Molecular cloning, reconstruction and expression of the gene encoding the alpha-chain of the bovine CD8—definition of three peptide regions conserved across species

P. LALOR,* C. BUCCI,† M. FORNARO,‡ M. C. RATTAZZI,* H. NAKAUCHI,* L. A. HERZENBERG* & S. ALBERTI‡
*Department of Genetics, Stanford University, Stanford, California, U.S.A.; †Walter and Eliza Hall Institute, Melbourne, Australia; ‡Department of Cellular and Molecular Biology and Pathology, University of Napoli, Napoli, Italy; ¶Mario Negri Institute for Pharmacological Research, Milano, Italy and \ Laboratory of Molecular Regulation of Aging, Riken Research Center, Tsukuba, Japan

Accepted for publication 22 December 1991

SUMMARY

We report the cloning of a cDNA encoding the alpha-chain of the bovine CD8 (BoCD8α). A bovine thymus cDNA library was hybridized at low stringency with a human CD8α cDNA clone. The first round of screening of 5 x 10⁶ independent colonies yielded 12 clones containing incomplete BoCD8α genes. Two further rounds of colony hybridization were conducted, each using as a probe the 5′ fragment from the longest BoCD8α clone previously isolated. The final screening yielded a clone containing a 2 kilobase (kb) insert. We mapped and sequenced the 2 kb BoCD8α clone and compared it with the published sequences of the genes encoding the human, mouse and rat CD8α. Sequence analysis confirmed that the clone under study encoded the BoCD8α. The overall similarity of the BoCD8α coding region with the human CD8α coding sequence is 74.7% at the nucleotide level and 62.1% at the protein level. Lower levels of similarity are found with the mouse and rat CD8α.

Interestingly, three separate highly homologous regions are clearly defined at the peptide level in bovine versus human and mouse versus rat comparisons. Two of the regions are highly conserved among all species analysed, while the most 5′ region is not. We speculate that the latter region may contain the binding site of CD8α to the α, domain of major histocompatibility complex (MHC) class I molecules. Sequence analysis showed that the 2 kb BoCD8α clone contains an incomplete coding region, i.e. lacks six bases corresponding to the first two amino acids of the leader region. To allow efficient translation and processing of the BoCD8α gene, we constructed a chimeric gene containing the coding sequence of the BoCD8α clone and synthetic sequences corresponding to the first two amino acids of the human CD8α leader sequence. The chimeric gene was subcloned in the pKSV10 expression vector. The pKSV10-BoCD8α construct is efficiently expressed both transiently in COS cells and stably in L cells, as determined by Northern blot and by FACS analysis, using the ILA-51 monoclonal antibody to BoCD8α. The latter result formally proves that the ILA-51 antibody does indeed recognize the product of the BoCD8α gene, as previously suggested on serological grounds.

INTRODUCTION

Class I restricted T cells are characterized by the expression of the lymphocyte differentiation antigen CD8. CD8 is thought to bind class I molecules and to function both as an adhesion molecule, stabilizing the interaction of T lymphocytes with their targets, and as a signal transduction structure. CD8 exists either as homodimer or multimer of the alpha chain (CD8α) or as heterodimer of the alpha (α) and beta (β)-chains (CD8αβ). CD8α sequences are highly homologous in man, mouse and rat at both the nucleotide and the protein level, indicating that they exert similar conserved functions.

The CD8 molecule of domestic cattle (BoCD8) has primarily been studied by serological methods, in particular using CD8-specific monoclonal antibodies. The development of molecular probes for BoCD8α should help to study the role of this molecule in the bovine immune response. Here we use a human probe to isolate and begin the molecular characterization of the gene encoding the bovine CD8α chain.
MATERIALS AND METHODS

Cells
Mouse I. thymidine kinase (TK-) cells12 and monkey COS-7 cells14 were maintained in Dulbecco's modified Eagle's minimal essential medium (DMEM). L cells transfected with genomic human DNA and selected for CD8s or human leucocyte antigen (HLA) class I expression were obtained as described previously.12 L cells transfected with whole genomic bovine DNA and selected for CD8s expression were kindly provided by Niall McHugh (ILRAD, Nairobi, Kenya). All transfectants were maintained in DMEM supplemented with hypoxanthine, aminopterin and thymidine (HAT).

Cell DNA extraction
High molecular weight DNA was extracted from Igrov-1 human ovarian cancer cells, from bovine and human peripheral blood leucocytes (PBL), from mouse liver and from L-cell transfectants. To obtain higher transfection efficiency, DNA was extracted following a guanidinium thiocyanate-based procedure.13 DNA to be used solely for Southern blotting was extracted following the procedure of Wigler et al.14

DNA transfection
The calcium phosphate co-precipitation technique was followed.15 Briefly, the day before the transfection 106 L TK- cells were seeded in each dish. Transfections were performed with 20 μg of genomic DNA co-precipitated with 1 μg of plasmid containing the TK gene. Culture in HAT medium selected the cells expressing the transfected DNA. For transient transfections the same procedure was applied to COS-7 cells. DNA from the human ovarian cancer cell line Igrov-1 was used as carrier DNA. Transfected COS-7 cells were analysed in flow cytometry 48 hr after the transfection.

Antibodies
The ILA-51 anti-BoCD8s antibody was kindly provided by Ivan Morrison and Niall McHugh (ILRAD). The antibody was used as ascites in a two-step staining. A number of isotype- and allotype-matched antibodies with irrelevant specificities were used as controls for staining (data not shown).

Immunofluorescence
Fluorescence analyses and sorts were made on a modified FACS IV (Becton Dickinson, Sunnyvale, CA), used as described previously.16 To improve the detection of transfectants stained with fluorescein isothiocyanate (FITC) antibodies, subtraction of cell autofluorescence was utilized.17

Cloning of the BoCD8s cDNA gene
The bovine thymocyte cDNA library used for the screening was kindly supplied by Dr Ken-ichi Arao (DNAX, Palo Alto, CA). The library constructed in the pCDV expression vector18 was amplified in Escherichia coli HB-101 and plated. Replica filters were hybridized with the human CD8s cDNA.1 Low stringency hybridization was performed at 50° C in 6 × SSPE 0.1% SDS with washings at 55° C in 4 × SSPE (1 × SSPE = 180 mM NaCl, 10 mM NaPO4, 1 mM EDTA, pH 7.7).1

DNA sequencing
Nucleotide sequences were determined by the di-deoxy-nucleotide sequencing method19 after subcloning restriction fragments into M13 mp18 and mp19 phage vectors. The T6 Sequencing kit by Pharmacia (Uppsala, Sweden) was used. Ambiguities in the sequence ladder were resolved by the combined use of Taq polymerase and of de-oxa-guanosin tri-phosphate (instead of dGTP) in the sequencing reaction (TaqFenz sequencing kit, USB, Cleveland, OH). The elongation temperature when using Taq polymerase was 70°.

DNA and RNA hybridization
DNA was digested and analysed following the method of Southern.20 Cellular RNA was prepared by the guanidinium thiocyanate procedure13 and analysed by RNA blot hybridization.11 Restriction enzymes were purchased from New England Biolabs (Boston, MA), Toyobo and Boehringer Mannheim (Mannheim, Germany).

Computer analysis
DNA sequences were analysed on a VAX 6410 computer using the Genetics Computer Group (GCC) programs.21

RESULTS
Screening of a bovine thymocyte cDNA library
We screened a bovine thymus cDNA library with the 1-7 kilobase (kb) human CD8s cDNA probe under conditions of low stringency hybridization.1 The first round of screening of 5 × 106 independent colonies yielded 18 clones which showed strong hybridization with the probe. DNA from 12 of these clones showed a similar pattern of restriction cutting, as expected from clones containing a bona fide BoCD8s gene. Only one restriction fragment varied in length among the latter clones, presumably reflecting the length heterogeneity of the 5' portion of the various cDNA. The frequency of detection of the above cDNA clones indicates an abundance of BoCD8s mRNA in cattle thymocytes of about 0.04%. Two further rounds of colony hybridization with a total of 5 × 1010 independent colonies were conducted, each using as a probe the 5'-fragment from the longest BoCD8s clone previously isolated. The final screening yielded the clone 151, which contains a 2 kb insert.

Primary protein structure
The 2 kb insert of clone 15-1 was completely sequenced. The sequencing strategy is presented in Fig. 1a. The nucleotide sequence and the amino acid sequence corresponding to the longest reading frame of the clone 15-1 are presented in Fig. 1b. Hydrophobicity analysis of the predicted protein sequence shows that the first 25 residues have a profile characteristic of hydrophobic leader sequences (data not shown). Sequence comparison with the human, mouse and rat CD8β18,22 confirmed this finding (Figs 2, 3). These data indicate that the mature protein is 217 amino acid long with a predicted size of 22,000 MW. Analysis of the hydrophobicity plot reveals a second hydrophobic region near the COOH terminus, which is followed by a hydrophilic region. We predict that the two regions correspond to the transmembrane and cytoplasmic domain, respectively. No N-linked glycosylation sites were apparent in the sequence of BoCD8β. In comparison, the rat and human CD8α possess one N-glycosylation site, whereas the mouse CD8α has three potential N-glycosylation sites.
Figure 1 (a) (upper) Restriction map of the BoCDHs clone. Restriction sites used for subcloning are shown: B, BamHI; E, EcoRI; Ps, PstI; N, NdeI; Pv, PvuII; S, SalI; Sa, SacI; X, XhoI. Thick line: BoCDHs. Thin line: pCDV vector. (lower) Sequencing strategy. Horizontal arrows indicate the sequenced fragments and their orientation. (b) Nucleotide sequence of the BoCDHs gene of the corresponding translation product. Nucleotides from 1 to 21 (lower case letters) were added to the 15.1 clone following the sequence of the human CDHs gene.
Comparison of the BoCD8x with the human, mouse, and rat homologues

Using the GAP computer program\(^2\) we determined that the identity between the bovine and the human CD8x nucleotide sequences is 74.7% in the coding region and 69.2% in the non-coding region. Identity with the coding regions of more distantly related species like mouse and rat is 64.4% and 65.9%, respectively. For comparison, the DNA coding regions of mouse and rat CD8x genes are 81% identical. At the protein level the bovine and human genes are 62.1% identical (74.9% similar, taking into account conservative amino acid substitutions). Much lower levels of identity were found between bovine and rat (47%) or mouse (48.5%) CD8x peptides. In comparison, protein identity between mouse and rat is 68.7%.

Interestingly, three highly conserved regions are clearly defined by dot-plot matrix analysis\(^2\) of the bovine and human CD8x sequences (Fig. 3). The existence of three separate regions of similarity is particularly clear at the peptide level (Fig. 3a versus d). Similar results were obtained comparing mouse and rat CD8x peptide sequences (Fig. 3b). Interestingly, homology between bovine and rodent peptide sequences (Fig. 3c; P. Lalor, C. Bucci, M. Fornaro, M. C. Rattazzi, H. Nakazaki, L. A. Herzenberg and S. Alberi, unpublished observations) is high in the middle (region 2) and in the 3′ portion of the gene (region 3), but is much lower in the most 5′ segment of the CD8x (region 1).

Similar results were obtained comparing human with rodent peptide sequences (data not shown).

We and others previously reported that CD8x is a member of the immunoglobulin superfamily.\(^3\) By comparison with immunoglobulin sequences it is possible to define variable-like, hinge-like, transmembrane and cytoplasmic regions. The three conserved regions cited above do not strictly correspond to specific immunoglobulin-like regions. In fact, region 1 falls in the middle of the variable-like region. The second region of homology spans across the variable-like and hinge-like regions and the third one includes both the transmembrane and cytoplasmic domains.

Construction of the chimeric BoCD8x gene

Sequence analysis showed that BoCD8x clone 1-16 lacks six bases corresponding to the first two amino acids of the leader region (Fig. 1b). To allow efficient translation of the BoCD8x gene and transport of the protein to the cell surface, we decided to add the missing 5′ region to the clone 1-16, following the sequence of the human CD8x gene. Complementary oligonucleotides were synthesized on an Applied Biosystem Oligosynthesizer (Figs 1, 4a), taking care to include convenient restriction sites at both ends and a strong Kozak consensus sequence\(^4\) around the methionine start codon. Since the distribution of
restriction sites in the pCDV vector did not allow efficient subcloning procedures, we performed a stepwise reconstruction process in the pSP65 plasmid (Fig. 4b). For the final subcloning in the pKS10 cDNA expression vector (Pharmacia) the BamHI-PvuII fragment of the chimeric BoCD8x gene was inserted in the BgIII site of the pKS10 plasmid by successive sticky end ligation, fill in with Klenow polymerase and blunt and ligation (Fig. 4b). Correct orientation was assessed by restriction mapping. Insertion of the BamHI-XhoII fragment of the chimeric BoCD8x gene in the BgIII site of the pKS10 plasmid yielded an expressible construct but with lower and more unstable expression than the above. The latter construct was not studied further.

Transfection of the chimeric BoCD8x gene

Using the calcium phosphate co-precipitation technique, we obtained efficient expression of the chimeric BoCD8x gene after both transient transfection in COS cells and stable transfection in L cells. Flow cytometric analysis of two independent BoCD8x cDNA L-cell transfecants is shown in Fig. 5a and b. Staining controls were a genomic BoCD8x transfectant obtained by FACS selection of L cells transfected with whole bovine genomic DNA (Fig. 5c) and L cells transfected with mouse liver DNA, but unselected for the expression of surface antigens (Fig. 5d).

Southern and Northern blot analysis of BoCD8x transfectants

DNA extracted from cattle PBL, a genomic BoCD8x transfectant, two independent BoCD8x cDNA transfectants and L cells transfected with control DNA were hybridized at high stringency with the purified insert of the chimeric BoCD8x cDNA (Fig. 6a). Bovine PBL and the genomic BoCD8x transfectant show quite a similar pattern of hybridization. The intensity of hybridization is consistent with the presence of a single copy of the BoCD8x gene per haploid genome in both cell types. The two BoCD8x cDNA transfectants show an intense pattern of hybridization, indicating the presence of multiple copies of the BoCD8x gene in the transfected cells. The complex array of bands observed presumably stems from the heterogeneity of the integration sites in different cells and/or rearrangement and amplification of the BoCD8x cDNA during the process of transfection and selection by sorting. A major band is shared by both BoCD8x cDNA transfectants and corresponds in molecular weight (9 kb) to the pKS10-BoCD8x plasmid. In long exposures a faint signal is evident in the lane of the human CD8x amplified transfectant (data not shown), whereas no hybridization is detectable in L cells transfected with unrelated genes.

Similar results were obtained by Northern blot analysis of RNA from the same cells (Fig. 6b). The genomic BoCD8x transfectants show a single broad band with an average molecular weight of approximately 2 kb, in good agreement with the size of the human CD8x gene message \(^{26,36}\) and with the size of the clone 15-1, the longest cDNA we had isolated. The
Figure 4. (a) Sequence of the oligonucleotides used for the construction of the chimeric BoCD8s gene. Underlined is the Kozak consensus sequence. Restriction sites used in the process of construction are indicated. Nucleotides added to the BoCD8s gene are typed in lower case. (b) Scheme of the construction of the chimeric BoCD8s gene. Restriction sites used in the process of construction are indicated: B, BamHI; Bg, BglII; E, EcoRI; K, KpnI; Ps, PstI; Pv, PvuII; S, Sau3A I; X, XhoI. pCDV-C1 15.1, longest isolate of the BoCD8s gene; pSP65/a, pSP65/b, pSP65/c. Intermediates in the reconstruction procedure engineered in the pSP65 vector, pKS10-BoCD8, final construct.

Figure 5. FACS profiles of L cells transfected with the chimeric BoCD8s gene after three rounds of sorting. (a) BoCD8s transfectant no. 6; (b) BoCD8s transfectant no. 2; (c) LTK-BoT8, BoCD8s transfectant obtained after transfection with whole bovine genomic DNA and FACS selection with the ILA-51 mAb; (d) L cells transfected with mouse liver DNA and the chicken TK plasmid, unselected for the expression of surface antigens. Unstained cells (-----); cells stained with the ILA-51 mAb (-----).
two cDNA transfectants analysed show strong hybridization signals (Fig. 6b and data not shown), roughly in proportion to the respective levels of surface expression (Fig. 5a, b) and to the gene copy number of the transfected BoCD8x cDNA (Fig. 6a). Two bands are present. The low molecular weight band (1.2 kb) presumably corresponds to the correctly spliced message, the higher molecular weight band (2.7 kb) is likely to contain the BoCD8x insert plus the unspliced SV40 intron of the pKSV10 plasmid.

**DISCUSSION**

In this article we report the cloning of a cDNA encoding the BoCD8x chain. The cDNA was isolated by low stringency hybridization of a cDNA library constructed using cattle thymocyte messenger (mRNA) and using as a probe a human clone encoding CD8a. After multiple rounds of screening, the longest isolate was 2 kb in length. Sequence analysis confirmed that it does encode the BoCD8x gene. The identity between the bovine and the human BoCD8a at the nucleotide level is quite high (74.7% in the coding region and 71.2% overall) indicating a high level of conservation of the CD8a primary structure. This finding is confirmed by the high level of identity with more distantly related species like mouse and rat, which is about 65% in the DNA coding region. At the protein level the bovine and human genes are 62.1% identical. Lower but still high levels of identity are found between bovine and rat or mouse CD8a peptides (about 48%).

Interestingly, three highly homologous regions are clearly defined by stringent matrix-dot plot comparison of the bovine and human CD8a peptide sequences (Fig. 3a). Similar results are obtained comparing the mouse and rat peptide sequences (Fig. 3b). Interestingly, the homology between the bovine and rodent peptide sequences (Fig. 3c; P. Lalor, C. Bucci, M. Forneris, M. C. Rattazzi, H. Nakashi, L. A. Herzenberg and S. Alberti, unpublished observations) is high in regions 2 and 3, but is much lower in region 1. Similar results are obtained comparing human and rodent peptide sequences (data not shown). Taken together, these data indicate the existence of three conserved regions, two of which are quite conserved even among distantly related species. It is tempting to postulate functional significance for this conservation. In particular, we hypothesize that region 1 is responsible for the interaction of CD8a with the major histocompatibility complex (MHC) class I gene products: the need for co-evolution with a polymorphic set of molecules such as the MHC may explain the lower amount of conservation of region 1 compared with the second and third region. The known lack of efficient interaction between the MHC and CD8 of different species is consistent with this speculation. The two most homologous domains of CD8a may serve more conserved interactions, e.g. with non-polymorphic regions of the CD8β chain and/or with the T-cell receptor and the p56lck. CD8a is a member of the immunoglobulin gene superfamily and domains homologous to the variable, hinge, transmembrane and cytoplasmic regions of immunoglobulins can be identified.

Interestingly, the three regions of homology do not correspond to specific Ig-like domains. Indeed, the most 5’ (region 1) falls in the middle of the putative variable region. Region 2 spans across the variable-like and hinge-like regions and region 3 includes the transmembrane and cytoplasmic segments. This lack of precise correspondence possibly reflects evolutionary pressure on specific functions of CD8a evolved after divergence from the immunoglobulin genes. Consistent with the latter interpretation is the finding that the conserved CD8a regions that have been defined span across exon/intron boundaries of the human and mouse genomic clones.

The BoCD8x clone we have isolated lacks six bases corresponding to the first two amino acids of the leader sequence.
hampers both efficient translation of the BoCD8a mRNA and processing/transport of the mature BoCD8a molecule. To allow expression of the BoCD8a gene and make direct functional studies of this molecule possible, we decided to synthetically add the missing six bases following the sequence of the human CD8a gene.1 The chimeric BoCD8a gene is efficiently expressed after both transient and stable transfection as indicated by Northern blot and immunofluorescence analysis. The latter result formally proves that the ILA-S1 antibody is able to recognize the product of the BoCD8a gene. The two evidences that support this conclusion are the specific recognition of our BoCD8a cDNA transfectants by the ILA-S1 antibody and the specific hybridization of our BoCD8a clone with DNA and RNA from the bovine genomic DNA transfectant selected with the ILA-S1 antibody.

ACKNOWLEDGMENTS

We are grateful to Dr. Kes-ichi Arai for the gift of the bovine thymus cDNA library in the pCDV vector used for the cloning of the BoCD8a gene. This work was supported by the Italian Association for Cancer Research and by the Italian National Research Council. P.L. was a recipient of a Damon Runyon-Walter Winchell fellowship during the course of this study. M.F. is a recipient of a fellowship from the Italian Association for Cancer Research.

REFERENCES