Mutational Analyses and Natural Variability of the gp41 Ectodomain

Rogier W. Sanders[1], Bette Korber[2], Min Lu[3], Ben Berkhout[1], and John P. Moore[4]

[1] *Department of Human Retrovirology, Academic Medical Center, University of Amsterdam, 1105 AZ Amsterdam, The Netherlands; r.w.sanders@amc.uva.nl*
[2] *Theoretical Biology and Biophysics, MS K710, Los Alamos National Laboratory, Los Alamos, New Mexico 87545*
[3]*Department of Biochemistry and* [4]*Department of Microbiology and Immunology, Weill Medical College, Cornell University, New York, New York 10021*

The HIV-1 envelope glycoproteins mediate viral attachment and release of the viral core in susceptible target cells. A single gp160 precursor protein is processed intracellularly to yield the native form of the envelope complex, consisting of three gp120 and three gp41 molecules associated through non-covalent interactions. Upon receptor and co-receptor binding to the surface subunit gp120, conformational changes within the envelope glycoprotein complex enable the insertion of the hydrophobic fusion peptide of the transmembrane subunit gp41 into the target membrane. Subsequent rearrangements within gp41 allow fusion of viral and cellular membranes. These late structural alterations are targeted by the entry inhibitor T-20 (for reviews see 13, 20, 21, 24, 46, 75).

A considerable body of mutagenesis data on structure-function relationships within the HIV-1 gp41 ectodomain (gp41e) has been published over the years. The value of this data-set has been increased considerably by the determination of the structure of the gp41e core, allowing some of the mutational effects to be interpreted and at least partially understood (9, 12, 38, 41, 68, 71). The native, pre-fusion structure of gp41e in the trimeric gp120-gp41 complex on the virion surface prior to receptor engagement is not known, however, and the various transitional structures of gp41 during the virus-cell fusion process are still ill-defined. Consequently, the structural and functional consequences of many amino acid substitutions in gp41e remain unclear.

Here, we have summarized the results of published mutagenesis studies on gp41e (see the accompanying table). The HXB2 reference strain has been used as a basis for numbering individual amino acid residues (Figure 1). This information should facilitate the research of those who study the HIV-1 envelope glycoproteins as fusogens or vaccine antigens. In general, we have tabulated only data for single mutants, but several publications contain information on the effects of multiple amino acid substitutions (25, 43, 44, 49, 56, 57, 62). The table does not include information on every naturally occurring gp41e sequence variant, as the variation is extensive. However, a summary of natural variability in clades B and C is presented in Figure 2. Also, the last two columns in the table present the entropy scores for gp41e positions that have a defined impact on Env function, for both the B clade and the C clade. Not surprisingly, positions identified through mutational analysis as those where substitutions can abrogate key functions, also tend to be highly conserved among the natural variants. The clearest example is provided by positions where substitutions essentially eliminate cell-cell fusion (*i.e.*, where fusion efficiencies in syncytium assays or reporter gene assays have been reduced to less than 3% of the wild-type value). Sites at which substitutions can abrogate cell-cell fusion tended to be more invariant among 123 B clade sequences (26/44, 59%), compared to those sites where amino acid changes did not dramatically reduce fusion (11/39, 28%, Fisher's exact test $p = 0.004$). Some unusual gp41e variants found in neutralization-resistant isolates are also included in the table, as are variants that arise in response to selection pressure, both *in vitro* and *in vivo*, from the entry inhibitor T-20, which targets gp41e.

The precision with which the available data could be analyzed was sometimes limited because different viral clones, isolates and assays were used to obtain the experimental data. We have therefore chosen to summarize quantitative parameters using the grading system –, +, ++ and +++, as indicated in the footnotes. In some cases these grades had to be deduced from the primary reports, so readers are encouraged to consult the original papers for quantitative details; we regret any errors of interpretation we may have made during this estimation process. Not surprisingly, perhaps, different studies sometimes yielded conflicting results. We have recorded the conflicting data sets but shall leave it to the readers to judge which are the more plausible.

The natural variability of residues in clade B and C isolates was analyzed and mapped on the structure of gp41 (see Figures 2 and 3). A focus of variable residues in clade B sequences is located in the upper part of the C-terminal helix centered around the highly variable leucine-glutamate-glutamine (LEQ) triplet, indicating that this region is under selective pressure. However, it is also possible that certain changes in residues in this region have little impact on Env function, particularly if there is some flexibility in Env structure(s) around this region. This relatively variable region also contains four glycosylation sites, which could be involved in immune evasion (30). Indeed, mutations that affect

glycosylation in this region can modulate neutralization sensitivity (65). Of note is that no CTL or antibody epitopes have been mapped to this region despite the intense positive selection. One interpretation of this observation is that the selection pressure is exerted indirectly on distant antibody epitopes elsewhere in gp41e or even in gp120 (32). Another is that some neutralizing antibodies remain as yet undiscovered in this region of gp41e. In clade C viruses the variability is somewhat shifted towards the 2F5 epitope, compared to clade B. Furthermore, certain residues are significantly more variable in clade C viruses compared to clade B, and vice versa, suggesting that subtly different selection pressures may operate on viruses from the two clades.

Acknowledgments. We thank Brian Foley and Charles Calef for their help with graphical presentation of Figures and Tables. Financial support was obtained from the Dutch AIDS Fund, Amsterdam.

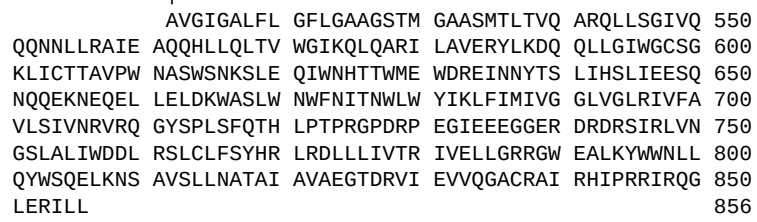

Figure 1. The HXB2 reference strain and the numbering of positions in the gp41 sequence. Only information on the ectodomain (residue 512–684) is incorporated in subsequent analyses.

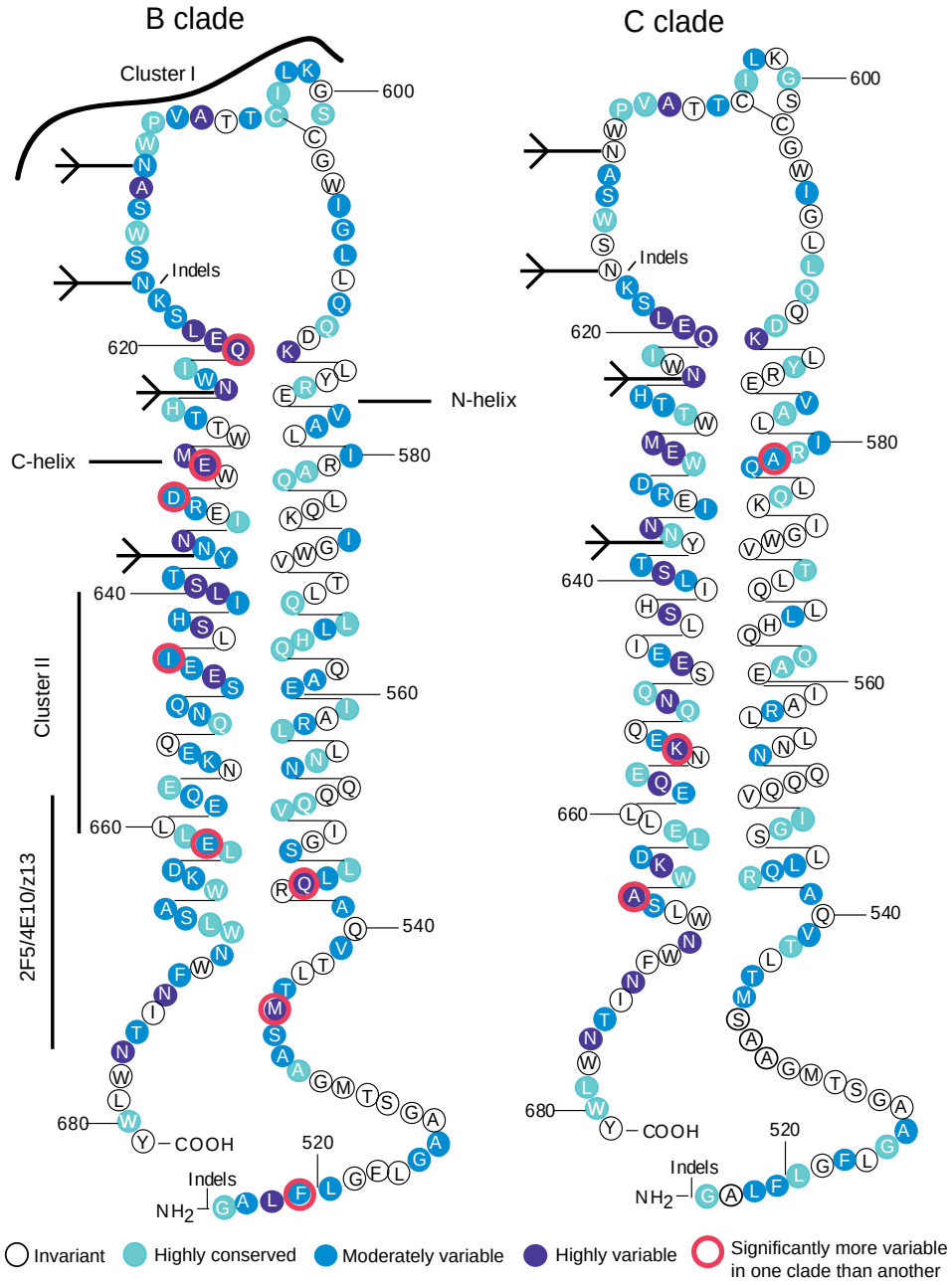

Figure 2. Variability of gp41e. The relative entropies of residues were mapped onto a 2D representation of the HXB2 gp41e (adapted from 29, 61). The variability of residues in clade B isolates (left panel) and clade C isolates (right panel) is indicated according to their entropy values. The entropy is a simple measure of variation in each position based on a sequence alignment (33). Not surprisingly, entropy values for each amino acid were highly correlated with the ratio of the nonsynonymous/synonymous substitution rates, a measure which is indicative of selective pressure, calculated using PAML (76) (Spearman's rank correlation tests gave $z = 7.3, p = 2 \times 10^{-13}$ for the B clade, and $z = 7.5, p = 5 \times 10^{-14}$ for the C clade). We used the entropy scores as our measure of variability here because they lent themselves to testing for differences in variability between the B clade and C clade (33). The color coding for the sites is as follows: white, invariant (entropy score of zero); light blue, very conserved (entropy score below the median, corresponding to only one observed substitution); medium blue, variable (entropy score above the median: 2 or more observed substitutions); dark blue, highly variable (highest 10% of entropy scores: > 0.8 for clade B and > 0.75 for clade C). Residues that are significantly more variable in clade B than in clade C or vice versa (p value ≤ 0.03 after a Bonferroni correction for multiple tests, using a Monte Carlo scheme and randomizing the B and C clade data 10,000 times) are indicated by red circles. 123 clade B sequences and 48 clade C sequences were used for the analyses. The four glycans and the major antibody epitopes (non-neutralizing clusters I and II and the neutralizing 2F5/4E10/z13 cluster) are also indicated, as are regions labelled "indel" where insertions and deletions are frequently observed in natural variants.

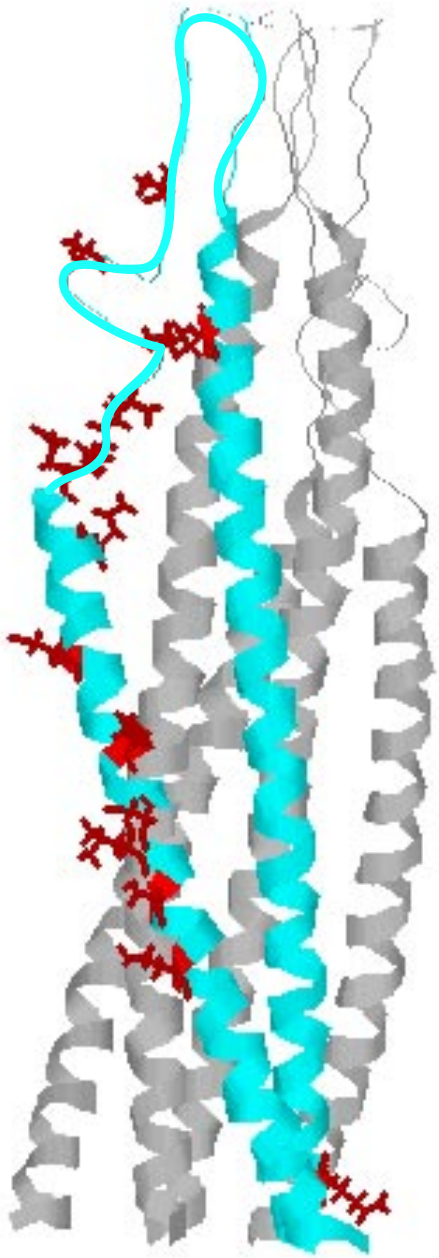

Figure 3. The residues with the highest 10% of entropy scores in clade B are indicated in red on the 3D structure model of Caffrey (pdb accession number 1IF3, (8)). These residues are only indicated in one monomer. The other two monomers are shown in grey for orientation purposes.

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
WT					++	++	++	++	++	++		++	++	++	++	+	++		
A512		V[16]	NL4-3	Freed90	++		++	++	++		+							0.136	0
		E	NL4-3	Freed90	++		++	++	++		+								
V513		E	NL4-3	Freed90	++		++	++	++		–							0.326	0.44
		A	NL4-3	Buchschacher95							++								
		G	NL4-3								++								
		R	NL4-3		++		++				–								
G514		V	NL4-3	Delahunty96	++		++	++			++							0.628	0.594
G516		V	NL4-3	Delahunty96	+++		++	++			++							0.047	0.101
A517		[17]	HXB2	Kowalski91	++	++	++	++	++					++				0.115	0
		[18]	HXB2	Kowalski91										–					
M518		V[19]	ELI1	Kozak97					+++					++				0.985	0.658
F519		L[16]	NL4-3	Freed90	++		++	++	++		+							0.19	0.473
		V	NL4-3	Delahunty96	+++		++	++			++			++					
L520		R	NL4-3	Freed90	++		++	++	++		–							0.13	0.101
G521		V	NL4-3	Delahunty96	+	++	++	++			–			–				0	0
F522		V	NL4-3	Delahunty96	+++		++	++			+							0	0.302
		G	BH8	Pritsker99	++		++				+								
G524		V	NL4-3	Delahunty96	+++	++	++	++			+			+				0.083	0.101
A525		T[20]	LAI	Bahbouhi01	++		++					++		++				0.115	0.202
A526		E	NL4-3	Freed90	++		+	+	++		–							0	0
G527		V	NL4-3	Delahunty96	+++		++	–			–							0	0
S528		T	HXB2	Cao93		+	+	–		+	–		+					0	0
M530		S	HXB2	Cao93		++	–	–		+	–		–					0	0
G531		V	NL4-3	Delahunty96	+++		++	++			++							0	0
L537		R	NL4-3	Freed90	++		+	+	++		–							0	0
V539		E	NL4-3	Freed90	++		++	++	++		+							0.083	0.334
Q540		L	NL4-3	Freed90	++		+	+	++		–							0	0

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
R542	e in heptad-repeat	G	NL4-3	Freed90	++		++	++	++		+							0	0.101
Q543	f in heptad-repeat	H	PI	Wei02, Kilby02										++				0.811	0.202
		R												++					
P543		L[28]	MN	Park00										++					
L544	g in heptad-repeat	S[22]	PI	Fikkert02										++				0.094	0.234
L545	a in heptad-repeat	F[21]	JR-FL	Sanders02		–										+		0.047	0
		N[21]	JR-FL			++										+			
		P[21]	JR-FL			++										+			
		G[21]	JR-FL			++										+			
G547	c in heptad-repeat	S[22]	NL4-3	Rimsky98										++				0	0.101
		D[22]	NL4-3											++					
		D[22]	PI	Baldwin03										++					
		V[22]	PI	Poveda02										++					
		D[22]	PI	Wei02										++					
I548	d in heptad-repeat	A	HXB2	Cao93	++	++	+	+++		++	+		++					0	0.101
		T[22]	NL4-3	Rimsky98										++					
		K[22]	PI	Baldwin03										++					
		V[22]	PI	Wei02, Kilby02										++					
		V[21]	JR-FL	Sanders02		++										+			
		L[21]	JR-FL			++										+			
		H[21]	JR-FL			++										+			
		N[21]	JR-FL			++										+			
		S[21]	JR-FL			++										+			
		G[21]	JR-FL			++										+			
		R[21]	JR-FL			++										+			

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
V549	e in heptad-repeat	M[22]	NL4-3	Rimsky98										++				0.047	0
		M[22]	PI	Wei02										++					
		A[22]	PI											++					
		A[22]	PI	Baldwin03										++					
		W[22]	PI											++					
		G[22]	PI											++					
		A	HXB2	Lu01		++	++	++			++						++		
Q551	g in heptad-repeat	A	HXB2	Lu01, Follis02		++	++	++			++		++				++	0	0
Q552	a in heptad-repeat	L	HXB2	Cao93	++		–	–			–							0	0
N554	c in heptad-repeat	K[22]	PI	Fikkert02										++				0.047	0
L555	d in heptad-repeat	G	HXB2	Cao93	++		–	–			–							0	0
		A	BH8	Poumbourios97	++		–		++		–				++				
		V[21]	JR-FL	Sanders02		–													
		W[21]	JR-FL			–													
		Y[21]	JR-FL			–													
		S[21]	JR-FL			–													
		P[21]	JR-FL			–													
L556	e in heptad-repeat	P	HXB2	Chen94	++		+	–					–					0.047	0
		R	HXB2	Weng98	–						–		–						
		E	HXB2		–						–		–						
		A	HXB2		+	–							–						
		D	HXB2	Weng00	++		++					++	–						
		G	HXB2		++		++					++	–						
		K	HXB2		–								–						
		N	HXB2		++		++					++	–						
		A	HXB2	Lu01, Follis02		++	+	++			–		+				++		
		P[21]	JR-FL	Sanders02		++										+			

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
R557	f in heptad-repeat	P[21]	JR-FL	Sanders02		++										+		0.237	0.334
		M	PI	Wei02										++					
A558	g in heptad-repeat	R	HXB2	Weng98	–								–					0	0
		E	HXB2		+								–						
		C	HXB2	Weng00	++		++					++	–						
		G	HXB2		++		++					++	–						
		T	HXB2		++		++					++	+						
		P[21]	JR-FL	Sanders02		++										+			
I559	a in heptad-repeat	P	HXB2	Chen93, Chen94	++	++	–	–	++		–		–		++			0.047	0
		A	BH8	Poumbourios97	++		–		++		–				++				
		V[21]	JR-FL	Sanders02		+										+			
		F[21]	JR-FL			–										+++			
		N[21]	JR-FL			–										+++			
		P[21]	JR-FL			++	++									+++	–		
		G[21]	JR-FL			+	++									+++	+		
		R[21]	JR-FL			+										+++			
		P	LAI/ JR-FL	Sanders03b										–			–		
		G	LAI/ JR-FL											–			+		
		L	LAI/ JR-FL											++			++		
E560	b in heptad-repeat	P[21]	JR-FL	Sanders02		+++										+		0.217	0
		G[19]	ELI1	Kozak97							+								
A561	c in heptad-repeat	P[21]	JR-FL	Sanders02		+++										+		0.094	0.101
S561		A[28]	MN	Park00										++					

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
Q562	d in heptad-repeat	L	HXB2	Cao93	++		+	–			–							0	0.101
		A	BH8	Poumbourios97	++		++	+	++		–				++				
		P[21]	JR-FL	Sanders02		+++										+			
Q563	e in heptad-repeat	A	HXB2	Weng00	++		++					++	++					0.047	0
		E	HXB2		++		++					++	–						
		M	HXB2		++		++					++	–						
		G	HXB2		++		++					++	++						
		R	HXB2		++		++					++	+						
		A	HXB2	Lu01, Follis02		++	++	++			++		++				++		
		P[21]	JR-FL	Sanders02		+++										+			
R564	f in heptad-repeat	P[21]	JR-FL	Sanders02		+++										+		0.047	0
H564		N[28]	MN	Park00										++					
		C[26]	HXB2	Rabenstein95											++				
L565	g in heptad-repeat	P	HXB2	Chen94	++	++	+	++	++		–							0.402	0.584
		A	HXB2	Lu01, Follis02		++	++	++			–		–				+		
		P[21]	JR-FL	Sanders02		++										+			
L566	a in heptad-repeat	G	HXB2	Cao93		++	+	+		–	–		+					0.047	0
		P	HXB2	Chen93, Chen94	++	++	+	+	++		–		–		++				
		A	BH8	Poumbourios97	++		–		++		–				++				
		V[23]	BH8	Earl93		++	++		++						++				
		V[21]	JR-FL	Sanders02		+	++									++	++		
		I[21]	JR-FL			–										+			
		N[21]	JR-FL			+										++			
		T[21]	JR-FL			+										++			
		P[21]	JR-FL			+	++									+			
		K[21]	JR-FL			–										+			

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
Q567	b in heptad-repeat	R	LAI	Sanders03a										++			++	0.177	0
L568	c in heptad-repeat	A	HXB2	Cao93	++	++	+	++	++	++	–		+					0	0
		P	HXB2	Chen94	++	++	+	+	++		–								
		A	HXB2	Ji00													+		
T569	d in heptad-repeat	A	BH8	Poumbourios97	++		–		++		–				++			0	0.101
		C	HXB2	Farzan98			–												
		S[21]	JR-FL	Sanders02		+										+			
		P[21]	JR-FL			+	++									++	+		
		K[21]	JR-FL			+										++			
		E[21]	JR-FL			–													
V570	e in heptad-repeat	R	HXB2	Weng98	++	++			++		–	++	–					0	0
		E[35]	HXB2		++	++			++			++							
		A	HXB2	Weng00	++		++					++	–						
		D	HXB2		++		++					++	–						
		E	HXB2		++		++					++	–						
		G	HXB2		++		++					++	–						
		I	HXB2		++		++					++	++						
		A	HXB2	Lu01, Follis02		++	++	++			–		–				+		
W571	f in heptad-repeat	R	HXB2	Cao93	++	++	+	++	++	–	–		–					0	0
		R	HXB2	Ji00													++		
		C[26]	HXB2	Rabenstein95											++				
G572	g in heptad-repeat	G	HXB2	Weng98	++	–			++		–	–	–					0	0
		A	HXB2	Lu01		++	++	++			–						+++		

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
I573	a in heptad-repeat	L	HXB2	Dubay92	++		++	++			++	++		++	++			0.083	0
		V	HXB2		++	++	++	++			++	++		++	++				
		A	HXB2		++		++	++			+	++		+	++				
		G	HXB2		++	++	++	++			–	++		–	++				
		E	HXB2		++	++	++	++			–	++		–	++				
		D	HXB2		++		++	++			–	++		–	++				
		S	HXB2		++		++	++			–	++		–	++				
		P[24]	HXB2	Bernstein95											–				
		A[24]	HXB2												+				
		D[24]	HXB2												–				
		A[25]	HXB3	Shugars96											++				
		S[25]	HXB3												–				
		P	HXB2	Chen93, Chen94	++	+++	+	+	++		–		–		++				
		P[26]	HXB2, LAI	Wild94	++		++	+			–						–		
		A[26]	HXB2, LAI		++	++	++	++			+			+			–		
		S[26]	HXB2, LAI		++	++	++	++			–			–			–		
		P[26]	HXB2	Rabenstein95											–				
		D[26]	HXB2												–				
		S[26]	HXB2												–				
		S	168P	Liu01													–		
		T	168P			++	++	++			++	++	++				+		
		V	LAI	Sanders03a										++			++		
		A	BH8	Poumbourios97	++		++	++	++		–				++				
		V	HXB2	Markosyan02													++		
		A	HXB2														+		
		S	HXB2														+		

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
I573 cont.		P	HXB2														–		
		L[21]	JR-FL	Sanders02		++										+			
		F[21]	JR-FL			++										+			
		Y[21]	JR-FL			++										+			
		Q[21]	JR-FL			++										+			
		N[21]	JR-FL			++										+			
		T[21]	JR-FL			++										+			
		P[21]	JR-FL			++										+			
		G[21]	JR-FL			++										+			
		K[21]	JR-FL			++										+			
K574	b in heptad-repeat	R	BH8	McInerney98	++		++	++	++	++	+							0	0
L576	d in heptad-repeat	P	HXB2	Chen94	++		+	+	++		–							0	0
		A	BH8	Poumbourios97	++		–		++		–				++				
		C[27]	HXB2	Farzan98	++	+	–								+++				
		V[21]	JR-FL	Sanders02		–										+			
		F[21]	JR-FL			–										+			
		Y[21]	JR-FL			–										+			
		Q[21]	JR-FL			–										+			
		N[21]	JR-FL			–										+			
		G[21]	JR-FL			–										+			
		K[21]	JR-FL			–										+			

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
Q577	e in heptad-repeat	R	HXB2	Weng98	++	++			++			++	–					0.047	0.173
		E	HXB2		++	++			++		+	+	+						
		A	HXB2	Weng00	++		++					++	+						
		D	HXB2		++		++					++	++						
		E	HXB2		++		++					++	+						
		G	HXB2		++		++					++	++						
		M	HXB2		++		++					++	+						
		C[27]	HXB2	Farzan98	++	+	–								+++				
		A	HXB2	Lu01		++	++	++			++						++		
A578	f in heptad-repeat	G[27]	HXB2	Farzan98	++	+	–								+++			0.047	0.483
R579	g in heptad-repeat	G	HXB2	Weng00	++		+					++	–					0	0.101
		A	HXB2	Lu01		++	+	++			–						++		
I580	a in heptad-repeat	P	HXB2	Chen94	++	++	+	–	++		–							0.432	0.173
		A	BH8	Poumbourios97	++		++	++	++		–				++				
		L[21]	JR-FL	Sanders02		++										+			
		H[21]	JR-FL			++										+			
		T[21]	JR-FL			++										+			
		P[21]	JR-FL			++										+			
		G[21]	JR-FL			++										+			
L581	b in heptad-repeat	Q[28]	MN	Park00										++				0	0
A582	c in heptad-repeat	T[28]	PI	Reitz88										++				0.094	0.101
		C[26]	HXB2	Rabenstein95											++				

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
V583	d in heptad-repeat	A	BH8	Poumbourios97	++		+	++	++		–				++			0.244	0.503
		C	HXB2	Farzan98			–												
		L[21]	JR-FL	Sanders02		++										+			
		Q[21]	JR-FL			++										+			
		N[21]	JR-FL			++										+			
		S[21]	JR-FL			++										+			
		P[21]	JR-FL			++										+			
		R[21]	JR-FL			++										+			
		K[21]	JR-FL			++										+			
E584	e in heptad-repeat	A	HXB2	Cao93		++	–	–		+	–		–					0	0
		Q	BH8	Maerz01	++	++	++	+			+								
		D	BH8		++		++				+								
		N	BH8		++		++				–								
Y586	f in heptad repeat	R	HXB2	Weng98	++	+			++			+	–					0	0.101
		E	HXB2		++	+			++			+	–						
		C[29]	HXB2	Farzan98			–												
L587	a in heptad-repeat	P	HXB2	Chen93, Chen94	++	++	++	–	++		–		–		++			0	0
		A	BH8	Poumbourios97	++		++	++	++		–				++				
		C[29]	HXB2	Farzan98			–												
		A[21]	JR-FL	Sanders02		–										+			
		P[21]	JR-FL			–										+			
		R[21]	JR-FL			–										+			
		D[21]	JR-FL			–										+			
		E[21]	JR-FL			–										+			
K588		R	BH8	McInerney98	++		++	++	++	++	++							1.112	0.775
D589		L	HXB2	Cao93	++	++	+++	+	++	+	–		+					0	0.101
		C[30]	JR-FL	Binley00		++													
		K	BH8	Maerz01	++	++	++	+			–								

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
Q591		A	BH8	Maerz01	++	++	++	++			++							0.083	0.101
		K	BH8		++		++				++								
		L	LAI	Sanders03c										+					
L592		V	BH8	Maerz01	++		++				++							0	0.101
		A	BH8		++		++				++								
L593		V	BH8	Maerz01	++		++				+		–					0.143	0
		A	BH8		++	++	++	–			+		–						
		Q	LAI	Sanders03c										+/–					
I595		F[31]	PI	Moore93										++				0.162	0.555
W596		M	HXB2	Cao93, Cao94	++	++	++	+	++		–		++	++	++			0	0
		Y	LAI, NL4-3	Rovinski99			++					++							
		A	LAI, NL4-3				–					+							
		C[30]	JR-FL	Binley00		++													
		F	BH8	Maerz01	++	++	++	+			++		++						
		H	BH8		++		++				+								
		L	BH8		++	++	++	+			+								
G597		P	BH8	Maerz01	++	++	++	–			–							0	0
		A	BH8		++	++	++	–			–								
		S	BH8		++	++	++	–			–								
C598		S	HXB2	Dedera92a	++		–				–							0	0
		S[23]	BH8	Earl93		++	++		++						++				
		G	HXB2	Syu91	++		–							–					
		A	LAI	Van Anken03										–					
G600		A	LAI, NL4-3	Rovinski99			++					++						0	0.101

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
K601		R	BH8	McInerney98	++		++	++	++	++	++							0.218	0
		R	LAI, NL4-3	Rovinski99			++					++							
		E	LAI, NL4-3				++					++							
		E	BH8	Merat99	++		+				++								
		E	BH8	Maerz01	++	++	++	+			++								
		H	BH8		++		++	+			++								
		Q	BH8		++		++	+			++								
		A	BH8		++		++				++								
C604		S	HXB2	Dedera92a	++		–				–							0.047	0
		S[23]	BH8	Earl93		++	++		++						++				
		G	HXB2	Syu91	++		–							–					
		A	LAI	Van Anken03										–					
T605		C[30]	JR-FL, HXB2, DH123, 89.6, GUN1-wt	Binley00		++	++	+++	++						+			0.177	0.173
		C	LAI	Sanders03c										++					
		Y	LAI											++					
V608		S	HXB2	Cao93			–	–			–							0.094	0.101
		C[30]	JR-FL	Binley00		++													
P609		C[30]	JR-FL	Binley00		++												0.047	0.101
W610		C[30]	JR-FL	Binley00		++												0.047	0
		F	BH8	Maerz01	++	++	++	–			–								
		H	BH8		++	++	++	–			–								

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
N611	Glycosylation site	Q	HXB2	Dedera92b	++		++	++			+			++	++			0.141	0
		H	HXB2	Lee92	++		++					++		+					
		S	NL4-3	Dash94	++	++	++				++								
		Q	SHIV-KB9	Johnson01	++		++							++					
S613	Glycosylation site N611	A	HXB2	Lee92	++		++							+				0.94	0.274
N616	Glycosylation site	Q	HXB2	Dedera92b							+			++				0.237	0
		Q[23]	BH8	Earl93		++	++		++						++				
		H	HXB2	Lee92	++		++					++		++					
		S	NL4-3	Dash94	++	++	++				++								
		Q	BH10	Perrin98	++		++				+								
		Q	SHIV-KB9	Johnson01	++		++							++					
K617		R	BH8	McInerney98	++		++	++	++	++	++							0.348	0.658
S618	Glycosylation site N616	A	HXB2	Lee92	++		++	–						–				0.495	0.483
N624	d in heptad-repeat	H	HXB2	Lee92	++		++					++		+				1.153	1.305
	Glycosylation site (N625 in most isolates)	Q	BH10	Perrin98	++		++				++								
		Q	SHIV-KB9	Johnson01	++		++							++					
N625	e in heptad-repeat Glycosylation site	Q[23]	BH8	Earl93		++	++		++						++			0.047	0.274
T626	f in heptad-repeat Glycosylation site N624	M	HXB2	Cao93	++	–	–	–		–	–		–					0.244	0.444
		M[28]	SHIV-HXBc2P	Si01										++					

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
W628	a in heptad-repeat	M	HXB2	Cao93			–	–		–	–							0	0
		A	HXB2	Weng00	++		–					++	–						
		F	HXB2		++		–					++	–						
		A	HXB2	Wang02		++	–	++			–						+		
W631	d in heptad-repeat	A	HXB2	Wang02		++	–	++			–						–	0	0.101
D632	e in heptad-repeat	N[32]	BH10	Perrin98	++		++				–							0.591	0.287
R633	f in heptad-repeat	G	PI	Wei02										++				0.55	0.451
I635	a in heptad-repeat	A	HXB2	Wang02		++	–	++			–						+	0.047	0.173
N637	c in heptad-repeat	K[22]	PI	Baldwin03										++				0.141	0.101
	Glycosylation site	Q	HXB2	Dedera92b							++			++					
		Q[23]	BH8	Earl93		++	++		++						++				
		H	HXB2	Lee92	++		++					++		–					
		S	NL4-3	Dash94	++	–	–				–								
		Q	BH10	Perrin98	++		++				++								
		Q	SHIV-KB9	Johnson01	++		++							++					
Y638	d in heptad-repeat	A	HXB2	Wang02		++	++	++					++				++	0.13	0
T639	e in heptad-repeat Glycosylation site	V	HXB2	Lee92	++		++							+				0.083	0.202
	N637	A	HXB2	Cao93	++	–	–			–	–								
I642	a in heptad-repeat	A	HXB2	Wang02		++	–	++			–						++	0.094	0
		A	HXB2	Markosyan02													++		
		S	HXB2														++		
H643	b in heptad-repeat	Y[20]	LAI	Bahbouhi01	++		++					++		++				0.115	0
		Y	LAI	Sanders03a										++					
L645	d in heptad-repeat	A	H64333	Wang02		++	++	++					++				++	0	0

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
E647	f in heptad-repeat	L	HXB2	Cao93		++	+++	+			++		++					0.188	0.173
S649	a in heptad-repeat	A	HXB2	Wang02		++	++	++					++				++	0.401	0
Q652	d in heptad-repeat	L	HXB2	Cao93		++	++	+			++		++					0.047	0.101
		L	HXB2	Shu00													++		
		A	HXB2	Wang02		++	++	++					++				++		
K655	g in heptad-repeat	R[33]	BH8	Poumbourios95	++		++	++			++							0.213	1.093
N656	a in heptad-repeat	L	HXB2	Cao93	++	++	+	++	++	++	–		+					0	0
L663	2F5 epitope	F	HXB2	Cao93		++	+++	++			++		++					0.047	0.101
K665	2F5 epitope	R[33]	BH8	Poumbourios95	++		++	++			++							0.451	0.922
W666	2F5 epitope	P	HXB2	Cao93		++	++	++			++		++					0.047	0.101
		A	HXB2, NL4-3	Salzwedel99	++		++	++			++								
S668	2F5 epitope	N[28]	HXB2	Back93										++				0.497	0.573
L669		P	HXB2	Cao93		+++	++	+		+++	++		++					0.047	0
W670		A	HXB2, NL4-3	Salzwedel99	++		++	++			++							0.047	0
N671	4E10/z13 epitope	P	HXB2	Cao93		++	++	++			++		++					0.713	0.945
W672	4E10/z13 epitope	S	HXB2	Cao93	++	++	++	+		+++	++		++					0	0
		S	HXB2, NL4-3	Salzwedel99	++		++	++			++								
		P	HXB2, NL4-3		++		++	++			++	–	+						
		F	HXB2, NL4-3		++		++	++			++	+	+						
F673	4E10/z13 epitope	P	HXB2	Cao93	++	++	++	+			++		++					0.94	0
		S[34]	HXB2	Stern95							++			++					

Residue[1]	Comments	Substitution	Isolate[2]	Reference	Expression (cell lysate)[3]	Expression (cell surface)[4]	gp160 processing[5]	gp120 association[6]	CD4-binding[7]	CD4-induced shedding[8]	Cell-cell fusion[9]	Virion incorporation[10]	Virus entry[11]	Viral Replication[12]	Oligomerization (gp160/gp140)[13]	Trimerization (SOS gp140)[14]	Thermal stability (gp41 core)[15]	B-clade entropy	C-clade entropy
N674	4E10/z13 epitope	H	HXB2	Lee92	++		++					++		++				1.038	1.375
		S	NL4-3	Dash94	++	++	++				++								
		D[28]	SHIV-HXBc2P	Si01										++					
I675	4E10/z13 epitope	S	HXB2	Cao93		++	++	+			++		++					0	0
		M[28]	HXB2	Back93										++					
N677		R	HXB2	Cao93		++	++	+			++		++					1.237	0.769
W678		A	HXB2	Cao93		++	++	+			++		++					0	0
		A	HXB2, NL4-3	Salzwedel99	++		++	++			++								
W680		A	HXB2, NL4-3	Salzwedel99	++		++	++			++							0.047	0.101
Y681		P	HXB2	Cao93			++	++			++		++					0	0
K683		R	BH8	McInerney98	++		++	++	++	++	++							0.375	0.325

Table footnotes:

[1]Residue numbering is based on HXB2 gp160, although the amino-acids studied may be different in the isolate used. The one-letter code for amino acids is used

[2]PI: primary isolate

[3]As assessed by western blot or immunoprecipitation. –, minimal or no expression; +, reduced expression; ++, expression similar to WT; +++, increased expression

[4]As assessed by surface biotinylation, iodination or FACS. When soluble gp140 constructs were used, the relative secretion levels (western blot or immunoprecipitation) are given. –, minimal or no expression; +, reduced expression; ++, expression similar to WT; +++, increased expression

[5]As assessed by western blot or immunoprecipitation in combination with densitometric measurements. –, minimal or no processing; +, reduced processing; ++, processing similar to WT; +++, increased processing

[6]As assessed by western blot or immunoprecipitation in combination with densitometric measurements. –, minimal or no association; +, reduced association; ++, association similar to WT; +++, increased association

[7]As assessed by immunoprecipitation with CD4-based reagents. ++, similar to WT; +++, increased CD4 binding

[8]As assessed by immunoprecipitation. –, no shedding; +, reduced shedding; ++, shedding similar to WT; +++, increased shedding. Note that CD4-induced shedding and to a lesser extent gp120 association (*i.e.*, the reverse of shedding), when measured in laboratory isolates, might be diminished in primary isolates that can retain gp120 more efficiently.

[9]As assessed by syncytium formation or reporter gene assays. –, fusion lower than 3% of WT; +, fusion between 3 and 30% of WT; ++, fusion greater than 30% of WT

[10]As assessed by western blot or immunoprecipitation. –, minimal or no incorporation; +, reduced incorporation; ++, incorporation similar to WT

[11]As assessed by various assays (replication complementation, use of reporter genes, p24 production). –, entry lower than 3% of WT; +, entry between 3 and 30% of WT; ++, entry

greater that 30% of WT

[12]–, no apparent replication; +, replication with a delay of more than 2 days compared to WT; ++ replication similar to WT

[13]As assessed by sucrose gradient fractionation, immunoprecipitation, velocity sedimentation or FPLC, unless indicated otherwise. –, oligomerization below 25% of WT; +, oligomerization between 25% and 50% of WT; ++, oligomerization similar to WT. No distinction between dimerization, trimerization or tetramerization is made.

[14]As assessed by Blue Native-PAGE. +, trimerization similar to WT SOS gp140 (occasional trimerization); ++, slightly more trimerization than in WT; +++, significantly more trimerization than in WT.

[15]As analyzed using the N34(L6)C28 or N36(L6)C34 peptide model, unless indicated otherwise. –, melting temperature (T_m) below 40°C; +, T_m between 40°C and 60°C; ++, T_m between 60°C and 80°C; +++, T_m over 80°C

[16]Analyzed in a double mutant, A512V + F519L

[17]Four amino-acid insertion GIPA

[18]Six amino-acid insertion IHRWIA

[19]Involved in cell line adaptation

[20]Identified in an isolate which is resistant to the furin inhibitor (α1-PDX)

[21]Analyzed in soluble SOS gp140 constructs and so also contain the A501C and T605C substitutions

[22]Involved in T-20 resistance

[23]Analyzed in soluble gp140

[24]Analyzed in an N-peptide/Protein A fusion protein

[25]Analyzed in an N-peptide/maltose binding protein (MBP) fusion protein

[26]Thermal stability (74) or oligomerization (53) of N-peptides analyzed in the absence of C-peptides

[27]Analyzed in a triple mutant L576C + Q577C + A578G

[28]Involved in neutralization resistance

[29]Analyzed in a double mutant Y586C + L587C

[30]Analyzed in combination with gp120 cysteine substitutions in the context of soluble gp140

[31]Involved in resistance to soluble CD4

[32]Generates a new glycosylation site

[33]Analyzed in a double mutant K655R + K665R

[34]Analyzed in a double mutant A582T + F673S

[35]Data on this mutant were corrected in reference 73

REFERENCES

[1] Van Anken, E., R. W. Sanders, B. Berkhout, and I. Braakman. 2003. unpublished results.

[2] Back, N. K., L. Smit, M. Schutten, P. L. Nara, M. Tersmette, and J. Goudsmit. 1993. Mutations in human immunodeficiency virus type 1 gp41 affect sensitivity to neutralization by gp120 antibodies. *J. Virol*. **67**:6897–6902.

[3] Bahbouhi, B., N. G. Seidah, and E. Bahraoui. 2001. Replication of HIV-1 viruses in the presence of the Portland α1- antitrypsin variant (α1-PDX) inhibitor. *Biochem. J*. **360**:127–134.

[4] Baldwin, C. E., R. W. Sanders, S. Jurriaans, J. M. Lange, and B. Berkhout. 2003. Two step evolution of HIV-1 gp41 during therapy with the T20 entry inhibitor. Submitted.

[5] Bernstein, H. B., S. P. Tucker, S. R. Kar, S. A. McPherson, D. T. McPherson, J. W. Dubay, J. Lebowitz, R. W. Compans, and E. Hunter. 1995. Oligomerization of the hydrophobic heptad repeat of gp41. *J. Virol*. **69**:2745–2750.

[6] Binley, J. M., R. W. Sanders, B. Clas, N. Schuelke, A. Master, Y. Guo, F. Kajumo, D. J. Anselma, P. J. Maddon, W. C. Olson, and J. P. Moore. 2000. A recombinant HIV-1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated structure. *J. Virol*. **74**:627–643.

[7] Buchschacher, G. L., Jr., E. O. Freed, and A. T. Panganiban. 1995. Effects of second-site mutations on dominant interference by a human immunodeficiency virus type 1 envelope glycoprotein mutant. *J. Virol*. **69**:1344–1348.

[8] Caffrey, M. 2001. Model for the structure of the HIV gp41 ectodomain: insight into the intermolecular interactions of the gp41 loop. *Biochim. Biophys. Acta*. **1536**:116–122.

[9] Caffrey, M., M. Cai, J. Kaufman, S. J. Stahl, P. T. Wingfield, D. G. Covell, A. M. Gronenborn, and G. M. Clore. 1998. Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. *EMBO J*. **17**:4572–4584.

[10] Cao, J., L. Bergeron, E. Helseth, M. Thali, H. Repke, and J. Sodroski. 1993. Effects of amino acid changes in the extracellular domain of the human immunodeficiency virus type 1 gp41 envelope glycoprotein. *J. Virol*. **67**:2747–2755.

[11] Cao, J., B. Vasir, and J. G. Sodroski. 1994. Changes in the cytopathic effects of human immunodeficiency virus type 1 associated with a single amino acid alteration in the ectodomain of the gp41 transmembrane glycoprotein. *J. Virol*. **68**:4662–4668.

[12] Chan, D. C., D. Fass, J. M. Berger, and P. S. Kim. 1997. Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89**:263–273.

[13] Chan, D. C. and P. S. Kim. 1998. HIV entry and its inhibition. *Cell* **93**:681–684.

[14] Chen, S. S. 1994. Functional role of the zipper motif region of human immunodeficiency virus type 1 transmembrane protein gp41. *J. Virol*. **68**:2002–2010.

[15] Chen, S. S., C. N. Lee, W. R. Lee, K. McIntosh, and T. H. Lee. 1993. Mutational analysis of the leucine zipper-like motif of the human immunodeficiency virus type 1 envelope transmembrane glycoprotein. *J. Virol*. **67**:3615–3619.

[16] Dash, B., A. McIntosh, W. Barrett, and R. Daniels. 1994. Deletion of a single N-linked glycosylation site from the transmembrane envelope protein of human immunodeficiency virus type 1 stops cleavage and transport of gp160 preventing env-mediated fusion.*J. Gen. Virol*. **75**(Pt 6):1389–1397.

[17] Dedera, D., R. L. Gu, and L. Ratner. 1992a. Conserved cysteine residues in the human immunodeficiency virus type 1 transmembrane envelope protein are essential for precursor envelope cleavage. *J. Virol*. **66**:1207–1209.

[18] Dedera, D. A., R. L. Gu, and L. Ratner. 1992b. Role of asparagine-linked glycosylation in human immunodeficiency virus type 1 transmembrane envelope function. *Virol*. **187**:377–382.

[19] Delahunty, M. D., I. Rhee, E. O. Freed, and J. S. Bonifacino. 1996. Mutational analysis of the fusion peptide of the human immunodeficiency virus type 1: identification of critical glycine residues. *Virol*. **218**:94–102.

[20] Doms, R. W. 2000. Beyond receptor expression: the influence of receptor conformation, density, and affinity in HIV-1 infection. *Virol*. **276**:229–237.

[21] Doms, R. W. and J. P. Moore. 2000. HIV-1 membrane fusion: targets of opportunity. *J. Cell Biol*. **151**:F9–14.

[22] Dubay, J. W., S. J. Roberts, B. Brody, and E. Hunter. 1992. Mutations in the leucine zipper of the human immunodeficiency virus type 1 transmembrane glycoprotein affect fusion and infectivity. *J. Virol*. **66**:4748–4756.

[23] Earl, P. L. and B. Moss. 1993. Mutational analysis of the assembly domain of the HIV-1 envelope glycoprotein. *AIDS Res. Hum. Retroviruses* **9**:589–594.

[24] Eckert, D. M. and P. S. Kim. 2001. Mechanisms of viral membrane fusion and its inhibition. *Annu. Rev. Biochem*. **70**:777–810.

[25] Farzan, M., H. Choe, E. Desjardins, Y. Sun, J. Kuhn, J. Cao, D. Archambault, P. Kolchinsky, M. Koch, R. Wyatt, and J. Sodroski. 1998. Stabilization of human immunodeficiency virus type 1 envelope glycoprotein trimers by disulfide bonds introduced into the gp41 glycoprotein ectodomain. *J. Virol*. **72**:7620–7625.

[26] Fikkert, V., P. Cherepanov, K. van Laethum, A. Hantson, B. van Remoortel, C. Pannecouque, E. de Clercq, Z. Debyser, A. Vandamme, and M. Witvrouw. 2002. Env chimeric virus technology for evaluating human immunodeficiency virus susceptility to entry inhibitors. *Antimicrob. Agents Chemother*. **46**:3954–3962

[27] Follis, K. E., S. J. Larson, M. Lu, and J. H. Nunberg. 2002. Genetic evidence that interhelical packing interactions in the gp41 core are critical for transition of the human immunodeficiency virus type 1 envelope glycoprotein to the fusion-active state. *J. Virol*. **76**:7356–7362.

[28] Freed, E. O., D. J. Myers, and R. Risser. 1990. Characterization of the fusion domain of the human immunodeficiency virus type 1 envelope glycoprotein gp41. *Proc. Natl. Acad. Sci. U. S. A*. **87**:4650–4654.

[29] Gallaher, W. R., J. M. Ball, R. F. Garry, A. M. Martin-Amedec, and R. C. Montelaro. 1989. A general model for the transmembrane proteins of HIV and other retroviruses. *AIDS Res. Hum. Retroviruses*. **5**:431–440.

[30] Johnson, W. E., J. M. Sauvron, and R. C. Desrosiers. 2001. Conserved, N-linked carbohydrates of human immunodeficiency virus type 1 gp41 are largely dispensable for viral replication. *J. Virol*. **75**:11426–11436.

[31] Kilby, J. M., J. P. Lalezari, J. J. Eron, M. Carlson, C. Cohen, R. C. Arduino, J. C. Goodgame, J. E. Gallant, P. Volberding, R. L. Murphy, F. Valentine, M. S. Saag, E. L. Nelson, P. R. Sista, and A. Dusek. 2002. The Safety, Plasma Pharmacokinetics, and Antiviral Activity of Subcutaneous Enfuvirtide (T-20), a Peptide Inhibitor of gp41-Mediated Virus Fusion, in HIV-Infected Adults. *AIDS Res. Hum. Retroviruses* **18**:685–693.

[32] Klasse, P. J., J. A. McKeating, M. Schutten, M. S. Reitz Jr., and M. Robert-Guroff. 1993. An immune-selected point mutation in the transmembrane protein of human immunodeficiency virus type 1 (HXB2-Env:Ala 582(→Thr)) decreases viral neutralization by monoclonal antibodies to the CD4-binding site. Virol. **196**:332–337.

[33] Korber BT, K Kunstman, BK Patterson, M Furtado, M McEvilly, R Levy, and SM Wolinsky. 1994. Genetic differences between blood- and brain-derived viral sequences from human immunodeficiency virus type 1-infected patients: evidence of conserved elements in the V3 region of the envelope protein of brain-derived sequences. *J. Virol*. **68**(11):7467–81

[34] Kowalski, M., L. Bergeron, T. Dorfman, W. Haseltine, and J. Sodroski. 1991. Attenuation of human immunodeficiency virus type 1 cytopathic effect by a mutation affecting the transmembrane envelope glycoprotein. *J. Virol*. **65**:281–291.

[35] Kozak, S. L., E. J. Platt, N. Madani, F. E. Ferro, K. Peden, and D. Kabat. 1997. CD4, CXCR-4, and CCR-5 dependencies for infections by primary patient and laboratory-adapted isolates of human immunodeficiency virus type 1. *J. Virol*. **71**:873–882.

[36] Lee, W. R., X. F. Yu, W. J. Syu, M. Essex, and T. H. Lee. 1992. Mutational analysis of conserved N-linked glycosylation sites of human immunodeficiency virus type 1 gp41. *J. Virol*. **66**:1799–1803.

[37] Liu, J., W. Shu, M. B. Fagan, J. H. Nunberg, and M. Lu. 2001. Structural and functional analysis of the HIV gp41 core containing an Ile573 to Thr substitution: implications for membrane fusion. *Biochem*. **40**:2797–2807.

[38] Lu, M., H. Ji, and S. Shen. 1999. Subdomain folding and biological activity of the core structure from human immunodeficiency virus type 1 gp41: implications for viral membrane fusion. *J. Virol*. **73**:4433–4438.

[39] Lu, M., M. O. Stoller, S. Wang, J. Liu, M. B. Fagan, and J. H. Nunberg. 2001. Structural and functional analysis of interhelical interactions in the human immunodeficiency virus type 1 gp41 envelope glycoprotein by alanine-scanning mutagenesis. *J. Virol*. **75**:11146–11156.

[40] Maerz, A. L., H. E. Drummer, K. A. Wilson, and P. Poumbourios. 2001. Functional analysis of the disulfide-bonded loop/chain reversal region of human immunodeficiency virus type 1 gp41 reveals a critical role in gp120-gp41 association. *J. Virol*. **75**:6635–6644.

[41] Malashkevich, V. N., D. C. Chan, C. T. Chutkowski, and P. S. Kim. 1998. Crystal structure of the simian immunodeficiency virus (SIV) gp41 core: conserved helical interactions underlie the broad inhibitory activity of gp41 peptides. *Proc. Natl. Acad. Sci. U.S.A.* **95**:9134–9139.

[42] Markosyan, R. B., X. Ma, M. Lu, F. Cohen, and G. Melikyan. 2002. The mechanism of HIV-1 Env-mediated cell-cell fusion by recombinant cores of gp41 ectodomain. *Virol.* **302**:174–184.

[43] McInerney, T. L., W. El Ahmar, B. E. Kemp, and P. Poumbourios. 1998. Mutation-directed chemical cross-linking of human immunodeficiency virus type 1 gp41 oligomers. *J. Virol.* **72**:1523–1533.

[44] Merat, R., H. Raoul, T. Leste-Lasserre, P. Sonigo, and G. Pancino. 1999. Variable constraints on the principal immunodominant domain of the transmembrane glycoprotein of human immunodeficiency virus type 1. *J. Virol.* **73**:5698–5706.

[45] Moore, J. P., L. C. Burkly, R. I. Connor, Y. Cao, R. Tizard, D. D. Ho, and R. A. Fisher. 1993. Adaptation of two primary human immunodeficiency virus type 1 isolates to growth in transformed T cell lines correlates with alterations in the responses of their envelope glycoproteins to soluble CD4. *AIDS Res. Hum. Retroviruses* **9**:529–539.

[46] Moore, J. P. and M. Stevenson. 2000. New targets for inhibitors of HIV-1 replication. *Nat. Rev. Mol. Cell Biol.* **1**:40–49.

[47] Park, E. J., M. K. Gorny, S. Zolla-Pazner, and G. V. Quinnan, Jr. 2000. A global neutralization resistance phenotype of human immunodeficiency virus type 1 is determined by distinct mechanisms mediating enhanced infectivity and conformational change of the envelope complex. *J. Virol.* **74**:4183–4191.

[48] Perrin, C., E. Fenouillet, and I. M. Jones. 1998. Role of gp41 glycosylation sites in the biological activity of human immunodeficiency virus type 1 envelope glycoprotein. *Virol.* **242**:338–345.

[49] Poumbourios, P., W. El Ahmar, D. A. McPhee, and B. E. Kemp. 1995. Determinants of human immunodeficiency virus type 1 envelope glycoprotein oligomeric structure. *J. Virol.* **69**:1209–1218.

[50] Poumbourios, P., K. A. Wilson, R. J. Center, W. El Ahmar, and B. E. Kemp. 1997. Human immunodeficiency virus type 1 envelope glycoprotein oligomerization requires the gp41 amphipathic alpha-helical/leucine zipper-like sequence. *J. Virol.* **71**:2041–2049.

[51] Poveda, E., B. Rodes, C. Toro, L. Martinez-Carbonero, J. Gonzalez-Lahoz, and V. Soriano. 2002. Evolution of the gp41 env region in HIV-infected patients receiving T-20, a fusion inhibitor. *AIDS* **16**:1959–1980.

[52] Pritsker, M., J. Rucker, T. L. Hoffman, R. W. Doms, and Y. Shai. 1999. Effect of nonpolar substitutions of the conserved Phe11 in the fusion peptide of HIV-1 gp41 on its function, structure, and organization in membranes. *Biochem.* **38**:11359–11371.

[53] Rabenstein, M. and Y. K. Shin. 1995. A peptide from the heptad repeat of human immunodeficiency virus gp41 shows both membrane binding and coiled-coil formation. *Biochem.* **34**:13390–13397.

[54] Reitz, M. S., Jr., C. Wilson, C. Naugle, R. C. Gallo, and M. Robert-Guroff. 1988. Generation of a neutralization-resistant variant of HIV-1 is due to selection for a point mutation in the envelope gene. *Cell* **54**:57–63.

[55] Rimsky, L. T., D. C. Shugars, and T. J. Matthews. 1998. Determinants of human immunodeficiency virus type 1 resistance to gp41- derived inhibitory peptides. *J. Virol.* **72**:986–993.

[56] Rovinski, B., G. A. Dekaban, S. X. Cao, F. L. Yao, R. Persson, T. J. Matthews, and M. H. Klein. 1999. Engineering of noninfectious HIV-1-like particles containing mutant gp41 glycoproteins as vaccine candidates that allow vaccinees to be distinguished from HIV-1 infectees. *Virol.* **257**:438–448.

[57] Salzwedel, K., J. T. West, and E. Hunter. 1999. A conserved tryptophan-rich motif in the membrane-proximal region of the human immunodeficiency virus type 1 gp41 ectodomain is important for Env-mediated fusion and virus infectivity. *J. Virol.* **73**:2469–2480.

[58] Sanders, R. W., E. Van Anken, M. Lu, I. Braakman, and B. Berkhout. 2003a. unpublished results.

[59] Sanders, R. W., E. Busser, J. P. Moore, M. Lu, and B. Berkhout. 2003b. unpublished results.

[60] Sanders, R. W., M. Dankers, J. P. Moore, and B. Berkhout. 2003c. unpublished results.

[61] Sanders, R. W., M. Vesanen, N. Schuelke, A. Master, L. Schiffner, R. Kalyanaraman, M. Paluch, B. Berkhout, P. J. Maddon, W. C. Olson, M. Lu, and J. P. Moore. 2002. Stabilization of the soluble, cleaved, trimeric form of the envelope glycoprotein complex of human immunodeficiency virus type 1. *J. Virol.* **76**:8875–8889.

[62] Schaal, H., M. Klein, P. Gehrmann, O. Adams, and A. Scheid. 1995. Requirement of N-terminal amino acid residues of gp41 for human immunodeficiency virus type 1-mediated cell fusion. *J. Virol*. **69**:3308–3314.

[63] Shu, W., J. Liu, H. Ji, L. Radigen, S. Jiang, and M. Lu. 2000. Helical interactions in the HIV-1 gp41 core reveal structural basis for the inhibitory activity of gp41 peptides. *Biochem*. **39**:1634–1642.

[64] Shugars, D. C., C. T. Wild, T. K. Greenwell, and T. J. Matthews. 1996. Biophysical characterization of recombinant proteins expressing the leucine zipper-like domain of the human immunodeficiency virus type 1 transmembrane protein gp41. *J. Virol*. **70**:2982–2991.

[65] Si, Z., M. Cayabyab, and J. Sodroski. 2001. Envelope glycoprotein determinants of neutralization resistance in a simian-human immunodeficiency virus (SHIV-HXBc2P 3.2) derived by passage in monkeys. J. *Virol*. **75**:4208–4218

[66] Stern, T. L., M. S. Reitz, Jr., and M. Robert-Guroff. 1995. Spontaneous reversion of human immunodeficiency virus type 1 neutralization-resistant variant HXB2thr582: in vitro selection against cytopathicity highlights gp120-gp41 interactive regions. *J. Virol*. **69**:1860–1867.

[67] Syu, W. J., W. R. Lee, B. Du, Q. C. Yu, M. Essex, and T. H. Lee. 1991. Role of conserved gp41 cysteine residues in the processing of human immunodeficiency virus envelope precursor and viral infectivity. *J. Virol*. **65**:6349–6352.

[68] Tan, K., J. Liu, J. Wang, S. Shen, and M. Lu. 1997. Atomic structure of a thermostable subdomain of HIV-1 gp41. *Proc. Natl. Acad. Sci. U.S.A.* **94**:12303–12308.

[69] Wang, S., J. York, W. Shu, M. O. Stoller, J. H. Nunberg, and M. Lu. 2002. Interhelical interactions in the gp41 core: implications for activation of HIV-1 membrane fusion. *Biochem*. **41**:7283–7292.

[70] Wei, X., J. M. Decker, H. Liu, Z. Zhang, R. B. Arani, J. M. Kilby, M. S. Saag, X. Wu, G. M. Shaw, and J. C. Kappes. 2002. Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20) monotherapy. *Antimicrob. Agents Chemother*. **46**:1896–1905.

[71] Weissenhorn, W., A. Dessen, S. C. Harrison, J. J. Skehel, and D. C. Wiley. 1997. Atomic structure of the ectodomain from HIV-1 gp41. *Nature* **387**:426–430.

[72] Weng, Y. and C. D. Weiss. 1998. Mutational analysis of residues in the coiled-coil domain of human immunodeficiency virus type 1 transmembrane protein gp41. *J. Virol*. **72**:9676–9682.

[73] Weng, Y., Z. Yang, and C. D. Weiss. 2000. Structure-function studies of the self-assembly domain of the human immunodeficiency virus type 1 transmembrane protein gp41. *J. Virol*. **74**:5368–5372.

[74] Wild, C., J. W. Dubay, T. Greenwell, T. Baird, Jr., T. G. Oas, C. McDanal, E. Hunter, and T. Matthews. 1994. Propensity for a leucine zipper-like domain of human immunodeficiency virus type 1 gp41 to form oligomers correlates with a role in virus- induced fusion rather than assembly of the glycoprotein complex. *Proc. Natl. Acad. Sci. U. S. A.* **91**:12676–12680.

[75] Wyatt, R. and J. Sodroski. 1998. The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science* **280**:1884–1888.

[76] Yang Z, Nielsen R, Goldman N, Pedersen AM. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**(1):431–49