicated the presence of dermal dendritic cells in the skin, a subpopulation of which were CD1+.

No staining of Langerhans cells was observed in the epidermis with BCD1b3.1. We also observed staining of cells with dendritic morphology in the paracortex of the lymph node consistent with expression of CD1 on interfollicular dendritic cells (Fig. 8F). These cells exhibit extended processes that penetrate between the surrounding cells. In addition, staining of lung tissue revealed sporadic clusters of cells exhibiting lymphoid morphology consistent with bronchiolar-associated lymphoid tissue (data not shown). We have not detected CD1 in the liver or intestine of guinea pigs with this Ab. These preliminary results show that CD1 proteins are constitutively expressed on cells of hemopoietic lineages in a pattern similar to that described for several human CD1 proteins.

Discussion
We have identified eight distinct full-length CD1 genes in the guinea pig. Based on comparison with the human CD1 isoforms, the eight predicted guinea pig CD1 proteins encoded by these genes include four CD1b-like, three CD1c-like, and one CD1e-like proteins. To date, this is the largest number of complete CD1 sequences identified from a single animal species. The predicted guinea pig CD1 protein sequences demonstrate a high degree of homology with CD1 protein sequences from other animal species, indicative of the highly conserved nature of the CD1 gene family.

CD1 genes and proteins have been described in a variety of mammalian species including human (2, 34), mouse (17), rat (18), rabbit (22), cow (40), sheep (41), pig (42), cat (43), and dog (44). Nevertheless, there is significant heterogeneity with respect to the complement of CD1 family members that have been preserved and, in some cases, expanded during the evolution of different mammalian species. The multiple isoforms of the CD1B and CD1C genes in the guinea pig suggest that relatively recent gene duplication events have occurred to create an extended family of CD1 genes in this species. This process has generated at least five CD1B-like genes in the guinea pig, four of which possess full-length coding sequences, with the fifth being a pseudogene. A similar phenomenon is seen in sheep (Ovis aries), which has at least four genes encoding CD1b-like proteins with only two of these being full-length CD1-coding sequences (41). In addition, sheep possess at least one homologue of the group 2 CD1d isoform.
Thus, both sheep and guinea pigs have undergone more extensive amplification of specific CD1 isoforms than primates. The purpose of these multiple versions of specific CD1 isoforms is unknown, but suggests an ongoing important role and evolutionary pressure to expand the opportunities for CD1 based immune responses.

Mice, in contrast to guinea pigs, lack group 1 CD1 homologues and have only two highly homologous group 2 CD1 genes (CD1D1 and CD1D2). The rat, a close relative of mice, appears to have only a single copy of the CD1D gene (46). Members of the order Lagamorpha (rabbits and hares) and Rodentia (mice, rats and guinea pigs) are closely related and together form the cohort Glires (47). The monophyly of rodents and lagamorphs implies a common ancestor. Because both group 1 and group 2 CD1 genes have been described in rabbits (22), the most likely explanation for the lack of group 1 CD1 in muroid rodents is that these genes were deleted during evolution. Our finding of group 1 CD1 homologues in guinea pigs further suggests that the loss of group 1 from muroid rodents occurred after the division of the rodent order into two major suborders: Sciurognathi containing rats and mice and Hystrixognathi containing guinea pigs. This split is estimated to have occurred ~55 million year ago during the early Eocene (R. Honeycutt, unpublished observations).

Recent molecular phylogenetic data suggest that the guinea pig is more distantly related to muroid rodents than was previously indicated by morphological criteria. It has been postulated that guinea pigs and their relatives (Hystrixognathi) may even represent a distinct order separate from the order Rodentia (32, 33). While this classification is not generally accepted (48), the evolutionary divergence between guinea pigs and other members of the order Rodentia may contribute to the differences we have described between the CD1 gene families of muroid rodents and guinea pigs. More likely, these differences reflect diverse evolutionary pressures between closely related species that may be acting to remove or duplicate existing CD1 genes. Given the ability of CD1 proteins to function as Ag-presenting molecules, environmental influences, such as exposure to different pathogens, may have exerted selective pressures to alter the number or expression of different CD1 genes in different animal species.

One hypothesis for the presence of multiple isoforms of CD1 in a given species is to facilitate sampling of Ags from different compartments within the APC. The human CD1b protein, for example, has been shown to traffic to late endosomes, colocalizing with MHC class II and HLA-DM molecules (37). This was shown to be dependent upon the presence of an intact YXXZ endosomal targeting motif (37, 38). In contrast, human CD1a protein lacks the YXXZ targeting motif and shows substantially lower steady-state accumulation in endosomes, suggesting that human CD1a and CD1b mediate distinct pathways for Ag presentation (M. Sugita and M. B. Brenner, unpublished observations). We have not identified a guinea pig homologue of human CD1a. However, the gpCD1b3 protein also lacks the critical YXXZ endosomal targeting motif at the carboxyl terminus of the protein. Based on previous data, this suggests that the gpCD1b3 protein may have a trafficking pattern similar to human CD1a. With the exception of gpCD1b3 and gpCD1e, all of the guinea pig CD1 proteins possess

FIGURE 8. Expression of CD1 in guinea pig tissues. Frozen sections of thymus (A), spleen (B), trunk skin (C and D), and lymph node (E and F) were stained with the mAb BCD1b3.1. Stained sections were visualized with an anti-mouse avidin-biotin-peroxidase complex. Positive staining appears red in these sections. A shows staining of the cortex in a thymic lobule from a 3-wk-old guinea pig (×100). B shows a cross section of spleen with staining in the white pulp (×100). C shows cells of dendritic morphotype staining in the dermal layer of the trunk skin (×100). D shows the surface staining of CD1 on cells with dendritic morphology in the trunk skin dermis (×600). E shows an adult inguinal lymph node with staining of the lymphocyte corona of two secondary B cell follicles and negative staining in the associated germinal centers (×100). F is centered on the paracortex of the lymph node and shows staining of numerous cells with dendritic morphology. The edge of a secondary B cell follicle evident in the upper right of F (×200). All tissues were counterstained with hematoxylin except A.
a YXXZ motif at the carboxyl terminus. Differences in the subcellular localization of human CD1a and CD1b isoforms potentially allows Ags in different subcellular compartments to be sampled by the APC and subsequently presented to T cells. This may be a particularly important mechanism for the presentation of Ags following infection by intracellular pathogens, such as *M. tuberculosis*, that are capable of parasitizing macrophages. The presence of multiple CD1b and CD1c isoforms in the guinea pig provides a unique opportunity to examine the fine specificity of CD1 Ag presentation and may provide additional insight into the specific function of the individual isoforms.

It is currently unclear whether the guinea pig harbors a CD1d homologue. Although the available data on CD1 in other species suggests that this member of the family is frequently conserved, our methods have so far failed to demonstrate CD1d in the guinea pig. One possible explanation for this is that a putative gpCD1d sequence may significantly differ from the degenerate primers used in the initial amplification of the α3 exon from genomic DNA. This may also explain the absence of cross-hybirdizing bands in Southern blot analyses of guinea pig genomic DNA when high stringency wash conditions were used with a probe containing the α1 and α2 exons of human CD1d cDNA. In contrast, identical Southern blots hybridized with a human CD1b α1 and α2 exon probe in the same experiment showed multiple cross-hybirdizing bands. Thus, genes encoding guinea pig homologues of CD1d may be significantly divergent from the human CD1d probes used here, or these genes may simply be absent in the guinea pig.

Characterization of the expression of guinea pig CD1 proteins revealed many features similar to CD1 expression in humans. The apparent molecular mass of the precipitated CD1 protein was almost identical with that of the glycosylated human CD1 protein and revealed a similar association with β2m (Fig. 5). CD1 expression was detected by FACS analysis of normal guinea pig PBMC, thymocytes, splenocytes, and lymph node cells (Fig. 7). In addition, FACS analysis of a guinea pig CD1 transfectant demonstrated that the BCD1b3.1 mAb cross-reacted with the gpCD1b3 gene product (Fig. 6). The expression of CD1 proteins detected by immunohistochemical staining using the BCD1b3.1 mAb revealed high levels of CD1 on thymocytes (Figs. 5A and 8A). Cytotoxic thymocytes, which are predominantly immature T cells, express high levels of CD1 protein during thymic development (36). CD1 protein expression appeared to be lost following maturation of guinea pig T cells, as evidenced by the absence of CD1 staining on medullary thymocytes (Fig. 8A). In addition, the CD3+ T cell regions in the spleen were negative for CD1 staining, and T cells propagated in vitro also lacked expression of CD1. More recent data now demonstrate that, in addition to gpCD1b3, the BCD1b3.1 mAb also reacts with the gpCDb2, -b4, and -c3 isoforms (our unpublished observations). Because of the lack of monospecific mAb reagents, we cannot identify the specific guinea pig CD1 isoforms expressed on a particular cell type within tissues.

The staining of CD1+ CD3- IgG2 cells with lymphoid morphology closely associated with periarteriolar lymphoid sheaths is consistent with CD1 expression on B cells in the guinea pig (Fig. 8, B and E). It has been shown previously that B cells derived from human spleen and PBMC express the CD1c protein (49). In addition to the spleen, strong staining of secondary lymphoid follicles in guinea pig lymph nodes is also consistent with CD1 expression on B cells (Fig. 8E). We have also detected a subpopulation of CD1+ cells in the PBMC of guinea pigs, which, as indicated by two-color FACS using Abs specific for CD1 and IgG, are most likely circulating B cells (Fig. 7B and unpublished observations). These data indicate that B cells are a major population of CD1-bearing cells in the guinea pig. One possible function of CD1 on B cells may be to present lipid or glycolipid Ags to T cells for the purpose of eliciting T cell help for the production of Abs to glycolipid Ags (11, 50).

Previous data have shown that the human dendritic cells found in lymphoid and nonlymphoid tissues express high levels of the group 1 CD1 isoforms. Immunohistochemistry of guinea pig lymph nodes also revealed the presence of CD1 on cells with dendritic morphology in the paracortex, which is consistent with interfollicular dendritic cells found in various lymphoid tissues (Fig. 8F). Dendritic cells in the dermal layer of the skin were also observed in guinea pig (Fig. 8, C and D). Dermal dendritic cells in human skin express CD1a and CD1c with a subpopulation of dermal dendritic cells also expressing CD1b (51). The CD1a isoform is highly expressed on epidermal Langerhans cells in human skin. So far, we have failed to identify CD1 staining of epidermal Langerhans cells in guinea pigs, although we have observed these cells in serial sections using an anti-MHC class II mAb. It is well established that dendritic cells are potent APCs and play a crucial role in priming specific T cell responses (52). Therefore, the expression of CD1 on interfollicular dendritic cells and dermal dendritic cells in guinea pigs increases the plausibility that these cells may be involved in the generation of CD1-restricted T cell responses.

The recent appreciation of the Ag-presenting capacity of CD1 now requires assessment of these molecules in the larger framework of the immune response to infection and other disease processes. Human T cells that respond to CD1-presented lipid and glycolipid Ags secrete high levels of IFN-γ (5), lyse infected cells, and also have bactericidal effects on *M. tuberculosis* localized within macrophages (53). These data support a possible role for CD1-restricted T cells in control or clearance of mycobacterial infections such as tuberculosis or leprosy (54). However, direct in vivo proof of a significant role for the group 1 CD1 proteins in host response to infection has yet to be obtained. Our demonstration that guinea pigs, like humans, have clearly preserved and expanded the group 1 subset of CD1 Ag-presenting molecules suggests that these proteins may confer an important evolutionary advantage in this species that may be relevant to their role in the human immune system. The findings presented here thus indicate that the guinea pig will be useful as a small animal model for examining the role of group 1 CD1 in tuberculosis and other infections relevant to human disease.

Acknowledgments

We thank Dr. Rodney Honeycutt for valuable comments on rodent evolution, Dr. David McMurray for discussions about the guinea pig animal model, Dr. Masahiko Sugita for sharing unpublished data, and Robert Carpenter for support of this project. We also thank Drs. Cheryl Murphy, Charlotte Kensil, and Gerald Belze for critical review of the manuscript.

References


