Distinctive Polymorphism at the HLA-C Locus: Implications for the Expression of HLA-C

By Jacqueline Zemmour and Peter Parham

From the Departments of Cell Biology, and Microbiology and Immunology, Stanford University, Stanford, California 94305

Summary

The HLA-C locus remains an enigma. The serological polymorphism is poorly defined, HLA-C molecules are expressed at the cell surface at about 10% the levels of HLA-A and -B, and their importance for antigen presentation to either CD8-bearing T cells or natural killer cells is unclear. Our understanding of HLA-C polymorphism has also lagged behind that of HLA-A and -B. We have applied the polymerase chain reaction to the characterization of cDNA encoding HLA-C antigens. Combining the recent results with previously characterized HLA-C alleles gives a data base of 26 sequences, which was used to analyze the nature of HLA-C polymorphism and compare it to the variation seen in HLA-A and -B. The sequences form 10 families of alleles that correlate well with the patterns of serological crossreactivity, including the C blanks, and all major HLA-C allelic families appear to have been sampled. The families further divide into two groups of HLA-C alleles defined on the basis of linked substitutions in the 3′ exons. In comparison with HLA-A and -B, HLA-C alleles are more closely related to each other, there being less variation in residues of the antigen recognition site and more variation at other positions. In particular, the helix of the α_1 domain of HLA-C molecules is unusually conserved. Despite the reduced diversity in the antigen recognition site, it is evident that HLA-C genes have been the target of past selection for polymorphism. Within the antigen recognition site, it is the α_1 domain that is most diagnostic of HLA-C, whereas the α_2 domain is similar to that of HLA-B, the locus to which HLA-C is most closely related. In particular, conserved motifs in the α_1 helix and the conserved glycine at the base of the B pocket (position 45) provide a combination of features that is uniquely found in HLA-C molecules. We hypothesize that these features restrict the peptides bound by HLA-C molecules and in this manner reduce the efficiency of HLA-C assembly and expression at the cell surface. The overall picture of HLA-C polymorphism obtained from this sampling of HLA-C alleles is unlikely to change as further alleles are characterized.

HLA-A and -B heavy chains are polymorphic glycoproteins that associate noncovalently with β_2-microglobulin (β_2-m).[1] The function of these molecules is to bind and present antigenic peptides to CTL (reviewed in reference 1). Although many examples of antigenic peptides presented in an allele-specific manner by HLA-A or -B molecules have been described, the role of the homologous HLA-C molecules in the immune response is still poorly understood. Due to their weak immunogenicity and the lack of specific reagents, serological typing of the products of the HLA-C locus has been persistently difficult, and some 20% of HLA-C alleles type blank in most populations (2–5). Furthermore, HLA-C antigens are expressed on cell surfaces to a much lesser extent (~10%) than either HLA-A or -B (6–8), and exhibit heterogeneity in their carbohydrate moiety (8). In addition, analysis of HLA-C heavy chains by IEF reveals an unexplained heterogeneity that is not due to sialic acid (7, 9). HLA-C molecules are also distinguished by inefficient assembly with β_2-m and slower rate of exocytosis, which presumably contribute to the lower cell surface expression (10). These differences are not the result of lower levels of transcription or translation as HLA-C heavy chains are synthesized in amounts similar to HLA-A and HLA-B heavy chains (10, 11).

Analysis of alloreactive responses has shown HLA-C molecules can be recognized by the receptors of human CTL, both in in vitro culture (12–14) and in allograft rejection (15). Furthermore, EBV-specific cytotoxic T cells restricted by HLA-C have been described (16), and a peptide derived from the gag protein of HIV-1 was shown recently to be presented by HLA-Cw3 (17). The functional capacity of the Cw3 molecule has also been examined in transgenic mice, and examples of mouse CTL recognizing both viral peptides presented by

[1] *Abbreviation used in this paper:* β_2-m, β_2-microglobulin.

J. Exp. Med. © The Rockefeller University Press • 0022-1007/92/10/0937/14 $2.00
Volume 176 October 1992 937–950

Cw3, and Cw3 as an alloantigen, were found (18). These results demonstrate that HLA-C molecules can participate in the immune response, a concept also supported by the association of HLA-C antigens with susceptibility to disease (19–23). Güssow et al. (7), however, have suggested that "HLA-C may be dispensable for proper functioning of the immune system," and concerns as to the significance of the contribution of HLA-C to protective immunity compared with HLA-A and B remain (24, 25).

The feature that clearly distinguishes HLA-C from HLA-A and -B is its low cell surface expression, and in this regard, HLA-C is similar to the H-2L^d molecule (26), which is "expressed on the cell surface at levels three to four times lower than D^d or K^d" (27). In the case of H-2L^d, reduced expression correlates with slower transport to the cell surface, and weaker association with β_2-m (27). Analysis of chimeric molecules formed by exon shuffling of the H-2D^d and H-2L^d genes showed the sequence of the α_1 domain determined the level of cell surface expression (28). Furthermore, expression of H-2L^d at the surface can be increased fourfold by feeding cells with an appropriate peptide ligand (29). That all HLA-C alleles appear to give low expression at the cell surface (5, 10) suggests HLA-C heavy chains share sequence motifs that confer this property. In this paper, we analyze the sequences of HLA-C alleles to identify such features.

Molecular analysis of HLA-C polymorphism has been impeded by the difficulties in serological HLA-C typing and in purifying HLA-C proteins. Similarly, the isolation and molecular characterization of HLA-C alleles has lagged behind that of HLA-A and -B alleles. Application of the PCR to the specific cloning of HLA-C alleles circumvents these problems (30, 31) and has facilitated comparison here of 26 HLA-C alleles. This analysis has permitted determination of the nature of HLA-C polymorphism, its comparison with that of HLA-A and -B, and the identification of features that might determine low cell surface expression.

Materials and Methods

The isolation and characterization of HLA-C cDNA from human B cell lines was essentially as described by Ennis et al. (30), except that the 3′ oligonucleotide primer used in amplification by the PCR was derived from a region of the 3′ untranslated region that contained HLA-C-specific substitutions. The sequences of the two primers used in amplification were HLA-5P2: 5′ GCC CGT CGA CGG ACT CAG AAT CTC CCC AGA CGC CGA G 3′ (5′ primer), and HLA-3pC: 5′ CCG CAA GCT TTC GGG GAG GGA ACA CAG GTC AGT GTG GGG AC 3′ (3′ primer) (31). This strategy, which yields the entire coding region, has been used to isolate alleles encoding the previously uncharacterized Cw4 and Cw8 antigens and novel subtypes of Cw1 and Cw3.

Previously published HLA-C sequences used in this analysis were Cw*0101, Cw*0201 (7); Cw*0301 (32); Cw*0102, Cw*0302 (33); Cw*02021 (34); Cw*02022 (35); BeWo C.1 (36); Cl.9 and Cl.10 (37); CW6P, Cw*0701 (38); Cw*0501 and Cw*1301 (3); Cw*0601 (39); Cw*0702 (40); Cw*0801 and Cw*0802 (31); Cw*0803 (41); Cw*1201 (42); Cw*1202 and Cw*1401 (4); HLA-4 (43); and C*X (44). World Health Organization nomenclature for the alleles and their previous designations are described in references 45 and 46.

In comparison with the other 25 HLA-C alleles, Cw*0301 (the first C allele sequenced [32]) has an unusually high number of unique substitutions. For HLA-A and -B, a number of such substitutions have proved to be the result of errors or artifacts introduced in cloning and sequencing. Supporting this contention for Cw*0301 are the results we obtained from sequencing Cw3 from AP, a cell line of Korean origin (33). This allele (Cw*0302) differs from Cw*0301 by 16 nucleotide substitutions in the 1,101 base pairs of the coding region, which produce a 10–amino acid substitution (six in the α_1 domain and four in the α_2 domain). At all six of the substitutions in the α_1 domain, Cw*0301 has a unique residue, whereas Cw*0302 is identical to the consensus. Four of the positions occur in two pairs (40,41 and 54,55), which could be due to localized misordering of the sequence; one of these substitutions (55) would destroy the conserved salt bridge that links the α_1 and α_2 helices; a fifth substitution (69) places a proline in the α_1 helix providing additional reason for concern. Within the α_2 domain, the published Cw*0301 sequence placed a unique glycine at 181, which was subsequently corrected to the consensus residue arginine (47). Of the remaining three positions of difference, 95 and 116 are consistent with the general pattern of diversity of class I molecules and are those most likely to represent real differences. Asparagane 137 is unique to Cw*0301. The nucleotides at these positions in both Cw3 sequences are seen in other class I HLA sequences and two of the positions point into the peptide binding groove. Thus, in our analyses, the Cw*0301 sequence has not always been included.

Results

The sequences of 26 alleles of the HLA-C locus have now been determined. This base of data is the result of independent research from 15 laboratories (2, 3, 6, 7, 31–44, 48, 49), and in some instances, clones encoding HLA-C alleles have emerged from the study of immunological phenomena not immediately related to class I HLA molecules (36, 49). Determination of serologically well-characterized HLA-C alleles now permits identification of these sequences of unknown HLA-C alleles.

HLA-Cw4 Is Identical to the PL208 Clone Isolated on the Basis of Serological Crossreactivity with gp120. The Cw4-encoding allele Cw*0401 was characterized from the mutant B cell line C1R of Caucasian origin, which expresses Cw4 as the only serologically detected class I antigen (48). Subsequently, we and Watkins et al. (41, 44) cloned the identical Cw4 allele from cell lines derived from Native South Americans and poorly characterized for HLA-C alleles. Thus, the Cw*0401 allele is widely dispersed in human populations, and in the C1R mutant, the Cw4 gene has not undergone mutation.

Beretta et al. (49) identified a mAb (M38) with specificity for both the gp120 of HIV-1 and a surface protein of activated lymphocytes and monocytes. This group subsequently cloned a cDNA encoding the M38 antigen from a library made from PHA-activated lymphocytes of unknown HLA type (50). This cDNA "revealed a high degree of homology with the class I MHC gene family." The sequence being "not identical to any of the already sequenced alleles" but having maximal homology to alleles of the HLA-C locus. We discovered the sequence of this PL208 clone to be identical to Cw*0401. Grassi et al. (50) found the M38 antibody binds to cells from individuals of different HLA type, showing that

Cw4 is one of a number of class I molecules that bind the antibody (49). In their comparison of the PL208 sequence Grassi et al. (50) mistakenly labeled the HLA-4 sequence (43), which is closely related to Cw7, as Cw4, and thus from their analysis, it appeared PL208 was not identical to Cw4.

HLA-Cw4 Is Related to the BeWo C.1 Clone Isolated from a Choriocarcinoma Cell Line. The BeWo C.1 sequence reported by Ellis et al. (36), which was obtained from the "HLA-A,B,C-negative" choriocarcinoma cell line BeWo, is now seen to be most closely related to Cw*0401. Amino acid substitutions at five positions, 49, 50, and 68 in the α_1 domain and 155 in the α_2 domain, and 340 in the cytoplasmic domain, distinguish the two sequences (Fig. 1). Thus, BeWo C.1 appears to be a subtype of Cw4. Ellis et al. (36) found it "impossible to obtain a tissue type for BeWo cells," suggesting the amino acid differences between the subtypes may have an affect on serological determinants. Alternatively, the difficulties in typing this nonlymphoid cell may be due to low cell surface density of the class I molecules.

HLA-Cw8 and Cw3 Sequences. We have cloned alleles corresponding to the Cw8 antigen from cell lines of Caucasian and Native American origin, and partial sequences of two alleles were previously reported (33) in an analysis of the serological properties of the HLA-B46 molecule. Three subtypes have now been defined, which differ by one to four amino acid substitutions. The Caucasian subtype of Cw8 (Cw*0802) differs by just a single amino acid from the HLA-C sequence reported by Bronson et al. (51), and which was derived from a yeast artificial chromosome (YAC) library made from the CGM1 cell line. The Cw*0801 subtype, which was isolated from a Native North American, is identical to the sequence that we erroneously thought corresponded to the Cw11 antigen (33, 52). The HLA-Cw11 antigen is now known not to define a unique allele but an epitope shared by B46 and Cw1 molecules. The third subtype Cw*0803 was obtained from a Native South American of the Kaingang tribe (41).

Two Groups of HLA-C Sequences. Comparison of the 26 HLA-C sequences reveals a group of three that stand out from the rest: Cw*0701, HLA-4, and Cw*0702 (previously called JY328 [40]). The HLA-4 and Cw*0702 sequences were derived from serologically uncharacterized alleles (40, 43), but their striking sequence similarity to HLA-Cw*0701 shows they comprise a group of Cw7 subtypes. Frequency histograms representing the distribution of nucleotide differences between pairs of HLA-C alleles clearly show the Cw7 sequences form a distinct group (Fig. 2), and this can also be seen with a dendogram calculated using the unweighted pair-group method using arithmetic averages (53) from the HLA-C sequences (Fig. 3). Characteristic of the Cw7 group is a series of 16 nucleotide substitutions spread through exons 4–8, which produce 11 amino acid substitutions (Fig. 1). These substitutions are common to Cw*0701, Cw*0702, and HLA-4, but are found in no other HLA-C sequences. An additional three nucleotide substitutions, of which two are coding changes, are unique to two of the three sequences. The independent determination of the JY328, HLA-4, and Cw7 sequences in different laboratories (38, 40, 43) gives confidence to the validity of this unique set of linked substitutions; furthermore, they have also been found in gorilla homologues of HLA-C (54). HLA-Cw7 is the most common HLA-C antigen in many human populations (2), having a frequency of ~40% in Caucasians, for example (22). Thus, the representation at the population level of the unusual Cw7 sequence is much greater than appreciated from simple comparisons in which alleles are given equal weight irrespective of frequency (55).

Locus-specific substitutions are found in the 3′ exons and the 3′ untranslated regions of class I HLA genes (43, 52, 56, 57), suggesting the Cw7 subtypes may represent a locus distinct from that encoding the other HLA-C antigens. Indeed, an RFLP analysis by Duceman et al. (58) suggested that was the case for the JY328 gene. A more recent study, however, favors the interpretation that all HLA-C antigens are encoded by the alleles of a single locus (59). In that case the linked substitutions in the Cw7 alleles may have originated from another locus and been introduced into a C allele by gene conversion.

Correlation of Structure and Serology. For the most part, the structural relationships defined by the dendogram (Fig. 3) are consistent with the patterns of HLA-C serological cross-reactivity. All alleles serologically typed as the same antigen group together, for example, the various subtypes of the serologically defined antigens. Similarly, the Cw5 and Cw8 antigens that are serologically crossreactive (60) are found to be most closely related in sequence. An exception to such correlations is provided by the Cw4 and Cw6 antigens, which are serologically crossreactive (61) but quite divergent in se-

Leader Peptide

```
CONSENSUS  MRVMAPRTLILLLSGALALTETWA
  Cw*0101  ------------------------
  Cw*0102  ------------------------
  Cw*0201  ------------------------
 Cw*02021  ---------L--------------
 Cw*02022  ....................-----
  Cw*0301  ------------------------
  Cw*0302  ------------------------
  Cw*0401  ------------------------
 BeWo C.1  ----E-------------------
  Cw*0501  ------------------------
  Cw*0601  ------------------------
     CW6P  ....................X---
  Cw*0701  ........................
  Cw*0702  ...------L---LAPR---Q---
    HLA-4  ........................
  Cw*0801  ------------------------
  Cw*0803  ------------------------
  Cw*0802  ------------------------
     CGM1  ........................
  Cw*1201  ........................
  Cw*1202  ------------------------
  Cw*1301  ------------------------
  Cw*1401  ------------------------
    Cl.10  ------------------------
     Cl.9  ........................
      C*X  ........-L--------------
CONSENSUS  MRVMAPRTLILLLSGALALTETWA
```

continued on next page

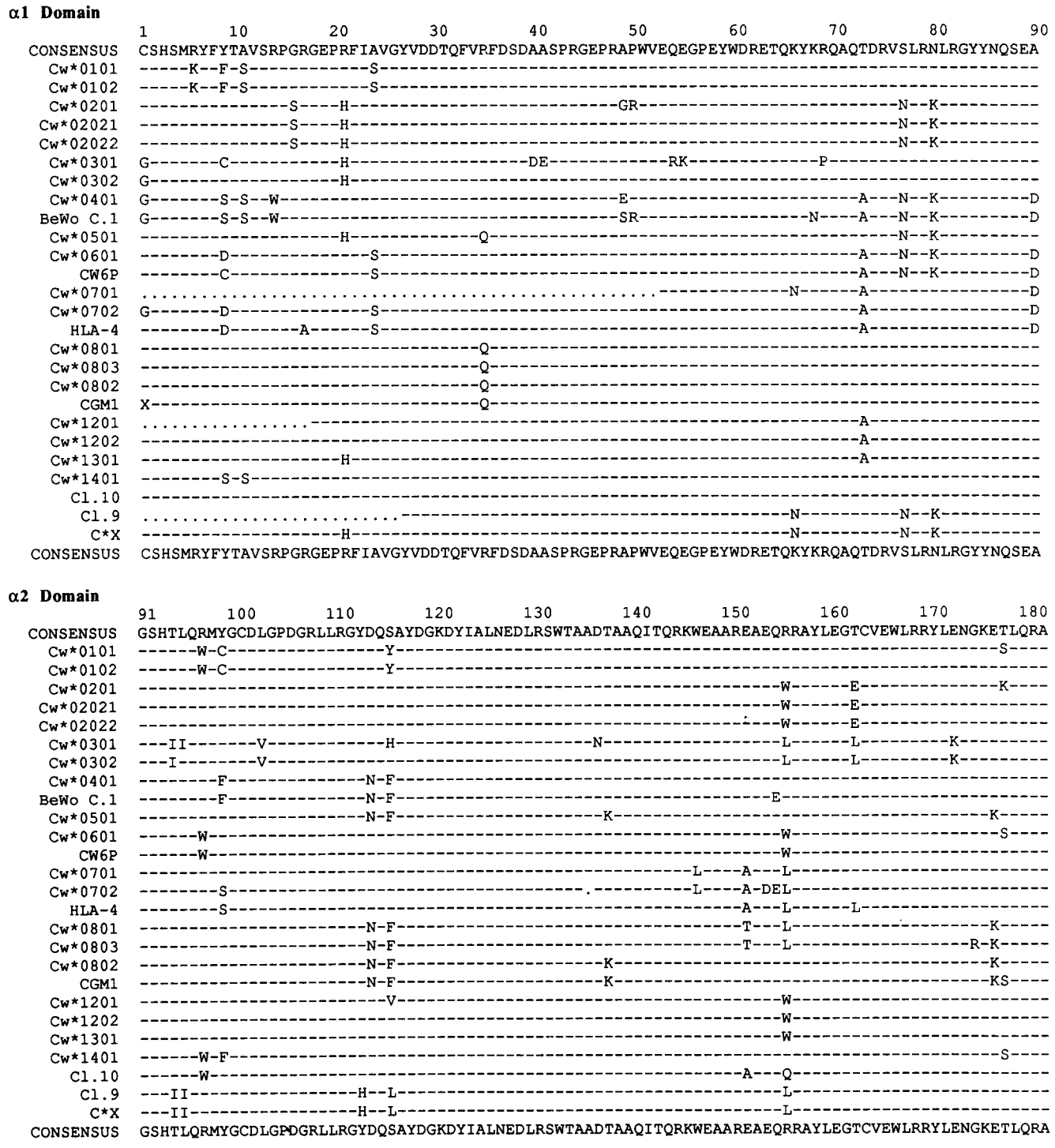

Figure 1. *(continued)*

quence. Amino acid substitutions that correlate with known serological reactivities are shown in Table 1. The Cw1, Cw2, and Cw3 proteins have the greatest numbers of unique substitutions, probably explaining why they were most easy to define by serology.

Approximately 20% of HLA-C alleles cannot be serologically typed (2), and structures for a number of these blank alleles have been determined (3, 4, 42). That designated as Cw*1401 (Cb1) groups with the Cw1 subtypes, while three others, Cw*1201 (Cx52), Cw*1202 (Cb2), and Cw*1301

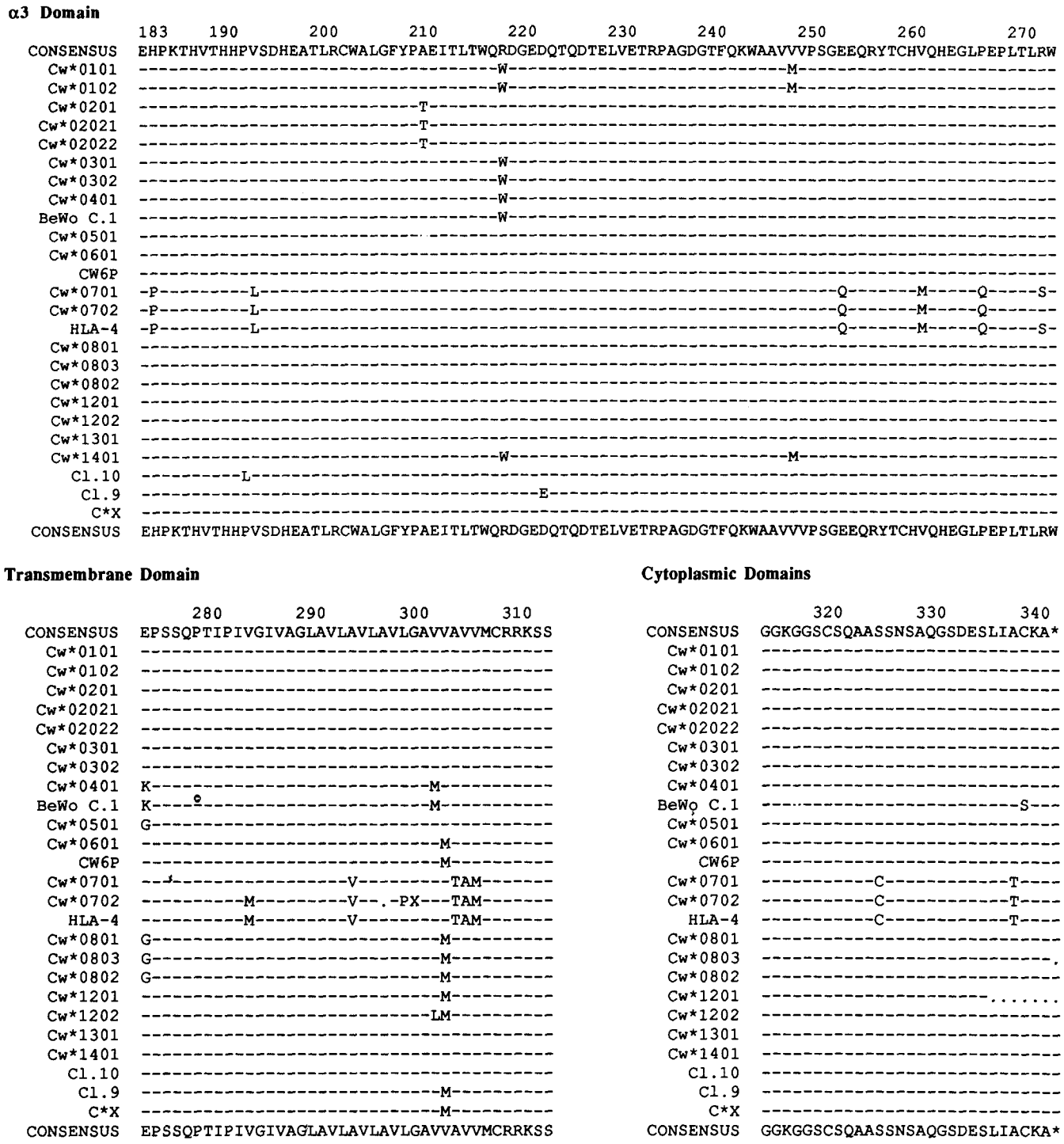

Figure 1. Amino acid sequences of the heavy chains encoded by 26 HLA-C alleles. WHO nomenclature is used for those sequences assigned names. A compete listing of previous names can be found in references 45 and 46.

(CwBL18), form a separate group with similarity to the Cl.10 and Cw6 sequences (Fig. 3). That Cl.10 groups with Cw6 is not surprising, as Cianetti et al. (37) typed the GM637 fibroblast cell line from which this clone was derived as either Cw6 or Cw7. The C1.9 and Cw*X sequences form separate group that is not closely associated with any serologically defined antigens. That the blanks assort into one of two groups indicates that most, if not all, of the major allelic motifs at the HLA-C locus have been defined.

Nature of HLA-C Polymorphism. The total number of poly-

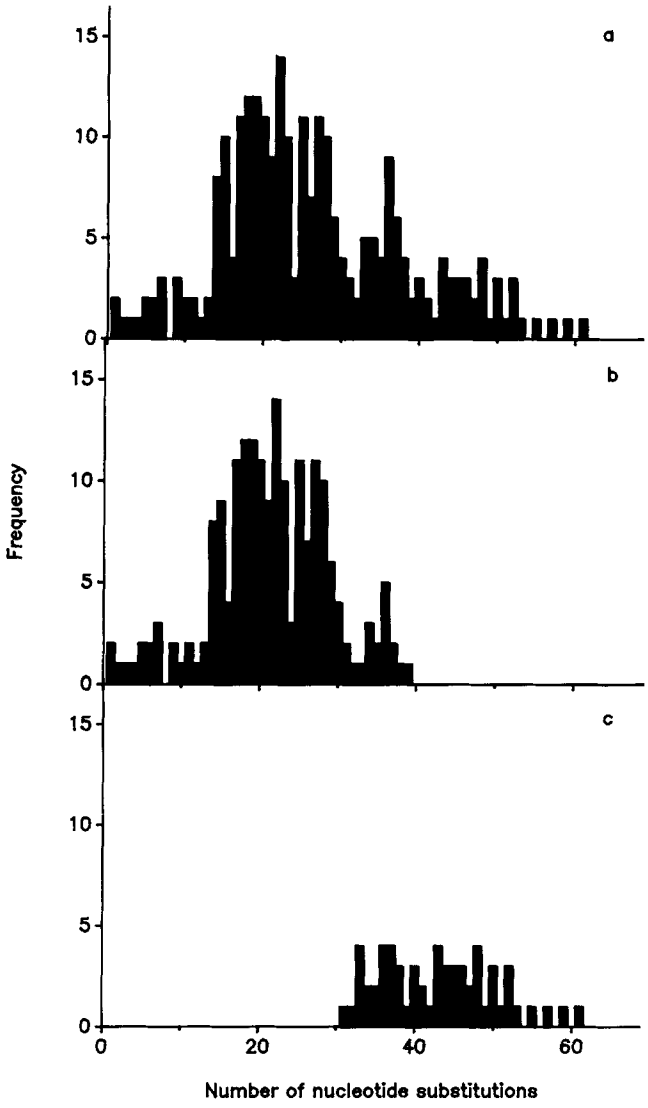

Figure 2. Two groups of HLA-C alleles. The number of nucleotide substitutions between pairs of HLA-C alleles were calculated. Histograms of the frequency of these differences are shown. (*a*) Distribution when all 26 HLA-C alleles are included. (*b*) Distribution when the group of three Cw7 subtypes are omitted from the analysis. (*c*) Distribution of differences when one member of the pair is from the Cw7 group and when the other is not.

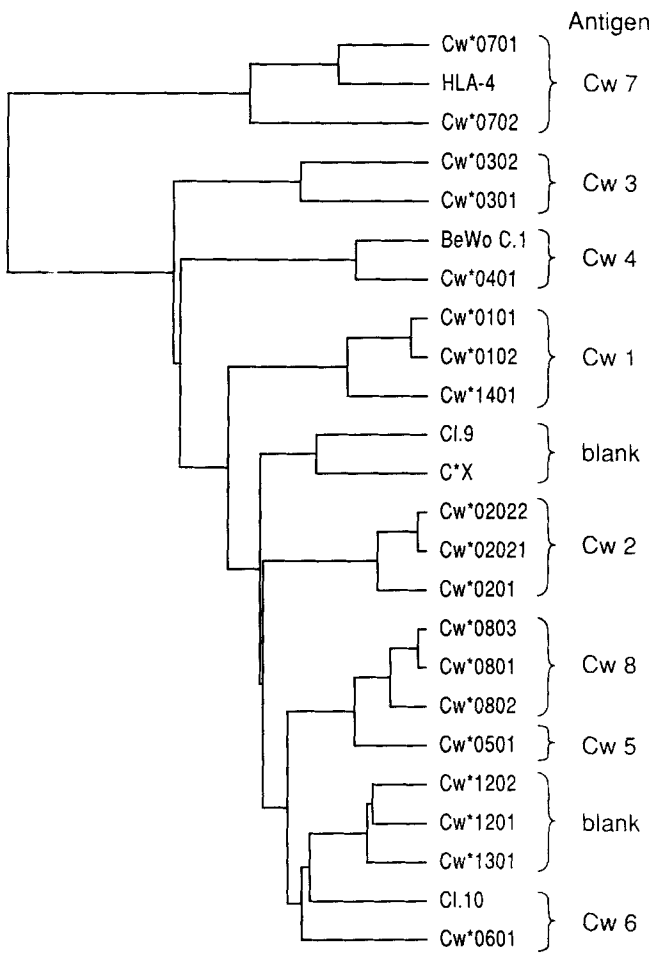

Figure 3. A dendogram constructed from the nucleotide sequences of the coding regions of HLA-C alleles, using the unweighted pair group method using arithmetic averages (48). The serological antigens corresponding to the different alleles are shown on the right.

morphic positions is comparable for HLA-B and -C, with HLA-A being somewhat higher. However, when just the α_1 and α_2 domains are considered, the number is significantly less for HLA-C than for either HLA-A or B (Table 2). This trend is even stronger if just the functional positions of the antigen recognition sites are considered Thus, in comparison with HLA-B, the locus to which it is most closely related in evolution, HLA-C has less variation in the antigen recognition site and more variation elsewhere.

A second measure of relative variation was to compare the range and distribution of nucleotide differences between pairs of HLA-A, -B, and -C alleles. Although the range of differences between alleles at all three loci is remarkably similar, the median and mode are at significantly lower values for the HLA-C distribution (Fig. 4). The part of the HLA-C distribution at higher values, >40, is entirely due to the Cw7 group (Figs. 2, 4). Thus, HLA-C alleles are, on average, more

Table 1. *Serological Epitopes*

Potential epitope specificity	Amino acid substitutions
Cw1	K6, F9, C99
Cw2	S16, T211, W156, E163
Cw3	L156, L163, K173
Cw4	S9, S11, W14, E49
Cw7	L147, A152, L156
Cw5 × Cw8	Q35, K138, K177
Cw4 × Cw6	A73, N77, K80
Cw6 × Cw7	D9, S24

Table 2. *Numbers of Polymorphic Positions*

	α1 domain	α2 domain	α3 domain	External domains	T.M domain	Cyt. domains	All domains	α1 α2 domains	Functional positions
HLA-A	19*	23	8	50	6	2	58	42	31
HLA-B	23	20	2	45	3	1	49	43	27
HLA-C	13	17	8	38	7	2	47	30	15

Total number of functional positions = 54.
* Number of residues that show some polymorphism.

similar to each other than are HLA-A or -B alleles. This difference is more pronounced when just the domains (α_1 and α_2) forming the antigen recognition site are analyzed in this fashion (Fig. 5). The allelic differences are greatest for HLA-B, intermediate for HLA-A, and lowest for HLA-C.

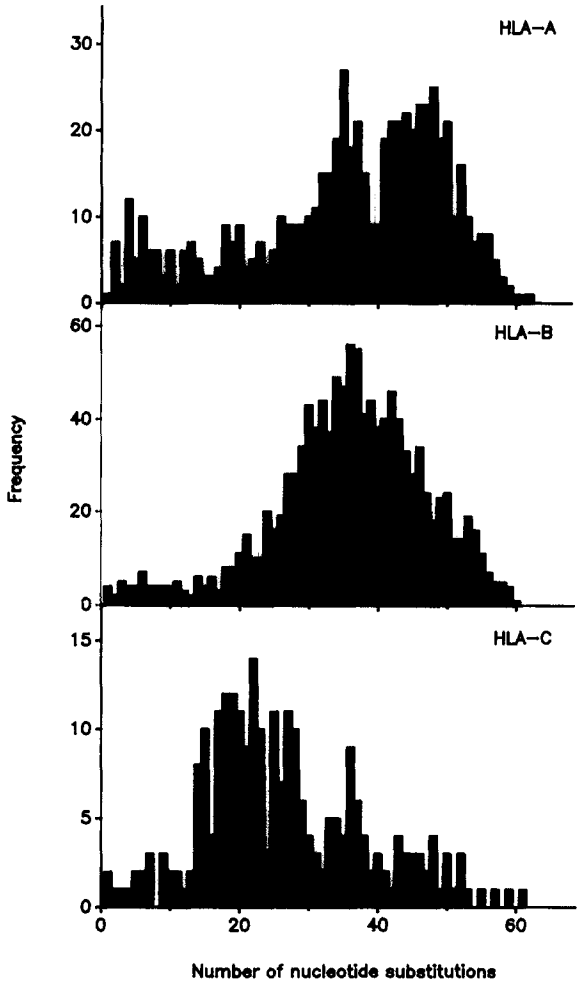

Figure 4. Pairwise comparison of nucleotide differences between alleles of the HLA-A (*top*), HLA-B (*middle*), and HLA-C (*bottom*) loci. For each locus the number of nucleotide substitutions between all pairs of alleles are calculated and their frequency distributions are shown. The sequences of the entire coding regions are compared.

For HLA-A, -B, and -C, analysis of sequence variability by the method of Wu and Kabat (62) shows that positions of high variability (equal to or greater than four) are, with one exception, at functional positions of the antigen recognition site. 11, 15, and 7 positions of high variability are found for HLA-A, -B, and -C, respectively (Table 3). Moreover, position 49 in HLA-C is the exceptional position of high variability that is not in the antigen recognition site. Also ap-

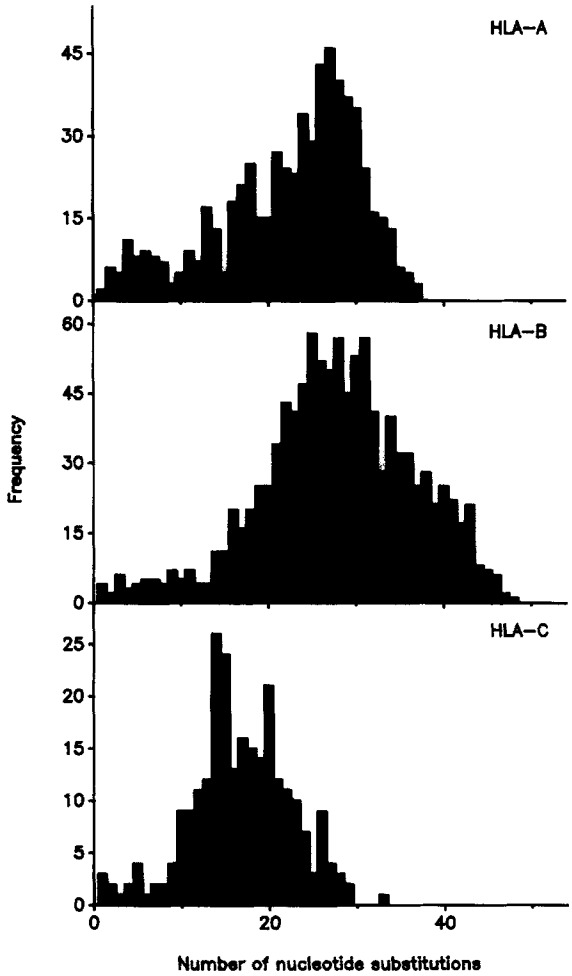

Figure 5. A similar analysis to that shown in Fig. 4, with the difference that only the sequences of exons 2 and 3 encoding the α_1 and α_2 domains are compared.

Table 3. *Positions of High Variability (>4.0) in HLA-A, -B, and -C Molecules*

Position	Location on structure	Variability			Potential contact
		HLA-A	HLA-B	HLA-C	
α1 domain					
9	β1	11	4	9	Peptide
24	β2	*	7		Peptide
45	β4	*	15	*	Peptide
49	Between β4 and short α helix	*	*	5	Silent?
62	α helix	13		*	TcR + peptide
63	α helix	4		*	Peptide
67	α helix		13	*	Peptide
69	α helix	*	4	*	TCR
70	α helix		6	*	Peptide
76	α helix	5		*	TCR
77	α helix	5	5		Peptide
80	α helix		5		Peptide
82	α helix		4	*	TCR
α2 domain					
95	β1	5	7		Peptide
97	β1	7	13		Peptide
99	β1			6	Peptide
114	β2	10	5		Peptide
116	β2	5	17	12	Peptide
152	α helix	7		4	Peptide
156	α helix	8	5	11	Peptide
163	α helix		6	4	TCR + peptide

Wu and Kabat variability (62) was calculated for 24 HLA-A, 24 HLA-B, and 24 HLA-C sequences. The variability at each position in a set of homologous sequences is defined as the number of different amino acids found at the position divided by the frequency of the most common residue.
* Conserved position.

parent is that eight positions that show high variability in either HLA-A or -B are conserved in all HLA-C sequences. These residues focus on the α_1 helix and result in this element of the antigen recognition site being highly conserved in HLA-C molecules. This property undoubtedly contributes to the difficulty in the serological definition of HLA-C alleles, as variation in the α_1 helix produces many of the epitopes that distinguish HLA-A and -B molecules (63).

Crystallography has defined six specificity pockets within the binding groove that accommodate peptide side chains and terminal groups (64–67). Overall, the sequence motifs found in these pockets of HLA-C molecules are distinct from those in HLA-A and -B. The peptides bound by HLA-C are therefore expected to be distinct. Although exhibiting a small number of positions of high variability, there is at least one such position in each of the specificity pockets for HLA-C (Fig. 6).

Analysis of the interaction between HLA-B27 and peptide shows the B pocket plays a critical "anchoring" role in binding the arginine at position 2, which is common to B27-binding peptides (68–70). In particular, the negatively charged glutamic acid at position 45 at the base of the pocket forms an electrostatic interaction with the arginine side chain of the peptide. Contrasting with HLA-B molecules for which the B pocket is highly diversified, the B pocket of HLA-C molecules is conserved. Of significance is that residue 45 is a conserved glycine that may contribute little to peptide interaction. Thus, the anchoring role for the B pocket in HLA-B could be absent or attenuated in HLA-C.

In the extracellular domains there are 12 positions at which HLA-C shows variability and where HLA-A and/or -B are conserved. None of these is in the antigen recognition site. Thus, no residues of the peptide binding groove are uniquely variable in HLA-C. In contrast, of 24 residues that are con-

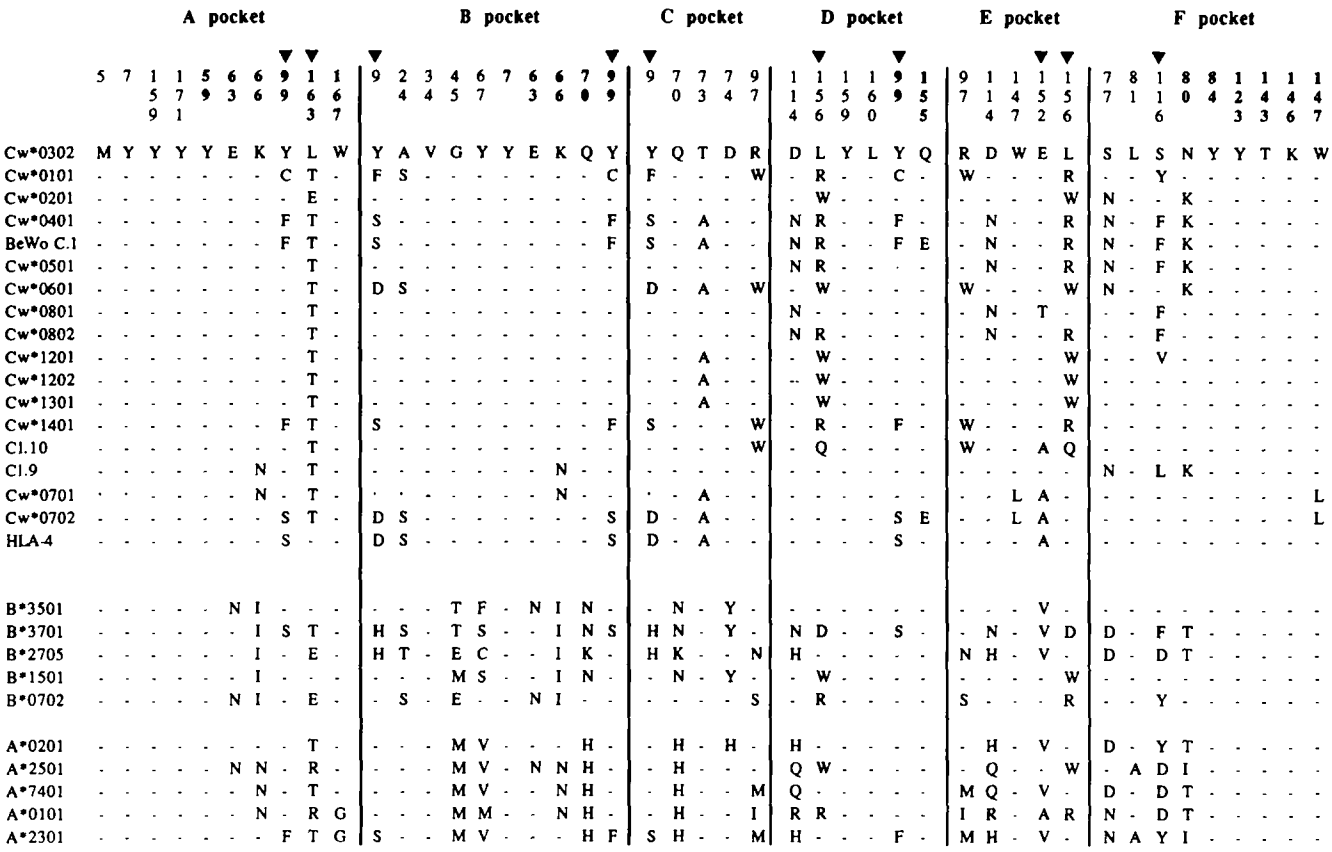

Figure 6. Comparison of the amino acid sequence motifs in the six specificity pockets of the antigen recognition site defined by Saper et al. (67). Positions with variability >4.0 are shown (▼). HLA-A and -B sequences are from references 30, 31, 38, 52, and 85–87.

served in HLA-C and variable in HLA-A and/or -B, 10 are residues that contribute to the specificity pockets of the binding groove.

The distribution of silent (synonymous) and replacement (nonsynonymous) nucleotide substitutions in HLA-C sequences was determined as described by Nei et al. (71, 72). In a previous analysis of 10 HLA-A and 6 HLA-B sequences, Hughes and Nei (73) showed replacement substitutions were nonrandomly focused on residues of the antigen recognition site, thus providing quantitative evidence for selection for the

Table 4. *Pattern of Nucleotide Substitutions within HLA Sequences*

	Antigen recognition site (n = 54)		α1-α2 (remainder) (n = 128)		α3 (n = 92)	
	d_S	d_N	d_S	d_N	d_S	d_N
C vs. C						
(21)*	2.3 ± 1.4	6.9 ± 1.1	5.2 ± 1.2	1.8 ± 0.4	3.4 ± 1.3	1.2 ± 0.4
(17)	2.2 ± 1.4	6.6 ± 1.1	4.7 ± 1.1	1.7 ± 0.4	2.9 ± 1.2	1.0 ± 0.3
B vs. B						
(31)	4.8 ± 2.0	15.1 ± 1.6	4.8 ± 1.2	1.8 ± 0.4	2.1 ± 1.1	0.2 ± 0.2
A vs. A						
(31)	3.8 ± 1.7	12.5 ± 1.6	3.3 ± 1.0	1.3 ± 0.4	6.6 ± 1.9	1.5 ± 0.5

Mean number of nucleotide substitutions per 100 synonymous sites (d_S) and per 100 nonsynonymous sites (d_N). n = Number of codons compared. d values are estimated using the Nei and Gojobori method (71). SEM d_S and d_N are estimated by Nei and Jin's method (72).
* Number of sequences compared.

polymorphism. From analysis of 31 HLA-A and 31 HLA-B sequences, we obtained results similar to those previously reported (Table 4). Analysis of 21 HLA-C sequences by this method shows the same trend as seen with HLA-A and -B, namely that residues of the antigen recognition site have a predominance of coding (nonsynonymous) substitutions, whereas noncoding (synonymous) substitutions predominate elsewhere. At nonfunctional positions of the α_1 and α_2 domains, the frequency of synonymous and nonsynonymous substitutions is comparable to that seen for HLA-C. In contrast, at the functional positions, the frequency of both types of substitution in HLA-C is about half that seen for HLA-B. Whereas the frequency of synonymous substitutions is similar in the functional and nonfunctional positions for HLA-A and -B, for HLA-C there is a relative suppression of such substitutions in the functional positions. Thus, there appear to be at least two factors operating on the functional sites of HLA-C antigen recognition site: one to select for amino acid substitution and the other to maintain nucleotide sequence homogeneity.

Dividing the functional positions into the α_1 and α_2 domains shows suppression of substitutions is focused on the α_1 domain (Table 5). Further subdivision into the six specificity pockets (67) reveals that in the A, B, and D pockets, there is little evidence for selection for amino acid diversity in HLA-C molecules. The C pocket is unusual in that silent substitutions are absent. The sequence motifs found for HLA-C molecules in the C pocket are not found in HLA-A or -B molecules (*Fig. 6*).

Shared Features of HLA-C Molecules. HLA-C molecules share the property of low cell surface expression, despite levels of heavy chain production comparable to HLA-A and -B (10, 11, 74). Therefore, common features of the HLA-C sequences, not shared with HLA-A or -B heavy chains, must underlie this behavior. Residues conserved in all HLA-C sequences but not found in HLA-A or -B sequences provide candidates for such features.

No residues of the α_1 and α_2 domains strictly fit this criterion. In particular, the α_2 sequences of HLA-C are very similar in their sequence motifs to those found in HLA-B α_2 domains. Characteristic C features exist in α_1, but exceptional HLA-B alleles also share some of these features. Glycine 45 and valine 52 are common to all HLA-C and to HLA-B54, but are found in no other HLA-A or B heavy chains (75). Similarly, the KYRV motif of residues 66, 67, 69, and 76 of the α_1 helix is conserved in HLA-C (proline 69 in Cw*0301 is probably an error) and absent in all HLA-A and -B molecules except B46 (reference 33). B54 and B46 both have clusters of substitutions that shared with HLA-C alleles and were formed by gene conversions between Cw1 and, respectively, B55 and B62 (33, 75). As both these B alleles exhibit normal expression, neither glycine 45, valine 52, nor the α_1 helical motif is sufficient to reduce expression. However, it is possible that the combination of these features is important.

Three positions in the α_3 domain, E183, G239, and E268,

Table 5. *Pattern of Nucleotide Substitutions in the Antigen Recognition Site of HLA Molecules*

	Pocket A		B		C		D		E		F	
	d_S	d_N	d_S	d_N	d_S	d_N	d_S	d_N	d_S	d_N	d_S	d_N
HLA-A	2.8 ± 4.4	12.2 ± 4.5	6.3 ± 6.5	15.0 ± 4.8	4.2 ± 9.0	27.8 ± 9.2	9.0 ± 8.5	16.7 ± 6.4	nd	33.9 ± 11.0	0 ± 0	11.0 ± 4.5
HLA-B	9.3 ± 9.6	15.2 ± 5.2	16.5 ± 10.0	31.9 ± 7.4	2.8 ± 5.0	31.6 ± 10.2	17.0 ± 13.0	15.2 ± 6.7	12.4 ± 13.2	29.7 ± 8.9	1.1 ± 2.6	18.6 ± 6.2
HLA-C	8.9 ± 9.3	8.1 ± 3.4	6.9 ± 7.1	8.5 ± 3.7	0 ± 0	13.4 ± 7.1	13.9 ± 13.3	17.2 ± 6.6	6.1 ± 13.1	22.1 ± 8.2	2.9 ± 3.4	10.6 ± 4.7

	α1 domain		α2 domain	
	d_S	d_N	d_S	d_N
HLA-A	3.5 ± 2.3	11.7 ± 2.4	4.2 ± 2.9	13.6 ± 2.6
HLA-B	4.0 ± 2.6	16.3 ± 2.7	5.8 ± 3.6	14.1 ± 2.4
HLA-C	1.0 ± 1.0	4.8 ± 1.5	3.8 ± 2.8	9.4 ± 2.0

Values are the mean number of nucleotide substitutions per 100 synonymous sites (d_S) and per 100 nonsynonymous sites (d_N) for the six pockets forming the antigen recognition site (*top*) and all positions of the α_1 and α_2 domains (*bottom*).

distinguish HLA-C from HLA-A and -B. Position 239 is part of a loop that joins two β strands, nos. 4 and 5, which are the sites of interaction with β_2-m (64, 67). The size and charge differences between the glycine at this position in HLA-C and the arginine found in HLA-B could have an influence on the association with β_2-m. In HLA-A2, arginine 181 forms a salt bridge with aspartate 183 and this combination of residues is conserved in HLA-B. Substitution of glutamate for aspartate at 183 in HLA-C might perturb this salt bridge and the interaction between the α_2 and α_3 domains. Glutamate 268, which is located on the seventh β strand of the α_3 domain, is surrounded by two proline residues. In all HLA-A and -B heavy chains this glutamate is replaced by the oppositely charged lysine. In the vicinity is glutamate 264, which contacts threonine 182 in the HLA-A2 structure (67).

Within the cytoplasmic domains are four HLA-C-specific residues. These comprise cysteine 321, asparagine 328, glutamate 335, and isoleucine 338. Of these, cysteine 321 replaces a tyrosine in HLA-A and -B molecules that can be phosphorylated in vitro (76). These substitutions could possibly affect interactions with cytoplasmic proteins or other membrane components.

Discussion

In comparison with HLA-A and -B, the HLA-C locus has been studied less and remains poorly understood. Here, we present the first extensive analysis of the polymorphism of HLA-C alleles. It reveals that the serological description of HLA-C antigens generally reflects the underlying structures of the alleles, which are variations on 10 basic motifs. Although some 20% of HLA-C alleles cannot be defined by serological HLA typing, these form two groups, suggesting the blank does not encompass a multitude of undiscovered alleles. Overall, there is good indication that the 26 alleles analyzed here are representative of the HLA-C alleles to be found in human populations and that few, if any, distinctive motifs remain to be found.

The hallmark of class I antigen-presenting molecules is their polymorphism, which crystallography has shown is concentrated in functional positions of the antigen recognition site. Pairs of HLA-C alleles show fewer differences than their HLA-A and -B homologues, and this property is reflected in a greater homogeneity of the antigen recognition site. Conservation of sequence in the α_1 helix and in the B pocket is particularly striking.

Despite similar levels of mRNA and heavy chain protein (10, 11, 74), cell surface expression of HLA-C molecules is $\sim$10% that of HLA-A and -B. This appears to be true for all HLA-C alleles, and one goal of our analysis was to identify features shared by HLA-C molecules that could account for this distinguishing property. Candidates are found in the α_1, α_3, and cytoplasmic domains, but it is those of the α_1 domain that are most impressive. These residues largely overlap with the conserved residues of the α_1 helix and the B pocket, raising the possibility that inefficient assembly and cell surface expression of HLA-C molecules is primarily a failure in the binding of peptides. This could be due to an intrinsic property of the antigen binding groove of HLA-C molecules, which have low affinity for the majority of peptides generated from endogenous proteins. The conserved features of the pockets formed by the α_1 domain could be instrumental in producing such binding properties. As a consequence, HLA-C may selectively bind a restricted set of peptides, perhaps those more commonly in the sequences of foreign, rather than self-proteins. Selectivity in the proteases that produce antigenic peptides or in the peptides delivered to the endoplasmic reticulum (ER) could act to restrict the supply of HLA-C binding peptides in the ER. The murine molecule H-2L^d exhibits inefficient cell surface expression, which can be enhanced by external provision of appropriate peptides (29). It will be important to see if the same is true for HLA-C and if viral infection acts to increase HLA-C expression through this mechanism.

Low cell surface expression of HLA-C may not result from a deficiency in binding peptides but from interactions of the HLA-C heavy chain with other components that contribute to the assembly and transport of class I molecules. For example, if HLA-C were to have a weaker interaction with β_2-m, or a stronger interaction with the heavy chain binding chaperonin (77), than HLA-A or -B, then expression would be predicted to be reduced.

Previous analysis of class I HLA genes and pseudogenes has correlated function with polymorphism. Thus, the HLA-A and -B genes, which function in antigen presentation to T cells, are highly polymorphic, whereas the pseudogenes, HLA-H and J, are conserved (25, 78, 79). HLA-C is intermediate in character both in level of expression and diversification. This property led to suggestions that HLA-C is less functional than HLA-A or -B, a dispensable or declining locus (7, 24, 25). Alternatively, HLA-C could be a more recently formed locus that is still undergoing improvement.

From the evidence available it seems clear that HLA-C is not a defunct class I locus. HLA-C molecules bind peptides and interact with TCRs in a variety of experimental systems (12–18). Moreover, HLA-C alleles are correlated with susceptibility to diseases believed to be autoimmune in nature (*Psoriasis vulgaris* for example) (21, 80).

The low cell surface expression may represent an adaptation giving HLA-C molecules complementary functions to those of HLA-A and -B. For example, by not being saturated with endogenous peptides, HLA-C may be poised to present particular viral peptides. The low expression would be expected to influence thymic selection, and in this manner may enable distinctive sets of T cells (high affinity, perhaps) to be selected by HLA-C. In this regard, it is worth noting that residues of the α_3 domain important for CD8 interactions (81, 82) are preserved in HLA-C molecules, although a direct demonstration for the interaction between HLA-C molecules and CD8 has yet to be made.

Recent studies on NK cell specificity have implicated

HLA-C molecules in the negative regulation of allotypic NK clones (83). In this regard, the unusual features of the C pocket of HLA-C are of particular interest, as HLA-A2 has been shown to poorly inhibit NK cells due to a unique histidine at position 74 in the C pocket (84). Moreover, Asahina et al. (80) have shown that susceptibility to psoriasis correlates with the presence of alanine at position 73 in the C pocket of HLA-C molecules.

We thank Drs. M. Nei and T. Ota for providing the computer program "Nag" for the estimation of the d_S and d_N values, and Victoria Ogbuehi for preparation of the manuscript.

This research was supported by National Institutes of Health grant AI-24258, and a Shannon Award from the U.S. Public Health Service.

Address correspondence to Peter Parham, Department of Cell Biology, Sherman Fairchild Building, Room D-157, Stanford University, Stanford, CA 94305.

Received for publication 15 May 1992.

References

1. Townsend, A., and H. Bodmer. 1989. Antigen recognition by class I–restricted T lymphocytes. *Annu. Rev. Immunol.* 7:601.
2. Baur, M.P., M. Neugebauer, and E.D. Albert. 1984. Reference tables of two-locus haplotype frequencies for all MHC marker loci. *In* Histocompatibility Testing. E.D. Albert, M.P. Baur, and W.R. Mayr, editors. Springer-Verlag, Berlin. 677–755.
3. Tibensky, D., R. DeMars, E.W. Holowachuk, and T.L. Delovitch. 1989. Sequence and gene transfer analysis of HLA-CwBL18 (HLA-C blank) and HLA-Cw5 genes: implications for the control of expression and immunogenicity of HLA-C antigens. *J. Immunol.* 143:348.
4. Takiguchi, M., I. Nishimura, H. Hayashi, S. Karaki, A. Kariyone, and K. Kano. 1989. The structure and expression of genes encoding serologically undetected HLA-C locus antigens. *J. Immunol.* 143:1372.
5. Hajek-Rosenmayr, A., L. Jungl, M. Stammler, and M. Kirnbauer. 1989. HLA-C "blank" alleles express class I gene products. Biochemical analysis of four different HLA-C "blank" polypeptides. *Immunogenetics.* 30:399.
6. Sodoyer, R., M. Damotte, T.L. Delovitch, J. Trucy, B.R. Jordan, and T. Strachan. 1984. Complete nucleotide sequence of a gene encoding a functional human class I histocompatibility antigen (HLA-CW3). *EMBO (Eur. Mol. Biol. Organ.) J.* 3:879.
7. Güssow, D., R.S. Rein, I. Meijer, W. de Hoog, G.H.A. Seemann, F.M. Hochstenbach, and H.L. Ploegh. 1987. Isolation, expression, and the primary structure of *HLA-Cw1* and *HLA-Cw2* genes: evolutionary aspects. *Immunogenetics.* 25:313.
8. Snary, D., C.J. Barnstable, W.F. Bodmer, and M.J. Crumpton. 1977. Molecular structure of human histocompatibility antigens: the HLA-C series. *Eur. J. Immunol.* 8:580.
9. Hajek-Rosenmayr, A., L. Jungl, M. Stammler, and M. Kirnbauer. 1989. HLA-C locus antigens seen by one-dimensional isoelectric focusing: definition of the so far known HLA-C specificities and of two subtypes. *Hum. Immunol.* 26:227.
10. Neefjes, J.J., and H.L. Ploegh. 1988. Allele and locus-specific differences in cell surface expression and the association of HLA class I heavy chain with β_2-microglobulin: differential effects of inhibition of glycosylation on class I subunit association. *Eur. J. Immunol.* 18:801.
11. Tibensky, D., F. Decary, and T.L. Delovitch. 1988. HLA-C genes are transcribed in HLA-C blank individuals. *Immunogenetics.* 27:220.
12. Grunnet, N., T. Kristensen, and F. Kissmeyer-Nielsen. 1976. Cell mediated lympholysis in man. The impact of HLA-C antigens. *Tissue Antigens.* 7:301.
13. Malissen, B., T. Kristensen, C. Goridis, M. Madsen, and C. Mawas. 1981. Clones of human cytotoxic T lymphocytes derived from an allosensitized individual: HLA specificity and cell surface markers. *Scand. J. Immunol.* 14:213.
14. Kariyone, A., M. Tanabe, T. Juji, K. Kano, and M. Takiguchi. 1990. Functional expression of HLA-C blank antigens on human blood lymphocytes. *J. Immunol.* 145:3714.
15. Bonneville, M., J.F. Moreau, E. Blokland, J. Pool, J.P. Moisan, E. Goulmy, and J.P. Soulillou. 1988. T lymphocyte cloning from rejected human kidney allograft: recognition repertoire of alloreactive T cell clones. *J. Immunol.* 141:4187.
16. Chen, B.P., V. Lam, E.E. Kraus, R. DeMars, and P.M. Sondel. 1989. Restriction of Epstein-Barr virus-specific cytotoxic T cells by HLA-A, -B, and -C molecules. *Hum. Immunol.* 26:137.
17. Littaua, R.A., M.B.A. Oldstone, A. Takeda, C. Debouck, J.T. Wong, C.U. Tuazon, B. Moss, F. Kievits, and F.A. Ennis. 1991. An HLA-C-restricted CD8+ cytotoxic T-lymphocyte clone recognizes a highly conserved epitope on human immunodeficiency virus type 1 *gag*. *J. Virol.* 65:4051.
18. Dill, O., F. Kievits, S. Koch, P. Ivanyi, and G.J. Hämmerling. 1988. Immunological function of HLA-C antigens in HLA-Cw3 transgenic mice. *Proc. Natl. Acad. Sci. USA.* 85:5664.
19. Groop, L., S. Koskimies, R. Pelkonen, and E.-M. Tolppanen. 1983. Increased frequency of HLA-Cw4 in type 2 diabetes. *Acta. Endocrinol.* 104:475.
20. D'Amaro, J., J.J. Van Rood, F.H. Bach, A.A. Rimm, and M.M. Bortin. 1984. HLA-C associations with acute leukaemia. *Lancet.* 2:1176.
21. Green, J., M. Montasser, H.C. Low, and J.C. Woodrow. 1988. Investigation of the associations of a number of HLA antigens with Psoriasis and Psoriatic arthritis. *Stat. Med.* 7:443.
22. Müller, C.A., R. Hasmann, H. Grosse-Wilde, U. Vögeler, C. Bei-Jun, R. Dopfer, and H.D. Waller. 1988. Significant association of acute lymphoblastic leukemia with HLA-Cw7.

Genet. Epidemiol. 5:453.
23. Ozawa, A., M. Ohkido, H. Inoko, A. Ando, and K. Tsuji. 1988. Specific restriction fragment length polymorphism on the HLA-C region and susceptibility to Psoriasis Vulgaris. *J. Invest. Dermatol.* 90:402.
24. Parham, P., R.J. Benjamin, B.P. Chen, C. Clayberger, P.D. Ennis, A.M. Krensky, D.A. Lawlor, D.R. Littman, A.M. Norment, H.T. Orr, et al. 1989. Diversity of class I HLA molecules: Functional and evolutionary interactions with T cells. *Cold Spring Harbor Symp. Quant. Biol.* 54:529.
25. Lawlor, D.A., J. Zemmour, P.D. Ennis, and P. Parham. 1990. Evolution of class I MHC genes and proteins: from natural selection to thymic selection. *Annu. Rev. Immunol.* 8:23.
26. Dower, S.K., and D.M. Segal. 1985. Interaction of monoclonal antibodies with MHC class I antigens on mouse spleen cells. II. Levels of expression of H-2K, H-2D and H-2L in different mouse strains. *J. Immunol.* 134:431.
27. Beck, J.C., T.H. Hansen, S.E. Cullen, and D.R. Lee. 1986. Slower processing, weaker β_2-m association, and lower surface expression of H-2L^d are influenced by its amino terminus. *J. Immunol.* 137:916.
28. Weis, J.H., and C. Murre. 1985. Differential expression of H-2D^d and H-2L^d histocompatibility antigens. *J. Exp. Med.* 161:356.
29. Lie, W.-R., N.B. Myers, J. Gorka, R.J. Rubocki, J.M. Connolly, and T.H. Hansen. 1990. Peptide ligand-induced conformation and surface expression of the L^d class I MHC molecule. *Nature (Lond.).* 344:439.
30. Ennis, P.D., J. Zemmour, R.D. Salter, and P. Parham. 1990. Rapid cloning of HLA-A,B cDNA using the polymerase chain reaction: frequency and nature of errors produced in amplification. *Proc. Natl. Acad. Sci. USA.* 87:2833.
31. Zemmour, J., A.-M. Little, D.J. Schendel, and P. Parham. 1992. The HLA-A,B "negative" mutant cell line CIR expresses a novel HLA-B35 allele, which also has a point mutation in the translation initiation codon. *J. Immunol.* 148:1941.
32. Sodoyer, R., M. Damotte, T.L. Delovitch, J. Trucy, B.R. Jordan, and T. Strachan. 1984. Complete nucleotide sequence of a gene encoding a functional human class I histocompatibility antigen (HLA-Cw3). *EMBO (Eur. Mol. Biol. Organ.) J.* 3:879.
33. Zemmour, J., J.E. Gumperz, W.H. Hildebrand, F.E. Ward, S.G.E. Marsh, R.C. Williams, and P. Parham. 1992. The molecular basis for reactivity of anti-Cw1 and anti-Cw3 alloantisera with HLA-B46 haplotypes. *Tissue Antigens.* 39:249.
34. Parham, P., C.E. Lomen, D.A. Lawlor, J.P. Ways, N. Holmes, H.L. Coppin, R.D. Salter, A.M. Wan, and P.D. Ennis. 1988. Nature of polymorphism in HLA-A, -B, and -C molecules. *Proc. Natl. Acad. Sci. USA.* 85:4005.
35. Lutz, C.T., D.A. Jensen, J. Schiffenbauer, D.K. Didier, B.D. Schwartz, and C.S. Davis. 1990. Multiple mechanisms produce diversity of HLA-C alleles. *Hum. Immunol.* 28:27.
36. Ellis, S.A., T. Strachan, M.S. Palmer, and A.J. McMichael. 1989. Complete nucleotide sequence of a unique HLA class I C locus product expressed on the human choriocarcinoma cell line BeWo. *J. Immunol.* 142:3281.
37. Cianetti, L., U. Testa, L. Scotto, R. La Valle, A. Simeone, G. Boccoli, G. Giannella, C. Peschle, and E. Boncinelli. 1989. Three new class I *HLA* alleles: structure of mRNAs and alternative mechanisms of processing. *Immunogenetics.* 29:80.
38. Pohla, H., W. Kuon, P. Tabaczewski, C. Doerner, and E.H. Weiss. 1989. Allelic variation in HLA-B and HLA-C sequences and the evolution of the HLA-B alleles. *Immunogenetics.* 29:297.
39. Mizuno, S., S.H. Kang, H.W. Lee, J.A. Trapani, B. Dupont, and S.Y. Yang. 1989. Isolation and expression of a cDNA clone encoding HLA-Cw6: unique characteristics of HLA-C encoded gene products. *Immunogenetics.* 29:323.
40. Srivastava, R., B.W. Duceman, P.A. Biro, A.K. Sood, and S.M. Weissman. 1985. Molecular organization of the class I genes of human major histocompatibility complex. *Immunol. Rev.* 84:93.
41. Belich, M.P., J.A. Madrigal, W.H. Hildebrand, J. Zemmour, R.C. Williams, R. Luz, M.L. Petzl-Erler, and P. Parham. 1992. Unusual HLA-B alleles in two tribes of Brazilian Indians. *Nature (Lond.).* 357:326.
42. Takata, H., H. Inoko, A. Ando, M. Haranaka, B. Watanabe, K. Tsuji, and H. Iri. 1988. Cloning and analysis of HLA class I cDNA encoding a new HLA-C specificity Cx52. *Immunogenetics.* 28:265.
43. Davidson, W.F., M. Kress, G. Khoury, and G. Jay. 1985. Comparison of HLA class I gene sequences: derivation of locus-specific oligonucleotide probes specific for HLA-A, HLA-B and HLA-C genes. *J. Biol. Chem.* 260:13414.
44. Watkins, D.I., S.N. McAdam, X. Liu, C.R. Strang, E.L. Milford, C.G. Levine, T.L. Garber, A.L. Dogon, C.I. Lord, S.H. Ghim, G.M. Troup, A.L. Hughes, and N.L. Letvin. 1992. New recombinant HLA-B alleles in a tribe of South American Amerindians indicate rapid evolution of MHC class I loci. *Nature (Lond.).* 357:329.
45. Bodmer, J.G., S.G.E. Marsh, E.D. Albert, W.F. Bodmer, B. Dupont, H.A. Erlich, B. Mach, W.R. Mayr, P. Parham, T. Sasazuki, G.M.Th. Schreuder, J.L. Strominger, A. Svejgaard, and P.I. Terasaki. 1992. Nomenclature for factors of the HLA system, 1990. *Tissue Antigens.* 37:97.
46. Bodmer, J.G., S.G.E. Marsh, E.D. Albert, W.F. Bodmer, B. Dupont, H.A. Erlich, B. Mach, W.R. Mayr, P. Parham, T. Sasazuki, G.M.Th. Schreuder, J.L. Strominger, A. Svejgaard, and P.I. Terasaki. 1992. Nomenclature for factors of the HLA system, 1991. *Tissue Antigens.* 39:161.
47. Maryanski, J.L., P. Pala, G. Corradin, B.R. Jordan, and J.C. Cerottini. 1986. H-2 restricted cytotoxic T cells specific for HLA can recognize a synthetic HLA peptide. *Nature (Lond.).* 324:578.
48. Storkus, W.J., D.N. Howell, R.D. Salter, J.R. Dawson, and P. Cresswell. 1987. NK susceptibility varies inversely with target cell class I HLA antigen expression. *J. Immunol.* 138:1657.
49. Beretta, A., F. Grassi, M. Pelagi, A. Clivio, C. Parravicini, G. Giovinazzo, F. Andronico, L. Lopalco, P. Verani, S. Buttò, F. Titti, G.B. Rossi, G. Viale, E. Ginelli, and A.G. Siccardi. 1987. HIV *env* glycoprotein shares a cross-reacting epitope with a surface protein present on activated human monocytes and involved in antigen presentation. *Eur. J. Immunol.* 17:1793.
50. Grassi, F., R. Meneveri, M. Gullberg, L. Lopalco, G.B. Rossi, P. Lanza, C. De Santis, G. Brattsand, S. Buttò, E. Ginelli, A. Beretta, and A.G. Siccardi. 1991. Human immunodeficiency virus type I gp120 mimics a hidden monomorphic epitope borne by class I major histocompatibility complex heavy chains. *J. Exp. Med.* 174:53.
51. Bronson, S.K., J. Pei, P. Taillon-Miller, M.J. Chorney, D.E. Geraghty, and D.D. Chaplin. 1991. Isolation and characterization of yeast artificial chromosome clones linking the HLA-B and HLA-C loci. *Proc. Natl. Acad. Sci. USA.* 88:1676.
52. Parham, P., D.A. Lawlor, C.E. Lomen, and P.D. Ennis. 1989. Diversity and diversification of HLA-A,B,C alleles. *J. Immunol.* 142:3937.
53. Sneath, P.H.A., and R.R. Sokal. 1973. Numerical Taxonomy. W.H. Freeman, San Francisco. 230–234.
54. Lawlor, D.A., E. Warren, P. Taylor, and P. Parham. 1991. Gorilla

class I major histocompatibility complex alleles: comparison to human and chimpanzee class I. *J. Exp. Med.* 174:1491.
55. Hedrick, P.W., T.S. Whittam, and P. Parham. 1991. Heterozygosity at individual amino acid sites: extremely high levels for *HLA-A* and *-B* genes. *Proc. Natl. Acad. Sci. USA.* 88:5897.
56. Koller, B.H., B. Sidwell, R. DeMars, H.T. Orr. 1984. Isolation of HLA-locus specific DNA probes from the 3'-untranslated region. *Proc. Natl. Acad. Sci. USA.* 81:5175.
57. Strachan, T., A.B. Dodge, D. Smillie, P.A. Dyer, R. Sodoyer, B.R. Jordan, and R. Harris. An HLA-C-specific DNA probe. *Immunogenetics.* 23:115.
58. Duceman, B.W., D. Ness, R. Rende, M.J. Chorney, R. Srivastava, D.S. Greenspan, J. Pan, S.M. Weissman, and F.C. Grumet. 1986. HLA-JY328: mapping studies and expression of a polymorphic *HLA* class I gene. *Immunogenetics.* 23:90.
59. Nössner, E., and D.J. Schendel. 1991. HLA-Cw7-associated restriction fragment length polymorphism detected with an HLA-C locus-specific DNA probe. *Tissue Antigens.* 37:168.
60. Mayr, W.R., L. Contu, M. Kirnbauer, and H. Mervart. 1989. Antigen society #17 report (Cw5 and Cw8). *In* Immunobiology of HLA, vol. I. B. Dupont, editor. Springer-Verlag, New York. pg. 222.
61. Conighi, C., L. Contu, M.T. Grappa, E. Du Toit, M. Hammond, P. Lulli, W.R. Mayr, A. Menicucci, H. Mervart, and M. Pupura. 1989. Antigen society #18 report (Cw4 and Cw6). *In* Immunobiology of HLA, vol. I. B. Dupont, editor. Springer-Verlag, New York. 222–235.
62. Wu, T.T., and E.A. Kabat. 1970. An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J. Exp. Med.* 132:211.
63. Parham, P., D.A. Lawlor, R.D. Salter, C.E. Lomen, P.J. Bjorkman, and P.D. Ennis. 1989. HLA-A,B,C: patterns of polymorphisms in peptide binding proteins. *In* The Immunobiology of HLA, vol. II. Immunogenetics and Histocompatibility. B. Dupont, editor. Springer-Verlag, New York. 10–33.
64. Bjorkman, P.J., M.A. Saper, B. Samraoui, W.S. Bennett, J.L. Strominger, and D.C. Wiley. 1987. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature (Lond.).* 329:506.
65. Bjorkman, P.J., M.A. Saper, B. Samraoui, W.S. Bennett, J.L. Strominger, and D.C. Wiley. 1987. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature (Lond.).* 329:512.
66. Garrett, T.P.J., M.A. Saper, P.J. Bjorkman, J.L. Strominger, and D.C. Wiley. 1989. Specificity pockets for the side chains of peptide antigens in HLA-Aw68. *Nature (Lond.).* 342:692.
67. Saper, M.A., P.J. Bjorkman, and D.C. Wiley. 1991. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution. *J. Mol. Biol.* 219:277.
68. Madden, D.R., J.C. Gorga, J.L. Strominger, and D.C. Wiley. 1991. The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation. *Nature (Lond.).* 353:321.
69. Jardetsky, T.S., W.S. Lane, R.A. Robinson, D.R. Madden, and D.C. Wiley. 1991. Identification of self peptides bound to purified HLA-B27. *Nature (Lond.).* 353:326.
70. Buxton, S.E., R.J. Benjamin, C. Clayberger, P. Parham, and A.M. Krensky. 1992. Anchoring pockets in human histocompatibility complex leukocyte antigen (HLA) class I molecules: analysis of the conserved B ("45") pocket of HLA-B27. *J. Exp. Med.* 175:809.
71. Nei, M., and T. Gojobori. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* 3:418.
72. Nei, M., and L. Jin. 1989. Variances of the average numbers of nucleotide substitutions within and between populations. *Mol. Biol. Evol.* 6:290.
73. Hughes, A.L., and M. Nei. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature (Lond.).* 335:167.
74. Stam, N.J., H. Spits, and H.L. Ploegh. 1986. Monoclonal antibodies raised against denatured HLA-B locus heavy chains permit biochemical characterization of certain HLA-C locus products. *J. Immunol.* 137:2299.
75. Hildebrand, W.H., J.A. Madrigal, A.-M. Little, and P. Parham. 1992. HLA-Bw22: a family of molecules with identity to HLA-B7 in the α_1 helix. *J. Immunol.* 148:1155.
76. Guild, B.C., R.L. Erikson, and J.L. Strominger. 1983. HLA-A2 and HLA-B7 antigens are phosphorylated *in vitro* by Rous sarcoma virus kinase ($pp60^{v\text{-}src}$) at a tyrosine residue encoded in a highly conserved exon of the intracellular domain. *Proc. Natl. Acad. Sci. USA.* 80:2894.
77. Degen, E., and D.B. Williams. 1991. Participation of a novel 88-kD protein in the biogenesis of murine class I histocompatibility molecules. *J. Cell Biol.* 112:1099.
78. Zemmour, J., B.H. Koller, P.D. Ennis, D.E. Geraghty, D.A. Lawlor, H.T. Orr, and P. Parham. 1990. HLA-AR, an inactivated antigen presenting locus related to HLA-A: implications for the evolution of the MHC. *J. Immunol.* 144:3619.
79. Messer, G., J. Zemmour, H.T. Orr, P. Parham, E.H. Weiss, and J. Girdlestone. 1992. HLA-J: a second inactivated class I HLA gene related to HLA-G and HLA-A: implications for the evolution of the HLA-A related genes. *J. Immunol.* 148:4043.
80. Asahina, A., S. Akazaki, S. Nakagawa, S. Kuwata, K. Tokunaga, Y. Ishibashi, and T. Juji. 1991. Specific nucleotide sequence of HLA-C is strongly associated with Psoriasis Vulgaris. *J. Invest. Dermatol.* 97:254.
81. Connolly, J.M., T.H. Hansen, A.L. Ingold, and T.A. Potter. 1990. Recognition by CD8 on cytotoxic T lymphocytes is ablated by several substitutions in the class I α_3 domain: CD8 and the T-cell receptor recognize the same class I molecule. *Proc. Natl. Acad. Soc. USA.* 87:2137.
82. Salter, R.D., R.J. Benjamin, P.K. Wesley, S.E. Buxton, T.P.J. Garrett, C. Clayberger, A.M. Krensky, A.M. Norment, D.R. Littman, and P. Parham. 1990. A binding site for the T-cell co-receptor CD8 on the α_3 domain of HLA-A2. *Nature (Lond.).* 345:41.
83. Biassoni, R., D. Pende, O. Viale, C. Di Donato, S. Ferrini, E. Ciccone, A. Moretta, and L. Moretta. 1992. Analysis of B-EBV detection mutants at the MHC region for the susceptibility to lysis by 6 alloreactive NK clones. *J. Cell. Biochem.* 16D(Suppl.):O 401. 58. (Abstr.)
84. Storkus, W.J., R.D. Salter, J. Alexander, F.E. Ward, R.E. Ruiz, P. Cresswell, and J.R. Dawson. 1991. Class I-induced resistance to natural killing: identification of nonpermissive residues in HLA-A2. *Proc. Natl. Acad. Sci. USA.* 88:5989.
85. Bjorkman, P.J., and P. Parham. 1990. Structure, function and diversity of class I major histocompatibility molecules. *Annu. Rev. Biochem.* 59:253.
86. Little, A.-M., J.A. Madrigal, and P. Parham. 1992. Molecular definition of an elusive third HLA-A9 molecule: HLA-A9.3. *Immunogenetics.* 35:41.
87. Madrigal, J.A., M.P. Belich, W.H. Hildebrand, R.J. Benjamin, A.-M. Little, J. Zemmour, P.D. Ennis, F.E. Ward, M.L. Petzl-Erler, E.D. Du Toit, and P. Parham. 1992. Distinctive HLA-A,B antigens of black populations formed by interallelic microrecombination. *J. Immunol.* In press.