The Inter-Locus Recombinant HLA-B*4601 Has High Selectivity in Peptide Binding and Functions Characteristic of HLA-C

By Linda D. Barber,* Lucy Percival,* Nicholas M. Valiante,* Ling Chen,† Christopher Lee,‡ Jenny E. Gumperz,* Joseph H. Phillips,§ Lewis L. Lanier,∥ Judith C. Bigge,¶ Raj B. Parekh,§ and Peter Parham*

From the *Departments of Structural Biology and Microbiology & Immunology, and °Department of Chemistry, Stanford University, Stanford, California 94305; †Perkin Elmer Applied Biosystems Division, Foster City, California; ‡Department of Human Immunology, DNAX Research Institute of Molecular and Cellular Biology, Palo Alto, California; and °Oxford GlycoSystems Ltd., Abingdon, Oxford, OX14 1RG, UK

Summary

The vast majority of new human HLA class I alleles are formed by conversions between existing alleles of the same locus. A notable exception to this rule is HLA-B*4601 formed by replacement of residues 66-76 of the α1 helix of B*1501 by the homologous segment of Cw*0102. This inter-locus recombination, which brings together characteristic elements of HLA-B and HLA-C structure, is shown here to influence function dramatically. Naturally processed peptides bound by B*4601 are distinct from those of its parental allotypes B*1501 and Cw*0102 and dominated by three high abundance peptides. Such increased peptide selectivity by B*4601 is unique among HLA-A,B,C allotypes. For other aspects of function, presence of the small segment of HLA-C-derived sequence in an otherwise HLA-B framework converts B*4601 to an HLA-C-like molecule. Alloreactive cytotoxic T lymphocytes (CTL), natural killer (NK) cells, and cellular glycosidases all recognize B*4601 as though it were an HLA-C allotype. These unusual properties are those of an allotype which has frequencies as high as 20% in south east Asian populations and is associated with predisposition to autoimmune diseases and nasopharyngeal carcinoma.

In human populations, divergence of major histocompatibility complex class I loci has been facilitated by the rarity of inter-locus recombination (1). Thus HLA-B and -C which were formed by duplication of a common ancestor (1) now have distinctive polymorphism (2) and levels of expression (3) suggestive of functional specialization. In this context, B*4601 is an exceptional allele in which HLA-B-specific and HLA-C-specific sequences have been combined (4). B*4601 is the product of an inter-locus recombination involving B*1501 and Cw*0102 where an HLA-C-derived sequence encoding residues 66-76 of the α1 helix is placed within a typically HLA-B framework. In consequence, B*4601 differs from B*1501 by seven amino acid substitutions; Lys from Ile at 66, Tyr from Ser at 67, Arg from Thr at 69, Gln from Asn at 70, Ala from Thr at 71, Asp from Tyr at 74, and Val from Glu at 76. Residues within the segment 66-76 contribute both to formation of the peptide-binding site and the membrane-distal surface of the class I glycoprotein with which CTL and NK cell receptors interact. Residue differences between B*4601 and B*1501 are therefore likely to influence class I function.

The B*4601 allotype is common in south east Asian populations where it is found at frequencies up to 20% (5), and is known for its association with susceptibility to nasopharyngeal carcinoma (6-7) and the autoimmune diseases myasthenia gravis (8) and Graves' disease (9-12). Despite apparent association with predisposition to certain diseases, B46 haplotypes surpass the frequency of the parental B15-Cw1 haplotype by more than fourfold (13) in south east Asia. Ascendance of B*4601 suggests the residue substitutions introduced by the inter-locus recombination produce functional changes that confer selective advantage. To determine how the conversion event that created B*4601 affects the function of the class I glycoprotein, properties of B*4601 and the parental allotypes B*1501 and Cw*0102 were compared.
Materials and Methods

Cell Lines. Transfected versions of the HLA-A,B,C negative EBV-transformed B lymphoblastoid cell line 721.221 expressing B*4601, B*1501, or Cw*0102 were generated as described previously (14). Cell surface expression of the transfected class I alleles was monitored on a regular basis by immunofluorescent staining and flow cytometric analysis as described (14).

Purification and Analysis of Naturally Processed Peptides Bound by HLA Class I Allotypes. Class I glycoprotein was purified and bound peptides isolated and characterized essentially as described in (15). Class I glycoproteins were purified from detergent lysates of 1 X 10^10 cells using the HLA class I-specific monoclonal antibody W6/32 (16). Separate batches of antibody were used to purify each allotype. Purity of the class I protein was assessed by SDS-PAGE and Coomassie blue staining, and yields quantified by densitometry based on comparisons to known amounts of a protein standard. Bound peptides were released by denaturation of class I glycoprotein with 10% acetic acid and isolated using a 3-kD molecular mass cut off ultra filtration unit. Peptides were separated by HPLC using a reversed-phase Beckman Ultrasphere C-18 2.0 X 250-mm column (Beckman Instruments Inc., Fullerton, CA) using the gradient described in (15). Peptide sequence analysis was performed using an automated protein sequencer (model 476A; Perkin Elmer Applied Biosystems Division, Foster City, CA) operated in gas-phase mode. Cys was not derivitized and therefore not detectable. Sequence analysis of the pool of peptides bound by each allotype was performed using 30% of each HPLC fraction collected between 15 through 65 min. Residue preferences at each position are grouped into the following categories: dominant; (yield of a single amino acid increases several fold over the previous cycle); strong (amino acids with more than a twofold increase over the previous cycle); weak (amino acids with an increase of between 1.5- and 2-fold over the previous cycle). Residues at each position are listed in order of decreasing size of the picomole increase relative to the previous cycle. HPLC fractions containing well resolved peaks from the absorbance trace at 214 nm were selected for sequence analysis of individual peptides. Yields of individual peptide sequences are based on initial sequence levels. Sequences assignments for all peptides isolated from B*4601 were confirmed by mass measurement and MS/MS experiments. Mass spectra were acquired using a PE SCIEX API 300 triple quadrupole mass spectrometer (Perkin Elmer Applied Biosystems Division, Foster City, CA) with an ionspray source operating in the positive ion mode. Proteins from which the peptides could be derived were identified by searching the GenBank database.

Molecular Modeling. Residues 1-180 of B*4601 were modeled using self-consistent ensemble optimization (SCEO) (17,18), via the program LOOK (Molecular Applications Group, Palo Alto CA). A multiple structural alignment of class I crystal structures B*2705, A*0201 and H-2D^b was performed, and B*4601 modeled primarily from B*2705 (for residues 38-65 and 103-151), but also A*0201 (1-37, 68-76, 152-180), A*0201 (66-67, 82-102) and H-2D^b (77-81). Residues modeled ab initio in the peptide binding site included 7, 9, 45, 63, 66, 67, 69, 70, 97, 99, 114, 116, 152, and 163. To model the proteasome component C8 peptide bound to B*4601, the coordinates of A*0201 complexed with a decamer peptide from hepatitis B (19) were superimposed on the B*4601 model, and the proteasome component C8 residues 150-159 (YMIDPSGVSY) modeled on the peptide backbone using SCEO.

Preparation of CTL Lines and NK Cell Clones, and Cell-mediated Cytotoxicity Assays. Allogeneic CTL lines were generated by coculture of human PBL from healthy donors with irradiated (10,000 rads) 721.221 transfected cells expressing B*4601, B*1501, or Cw*0102 (ratio of 5:1). CTL lines were cultured in IMDM medium supplemented with 10% FCS, 2% human serum, 2 mM L-Gln, 100 U/ml penicillin, 100 µg/ml streptomycin and maintained at 37°C in a humidified atmosphere containing 5% CO_2. Two weeks after cultures were established, cells were stained with Simultest CD3/CD8 (Becton Dickinson and Company, Mountain View, CA) and cells expressing CD3 and CD8 were collected using a FACStarPLUS® flow cytometer (Becton Dickinson). CTL lines were restimulated every 7 d and, after sorting, were maintained in medium additionally supplemented with 10% Lymphocult-T (BioTest). Allogeneic NK cell clones (CD3−, CD56+) were prepared as described (14). All CTL lines and NK cell clones were tested for expression of appropriate cell surface markers by immunofluorescent staining and flow cytometry. Cytolytic activity was determined in a standard 4 h chromium release assay and results expressed as percent specific lysis calculated as follows: percent specific lysis = (experimental cpm - spontaneous cpm)/(total cpm - spontaneous cpm) X 100. Assays were performed in duplicate or triplicate and repeated 2-5 times.

Release and Analysis of Oligosaccharides from HLA Class I Glycoproteins. Oligosaccharides were released from purified desalted class I glycoprotein using the GlycoPrep 1000, labeled at the reducing terminus using the Signal Labeling method, and de-sialylated by digestion with neuraminidase as described (20). Fractionation
Figure 1. Analysis of the naturally processed peptides bound by B*4601, B*1501 and Cw*0102. Bound peptides were isolated from 350 pg B*4601, 500 pg B*1501, and 700 pg Cw*0102, respectively and separated by reversed phase HPLC (a, c, and e). Peptide sequences were determined by Edman degradation and mass spectrometry (b, d, and f).

Results and Discussion

To determine how the inter-locus recombination between B*1501 and Cw*0102 that created B*4601 modified peptide binding specificity we characterized the naturally processed peptides bound by these allotypes. Reversed phase HPLC of the peptide pools showed B*1501 and Cw*0102 (Fig. 1, c and e) had typically "humped" profiles indicative of heterogeneous peptides (15, 21) whereas peptides eluted from B*4601 gave a flatter profile having a few dominant peaks (Fig. 1 a) which contained high abundance peptides (Fig. 1 b). Most abundant was a peptide derived from proteasome subunit C8, and its sequence dominated when the pool of peptides eluted from B*4601 was sequenced (Fig. 1 b). Two of the three most abundant peptides derive from constitutively expressed proteins (22, 23) consequently the peptides displayed by B*4601 in most cell types may be dominated by a few peptides.

In marked contrast to B*4601, the peptides isolated from the parental allotypes, B*1501 and Cw*0102, comprise a complex mixture containing no predominating peptides (Fig. 1, c-f). Pool sequencing revealed residue preferences at two positions defining their peptide-binding motifs (Fig. 1, d and f). None of six individual B*4601-binding peptides (Fig. 1 b) possess either the B*1501 (Gln at position 2, Phe or Tyr at position 9) or the Cw*0102 (Pro at 3, Leu at 9) motif. Neither were the B*4601-binding peptides found amongst the peptides eluted from B*1501 or Cw*0102. Preferential binding of a subset of peptides by B*4601 is almost certainly a result of its unique structure and not other factors affecting antigen presentation as the peptides we isolated from B*4601, B*1501 and Cw*0102 derive from the identical cellular background.

Modeling suggests three features of the B*4601 peptide binding site combine to augment peptide selectivity (Fig. 2). The HLA-C-derived sequence introduces a basic patch (Arg 62, Lys 66, and Arg 69) into the peptide-binding site which may account for selective binding of peptides with acidic side chains at position 4; a feature shared by the dominant peptides bound by B*4601 (Fig. 1 b). Examination of the subsites within the peptide-binding site designated pockets A through F suggests additional constraints on peptide selectivity by B*4601 may reflect a truncated B pocket due to Tyr 67 and Glu 63 which likely restricts res-
Figure 2. Molecular model of the B*4601 peptide binding site. (Top left) Ribbon diagram of the α1 and α2 domains of B*4601 showing the HLA-C-derived sequence (red) present at residues 66–76 of the α1 helix. (Top right) Molecular surface of B*4601 showing the basic groups Arg 62, Lys 66, and Arg 69 (blue) immediately adjacent to the Asp at position 4 of the proteasome component C8-derived peptide (peptide is shown in bonds). (Bottom left) Cut-away view showing the B pocket truncated by Tyr 67 (pink) creating a shallow depression for accommodating the peptide position 2 Met side chain (purple, ball and stick). (Bottom right) Cut-away view showing the deep F pocket due to Ser 116 (olive) accommodating the peptide COOH-terminal Tyr side chain (red, ball and stick).

Figure 3. Alloreactive CTL and NK cells recognize B*4601 as HLA-C-like. Results shown in A are percent specific lysis of untransfected 721.221 cells and 721.221 cells transfected with B*4601, B*1501 or Cw*0102 by CTL lines raised from one donor against B*4601 (left), Cw*0102 (middle) or B*1501 (right). Assays shown were performed at an effector to target ratio of 2.5:1 for anti B*4601 and anti B*1501 CTL lines and 5:1 for the anti Cw*0102 CTL line. Cross reactive recognition of B*4601 and Cw*0102 but not B*1501 was observed for CTL lines raised from all three donors tested. Results shown in Figure 3 B are percent specific lysis of untransfected 721.221 cells and 721.221 cells transfected with B*4601, B*1501, Cw*0102 (group I) or Cw*1503 (group II) by a group I-specific NK cell clone (left) and a group II-specific clone (right). Expression of B*4601 or Cw*0102 but not B*1501 protects cells from lysis by group I HLA-C-specific NK cells. B*4601 inhibited lysis by all group I but not group II HLA-C-specific NK cell clones isolated and tested from three healthy donors (total 31 clones).
idues permissible at peptide position 2, and a large deep F pocket, held in common with the B*1501 parent, which selects for peptides with carboxyl-terminal aromatic residues.

Despite sharing greatest homology to B*1501, the HLA-C-derived sequence at residues 66-76 dominates the interaction of B*4601 with both alloreactive CTL and NK cells. Most alloreactive CTL lines were found to exhibit cross reactive recognition of B*4601 and Cw*0102, but no cross-reactivity between B*1501 and B*4601 was observed (Fig. 3 a). Class I receptors of NK cells also view B*4601 as HLA-C-like. HLA-C allotypes can be divided into two groups based upon their capacity to inhibit lysis by distinct subsets of allogenic NK cells (24, 25). B*4601 inhibited NK cell clones having specificity for group I HLA-C allotypes such as Cw*0102, but B*1501 was not inhibitory for NK cells of either specificity (Fig. 3 b).

The HLA-C-derived sequence also exerts an overriding influence on the N-linked glycosylation of B*4601. HLA-B and C allotypes have distinctive patterns of N-linked glycosylation at position 86 (20), close to the HLA-C-like sequence possessed by B*4601. Structural analysis of the glycans of B*4601 showed the typically HLA-C pattern of four oligosaccharides and not the more restricted HLA-B pattern of two oligosaccharides (Fig. 4). Substitutions at residues 66-76 therefore make B*4601 an HLA-C-like substrate for glycosyl transferases.

The inter-locus recombination which generated B*4601 created a molecule with novel peptide-binding properties and imposed features characteristic of HLA-C onto an HLA-B allotype. These results demonstrate the radical changes in class I function that can be effected by a single gene conversion. That B46 haplotypes are common in south east Asia (5) and surpass the frequency of the parental B15-Cw1 haplotype (13) suggests the unusual properties of B*4601 confer selective advantage. Indeed the functional competence of B*4601 is indicated by the discovery of a healthy individual who is homozygous for the B46-Cw1 haplotype and lacks expression of HLA-A (26). In contrast, the association of B46 with predisposition to autoimmune disease (8-12) and nasopharyngeal carcinoma (6-7), a prevalent cancer in southeast Asia involving Epstein-Barr virus (27), suggest negative effects may also be associated with the unusual properties of B*4601. Presentation of high amounts of a few peptides by B*4601 may impose biases on the T cell repertoire (as reported for patients possessing B*4601 with nasopharyngeal carcinoma [28]) that strengthen or weaken T cell responses to individual antigens with an intensity not seen with less discriminating allotypes.

This research was supported by National Institutes of Health grant AI-22039 to Peter Parham. Linda D. Barber is a post-doctoral research fellow of the American Heart Association, California Affiliate. Nicholas M. Valiante is a post-doctoral fellow of the Cancer Research Institute of New York. Jenny E. Gumperz is supported by an NIH training grant. DNAX Research Institute for Molecular and Cellular Biology is supported by Schering Plough Corporation.

Address correspondence to Peter Parham, Department of Structural Biology, Stanford University School of Medicine, Stanford, CA 94305.

Received for publication 9 April 1996 and in revised form 12 June 1996.

References