Immunogenetics (2001) 53:10–14
DOI 10.1007/s002510000299

ORIGINAL PAPER

Mark A. Luscher · Kelly S. MacDonald
Job J. Bwayo · Francis A. Plummer · Brian H. Barber

Sequence and peptide-binding motif for a variant of HLA-A*0214 (A*02142) in an HIV-1-resistant individual from the Nairobi Sex Worker cohort

Received: 15 September 2000 / Revised: 12 December 2000 / Published online: 6 February 2001
© Springer-Verlag 2001

Abstract As part of the ongoing study of natural HIV-1 resistance in the women of the Nairobi Sex Workers' study, we have examined a resistance-associated HLA class I allele at the molecular level. Typing by polymerase chain reaction using sequence-specific primers determined that this molecule is closely related to HLA-A*0214, one of a family of HLA-A2 supertype alleles which correlate with HIV-1 resistance in this population. Direct nucleotide sequencing shows that this molecule differs from *A*0214*, having a silent nucleotide substitution. We therefore propose to designate it *HLA-A*02142*. We have determined the peptide-binding motif of HLA-A*0214/02142 by peptide elution and bulk Edman degradative sequencing. The resulting motif, X-[Q,V]-X-X-X-K-X-X-[V,L], includes lysine as an anchor at position 6. The data complement available information on the peptide-binding characteristics of this molecule, and will be of use in identifying antigenic peptides from HIV-1 and other pathogens.

Keywords HLA-A*0214 · Peptide-binding motif · HIV-1 resistance · African alleles · Pool sequence

The sequence data reported in this paper have been submitted to the GenBank nucleotide sequence database and have been assigned the accession number AF305699

K.S. MacDonald (✉)
Department of Microbiology, Mount Sinai Hospital, Room 1484, 600 University Ave., Toronto, Ontario, M5G 1X5, Canada
E-mail: KMacDonald@MtSinai.on.ca
Phone: +1-416-5868879
Fax: +1-416-5868746

M.A. Luscher · K.S. MacDonald
Department of Medicine, University of Toronto, Toronto, Canada

M.A. Luscher
University Health Network, Toronto, Canada

J.J. Bwayo
Department of Microbiology, University of Nairobi, Nairobi, Kenya

F.A. Plummer
Department of Medical Microbiology University of Manitoba, Winnipeg, Canada

B.H. Barber
Department of Immunology, University of Toronto, Toronto, Canada

Introduction

Compelling epidemiological evidence of natural resistance to HIV-1 infection has emerged as one of the outstanding findings of the Nairobi Sex Workers' study (Fowke et al. 1996). This open cohort study has documented the health, sexual practices, and immune parameters of over 1500 women. It has become clear that a small fraction of these women, about 5%, remain HIV-1 negative despite high rates of exposure. A correlation has been observed between the persistently seronegative (PSN) phenotype and the presence of certain HLA alleles such as class II DRB1*0101/02, and the class I alleles of A2/6802 supertype (which in this population include A*0201, A*0202, A*0205, A*0214, and A*6802). Further understanding of these relationships will benefit from a detailed understanding of the structure and function of the HLA molecules involved (MacDonald et al. 2000).

We have sequenced a resistance-associated HLA allele from a PSN prostitute and determined its peptide-binding motif. The sequence varies from the previously reported sequence of *HLA-A*0214* by a single, silent nucleotide substitution. We therefore propose to designate it *HLA-A*02142*. The peptide-binding motif of A*0214/02142 was determined by Edman degradative sequencing and was found to include anchors at positions 2 (Q,V), 6 (K), and 9 (V,L). This information, together with published peptide-binding motif

data, should assist in the selection of candidate A*0214-binding peptides from HIV and other pathogens.

Materials and methods

Patient cells

The patient selected for this study, designated ML1260, was enrolled in the Nairobi Sex Workers' study in April 1990. She has remained persistently seronegative for HIV-1 and HIV-2. In addition to regular negative HIV seroassays, she was persistently negative by PCR for HIV-1 proviral DNA using methods as previously described (MacDonald et al. 2000). During her time in the study, she had three to five sex partners per day and several episodes of genital ulcers and gonococcal infection. Her class I phenotype by microlymphocytotoxicity testing was HLA-A2, A69, B18, B71, C4, C7 and by sequence-specific primer-PCR her typing at the *A* and *B* locus was *A*0214*, blank, *B*1801*, *B*71*, using methods as previously described (Bunce et al. 1995; Krausa et al. 1995b). Peripheral blood mononuclear cells taken from ML1260 were used as a source of B cells which were transformed using Epstein-Barr virus (EBV). The resulting cell line was cultured in a hollow-fiber cell culture system (Cell-Max MPS 30 kDa, catalog no. 41008-061; Life Technologies, Grand Island, N.Y.) using RPMI1640 medium supplemented with fetal calf serum, Glutamax II (Life Technologies), non-essential amino acids, and sodium pyruvate. Lactate production was monitored using a lactate assay kit (catalog no. 735-10; Sigma, St. Louis, Mo.) and was maintained below 1 mg/ml. Cells were harvested when the rate of lactate production exceeded 4 mg/h. Typical yields were about 5×10^8/week.

Direct sequencing of HLA-A

RNA was extracted from 5×10^6 patient EBV B-lymphoblastoid cells using the Oligotex Direct mRNA kit (Qiagen, Valencia, Calif.). The GeneAmp RNA PCR kit from Applied Biosystems (Foster City, Calif.) was used to perform reverse transcription of the isolated mRNA to cDNA (using murine leukemia virus reverse transcriptase and oligo dT primers) at 42 °C for 15 min, 99 °C for 5 min, and 5 °C for 5 min. Subsequent PCR amplification was done using AmpliTaq DNA polymerase (Roche) and class I locus-specific primers (5′ HLA-A Forward and 3′ HLA-3PG for locus *A*), at 50 pmol per reaction. Amplification was carried out on a Perkin Elmer GeneAmp 9600 thermocycler (Applied Biosystems) under the following conditions: 95 °C for 5 min, 20 cycles at 95 °C for 1 min, 60 °C for 1 min, and 72 °C for 50 s, 10 cycles at 95 °C for 1 min, 60 °C for 1 min, and 72 °C for 1.5 min, followed by 72 °C for 10 min. PCR-amplified products were TA cloned into either the vector PT7 (Novagen, Madison, Wis.) or PCR II (Invitrogen, Carlsbad, Calif.). Conventional ^{35}S-double-stranded DNA sequencing reactions were performed with the Sequenase T7 DNA Polymerase version 2.0 kit (US Biochemical, Cleveland, Ohio). Plasmid DNA (123 pg) and 1 pmol of sequencing primer were used for each reaction. The forward sequencing primers were the T7 promotor primer (TAATACGACTCACTATAGGG) and (from Ennis et al. 1990) Exon 2 236–255 (AGGGGCCGGAGTATTGGGAC), Exon 3 429–450 (CGGGAAGGATTACATCGCCCTG), and Exon 4 703–724 (GCGGAGATCACACTGACCTGGC). The reverse primers were (also from Ennis et al. 1990) Exon 2 255–236 (TCCCCGGCCTCATAACCCTG) and Exon 3 450–429 (CAGGGCGATGTAATCCTTGCCG). Other primers (after the method of Domena et al. 1993) were Exon 3 340–321 (CAGGACCTCGCTCTGGTTGTAGTAG), and Exon 4 606–582 (CAGGACGTCTCCTTCCCGTTCTCC). Samples were loaded onto 6% acrylamide sequencing gels (BRL). The gels were run on a BRL sequencing apparatus at 1700–1900 V for 2–4 h. Sequences were compared to the canonical *HLA-A*0214* sequence.

Preparation of cell lysates

Cells were harvested from hollow-fiber culture according to the manufacturer's directions and were pelleted. Lysis buffer [about 1 ml per 5×10^7 cells, prepared fresh for each use, 0.1 M Tris pH 8.1, 0.05 M NaCl, 1 mM $MgCl_2$, 1 mM phenylmethylsufonylfluoride, 10 mM iodoacetamide, and 0.5% high-purity Surfact-Amps NP-40 (stored under nitrogen or argon); Pierce, Rockford, Ill.] at 4 °C was added to the cell pellet which was then vigorously resuspended and disrupted by intermittent vortexing over the next 30 min and was otherwise kept on ice. The lysate was pelleted at 20,000 *g* and the supernatant stored at –70 °C.

Purification of HLA

Stored lysates (8.2×10^9 cells total) were thawed in a 4 °C water bath and pelleted at 100,000 *g* for 30 min. The supernatants were kept at 4 °C and were filtered (0.8 µm) just prior to use. The HPLC-protein A matrix TSK-gel AF-protein A Toyopearl 650M (Toso, Tokyo, Japan) was prepared according to the manufacturer's directions. The following buffers were prepared: HPLC Tris buffer (0.1 M Tris, pH 8.1, 0.05 M NaCl, 0.02% NaN_3) with or without high-purity NP-40 (0.5%); column wash buffer (sodium borate 0.05 M, pH 8.1 adjusted with HCl, 0.02% NaN_3); elution buffer (0.2 N acetic acid aldehyde-free, prepared just prior to use, catalog no. 1003-2; Caledon, Georgetown, Canada). The protein-A matrix was washed into HPLC Tris, and antibodies were added to portions of the matrix as follows: Control, 1.5 mg HB65 [purified monoclonal IgG anti-influenza nucleoprotein; American Type Culture Collection (ATCC), Rockville Md.]; Anti-A2, a mixture of 2 mg HB82 (BB7.2, monoclonal IgG anti-HLA-A2; ATCC) and 2 mg HB117 (PA2.1, purified monoclonal IgG anti-HLA-A2; ATCC). The mixtures were incubated on a rotator at 4 °C overnight. The matrix-antibody conjugates were packed into HPLC columns (Omega columns; Upchurch, Oak Harbor, Wash.) using HPLC Tris and a flow rate of 25 ml/min. Subsequently, HPLC Tris with NP-40 was run into each column at 0.8 ml/min until a steady absorbance (280 nm) was reached. The columns were connected in series – Control, followed by Anti-A2 – and the filtered cell lysate was run through at 0.7 ml/min, followed by HPLC Tris with NP-40 until a steady absorbance was reached (about 20 min). The columns were detached and the A2 column was washed with column wash buffer to background absorbance. Its orientation was reversed, and bound material was eluted with the elution buffer. A prominent peak was collected as a single fraction of 3 ml, and was stored at –70 °C.

Peptide isolation

Eluted HLA was thawed and made up to 10% aldehyde-free acetic acid. It was heated to 100 °C for 10 min, and ultrafiltered using prewashed Ultrafree CL filters (5000 NMWL, catalog no. UFC4LCC25; Millipore, Bedford, Mass.) at 2500 *g*. The flow-through (peptide fraction) was concentrated on a centrifugal concentrator to less than 20 µl, and 0.5 ml water was added. Salts, acetic acid, and nonpeptide impurities were removed by running on reversed-phase C18 HPLC, and washing to background absorbance with water-0.1% trifluoroacetic acid. The solvent flow in the C18 column was then reversed, and peptides were eluted using 50% acetonitrile-0.1% trifluoroacetic acid. The eluates were lyophilized, then resuspended in 30 µl 50% acetonitrile.

Amino acid sequencing

The sample was adsorbed to peptide support disks (catalog no. 290-110; Beckman, Fullerton, Calif.) pretreated with NaCl according to the manufacturer's directions, and was sequenced in bulk by Edman degradation on a Porton instruments 2090 microsequencer.

Results

*Nucleotide sequencing reveals a new variant of HLA-A*0214 differing by a silent substitution in the second codon of exon 2*

RT-PCR using *HLA-A* locus-specific primers was used to derive molecular clones of *HLA-A* from patient ML1260. Two independent PCR reactions were performed; five clones were sequenced from one, and four from another. Nucleotide sequencing revealed a consistent difference in the observed sequence from the canonical sequence of *HLA-A*0214*, namely a T→C substitution at position 78 in the coding sequence, located at the end of codon 2 of exon 2. The change does not alter the predicted amino acid sequence of the protein relative to HLA-A*0214 (Fig. 1).

*Pool sequencing of peptides eluted from HLA-A*0214/02142 protein reveals its peptide-binding motif*

HLA-A*02142 was purified from EBV-transformed patient cells by HPLC affinity chromatography. Bound peptides were eluted using an adaptation of the protocol of Falk and co-workers (1991) that included the use of aldehyde-free reagents to avoid end blocking of the peptides to be sequenced, and the use of C18 reversed-phase HPLC to aid the removal of salts and detergents. The data for the Edman degradation of this preparation are shown in Table 1. The strong peptide signal confirms minimal end blocking of the sample. The amount of each amino acid in each cycle is listed as a percentage of the total amino acid content eluted in that cycle. A residue was defined as a minor anchor if its relative concentration was at least double that of the previous cycle, and was at least 15% of the total. Major anchors were amino acids that were present at double the relative concentration of the previous cycle and, alone or with structurally related amino acids, accounted for at least 20% of the total. By these definitions, the amino acids Q and V emerged as major anchors at peptide position two, R as a minor anchor at position 3, K as a major anchor at position 6, and V and L as major anchors at position 9.

Discussion

HLA alleles can be functionally grouped by structural relatedness and peptide-binding similarity into groups called supertypes. An HLA-A2 supertype has been defined that includes members of the A*02 molecular type as well as A*6802 and A*6901 (del Guercio et al. 1995). Notably, the serologically related allele A*6801 is not included in this supertype and has been grouped in the A3 supertype based on motif (Sidney et al. 1996). Some heterogeneity in peptide binding does exist within the supertype, with evidence suggesting a close relationship in the peptide-binding specificity of A*0205 and A*0214 that is distinguishable from the peptide-binding specificity of A*0201 and A*0202 (Barouch et al. 1995). A difference in a polymorphic amino acid at position 9 of the alpha 1 domain influences the P2 binding pocket of the class I heavy-chain (F in A*0201/02 and Y in A*0205/14) and likely accounts for some of the difference.

The HLA-A*02 molecular group is more polymorphic in Africans than in Caucasians or Orientals

Fig. 1 Partial alignment of *HLA-A*0214* (from Krausa et al. 1995a), and *HLA-A*02142*, isolated from a Kenyan sex worker. A single nucleotide substitution at position 78 (T→C) occurs in the latter (*boxed*). *HLA-A*02142* was sequenced to residue 604 (amino acid 177, near the end of exon 3) and is identical to *A*0214* at all other positions

	Exon 1									Exon 2							
Nucleotide #				60											90		
A*02014	**	***	***	***	***	***	***	***	*	GC	TCT	CAC	TCC	ATG	AGG	TAT	TTC
A*020142	**	***	***	***	CAG	ACC	TGG	GCG	G	--	--C	---	---	---	---	---	---
Protein		X	X	X	Q	T	W	A		G	S	H	S	M	R	Y	F
Amino Acid #								-1		1							

	Exon 2															
Nucleotide #								120								144
A*02014	TTC	ACA	TCC	GTG	TCC	CGG	CCC	GGC	CGC	GGG	GAG	CCC	CGC	TTC	ATC	GCA
A*020142	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Protein	F	T	S	V	S	R	P	G	R	G	E	P	R	F	I	A
Amino Acid #		10										20				

Table 1 Pool sequencing of peptides eluted from HLA-A*0214. Peptides were sequenced by Edman degradation and the content of each amino acid was calculated as a fraction of the total at each cycle. Values greater than 10% and at least twice the previous cycle are considered significant. Preferred residues are *underlined*, anchors are *underlined* and in *bold*

Cycle	Amino acids (pmol)	Amino acid residue (% of total)																		
		R	K	H	D	E	G	S	T	N	Q	Y	A	P	M	V	L	I	F	W
1	1121	7.4	8.3	0.3	5.4	1.1	13.0	7.8	4.5	2.9	1.8	8.4	7.1	1.3	1.4	4.5	3.5	11.2	9.8	0.3
2	685	5.3	6.0	0.7	4.2	3.2	4.9	6.9	5.8	1.6	**17.5**	2.9	7.2	2.2	0.7	**12.4**	6.6	8.1	3.7	0.0
3	594	16.3	3.6	1.3	5.9	2.8	3.9	5.4	5.0	1.7	10.1	3.6	4.7	3.4	2.4	8.9	11.7	6.3	2.5	0.5
4	358	9.8	3.5	3.1	10.0	5.6	5.1	6.5	14.3	2.7	4.7	2.9	2.8	5.0	5.2	4.1	8.0	4.2	2.1	0.5
5	313	3.8	4.2	1.9	10.1	5.8	5.3	6.5	10.9	4.6	3.1	2.6	2.9	14.8	7.2	3.5	5.7	4.1	2.1	0.8
6	323	6.2	**20.3**	1.5	6.2	6.3	5.3	4.6	7.8	3.4	5.0	3.4	3.4	8.3	3.9	2.9	4.9	3.3	2.1	0.9
7	372	4.8	16.8	0.7	6.9	4.2	3.9	5.4	8.0	4.1	3.7	5.8	3.1	5.2	2.6	3.3	5.2	12.4	2.5	1.4
8	333	3.0	10.3	1.0	6.4	5.0	4.2	6.3	7.4	3.6	8.6	5.7	3.0	8.7	2.5	3.6	4.8	9.4	5.1	1.3
9	164	4.6	10.9	0.6	4.3	3.5	5.3	3.4	4.3	2.0	4.3	6.8	3.1	12.2	1.2	**10.3**	**10.0**	6.5	5.0	1.8
10	273	4.7	7.6	1.5	4.4	4.0	4.6	5.9	5.9	3.2	5.9	9.9	4.5	5.1	0.9	9.1	10.4	5.4	5.6	1.5

(Krausa et al. 1995b; MacDonald et al. 2000). HLA-A*0214 was originally discovered in Africa (Krausa et al. 1995a). It is relatively rare, with a frequency of 1.6% in the Nairobi Sex Workers' cohort (MacDonald et al. 2000; unpublished data). Here, we extend the A2 molecular family with the discovery of a variant of HLA-A*0214 which we propose to call HLA-A*02142. This molecule is identical in protein sequence to HLA-A*0214 but includes a nucleotide substitution at position 78 (T→C).

We determined the peptide-binding motif for this molecule by pool sequencing of eluted amino acids. The motif, X-[Q,V]-X-X-X-K-X-X-[V,L], has major anchors at positions 2, 6, and 9. The presence of QV as motif residues at position 2 is consistent with the motif for the related molecule A*0205 (VQL), and differs from the motif for A*0201 (LM) and A*0202 (L). We do not see in this analysis a preference for leucine at position 2 as previously reported (Barouch et al. 1995). At position 6, we see a preference for lysine (K). This residue is present in two of nine known A*0124 epitopes (Barouch et al. 1995). The preference for valine and leucine at position 9 is common across the A2/6802 supertype (Sette and Sidney 1999).

We have presented the entire data set for the pool sequencing of HLA-A*02142-bound peptides. Such data may be used in the construction of matrix or neural-net algorithms for the prediction of HLA-binding peptides, both of which have superior predictive value to peptide motifs (Gulukota et al. 1997). Matrix and neural-net algorithms reveal both positive and negative binding interactions, and De Groot and coworkers (1997) have pointed out that data from multiple sources are required for their refinement. Thus, this information may be used to improve predictive algorithms for HLA-A*0214, and the understanding of peptide binding in the A2 supertype.

Acknowledgements The authors gratefully acknowledge the support of the Toronto Hospital Foundation Skate the Dream Fund. K.S.M. is a Career Scientist of the Ontario HIV Treatment Network.

References

Barouch D, Friede T, Stevanovic S, Tussey L, Smith K, Rowland-Jones S, Braud V, McMichael A, Rammensee HG (1995) HLA-A2 subtypes are functionally distinct in peptide binding and presentation. J Exp Med 182:1847–1856

Bunce M, O'Neill C, Barnardo MC, Krausa P, Browning MJ, Morris PJ, Welsh KI (1995) Phototyping: comprehensive DNA typing for HLA-A, B, C, DRB1, DRB3, DRB4, DRB5 & DQB1 by PCR with 144 primer mixes utilizing sequence-specific primers (PCR-SSP). Tissue Antigens 46:355–367

De Groot AS, Jesdale BM, Szu E, Schafer JR, Chicz RM, Deocampo G (1997) An interactive Web site providing major histocompatibility ligand predictions: application to HIV research. AIDS Res Hum Retroviruses 13:529–531

Domena JD, Little AM, Madrigal AJ, Hildebrand WH, Johnston-Dow L, Toit E du, Bias WB, Parham P (1993) Structural heterogeneity in HLA-B70, a high-frequency antigen of black populations. Tissue Antigens 42:509–517

Ennis PD, Zemmour J, Salter RD, Parham P (1990) Rapid cloning of HLA-A,B cDNA by using the polymerase chain reaction: frequency and nature of errors produced in amplification. Proc Natl Acad Sci USA 87:2833–2837

Falk K, Rötzschke O, Stevanovic S, Jung G, Rammensee HG (1991) Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. Nature 351:290–296

Fowke KR, Nagelkerke NJ, Kimani J, Simonsen JN, Anzala AO, Bwayo JJ, MacDonald KS, Ngugi EN, Plummer FA (1996) Resistance to HIV-1 infection among persistently seronegative prostitutes in Nairobi, Kenya. Lancet 348:1347–1351

Guercio MF del, Sidney J, Hermanson G, Perez C, Grey HM, Kubo RT, Sette A (1995) Binding of a peptide antigen to multiple HLA alleles allows definition of an A2-like supertype. J Immunol 154:685–693

Gulukota K, Sidney J, Sette A, DeLisi C (1997) Two complementary methods for predicting peptides binding major histocompatibility complex molecules. J Mol Biol 267:1258–1267

Krausa P, Barouch D, Bodmer JG, Hill AV, Mason C, McMichael AJ, Browning MJ (1995a) Characterization of a novel HLA-A2 variant, A*0214, by ARMS-PCR and DNA sequencing. Immunogenetics 41:50

Krausa P, Brywka M 3rd, Savage D, Hui KM, Bunce M, Ngai JL, Teo DL, Ong YW, Barouch D, Allsop CE (1995b) Genetic polymorphism within HLA-A*02: significant allelic variation revealed in different populations. Tissue Antigens 45:223–231

MacDonald KS, Fowke KR, Kimani J, Dunand VA, Nagelkerke NJ, Ball TB, Oyugi J, Njagi E, Gaur LK, Brunham RC, Wade J, Luscher MA, Krausa P, Rowland-Jones S, Ngugi E, Bwayo JJ, Plummer FA (2000) Influence of HLA supertypes on susceptibility and resistance to human immunodeficiency virus type 1 infection. J Infect Dis 181:1581–1589

Sette A, Sidney J (1999) Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. Immunogenetics 50:201–212

Sidney J, Grey HM, Southwood S, Celis E, Wentworth PA, Guercio MF del, Kubo RT, Chesnut RW, Sette A (1996) Definition of an HLA-A3-like supermotif demonstrates the overlapping peptide-binding repertoires of common HLA molecules. Hum Immunol 45:79–93