Published in final edited form as:

Tissue Antigens. 2008 January ; 71(1): 72–76.

Description of two new HLA-C alleles in a Black South African population

G. Minevich¹, L. Kuhn^{2,3}, C. Polistena¹, J. Nuara¹, and R. Winchester¹

¹Department of Medicine, College of Physicians and Surgeons, Columbia University, New York, NY, USA

²Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA

³Sergievsky Center, College of Physicians and Surgeons, Columbia University, New York, NY, USA

Abstract

Two new *HLA-C* alleles, *Cw*0333* and *Cw*0217*, were identified in a Black South African population. *HLA-Cw*0333* differs from *Cw*030201* by an A→G substitution at nucleotide 323, yielding an unusual missense substitution of Cys for the conserved Tyr-84 at the antigen cleft terminus. Molecular modeling suggests that this alters the predicted interactions of this critical residue with the opposite α_2 -helix, the peptide COOH terminus and possibly KIR2DL2. The second allele, *Cw*0217*, differs from *Cw*0205* by an A→T substitution at nucleotide 368, resulting in a Tyr→Phe substitution at residue 99 of the *HLA-C4/9/2008* β -pleated sheet that likely influences peptide side-chain binding. Both *Cw*0333* and *Cw*0217* appear to have arisen by missense mutations, respectively, from the *HLA-B*5801-Cw*0302* and *B*080101-Cw*0205* haplotypes.

Keywords

HLA class I genes; HLA-C; novel allele

As a consequence of evolutionary mechanisms that diversify the peptide-presenting function of the major histocompatibility complex (MHC), the *HLA-A*, *-B*, and *-C* loci are among the most polymorphic in the genome. In particular, the *HLA-C* locus currently contains 307 alleles as of IMGT/HLA Database release 2.17.0, 4 December 2007 (1). Here we describe two new *HLA-C* alleles, *Cw*0217* and *Cw*0333*, recently identified in a Black South African population through sequence-based typing (SBT).

Sequencing of exons 2 and 3 of both *HLA-B* and *HLA-C* was performed on 270 mother–child pairs, as part of an investigation into the role of *HLA-B* and *HLA-C* alleles in maternal to infant human immunodeficiency virus (HIV) transmission in a Black South African population. We followed a SBT strategy and the protocol previously described by Cereb et al. (2) for heterozygous amplification of exon 2, intron 2, and exon 3 of the *HLA-B* and *HLA-C* loci.

Nucleotide sequencing was performed on an ABI 3730 Genetic Analyzer using Big Dye Terminator v1.1 (Applied Biosystems, Foster City, CA). Allele assignment was performed using SeqScape v2.5 software (Applied Biosystems) and a library compiled from the 2.17.0 release of the IMGT/HLA Database. To verify the new *HLA-C* variants, the amplified polymerase chain reaction (PCR) product was cloned to separate the heterozygous alleles using a TOPO TA cloning kit (Invitrogen, Carlsbad, CA). DNA inserts from at least 12 clones for each sequence were amplified by PCR and subsequently sequenced in forward and reverse

directions as described by Delfino et al. (3), with the exception of the primers used to sequence exon 2 and exon 3 in the sense direction. Exon 2 was sequenced in the forward direction with a modified form of the DLFEX2 forward primer published by Delfino (5'-AGG GTC GGG CGG GTC TCA GC-3') (*DLFEX2_mod*). Exon 3 was sequenced in the forward direction with a primer designed in-house: 671GACT_CI2_5' (5'-GAC TGG GGC GGG GCT CGG GGG-3'). The novel alleles from each individual were distinctly identified once separated by this technique.

The names *Cw*0333* (IMGT/HLA Database accession number HLA02715) and *Cw*0217* (IMGT/HLA Database accession number HLA02764) have been officially assigned to these new alleles by the World Health Organization Nomenclature Committee in November and December 2006, respectively. Their numerical designation follows the agreed policy that, subject to the conditions stated in the most recent Nomenclature Report (4), names will be assigned consecutively to new sequences in each allele subgroup as they are identified. The limited mother–child pedigrees allowed direct haplotype determination from the alleles shared between mother and child and the remaining non-shared alleles.

Figure 1 illustrates the nucleotide and inferred amino acid sequence in the region of exon 2 for the newly identified *HLA-Cw*0333* allele, showing that it differs from *Cw*030201* by one nucleotide substitution at nucleotide 323 (A→G). This translates to a substitution of cysteine for tyrosine at residue 84 (Tyr→Cys) in the HLA-C α heavy chain.

Figure 2a depicts the crystal structure of HLA-Cw3 complexed with the nonameric peptide GAVDPLLAL (GAV) using the coordinates of Boyington et al. (PDBID: 1EFX) (5) implemented in MacPyMOL (6) and illustrates the interactions of the GAV peptide with the conserved residues of both α -helices, Tyr-84, Lys-146, and Thr-143. These amino acids are found clustered near the end of the α -helical margin of the antigen cleft and have been proposed as features required for creating the three dimensional pattern of hydrogen bonding that is responsible for the correct orientation of the C-termini of peptides and for closing the end of the HLA-B peptide cleft (7). Similar interactions of these amino acids are evident in HLA-Cw3. Figure 2b illustrates the potential bonding interactions of Tyr-84 in the structure of HLA-Cw3. The predicted lengths of the hydrogen bond contacts between Tyr-84 and the nearest carboxylate oxygens of the ninth residue of the GAV peptide are both 3.5 Å. This bond length is consistent with the formation of a hydrogen bond (9), which is analogous to but slightly greater in length than that described for the hydrogen bonds between the equivalent peptide carboxylate oxygens and this residue in HLA-B*27 (7). Tyr-84 also has a predicted hydrogen bond contact with Lys-146 at a distance of 3.0 Å across to the opposite α_2 -helical margin, indicating that Tyr-84 participates in closing the C-terminal end of the peptide groove analogous to its function in HLA-B. Furthermore, a 3.6-Å hydrogen bond contact is predicted between Tyr-84 and Asp-183 of the killer cell immunoglobulin-like receptor (KIR), thus implicating Tyr-84 as an interface between the HLA-Cw3 complex with KIR2DL2. This interface occurs near one of the main stabilizing contacts between HLA-Cw3 and KIR2DL2, the 2.5-Å salt bridge between Lys-146 of HLA-Cw3 and Asp-183 of the KIR2DL2 (5). Tyr-84 is also adjacent to Asn-80 on the α_1 -helix, the HLA-C residue that determines C-1 allotype specificity of the KIR interaction (10).

Reflecting the conserved structural role of Tyr-84, this amino acid is encoded by all but four of the 1661 *HLA-A*, *HLA-B*, and *HLA-C* class I alleles whose sequence is determined at this position (1). Among the exceptions are a stop codon in *HLA-A*0253N*, the substitution of a large charged His or Asp in *HLA-B*0723* and *HLA-B*5112* and the newly identified *HLA-Cw*0333* allele with its substitution of the tiny uncharged Cys. Indeed the importance of these residues in the structure of this region of the class I molecule is supported by complete conservation of Lys-146, the hydrogen bonding partner of Tyr-84 in all class I molecules.

Similarly, Thr-143 is conserved in 1597 of 1661 (96%) class I molecules and the structurally homologous Ser is found in 63 of the remaining 64 alleles.

Figure 2c illustrates the situation in the HLA-Cw*0333 molecule, where the most important consequence of the change in the side chain at position 84 from the large aromatic Tyr with its phenolic hydroxyl group to the small Cys with its thiol group is that the small Cys-84 amino acid side chain does not project out as far from the α -helical backbone. Although both Tyr and Cys are polar and can engage in hydrogen bonding to the peptide, the different stereochemistry of the hydrogen bonds formed by these two donor/acceptors likely influences features of the molecule encoded by this allele (11, 12). The resulting increase in distances shown between the Cys-84 donor/acceptor pairs precludes effective bond interaction at each of the sites. For example, the hydrogen bond contact with Lys-146 that in the Tyr-84 molecule was 3.0 Å across to the opposite α_2 -helical margin is now 6.5 Å, a distance very unlikely to support bond formation.

It is unclear whether the conformation enforced by Cys-84 will cause the α_2 -helix to narrow the peptide-binding cleft by forming a Cys-84–Lys-146 H bond or whether the helices will remain apart and less tethered. Furthermore, the direction of the H bonds with Cys-84 are no longer constrained by the plane of the phenol moiety in Tyr-84 (13), and water molecules will not be excluded by the hydrophobic aromatic ring. No polar contacts are predicted between Cys-84 and the nearest carboxylate oxygen of the peptide because of the relatively great distance of 5.9 Å and 6.1 Å between the candidate-binding partners, as shown. While additional water molecules may fill the void and stabilize the structure, there are several other potential consequences to the peptide-binding function of the HLA-Cw*0333 molecule: First, because of the additional space, the bound peptide could be one amino acid longer. Second, despite the conserved hydrogen bond between Tyr-85 and Asp-137 that connects the two α -helices, the loss of the hydrogen bond contact between Tyr-84 and Lys-146 may result in an unstable MHC configuration that is poorly expressed. Third, if Cys-84 forms a strong hydrogen bond with Lys-146, this would distort the peptide groove. Fourth, the thiol group of Cys is the most reactive of any amino acid side chain (14), and it lies relatively exposed in the class I molecule so it may be free to react with peptides or other intra- or extracellular molecules in such a way as to alter MHC function, a mechanism postulated to occur with the Cys-67 in HLA-B*270502 (15). Lastly, the conformational change on the MHC–peptide complex by virtue of the absence of three Cys-84 interactions (with 146-Lys on the MHC, with the peptide carboxylate residues, and with Asp-183 on KIR2DL2) could weaken the interaction between HLA-Cw*0333 and KIR2DL2.

In the case of allele *HLA-Cw*0333*, the mother was HIV infected and transmitted both the novel allele and HIV infection to the child. This and the structural features of the allele raise the possibility of a decreased ability to elicit an antiviral response and directs attention to delineating the expression and peptide-binding features of this novel molecule.

Figure 3 illustrates the observed haplotype bearing this novel *HLA-C* allele. *HLA-Cw*0333* is present on the *HLA-B*5801* haplotype shared by Mother_268N and Child_268N. The almost identical *B*5801-Cw*0302* haplotype is frequent in the sub-Saharan population (~1%, n = 1520) (16). Interestingly, *Cw*0302* appears almost exclusively with *B*5801* on this haplotype within our study population of the mostly Zulu sub-Saharans (10/11 instances, 91%). Within the National Center for Biotechnology Information (NCBI) data set of all sub-Saharan haplotypes, *Cw*0302* is also commonly found with *B*5801* on the same haplotype (~50/75 instances, 67%) (16). We infer a point mutation with a transition of A to G likely accounts for the origin of *HLA-Cw*0333* from the prevalent *HLA-B*5801-Cw*0302* haplotype.

Figure 4 illustrates the nucleotide and inferred amino acid sequence of exon 3 for the newly identified *HLA-Cw*0217* allele, showing that this allele differs from the *Cw*0205* allele by one nucleotide substitution at nucleotide 368 (A→T). This translates to a difference of one amino acid at residue 99 (Tyr→Phe) of the HLA-C heavy chain β-pleated sheet.

Amino acid position 99 is highly polymorphic in HLA-C and is centrally positioned in the bottom of the peptide-binding cleft, potentially affecting peptide side chains 1, 2, and 3 binding in pockets A, B, and D (17). The large majority (258/299) of the polymorphisms at position 99 involve either Tyr or Phe at this position. Tyr→Phe is a semiconservative change with hydrophobic, polar Tyr being replaced with the non-polar Phe that lacks the OH moiety.

The *Cw*0217* allele was found in three individuals in this study (Figure 3). In one mother–infant pair (Mother_HLA_370 and Child_HLA_370), the novel allele was present on the unusual *HLA-B*5801-Cw*0217* haplotype transmitted to the child by the mother. No other known instances of an *HLA-B*5801-Cw*02* haplotype appear in the sub-Saharan population in our study of 270 mother–infant pairs or in the NCBI data set (16), suggesting that recombination or possibly gene conversion between *HLA-B* and *HLA-C* led to the development of the unique *HLA-B*5801-Cw*02* haplotype. The appearance of the novel *HLA-B*5801-Cw*0217* haplotype may be explained by the aforementioned recombination with *HLA-B*5801* and by a point mutation in the *Cw*0205* allele.

Child_HLA_493 also exhibited the *Cw*0217* allele, but it was present in the *HLA-B*080101-Cw*0217* haplotype inherited from the father (Figure 3). While an *HLA-B*080101-Cw*02* haplotype does not appear in the NCBI data set, it does appear in three additional individuals in the present study, suggesting that the *HLA-B*080101-Cw*0217* haplotype may have originated by a point transversion from the *B*080101-Cw*0205* haplotype present in the Black South African study population.

Acknowledgements

This study was funded in part by NICHD 42402. We would like to thank the clinical and laboratory teams in South Africa for recruitment and clinical management of the study participants and for collection and processing of the specimens, including Drs Glenda Gray, Caroline Tiemessen, Diane Schramm, and Stephen Meddows-Taylor.

References

1. Robinson, J.; Waller, MJ.; Parham, P., et al. IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex; Nucleic Acids Res. 2003. p. 311-4. <http://www.ebi.ac.uk/imgt/hla/>
2. Cereb N, Maye P, Lee S, Kong Y, Yang SY. Locus-specific amplification of HLA class I genes from genomic DNA: locus-specific sequences in the first and third introns of HLA-A, -B, and -C alleles. *Tissue Antigens* 1995;45:1–11. [PubMed: 7725305]
3. Delfino L, Morabito A, Longo A, Ferrara GB. HLA-C high resolution typing: analysis of exons 2 and 3 by sequence based typing and detection of polymorphisms in exons 1-5 by sequence specific primers. *Tissue Antigens* 1998;52:251–9. [PubMed: 9802605]
4. Marsh S, Albert ED, Bodmer WF, et al. Nomenclature for factors of the HLA system. *Tissue Antigens* 2005;65:301–69. [PubMed: 15787720]
5. Boyington JC, Motyka SA, Schuck P, Brooks AG, Sun PD. Crystal structure of an NK cell immunoglobulin-like receptor in complex with its class I MHC ligand. *Nature* 2000;405:537–43. [PubMed: 10850706]
6. DeLano, WL. MacPyMOL: A PyMOL-based Molecular Graphics Application for MacOS X. Palo Alto, California: DeLano Scientific LLC; 2007.
7. Madden DR, Gorga JC, Strominger JL, Wiley DC. The three-dimensional structure of HLA-B27 at 2.1 Å resolution suggests a general mechanism for tight peptide binding to MHC. *Cell* 1992;70:1035–48. [PubMed: 1525820]

8. Zhang W, Young AC, Imarai M, Nathenson SG, Sacchettini JC. Crystal structure of the major histocompatibility complex class I H-2Kb molecule containing a single viral peptide: implications for peptide binding and T-cell receptor recognition. *Proc Natl Acad Sci U S A* 1992;89:8403–7. [PubMed: 1325657]
9. Jeffrey, GA. *An Introduction to Hydrogen Bonding*. New York: Oxford University Press; 1997.
10. Mandelboim O, Reyburn HT, Vales-Gomez M, et al. Protection from lysis by natural killer cells of group 1 and 2 specificity is mediated by residue 80 in human histocompatibility leukocyte antigen C alleles and also occurs with empty major histocompatibility complex molecules. *J Exp Med* 1996;184:913–22. [PubMed: 9064351]
11. Fersht AR, Shi JP, Knill-Jones J, et al. Hydrogen bonding and biological specificity analysed by protein engineering. *Nature* 1985;314:235–8. [PubMed: 3845322]
12. Wilkinson AJ, Fersht AR, Blow DM, Winter G. Site-directed mutagenesis as a probe of enzyme structure and catalysis: tyrosyl-tRNA synthetase cysteine-35 to glycine-35 mutation. *Biochemistry* 1983;22:3581–6. [PubMed: 6615786]
13. Ippolito JA, Alexander RS, Christianson DW. Hydrogen bond stereochemistry in protein structure and function. *J Mol Biol* 1990;215:457–71. [PubMed: 2231715]
14. Creighton, TE. *Proteins: Structures and Molecular Properties*. New York: W.H Freeman; 1993.
15. Allen RL, O'Callaghan CA, McMichael AJ, Bowness P. Cutting edge: HLA-B27 can form a novel beta 2-microglobulin-free heavy chain homodimer structure. *J Immunol* 1999;162:5045–8. [PubMed: 10227970]
16. dbMHC. National Center for Biotechnology Information.
http://www.ncbi.nlm.nih.gov/projects/mhc/ihwg.cgi?cmd=page&user_id=0&probe_id=0&source_id=0&locus_id=0&locus_id=0
17. Marsh, S.; Parham, P.; Barber, LD. *The HLA Facts Book*. San Diego: Academic Press; 2000.

Exon 2

Nucleotide position	-311-	-314-	-317-	-320-	-323-
<i>Cw*010201</i> (Nuc.)	AAC	CTG	CGC	GGC	TAC
Amino acid position	80	81	82	83	84
<i>Cw*010201</i> (A.A.)	Asn	Leu	Arg	Gly	Tyr
<i>Cw*030201</i>	---	---	---	---	---
	Asn	Leu	Arg	Gly	Tyr
<i>Cw*0333</i>	---	---	---	---	-G-
	Asn	Leu	Arg	Gly	Cys

Figure 1.

Nucleotide alignment of reference *HLA-Cw*010201* and comparisons of reference *HLA-Cw*030201* and the new allele *HLA-Cw*0333*. *HLA-Cw*0333* differs from *Cw*030201* by an A→G substitution at nucleotide 323 encoding a Tyr→Cys substitution at amino acid 84. Amino acid and nucleotide positions are numbered according to IMGT release 2.17.0 (12 April 2007) (1).

Figure 2.

(A) Crystal structure 1EFX displaying HLA-Cw3 complexed with the nonameric peptide GAVDPLLAL (GAV) using the coordinates of Boyington et al. (PDBID: 1EFX) (5) implemented by MacPyMOL (6). The image is slightly rotated to better display the C terminus of the peptide and its proximity to three conserved residues of the α -helical margin of the antigen cleft, Tyr-84, Lys-146, and Thr-143. These three α -helical residues have been proposed as features required for creating the three dimensional pattern of hydrogen bonding that is responsible for the correct orientation of the C termini of peptides and for closing the end of the HLA-B peptide cleft (7). (B) Crystal structure 1EFX displaying HLA-Cw3 complexed with the nonameric peptide GAVDPLLAL (GAV) and NK cell receptor KIR2DL2 in the region

around Tyr-84 using the coordinates of Boyington et al. (PDBID: 1EFX) (5) implemented by MacPyMOL (6). The distances between the Tyr hydroxyl and potential hydrogen bond acceptors on the nonameric peptide GAVDPLLAL (GAV), the NK cell receptor KIR2DL2, and the opposite α_2 -helix forming the C terminus of the antigen cleft are illustrated. The anchor hydrogen bonding between the hydroxyl group of Tyr-84 and peptide that has been proposed as a general feature of tight peptide bonding to class I molecules and constrains the length of the bound peptide (7,8). Additional hydrogen bonds between the carboxyl terminus of the bound peptide and the amino group of Lys-146 on the opposite α_2 -helix are shown. (C) Crystal structure displaying the Cys-84 feature of new allele *HLA-Cw*0333* complexed with the nonameric peptide GAVDPLLAL (GAV) and NK cell receptor KIR2DL2 using the coordinates of Boyington et al. (PDBID: 1EFX) (5) implemented by MacPyMOL (6). The small Cys-84 amino acid side chain does not project out as far from the α -helical backbone as the corresponding Tyr-84 encoded in allele *HLA-Cw*030201* and thus likely influences features of the molecule encoded by this allele (11,12). The resulting increase in distances shown between the Cys-84 donor/acceptor pairs likely precludes effective bond interaction with the carboxylate oxygen of the GAV peptide, Lys-146 on the opposite α_2 -helix, and Asp-183 on KIR2DL2.

Allele	Sample_ID	Haplotype_1	Haplotype_2	Sex
Cw*0333				
	Mother_268N	B*5801-Cw*0333	B*530101-Cw*04010101	F
	Child_268N	B*5801-Cw*0333	B*5802-Cw*0602	F
Cw*0217				
	Mother_HLA_370	B*1510-Cw*030402	B*5801-Cw*0217	F
	Child_HLA_370	B*1510-Cw*030402	B*5801-Cw*0217	F
B*08-Cw*02				
	Mother_HIV_113	B*080101-Cw*0205	B*5802-Cw*0602	F
	Mother_HIV_152	B*080101-Cw*0210	B*4201-Cw*1701	F
	Child_HIV_152	B*080101-Cw*0210	B*5801-Cw*030201	F

Figure 3.
Haplotypes containing novel alleles.

Exon 3					
Nucleotide position	-356-	-359-	-362-	-365-	-368-
<i>Cw*010201</i> (Nuc.)	CTC	CAG	TGG	AGT	TGT
Amino acid position	95	96	97	98	99
<i>Cw*010201</i> (A.A.)	Leu	Gln	Trp	Met	Cys
 <i>Cw*0205</i>	 ---	 ---	 ---	 ---	 -A-
	Leu	Gln	Trp	Met	Tyr
 <i>Cw*0217</i>	 ---	 ---	 ---	 ---	 -T-
	Leu	Gln	Trp	Met	Phe

Figure 4.

Nucleotide alignment of reference *HLA-Cw*010201* and comparisons of *HLA-Cw*0205* and *HLA-Cw*0217*. *HLA-Cw*0217* differs from *Cw*0205* by an A→T substitution at nucleotide 368 and a Tyr→Phe substitution at amino acid 99. Amino acid and nucleotide positions numbered according to IMGT release 2.17.0 (12 April 2007) (1).