IV

HIV-1/SIVcpz proteins

Construction of HIV-1/SIVcpz protein alignments

The number of full-length gene sequences is still growing rapidly for all genes. The envelope master alignment now contains 307 full-length sequences. For the purposes of the printed alignments, we have had to limit the number of sequences dramatically. Here we list the criteria we have followed to make the selection.

First, we have decided to end the supremacy of the B clade sequences. More than half (198, to be precise) of the full-length envelope sequences are still subtype B, though the contribution of other subtypes is increasing. We have tried to balance the number of representatives of all subtypes in these alignments. For this, we had to make a heavy selection on subtype B sequences. We have tried to include as many "classical" sequences as possible. A lot of follow-up work has been done based on lab strains such as HXB2, MN, SF2, and JR-CSF/JR-FL, so these strains are included in the alignments. Furthermore, within subtype B we have tried to represent sequences from diverse geographical origins, so as to represent a broad spectrum of variants. In the case of subtype B, this means that we have included African, Asian and Brazilian variants along with the "Western" strains. For sequences from non-B subtypes, we have selected a few representative sequences from each dataset, again with an eye on maintaining geographical diversity. When possible we have left all representatives of group O in the alignment, as these sequences are much more genetically diverse that the subtypes.

Explanation of Symbols in Alignments

Symbol	Meaning
Alignment symbols	
? in consensus	no majority-rule consensus could be determined at this position
x	nucleotide missing from codon
#	frameshift, or codon contains N or illegal character
$	stop codon
Annotation symbols	
|- -|	domain boundaries
/	protein start point
\	protein end point
\/	splice site or exon join
->	start of overlapping coding region
<-	end of overlapping coding region
*	cysteine
^^^ [NxS, NxT]	glycosylation site
^*^ [NCS, NCT]	glycosylation site with cysteine
CD4	residue critical for CD4 binding
cds	coding sequence (indicates regions where two proteins overlap; the overlapping proteins use two different reading frames)
MHR	major homology region
nls	nuclear localization signal
phos site	phosphorylation site
PKC	protein kinase C binding
Zn-motif	Zinc finger binding motif

Sources of Annotation in the Alignments

Protein	Annotation	Reference
Gag	phos site Ser (111)	Yu, J Biol Chem 270:4792 (1995)
Gag	MHR, (284-302)	Otteken, J Virol 70:3407 (1996)
Gag	CyPa (205-241)	Braaten, J Virol 70:4220 (1996)
Gag	vpr packaging domain	Lu, J Virol 69:6873 (1995)
	LKSLFG, (489-494)	Kondo, J Virol 70:159 (1996)
Nef	myristylation, (1-7)	Huang, J Virol 69:93 (1995)
Nef	MHC downmodulation, PK recruitment (26-29)	Piguet, p. 448 Human Retroviruses and AIDS (1999)
Nef	heart of CD4 binding site (55-56)	Piguet, p. 448 Human Retroviruses and AIDS (1999)
Nef	acidic cluster, (60-64)	Piguet, p. 448 Human Retroviruses and AIDS (1999)
Nef	(PxxP)3, (67-76)	Huang, J Virol 69:93 (1995)
Nef	PKC, (75-80)	Huang, J Virol 69:93 (1995)
Nef	polypurine tract, (89-97)	Huang, J Virol 69:93 (1995)
Nef	PAK binding, (103-105)	Piguet, p. 448 Human Retroviruses and AIDS (1999)
Nef	Beta turn, (128-131)	Huang, J Virol 69:93 (1995)
Nef	PxxP, (145-148)	Huang, J Virol 69:93 (1995)
Nef	COP1 recruitment (152-153)	Piguet, p. 448 Human Retroviruses and AIDS (1999)
Nef	AP recruitment, (162-163)	Piguet, p. 448 Human Retroviruses and AIDS (1999)
Nef	V-ATPase and Raf-1 binding, (172-173)	Piguet, p. 448 Human Retroviruses and AIDS (1999)
Vpr	alpha helix, (16-34)	Cornelissen, ARHR 13:247 (1997)
Vpr	H(S/F)RIG motifs, (71-82)	Macreadie, PNAS USA 92:2770 (1995)
Vpu	all annotations	Cornelissen, ARHR 13:247 (1997)
Vpr	LR domain, (60-82)	Wang, Gene 178:7 (1996)

Table 1: Table of HIV-1/SIVcpz protein Alignments

Name	Accession	Region	Author	Reference
A.DE.AF200476	AF200476	VIF	Kuhn, J	Unpublished
A.FR.HIV232956	AJ232956	NEF	Jubier-Maurin, V	ARHR 15(1):23-32 (1999)
A.GB.MA246	Y13718	ENV	Douglas, NW	J Mol Biol 273(1):122-149 (1997)
A.KE.AF233689	AF233689	VIF	Kuhn, J	Unpublished
A.KE.K89	L22943	ENV	Louwagie, J	J Virol 69(1):263-271 (1995)
A.KE.Q23	AF004885	ENV GAG NEF POL REV TAT VIF VPR VPU	Poss, M	Unpublished
A.RW.PVPI	L07082	ENV	Bex, F	Unpublished (1992)
A.SE.SE6594	AF069672	GAG NEF POL REV TAT VIF VPR VPU	Carr, JK	AIDS 13(14):1819-1826 (1999)
A.SE.SE7253	AF069670	GAG POL REV TAT VIF VPR VPU	Carr, JK	AIDS 13(14):1819-1826 (1999)
A.SE.SE7535	AF069671	GAG POL REV TAT VIF VPR	Carr, JK	AIDS 13(14):1819-1826 (1999)
A.SE.SE8538	AF069669	GAG NEF POL REV TAT VIF VPR	Carr, JK	AIDS 13(14):1819-1826 (1999)
A.SE.SE8891	AF069673	GAG NEF REV TAT VIF VPR	Carr, JK	AIDS 13(14):1819-1826 (1999)
A.SE.UGSE8131	AF107771	ENV GAG NEF POL REV TAT VIF VPR VPU	Laukkanen, T	-
A.UA.ukr970063	AF082486	ENV REV VPU	Liitsola, K	AIDS 12(14):1907-1919 (1998)
A.UG.92UG037	U51190	ENV GAG NEF POL REV TAT VIF VPR VPU	Gao, F	J Virol 70(3):1651-1657 (1996)
A.UG.U13-2	X91354	VIF	Wieland, U	J Gen Virol 78:393-400 (1997)
A.UG.U455	M62320	ENV GAG NEF POL REV TAT VPR VPU	Oram, JD	ARHR 6(9):1073-1078 (1990)
A.UG.UG273A	L22957	REV TAT VPU	Louwagie, J	J Virol 69(1):263-271 (1995)
A.UG.UG275A	L22951	REV TAT VPU	Louwagie, J	J Virol 69(1):263-271 (1995)
A2.CD.97CDKFE4	AF286240	POL VIF VPR	Gao, F	ARHR 17(8):675-688(2001)
A2.CD.97CDKS10	AF286241	ENV REV TAT VIF VPR VPU	Gao, F	ARHR 17(8):675-688(2001)
A2.CD.97CDKTB48	AF286238	ENV GAG NEF REV TAT VIF VPR VPU	Gao, F	ARHR 17(8):675-688(2001)
A2.CY.94CY017-41	AF286237	ENV GAG NEF POL REV TAT VIF VPR VPU	Gao, F	ARHR 17(8):675-688(2001)

Table 1: cont.

Name	Accession	Region	Author	Reference
B.AU.MBC18	AF042102	ENV GAG POL REV VIF VPR VPU	Oelrichs, RB	ARHR 14(9):811-814 (1998)
B.AU.MBC200	AF042100	NEF TAT	Oelrichs, RB	ARHR 14(9):811-814 (1998)
B.AU.SC49	AF128998	GAG	Oelrichs, RB	Unpublished
B.AU.VH	AF146728	GAG REV VPR VPU	Oelrichs, R	Unpublished
B.CN.RL42	U71182	ENV GAG NEF POL REV TAT VIF VPR VPU	Graf, M	ARHR 14(3):285-288 (1998)
B.DE.D31	U43096	ENV GAG NEF POL REV TAT VIF VPR VPU	Kreutz, R	ARHR 8(9):1619-1629 (1992)
B.DE.HAN	U43141	ENV GAG NEF REV	Sauermann, U	ARHR 6(6):813-823 (1990)
B.ES.89SP061	AJ006287	ENV GAG NEF REV TAT VIF VPR VPU	Olivares, I	ARHR 14(18):1649-165 (1998)
B.FR.HXB2	K03455	ENV GAG NEF POL REV TAT VIF VPR VPU	Wong-Staal, F	Nature 313(6000):277-284 (1985)
B.FR.NE100	M58272	NEF	Delassus, S	J Virol 65:225-231 (1991)
B.FR.SWB884	M58206	NEF	Delassus, S	J Virol 65:225-231 (1991)
B.FR.vi02011A1H	AF143115	VIF	Hassaine, G	Virology 276(1):169-180 (2000)
B.GA.OYI	M26727	ENV GAG NEF POL REV TAT VIF VPR VPU	Huet, T	AIDS 3(11):707-715 (1989)
B.GB.CAM1	D10112	ENV GAG NEF POL REV TAT VIF VPR VPU	McIntosh, AA	Unpublished (1991)
B.GB.GB8	AJ271445	GAG	Farrar, GH	J Med Virol 34(2):104-113 (1991)
B.GB.I4663	Z68564	VPR	Kuiken, CL	J Gen Virol 77(Pt 4):783-792 (1996)
B.GB.I4663	Z68613	VPU	Kuiken, CL	J Gen Virol 77(Pt 4):783-792 (1996)
B.GB.MANC	U23487	GAG	Zhu, T	Nature 374(6522):503-504 (1995)
B.GB.WB	U36882	ENV	Douglas, NW	AIDS 10(1):39-46 (1996)
B.IN.HIVP35A	Y15122	NEF	Ahmad, KM	ARHR 14(16):1491-1493 (1998)
B.IT.B-IT-R5	AF147737	NEF	Catucci, M	J Med Virol 60(3):294-299 (2000)
B.JP.D70887	D70887	VIF	Tominaga, K	ARHR 12(16):1543-1549 (1996)
B.JP.ETR	D12582	ENV	Shimizu, H	Virology 189:534-546 (1992)
B.JP.JH31	M21137	GAG	Komiyama, N	ARHR 5:411-419 (1989)
B.JP.JH32	M21138	ENV VPU	Komiyama, N	ARHR 5:411-419 (1989)
B.JP.PT1-01	AB034578	VPU	Yamada, T	Arch Virol 145(5):1021-1027 (2000)
B.JP.PT1-4	AB034517	VPR	Yamada, T	Arch Virol 145(5):1021-1027 (2000)
B.JP.PT7-6	AB034474	VIF	Yamada, T	Arch Virol 145(5):1021-1027 (2000)

Table 1: cont.

Name	Accession	Region	Author	Reference
B.JP.nef<7>-a	AB034272	NEF	Yamada, T	Arch Virol 145(5):1021-1027 (2000)
B.KR.CSR9412d	AF238268	NEF	Cho, YK	Unpublished
B.KR.WK	AF224507	ENV GAG NEF POL REV TAT VIF VPR VPU	Cho, YK	Unpublished
B.NL.3202A21	U34604	ENV GAG NEF POL REV TAT VIF VPR VPU	Guillon, C	ARHR 11(12):1537-1541 (1995)
B.SE.AF047085	AF047085	NEF	Visco Comandini, U	J Hum Virol 1(5):320-327 (1998)
B.TH.28-19	U48917	NEF	Artenstein, AW	ARHR 12:557-560 (1996)
B.TH.93TH067	U39258	ENV	Penny, MA	ARHR 12(8):741-747 (1996)
B.TH.AF082839	AF082839	NEF	Vallejo, A	AIDS 13(4):532-534 (1999)
B.TT.QZ4589	U32396	ENV	Blattner, W	Unpublished (1995)
B.TW.TWB101	AF220464	VPU	Lee, CN	J Clin Microbiol 38(7):2468-2474 (2000)
B.TW.TWCYS	AF086817	ENV GAG NEF POL REV TAT VIF VPR VPU	Huang, LM	Unpublished
B.UA.UKR1216	AF193278	ENV REV VPU	Liitsola, K	ARHR 16(11):1047-1053 (2000)
B.US.1-2	U41181	VIF	Sova, P	J Virol 69(4):2557-2564 (1995)
B.US.85WCIPR54	U69584	GAG	Fang, G	J AIDS 12(4): 352-357 (1996)
B.US.AD8	AF004394	GAG	Theodore, TS	ARHR 12(3): 191-194 (1996)
B.US.AF019528	AF019528	VIF	Yedavalli, VR	J Virol 72(2):1092-1102 (1998)
B.US.BC	L02317	VIF	Ghosh, SK	Virology 194, 858-864 (1993)
B.US.DH123	AF069140	ENV GAG	Shibata, R	J Virol 69(7):4453-4462 (1995)
B.US.JRCSF	M38429	ENV GAG NEF POL REV TAT VIF VPR VPU	O'Brien, WA	Nature 348:69-73 (1990)
B.US.JRFL	U63632	ENV GAG NEF POL REV TAT VIF	O'Brien, WA	Nature 348:69-73 (1990)
B.US.LM1	U16909	NEF	Huang, Y	J Virol 69(1):93-100 (1995)
B.US.MNCG	M17449	ENV GAG NEF POL REV TAT VIF VPR VPU	Gurgo, C	Virology 164(2):531-536 (1988)
B.US.NC7	AF049495	NEF	Mwaengo, DM	J Virol 72(11):8976-8987 (1998)
B.US.NY5CG	M38431	GAG	Willey, RL	PNAS USA 83(14):5038-5042 (1986)
B.US.RF	M17451	ENV GAG NEF POL REV TAT VIF VPR VPU	Starcich, BR	Cell 45(5):637-648 (1986)
B.US.SC	M17450	REV	Gurgo, C	Virology 164(2):531-536 (1988)
B.US.SF2	K02007	ENV GAG NEF POL REV VIF VPR VPU	van Beveren, CP	RNA tumor viruses, 2nd edition, Vol 2: 1124-1141; Cold Spring Harbor Laboratory (1985)

Table 1: cont.

Name	Accession	Region	Author	Reference
B.US.WC001	AF003887	REV TAT VIF VPR VPU	Fang, G	J AIDS 12(4):352-357 (1996)
C.BI.BU910112	U39233	ENV	Penny, MA	ARHR 12(8):741-747 (1996)
C.BR.92BR025	U52953	ENV GAG NEF POL REV TAT VIF VPR VPU	Gao, F	J Virol 70(3):1651-1667 (1996)
C.BW.96BW01B03	AF110959	POL	Novitsky, VA	J Virol 73(5):4427-4432 (1999)
C.BW.96BW01B21	AF110960	NEF REV VIF VPR	Novitsky, VA	J Virol 73(5):4427-4432 (1999)
C.BW.96BW01B22	AF110961	GAG TAT	Novitsky, VA	J Virol 73(5):4427-4432 (1999)
C.BW.96BW0402	AF110962	ENV GAG NEF POL REV TAT VIF VPR VPU	Novitsky, VA	J Virol 73(5):4427-4432 (1999)
C.BW.96BW0502	AF110967	GAG NEF POL	Novitsky, VA	J Virol 73(5):4427-4432 (1999)
C.BW.96BW0504	AF110968	REV TAT VIF VPR	Novitsky, VA	J Virol 73(5):4427-4432 (1999)
C.BW.96BW1104	AF110969	GAG NEF POL REV VIF	Novitsky, VA	J Virol 73(5):4427-4432 (1999)
C.BW.96BW1210	AF110972	GAG REV	Novitsky, VA	J Virol 73(5):4427-4432 (1999)
C.BW.96BW15B03	AF110973	GAG REV	Novitsky, VA	J Virol 73(5):4427-4432 (1999)
C.BW.96BW1626	AF110978	GAG	Novitsky, VA	J Virol 73(5):4427-4432 (1999)
C.BW.96BW16B01	AF110976	REV	Novitsky, VA	J Virol 73(5):4427-4432 (1999)
C.BW.96BW17A09	AF110979	GAG REV	Novitsky, VA	J Virol 73(5):4427-4432 (1999)
C.CN.AF268277	AF268277	ENV	Chen, Z	J Virol 74(14):6501-6510 (2000)
C.DJ.DJ259A	L22940	REV TAT VPU	Louwagie, J	J Virol 69(1):263-271 (1995)
C.DJ.DJ373A	L23065	ENV REV TAT	Louwagie, J	J Virol 69(1):263-271 (1995)
C.ET.ETH2220	U46016	ENV GAG NEF POL REV TAT VIF VPR VPU	Salminen, MO	ARHR 12(14):1329-1339 (1996)
C.FR.HIV232980	AJ232980	NEF	Jubier-Maurin, V	ARHR 15(1):23-32 (1999)
C.FR.HIV232996	AJ232996	NEF	Jubier-Maurin, V	ARHR 15(1):23-32 (1999)
C.IN.93IN101	AB023804	ENV GAG NEF REV TAT VPR VPU	Mochizuki, N	ARHR 15(14):1321-1324 (1999)
C.IN.93IN904	AF067157	GAG NEF	Lole, KS	J Virol 73(1):152-160 (1999)
C.IN.93IN905	AF067158	GAG	Lole, KS	J Virol 73(1):152-160 (1999)
C.IN.93IN999	AF067154	GAG NEF POL	Lole, KS	J Virol 73(1):152-160 (1999)
C.IN.94IN11246	AF067159	GAG NEF POL REV TAT VIF VPR	Lole, KS	J Virol 73(1):152-160 (1999)
C.IN.95IN21068	AF067155	GAG NEF POL REV TAT VIF VPR	Lole, KS	J Virol 73(1):152-160 (1999)
C.IN.AF209990	AF209990	GAG	Gupta, S	Protein Epr Purif 21(7):378-385 (2001)
C.IN.HIVY15117	Y15117	NEF	Ahmad, KM	ARHR 14(16):1491-1493 (1998)

Table 1: cont.

Name	Accession	Region	Author	Reference
C.IN.HIVY17884	Y17884	NEF	Ahmad, KM	ARHR 14(16):1491-1493 (1998)
C.IN.HIVY17891	Y17891	NEF	Ahmad, KM	ARHR 14(16):1491-1493 (1998)
C.IN.HIVY17892	Y17892	NEF	Ahmad, KM	ARHR 14(16):1491-1493 (1998)
C.SN.SE364A	L22944	VPU	Louwagie, J	J Virol 69(1):263-271 (1995)
C.SO.SO145A	L22946	ENV REV VPU	Louwagie, J	J Virol 69(1):263-271 (1995)
C.TW.TWC2	AF220473	VPU	Lee, CN	J Clin Microbiol 38(7):2468-2474 (2000)
C.UG.UG268A2	L22948	ENV REV VPU	Louwagie, J	J Virol 69(1):263-271 (1995)
D.CD.84ZR085	U88822	ENV GAG NEF POL REV TAT VIF VPR VPU	Gao, F	J Virol 72(7):5680-5698 (1998)
D.CD.ELI	K03454	ENV GAG NEF POL REV TAT VIF VPR VPU	Alizon, M	Cell 46(1):63-74 (1986)
D.CD.JY1	J03653	ENV	Yourno, J	ARHR 4:165-173 (1988)
D.CD.NDK	M27323	ENV GAG NEF POL REV TAT VIF VPR VPU	Spire, B	Gene 81:275-284 (1989)
D.CD.Z2Z6	M22639	GAG POL REV TAT VIF VPR	Srinivasan, A	Gene 52:71-82 (1987)
D.CI.CI13	AJ277820	ENV	Beirnaert, E	Virology 281(2):305-314 (2001)
D.JP.PT14-4	AB034541	VPR	Yamada, T	Arch Virol 145(5):1021-1027 (2000)
D.SN.SE365A2	L22945	ENV REV TAT VPU	Louwagie, J	J Virol 69(1):263-271 (1995)
D.TZ.87TZ4622	U65075	ENV	Robbins, KE	ARHR 12(14):1389-1391 (1996)
D.TZ.TZ005	U12406	VPU	Siwka, W	ARHR 10(12):1753-1754 (1994)
D.UG.92UG024-D	U08805	ENV	WHO Global Programme	ARHR 10(11):1327-1343 (1994)
D.UG.94UG1141	U88824	GAG NEF POL REV TAT VIF VPR VPU	Gao, F	J Virol 72(7):5680-5698 (1998)
D.UG.U18-0	X91355	VIF	Wieland, U	J Gen Virol 78:393-400 (1997)
D.UG.U25-6	X91361	VIF	Wieland, U	J Gen Virol 78:393-400 (1997)
D.UG.U36-0	X91363	VIF	Wieland, U	J Gen Virol 78:393-400 (1997)
D.UG.UG266A2	L22947	VPU	Louwagie, J	J Virol 69(1):263-271 (1995)
D.UG.UG269A	L22949	REV	Louwagie, J	J Virol 69(1):263-271 (1995)
D.UG.UG274A2	L22950	REV TAT	Louwagie, J	J Virol 69(1):263-271 (1995)
D.UG.WHO15-474	U36886	ENV	Douglas, NW	AIDS 10(1):39-46 (1996)
D.ZR.AF233690	AF233690	VIF	Kuhn, J	Unpublished
F1.BE.VI850	AF077336	ENV GAG NEF POL REV TAT VIF VPR VPU	Carr, JK	Virology 269(1):95-104 (2000)

Table 1: cont.

Name	Accession	Region	Author	Reference
F1.BR.93BR020-1	AF005494	ENV GAG NEF POL REV TAT VIF VPR VPU	Gao, F	J Virol 72(7):5680-5698 (1998)
F1.BR.BZ126	L22082	ENV REV TAT VPU	Louwagie, J	ARHR 10(5):561-567 (1994)
F1.BR.BZ162	L11751	GAG	Louwagie, JJ	AIDS 7:769-780 (1993)
F1.BR.BZ163	L22085	REV TAT VPU	Louwagie, J	ARHR 10(5):561-567 (1994)
F1.CD.VI174	L11782	GAG	Louwagie, JJ	AIDS 7:769-780 (1993)
F1.DE.AF200475	AF200475	VIF	Kuhn, J	Unpublished
F1.FI.FIN9363	AF075703	ENV GAG NEF POL REV TAT VIF VPR VPU	Laukkanen, T	Unpublished
F1.FR.MP411	AJ249238	ENV GAG NEF POL REV TAT VIF VPR VPU	Peeters, M	ARHR 16(2):139-151(2000)
F1.RW.VI69	L11796	GAG	Louwagie, JJ	AIDS 7:769-780 (1993)
F2.CM.CA20	AJ277824	ENV	Nyambi, PN	J Virol 70(9):6235-6243 (1996)
F2.CM.HIM277819	AJ277819	ENV	Beirnaert, E	Virol 281(2):305-314 (2001)
F2.CM.HIV232985	AJ232985	NEF	Jubier-Maurin, V	ARHR 15(1):23-32 (1999)
F2.CM.MP255	AJ249236	ENV GAG POL REV TAT VIF VPR VPU	Peeters, M	ARHR 16(2):139-151(2000)
F2.CM.MP257	AJ232986	NEF	Jubier-Maurin, V	ARHR 15(1):23-32 (1999)
F2.CM.MP257	AJ249237	ENV GAG POL REV TAT VIF VPR VPU	Peeters, M	ARHR 16(2):139-151(2000)
G.BE.DRCBL	AF084936	ENV GAG NEF POL REV TAT VIF VPR VPU	Debyser, Z	ARHR 14(5):453-459 (1998)
G.CG.CNG30	AF056186	VIF VPR VPU	Harada, Y	Unpublished
G.FI.HH8793-1-1	AF061640	GAG NEF POL REV VIF VPR	Salminen, MO	ARHR 8(9):1733-1742 (1992)
G.FI.HH8793-12-1	AF061641	ENV TAT VPU	Salminen, MO	ARHR 8(9):1733-1742 (1992)
G.GA.LBV217	U09664	ENV	Janssens, W	ARHR 10:877-879 (1994)
G.ML.HIV232990	AJ232990	NEF	Jubier-Maurin, V	ARHR 15(1):23-32 (1999)
G.NG.92NG083	U88826	ENV GAG NEF POL REV TAT VIF VPR VPU	Gao, F	J Virol 72(7):5680-5698 (1998)
G.NG.IKCSW22	AJ232991	NEF	Jubier-Maurin, V	ARHR 15(1):23-32 (1999)
G.NG.MACSW39	AJ232992	NEF	Jubier-Maurin, V	ARHR 15(1):23-32 (1999)
G.NG.NG1937	AF069937	ENV	McCutchan, FE	Virology 254(2):226-234 (1999)
G.NG.NG1939	AF069935	ENV	McCutchan, FE	Virology 254(2):226-234 (1999)
G.SE.SE6165	AF061642	ENV GAG NEF POL REV TAT VIF VPR VPU	Carr, JK	Virology 247(1):22-31 (1998)

Table 1: cont.

Name	Accession	Region	Author	Reference
G.TW.TWG1	AF220486	VPU	Lee, CN	J Clin Microbiol 38(7):2468-2474 (2000)
H.BE.VI991	AF190127	ENV GAG NEF POL REV TAT VIF VPR VPU	Laukkanen, T	AIDS 14(11):1533-1543 (2000)
H.BE.VI997	AF190128	ENV GAG NEF POL REV TAT VIF VPR VPU	Laukkanen, T	AIDS 14(11):1533-1543 (2000)
H.CD.HIV232994	AJ232994	NEF	Jubier-Maurin, V	ARHR 15(1):23-32 (1999)
H.CD.HIV232995	AJ232995	NEF	Jubier-Maurin, V	ARHR 15(1):23-32 (1999)
H.CF.90CF056	AF005496	ENV GAG NEF POL REV TAT VIF VPR VPU	Murphy, E	ARHR 9(10):997-1006 (1993)
J.SE.SE7022	AF082395	ENV GAG NEF POL REV TAT VIF VPR VPU	Laukkanen, T	ARHR 15(3):293-297 (1999)
J.SE.SE7887	AF082394	ENV GAG NEF POL REV TAT VIF VPR VPU	Laukkanen, T	ARHR 15(3):293-297 (1999)
K.BE.VI325	L11789	GAG	Louwagie, JJ	AIDS 7:769-780 (1993)
K.CD.EQTB11C	AJ249235	ENV GAG NEF POL REV TAT VIF VPR VPU	Peeters, M	ARHR 16(2):139-151 (2000)
K.CM.MP535	AJ249239	ENV GAG NEF POL REV TAT VIF VPR VPU	Peeters, M	ARHR 16(2):139-151 (2000)
N.CM.YBF106	AJ271370	ENV GAG NEF POL REV TAT VIF VPR	Souquiere, S	Unpublished
N.CM.YBF30	AJ006022	ENV GAG NEF POL REV TAT VIF VPR VPU	Simon, F	Nature Med 4(9):1032-1037 (1998)
O.CM.ANT70	L20587	ENV GAG NEF POL REV TAT VIF VPR VPU	Vanden Haesevelde, M	J Virol 68(3):1586-1596 (1994)
O.CM.CM4974	AF009033	ENV	Korber, BT	Unpublished
O.CM.HIV1CA9EN	X96522	ENV	Janssens, W	AIDS 13:41-48 (1999)
O.CM.MVP5180	L20571	ENV GAG NEF POL REV TAT VIF VPR VPU	Gurtler, LG	J Virol 68:1581-1585 (1994)
O.FR.HIVY16019	Y16019	VIF VPR VPU	Bibollet-Ruche, F	ARHR 14(11):951-961 (1998)
O.FR.HIVY16020	Y16020	VIF VPR VPU	Bibollet-Ruche, F	ARHR 14(11):951-961 (1998)
O.FR.HIVY16021	Y16021	VIF VPR VPU	Bibollet-Ruche, F	ARHR 14(11):951-961 (1998)
O.FR.HIVY16022	Y16022	VIF VPR VPU	Bibollet-Ruche, F	ARHR 14(11):951-961 (1998)
O.FR.HIVY16023	Y16023	VIF VPR VPU	Bibollet-Ruche, F	ARHR 14(11):951-961 (1998)
O.FR.HIVY16024	Y16024	VIF VPR VPU	Bibollet-Ruche, F	ARHR 14(11):951-961 (1998)
O.FR.HIVY16025	Y16025	VPR	Bibollet-Ruche, F	ARHR 14(11):951-961 (1998)
O.FR.HIVY16026	Y16026	VPR VPU	Bibollet-Ruche, F	ARHR 14(11):951-961 (1998)

Table 1: cont.

Name	Accession	Region	Author	Reference
O.FR.HIVY16027	Y16027	VPR	Bibollet-Ruche, F	ARHR 14(11):951-961 (1998)
O.FR.HIVY16028	Y16028	VPR	Bibollet-Ruche, F	ARHR 14(11):951-961 (1998)
O.FR.HIVY16029	Y16029	VPR VPU	Bibollet-Ruche, F	ARHR 14(11):951-961 (1998)
O.FR.HIVY16030	Y16030	VPR	Bibollet-Ruche, F	ARHR 14(11):951-961 (1998)
O.FR.HIVY16031	Y16031	VPR VPU	Bibollet-Ruche, F	ARHR 14(11):951-961 (1998)
O.GA.VI686	X96526	ENV	Delaporte, E	AIDS 10(8):903-910 (1996)
O.GQ.193HA	U82990	ENV	Hunt, JC	ARHR 13(12):995-1005 (1997)
O.SN.MP1299	AJ302646	GAG NEF POL REV TAT VIF VPR VPU	Peeters, M	Unpublished (2000)
O.SN.MP1300	AJ302647	ENV GAG NEF POL REV TAT VIF VPR VPU	Peeters, M	Unpublished (2000)
CPZ.CD.CPZANT	U42720	ENV GAG NEF POL REV TAT VIF VPR VPU	Vanden Haesevelde, MM	Virology 221(2):346-350 (1996)
CPZ.CM.CAM3	AF115393	ENV GAG NEF POL REV TAT VIF VPR VPU	Corbet, S	J Virol 74:529-534 (2000)
CPZ.CM.CAM5	AJ271369	ENV GAG NEF POL REV TAT VIF VPR VPU	Souquiere, S	Unpublished
CPZ.GA.CPZGAB	X52154	ENV GAG NEF POL REV TAT VIF VPR VPU	Huet, T	Nature 345(6273):356-359 (1990)
CPZ.US.CPZUS	AF103818	ENV GAG NEF POL REV TAT VIF VPR VPU	Gao, F	Nature 397(6718):436-441 (1999)
01_AE.CF.90CF11697	AF197340	ENV GAG NEF POL REV TAT VIF VPR VPU	Anderson, JP	J Virol 74(22):10752-10765 (2000)
01_AE.CF.90CF402	U51188	ENV GAG NEF POL REV TAT VIF VPR VPU	Gao, F	J Virol 70(10):7013-7029 (1996)
01_AE.CF.90CF4071	AF197341	ENV GAG NEF POL REV TAT VIF VPR VPU	Anderson, JP	J Virol 74(22):10752-10765 (2000)
01_AE.CM.CA10	AJ277818	ENV	Beirnaert, E	Virology 281(2):305-314 (2001)
01_AE.DE.K08DE	AF215859	VIF	Kuhn, J	Unpublished
01_AE.FR.HIV232982	AJ232982	NEF	Jubier-Maurin, V	ARHR 15(1):23-32 (1999)
01_AE.TH.93TH057	AF197338	GAG NEF POL REV TAT VIF VPU	Anderson, JP	J Virol 74(22):10752-10765 (2000)
01_AE.TH.93TH065	AF197339	ENV GAG POL REV TAT	Anderson, JP	J Virol 74(22):10752-10765 (2000)
01_AE.TH.93TH253	U51189	POL REV TAT	Gao, F	J Virol 70(10):7013-7029 (1996)
01_AE.TH.93TH902	AF170549	GAG POL VPR	Chang, SY	ARHR 15(17):1591-1596 (1999)
01_AE.TH.94TH702	AF170545	POL VPR	Chang, SY	ARHR 15(17):1591-1596 (1999)

Table 1: cont.

Name	Accession	Region	Author	Reference
01_AE.TH.94TH7091	AF170546	GAG	Chang, SY	ARHR 15(17):1591-1596 (1999)
01_AE.TH.95TNIH047	AB032741	GAG POL REV TAT VIF VPR VPU	Auwanit, W	Unpublished (1999)
01_AE.TH.98-4	U48934	NEF	Artenstein, AW	ARHR 12:557-560 (1996)
01_AE.TH.CM240	U54771	ENV GAG NEF POL REV TAT VIF VPR VPU	Carr, JK	J Virol 70(9):5935-5943 (1996)
01_AE.TH.KH03	U48264	ENV	McCutchan, FE	J Virol 70(6):3331-3338 (1996)
01_AE.TH.TH022	AB032740	ENV GAG NEF POL REV TAT VIF VPR VPU	Auwanit, W	Unpublished (1999)
01_AE.TW.TWE13	AF220479	VPU	Lee, CN	J Clin Microbiol 38(7):2468-2474 (2000)
01_AE.TW.TWE6	AF220478	VPU	Lee, CN	J Clin Microbiol 38(7):2468-2474 (2000)
02_AG.CM.MP807	AJ286133	NEF POL REV TAT VIF VPR VPU	Montavon, C	J AIDS 23(5):363-374 (2000)
02_AG.DJ.DJ258A	L22939	VPU	Louwagie, J	J Virol 69(1):263-271 (1995)
02_AG.FR.DJ263	AF063223	ENV GAG NEF POL REV TAT VIF VPR VPU	Carr, JK	Virology 247(1):22-31 (1998)
02_AG.FR.DJ264	AF063224	ENV GAG NEF POL REV TAT VIF VPR VPU	Carr, JK	Virology 247(1):22-31 (1998)
02_AG.GH.G829	AF184155	GAG NEF POL REV TAT VIF VPR VPU	Candotti, D	J Med Virol 62(1):1-8 (2000)
02_AG.NG.IBNG	L39106	ENV GAG NEF POL REV TAT VIF VPR VPU	Howard, TM	ARHR 10(12):1755-1757 (1994)
02_AG.NG.NG1921	AF069941	ENV	McCutchan, FE	Virology 254(2):226-234 (1999)
02_AG.SE.SE7812	AF107770	ENV GAG NEF POL REV TAT VIF VPR VPU	Laukkanen, T	Unpublished
02_AG.SN.MP1211	AJ251056	ENV GAG NEF POL REV TAT VIF VPR VPU	Toure-Kane, C	ARHR 16(6):603-609 (2000)
02_AG.SN.MP1213	AJ251057	NEF POL REV TAT VIF VPR	Toure-Kane, C	ARHR 16(6):603-609 (2000)
03_AB.RU.KAL153-2	AF193276	ENV GAG NEF POL REV TAT VIF VPR VPU	Liitsola, K	AIDS 12(14):1907-1919 (1998)
03_AB.RU.KAL68-1	AF082485	ENV	Liitsola, K	AIDS 12(14):1907-1919 (1998)
03_AB.RU.RU98001	AF193277	ENV GAG NEF POL REV TAT VIF VPR VPU	Liitsola, K	ARHR 16(11):1047-1053 (2000)
04_cpx.CY.94CY032-3	AF049337	ENV GAG NEF POL REV	Gao, F	J Virol 72(12):10234-10241 (1998)

Table 1: cont.

Name	Accession	Region	Author	Reference
		TAT VIF VPR VPU		
04_cpx.GR.97PVCH	AF119820	ENV GAG NEF POL REV TAT VIF VPR VPU	Nasioulas, G	ARHR 15(8):745-758 (1999)
04_cpx.GR.97PVMY	AF119819	ENV GAG NEF POL REV TAT VIF VPR VPU	Nasioulas, G	ARHR 15(8):745-758 (1999)
05_DF.BE.VI1310	AF193253	ENV GAG NEF POL REV TAT VIF VPR VPU	Laukkanen, T	Virology 269(1):95-104 (2000)
05_DF.BE.VI961	AF076998	ENV GAG NEF POL REV TAT VIF VPR VPU	Carr, JK	Virology 269(1):95-104 (2000)
06_cpx.AU.BFP90	AF064699	ENV GAG NEF POL REV TAT VIF VPR VPU	Oelrichs, RB	ARHR 14(16):1495-1500 (1998)
06_cpx.ML.95ML127	AJ288982	ENV GAG NEF POL REV TAT VIF VPR VPU	Montavon, C	ARHR 15(18):1707-1712 (1999)
06_cpx.ML.95ML84	AJ245481	ENV GAG NEF POL REV TAT VIF VPR VPU	Montavon, C	ARHR 15(18):1707-1712 (1999)
06_cpx.NG.NG3670a	AF069934	ENV	McCutchan, FE	Virology 254(2):226-234 (1999)
06_cpx.SN.97SE1078	AJ288981	ENV GAG NEF POL REV TAT VIF VPR VPU	Montavon, C	ARHR 15(18):1707-1712 (1999)
10_CD.BFL061	AF289548	ENV GAG NEF POL REV TAT VIF VPR VPU	Koulinska, IN	ARHR 20(5):423-431(2001)
10_CD.BFL071	AF289549	ENV GAG NEF POL REV TAT VIF VPR VPU	Koulinska, IN	ARHR 20(5):423-431(2001)
10_CD.BFL110	AF289550	ENV GAG NEF POL REV TAT VIF VPR VPU	Koulinska, IN	ARHR 20(5):423-431(2001)
11_cpx.CM.CA1	AJ277823	ENV	McCutchan, FE	Virology 254(2):226-234 (1999)
11_cpx.CM.MP818	AJ291718	ENV GAG NEF POL REV TAT VIF VPR VPU	Peeters, M	Unpublished (2000)
11_cpx.FR.MP1298	AJ291719	ENV GAG NEF POL REV TAT VIF VPR VPU	Peeters, M	Unpublished (2000)
11_cpx.FR.MP1307	AJ291720	ENV GAG NEF POL REV TAT VIF VPR VPU	Peeters, M	Unpublished (2000)
11_cpx.GR.GR17	AF179368	ENV GAG NEF POL REV TAT VIF VPR VPU	Paraskevis, D	ARHR 16(9):845-855 (2000)
11_cpx.NG.NG3670b	AF069945	ENV	McCutchan, FE	Virology 254(2):226-234 (1999)

HIV-1/SIVcpz protein alignment: GAG

HIV-1/SIVcpz proteins

```
                    |-     membrane binding     -|                                                                                  phos site ^
                    / p17     |-         nls         -|                                                                                |-     nls     -|
B.FR.HXB2           MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEEEQNKSKKKA......QQAAADTGH    124
CONSENSUS A         mga---v---gkl-a---i------k-k-rl--lv--s---erfaL--s----te--Q--meq--pAlk--T--lr--y----t-y----r-dvk------d-i--iqn-skQkt......q--a-d--n    124
A.KE.Q23            -----------KF-A--------------RM--LI------D---L--S----AD--Q--MEK---A-K--T--IK--F-------------DV------------IK----Q-T................N    124
A.SE.SE6594         ..X--------K--A--------------R---L-----------L--S----T---Q--IE----A-N--T--IK--F-------------DV------------M-K---Q-T......S--T----S    122
A.SE.SE7253         ..X--------K--A--------------RM--L-----------L--S----T---Q--ME---SA-K--T------F----V----------------------I-K---Q--................N    122
A.SE.SE7535         -----------K--A------------Q-R---L-----------L--S----A---Q--MK----A-K--T--------------------NVT-----------I----EQ-T................N    124
A.SE.SE8538         ..X--------R--A--------------RM--L-------D---L--S----T---Q--I-----AFK--T---K---------W------DV---------L--I-K---Q-T.........-----V----S    122
A.SE.SE8891         ..X---K---EKK-A---M--------------M---R----KS-L--S----T---Q--MDK---A-R--T------------------K-AV------------I---N-Q-T................N    122
A.SE.UGSE8131       -----------K--A----------N---R---L--------K--L-------A---Q--IE---SAPK--T--------------------DV------------I-----Q-K......E-------N    124
A.UG.92UG037        -----------K--A--------------R---L-----------L--S----T---Q--ME---SA-R--T--------------------V------------I-K---Q-T................S    124
A.UG.U455           ----------KK--S----------N---R---L--------K-TL-------A---Q--------A----T-----------V--------DV-------N----M---N-QRT.............N--S    124
A2.CD.97CDKTB48     -----------K-EA--------------R---L--------K-SI--S----ET---R-F-----A-E--T---------I-V--F---K--V----------------C-Q-T................S    124
A2.CY.94CY017-41    ------I----K--A--------------R---L--------K-SI-------P------IR----A----T---K------VV--W----VDV----------------..Q-T.......-H------N    122
CONSENSUS B         ------v--g-e--r--ki----g---k-k---i----------av-pg---tse--r--lg--qps-q-----lr-ly-t--t---v-qr-evk---e--d-i----nksk--a?????.qqaaadtgn    124
B.AU.MBC18          ------I----K--------------------------------------------------------H----------K-V--A--V------N-D--------E--------C----QQAAA.-----G--N    129
B.AU.SC49           -----------K--K------------T-Q-----------------------GG--K---V-------------K----A---------G--VR--------------------------------N    124
B.AU.VH             --------------------------T--------------------------------V------P------K--F----------K---R------E-------------QQAAA.DA----A-N    129
B.CN.RL42           -----------Q----------------R---L---------------------------E----A----------F----------L-----------E-------------.......--T--G--N    124
B.DE.D31            ----------------------------R-----------------------------------------------F-------------V----------------------.......-P--R-A-N    124
B.DE.HAN            --------------K--------------Q-----------------------------M----------------------------K--V---------V-----N-----........-E---A-N    124
B.ES.89SP061        ------I----G----------------R---L-----------------------------A-----------------------K-DV-----------------------................N    124
B.GA.OYI            --------------K--------------Q---------------I----------------------K-----I---------------K--V----------------------........--T-----N    124
B.GB.CAM1           --------------K-----------------------------------------------------------------------K-DV-------E-------------.......------G--N    124
B.GB.GB8            ----------------------------R---V---------------------------E------------R---F-------------DV-------------S------..............-K-N    124
B.GB.MANC           -----------K-----------------------------------T--L-------------V-------------K-------V------G--V----------------------.......-P------N    124
B.JP.JH31           ------------------------------------------S-------------------------------K--F-------------V-------E-------------................N    124
B.KR.WK             ------I-------Q--------------R---L-----------------------------------------K--F-A--V-----------------E------S------.........---T----S    124
B.NL.3202A21        --------------K-------------R-----------------------A-----------A-----------F----V----------V-----------------------................N    124
B.TW.TWCYS          ------I--A----K--RV----------R---L-----------------------------S-----------------------I--K-------D--E-------------...........----TG--N    124
B.US.85WCIPR54      ------------------------------------------------------------A-----------H----V---------V-------E-------------................N    124
B.US.AD8            -----------K-----------------------------------------------------------------K--F-----------N-DVR----------------------.............-----AEK    124
B.US.DH123          -----------K--S-------------------------------------------------------------------------------E---V---------V----------................N    124
B.US.JRCSF          ----------------------------R---------------------S--------------K------T-----------------------E------T--M---................N    124
B.US.JRFL           -----------K--K--------------R---------------------S-----------L-A-K-----------------------V-------E------A-----E.................N    122
B.US.MN             -----------------N---------------V-----------------------------------------K--------------K----------E-------------...............N    124
B.US.NY5CG          --------------K------------Q-R-----------------------------R------------R---F----V--------DV----------------------................N    124
B.US.RF             -----------K--K--------R---R--------------------S----A------------A--------K----A---------N--VR----------------------................N    124
B.US.SF2            --------------K-----------------------------------------------------------------------DV-------E-------------.............-----AA-T    124
CONSENSUS C         ------IlR-gK--k--k-rlr-----h-ml--l---s---er--ln-g--etse--k--ik-l--A-q--Te-lr--y---at-y---ag-evr-----l-k-e----ksqqkt......q--keadg?    123
C.BR.92BR025        ------I-R--K--A--R-K-K-----H-MM--L-----------LD----------K--MK----A----TK--I--H----------EK-DVR---------K------QQ-T.........-E-ADKG    124
C.BW.96BW01B22      ------I-R--K--Q------------C-M---L-----------L-----------K--MK----A----T---K-------------EN--VR----------------QQ-T.........-KT-D-.    123
C.BW.96BW0402       ------I-R--K--A------------Q-RI--L-----------L-----------K--IQ----A----T------F----------KK--VR----------------CQQ-I.........-E-ADKG    124
C.BW.96BW0502       ------I-R-EK--K------------H-M---L--------G--L-------A---K--MK----A----T-----------------AG--VR----------------QQ-T.........-KEAD-.    123
C.BW.96BW1104       ------I-R--K--T------------R-MI--L-----------L-----G-----K--MK----A----T--I--------------EG--VQ-----V----------QQ-T.........-E-ADK.    123
C.BW.96BW1210       ------I-R-EK--T------------R-MM--L--------N--L--A----A---K--MK----A-K--T---K-------------EK--VR----------------CQQ-E.........-E-AAKG    124
C.BW.96BW15B03      ------ISR-EK--T-----S------C-M----------------L-----------K--MK----A----T------F----------EK-KV----------------QE-T.........-E-AA-.    123
C.BW.96BW1626       ------I-R--K--K------------R-M---L-----------L--------------MI-I--A-K--T---K-------V-F---AG--VQ----------------QQ-T.........-KEA--.    123
C.BW.96BW17A09      ------I-R--K--T------------H-M---L-----------L-------A---K--IK----A----T---K-------------G--VR----------------FQQ..........-K-AD-.    120
C.ET.ETH2220        ------I-R-EK--A----K-------H-M---L---N----K--L--D--D--A--K--IK----A----T---K--F-----------K-----------------E-QQ-T.........-G-ADRG    124
C.IN.93IN101        ------I-R--K--K--R---------H-M---L-----------L-------A---K--IK----A----T---K--H----------AG--VR----------------IQQ-T.........-KEAD-.    123
C.IN.93IN904        ------I-R-EK--K------------H-M---L-------D---L-----------K--IK----A----T------H---V------AG--VR----------------NQQ-T.........-KEADE.    123
C.IN.93IN905        ------I-R--K--K------------H-M---L-----------L-----------K--IK----A----T------F----------AE--VR----------------IQQ-T.........-KEAD-.    123
C.IN.93IN999        ------I-R-EK--K--R---------H-M---L-------D---L------AA---K--IK----A----T------H----------EG--VR----------------QQ-T.........-KEAD-.    123
C.IN.94IN11246      ------I-R--K--K------------H-M---L--------K--L-----------K--MK----A----T------F----------AE--VR----------------QQ-T.........-KEAD..    122
C.IN.95IN21068      ------I-R--K--K------------R-M---L-------D-----------A---K--IK----A----T------F----------AG--VR----------------I-Q-T.........-KE-D-.    123
C.IN.AF209990       ------I-R--K--K------------H-M---L-------------------A---K--IK----A----T-----------------AE--VR-------R--------QQ-T.........-NEAD-.    123
D.CD.84ZR085        -----------K--A-----------------------------L-----------K--ME----A-K--------I-----------K-----------------N---R-..........-G--N    124
D.CD.ELI            -----------K--K--------------R-------------Y-L-----------K--I-----AI---T-----------------KG-DV-------E-M----------................N    124
D.CD.NDK            -----------K--T--R-----------A---LI---------TL-----------K--I------I------I--------------E---V------VE-M----------T.............-----SS.    123
D.CD.Z2Z6           -----------K--A--------------R---L-----------L--------D--K--I-----AIR---------F----------E---V-------E-M--------N-KA.............-A-N    125
D.UG.94UG1141       -----------K--E--------------R---L-----------L--------------IR-----I------IK------V------E--KVAS------------A------.........---T---RN    124
CONSENSUS F1        ------v----K--a--k-r-----k-k-rm--Lv----------?d------s---?kiIg-l-p-l-------r--y--vaV--?---kvevk------e-L------sqQ-t......q-aa-dkg?    120
F1.BE.VI850         ------I----K--E----Q-------R--M--LI----------LD----------QK-IR-------------K--F----V--Y----AGVT--------L-------QQ-T.............-K-.    123
F1.BR.93BR020-1     -----------K--A--------------R---L-----------LD----------K-I--------------K-----I-V--Y---KV-V-------E-L------GRQ-T.........-T-EK-.    123
F1.BR.BZ162         -----------K--A--------------R---L-----------I-----------QKXI----------------V--F----V-V---------L-------QQ-T......K-----K-.    123
F1.CD.VI174         -----------K--A----Q---------RM--L-----------ID------P----K-I----T------------F--IVV--Y----V---------E-L-------QQ-T.........-.---K-.    122
F1.FI.FIN9363       -----------K--A----------Q-RI--L-----------ID----------QK-IA-I---I-------------I-V--F---K--V---------L-------QQ-T.............AADK    124
F1.FR.MP411         -----------K--A--R-----------RM--L-----------D------P---K--IR----------------F----V------K----------E-L------GQQ-T.............-K-.    123
F1.RW.VI69          -----------K--A----------R----M--LI----------LD----------K-I----------------------V--F---KV-V-------E-L-------QQ-K.......--V---K-.    123
F2.CM.MP255         -----------K--A------K-----R-R---L-----------L--S----T---KK-I----S----------K----A-VV--Y-----DVR--------LQ---D--QQ-E.........-K---KE.    123
F2.CM.MP257         -----------K--A--------------R-----------K---L-------T---KK-I--------------K--F--IVV--Y---K--VR--------LQ---D-HQQ-T.........-T--K-V    124
```

```
                    |-     membrane binding     -|                                                                  phos site ^
                    / p17     |-         nls         -|                                                                  |-     nls     -|
B.FR.HXB2           MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEEEQNKSKKKA......QQAAADTGH    124
G.BE.DRCBL          -----------K--A------------R-RM--L-------D---L-------A---QK-MA----A----T--IK--F-----------K--VR------EEV-KI-K--QQ-E.............NS    117
G.FI.HH8793         -----------K--A--------------R---L------M----L-------A---H--MS----AI---T--IK--F-----------K--V---T---EEV-KI-K--QQ-I...........----R-E-N    124
G.NG.92NG083        V----------K--S----------R---------------G---L-RD----A---V--MK----A-$--T------F-----------K--V------PEEV-KI-KN-QQEI..........----KNE-N    123
G.SE.SE6165         --------T--K--A----------R-S--I--L-----------L--D----A---Q--MR---------T--IK----A------------V-------EEV-KI-K--QE-I...........----M-K-N    124
H.BE.VI991          -----------K--A----------R---R---L-----------L--D----AD--Q--------A-K--T-D-Q-----I-V--------DV-------G----I---N-QRT..........-P-AADK    124
H.BE.VI997          -----------R--TL-------------R--------------L------SA---L--IE--R--IK--T---X--F---------L----V-------G----I---RQQ-T........---T-NKER    124
H.CF.90CF056        -----------K--A--------------R---L-----------L-------P---L--IE-I--AIK--T---K--F-L--V-----RK-DV------------I----QQ-T..............-KEK    124
J.SE.SE7022         ------I----K--D------------Q-RI--L-------D---L------SAK--Q---V----A----T--IK-----------------------------I---N-QQT........-K-ETDKK    123
J.SE.SE7887         ------I----K--D--------------RI--L-------D---L------SAK--Q---V----A----TQ-IK--------------------M---E----I---N-QQ........-K-ETDKK    123
K.BE.VI325          -----------K--T----------S---R---L-----------L--S----T------IR-------------K--F---V-----------R--------L---------Q--........-.ETANKG    123
K.CD.EQTB11C        -----------K--K----Q---------R---L-----------L--N----V------IR---------------F-------W---S-QVR--------L-----RTQQ-T........--GK--K-.    123
K.CM.MP535          -----------K--A------------------L-----------L-------T------IT-I---I------IK-----I-V--F---K--V---------L-------QR-T........-E---K-.    123
N.CM.YBF106         --------T--K--Q--A-Y---------R---L-----------C----MD-AN--A-LIN--E-A-K----G----X--L-V-----SN-PVHN-Q------K-K-EQH-SEP..KKPEAGT--AADS    128
N.CM.YBF30          --------T--K--Q--S-Y---------RM--L-----------C----MD-AD--AKL-N--E-A-K-----------AL-V-----S--Q-HN-Q------K-K-EQH-PEP..KNPEAG---A-DS    128
O.CM.ANT70          ---S----T-SK--A--Q---K--S----R---L-----------C--E----A--NEKL-Q--E-A-K---DS-Q--W-AIVV-W---N-YK-G--QQ-IQ-LK-VMGSR-SAD......AAKEDTSAR    124
O.CM.MVP5180        --------T-SK--A--R------S--A-R---L---------Y-C-------A--TE-L-Q--E-A-K----D-K--W-AI-V-W---N-FD-R--QQ-IQ-LK-VMASR-SAE......AAKEETSPR    124
O.SN.MP1299         --------T-SK--A--Q---K--C----R---L-------D-S-C--E----A--NEKL-Q--E-A-K----S-Q--W--I-V-W---N--KVE--QQ-IQ-LK-VMGSR-SAG......TAKEDTSAR    124
O.SN.MP1300         ----------SK--T--Q---K--C----R---L-----------C--E----A--NEEL-Q--E-A-K----S-Q--W--I-V-W---K-FKVE--QQ-IQ-LK-VMGSR-SAG......AAKEDTSAR    124
CONSENSUS 01        -g-r--v--ggK-dAw---------k-k-rm--l--------r--l--gl---a---Q-lie---StlK--s--lk--f--?a--W----r-eVk--ke-------v-nksqq-t............-g--S    123
01_AE.CF.90CF11697  -----------K--A--------------RM--L-----------L-------A---Q-LIE---SN-K-----IK--F--I---W------DV------------V---N-Q-T............-G--S    124
01_AE.CF.90CF402    ------I----K--A------------Q-RM--L-----------L-------A---Q-LIE---STIK------K--F--I---W------DV------------V-----Q--............-G--S    124
01_AE.CF.90CF4071   -----------K--A--------------R---L-----------L-------A---Q-L-E---ST-K------K--F--I---W------NVT-----------I---N-Q-T..................S    124
01_AE.TH.93TH057    -S--------EK-NA----------R-------L--------K--L-------A---Q-LIE---ST-K------K--F--I---W----K--V----K-------V-K--QQ-T............-G--S    124
01_AE.TH.93TH065    ---------X-K--AL--------------M--L-----------L-------A---Q--IE---SA-K------K------V--W------DV------------V----QQ-T............-G--S    124
01_AE.TH.93TH902    -----------K--A---------------M--L-----------I-----------Q--IE---ST-K------K--F------W-------V---R--------A-KR-QQ-T............-G--S    124
01_AE.TH.94TH7091   -----------K--A---------------M--L-----------I-----------Q--IE---ST-K------K--F------W-------V---R--------A-KR-QQ-T............-G--S    124
01_AE.TH.95TNIH047  ---K-------K--A----------R---H--------------L-------A---Q-LIE---ST-K---------F--IV--W-------V------------V----QQ-T............-G--S    124
01_AE.TH.CM240      -----------K--A----------R---R---L-----------L--SF---A---Q--IE---ST-K--L---K--F------W-------V------------V----QR-T............-G--S    124
01_AE.TH.TH022      ------I----K--A----------R---RM--L-----------L------A---Q-LI----SN-K---------F------W-------V------------A-K--QQ-T............-G--S    124
CONSENSUS 02        m---a---s--k--a--k-----g-----rl-hL-----------l-------a---Q-lme---sa-?t-----K--?--i---w---qr-dik------d-?--?-n---Q-t..........-aa-atgs    120
02_AG.FR.DJ263      -----------K--S--------------R---L-----------L-------A---Q-LIE---TA-G------K-----I---W---K--D--------E----V-----Q-T............-A--S    124
02_AG.FR.DJ264      -----------K--S--------A-----R---L-----------L-------A---Q-LMK---SA-G------K-----I---W------D------------V-----Q-T.........-E-A--S    124
02_AG.GH.G829       ----V---R--Q--T---------------M-LL-------------------T---Q---E----A-KA-----K--------------K-DVR--------L--I-----Q-T............-NS..    122
02_AG.NG.IBNG       -----------K--A--------------R---L-----------L-------A---Q-LME---SA-R------K--F--I---W------D----------M--I-----Q-V........-T--A--S    124
02_AG.SE.SE7812     -----------K--A--------------R---L-----------L-------A---Q-LME---ST-R------K--F--I---W------D----------V--V-K---Q-T..............-AA-S    124
02_AG.SN.MP1211     X----------K--A--R-----------R---L-----------L-------A---Q--IE---ST-K------K--F--------------V---------L--I-E---Q-T............-A--.    123
03_AB.RU.KAL153-2   -----------K--A-----------E--RI--L-----------L--S--------Q---E----T-K------K-----------------------------I-----Q-T..........-TG--S    124
03_AB.RU.RU98001    -----------K--A--------------RI--L-----------I--S--------Q---E----T-K------K-------------R---------------I-----Q-T..........-TG--S    124
04_cpx.CY.94CY032-3 -----------K--A--R-----------R---L-----------L-------A---Q-LME---ST-K------------IT--W------DVQ-----------I-S---Q-T............-AA-G    124
04_cpx.GR.97PVCH    -----------K--A--R-----------R---L-----------L-------A---Q-LME----T-R-----VK--F--L---W------DV---------V--V-K---Q-T..................G    124
04_cpx.GR.97PVMY    -----------R--A------------R-RI--LI----------L-------A---Q-LME---ST-K------K--F-LI---W-------V---------V--M-----Q-T.............--.N    122
05_DF.BE.VI1310     -----------K--A--------------R---L-----------L-----------K--IS----AI---T---K------------E--V-------E----------R-K......A-Q-EAGAG    124
05_DF.BE.VI961      -----------K--A--------------R-------------I-----------K--IA----AI--------------------EK--V-------E----------S-R.....A---E-G-KN    125
06_cpx.AU.BFP90     -----------K--E---------------M--L-----------I-------A---Q--IE---ST-K------K--F-------------KVT-----I-----I-K---Q--......H----.A--N    123
06_cpx.ML.95ML127   -----------K--E--------------R---L-----------L--S----A---Q--ME---SA-R------K---------------KVT-----------I-----Q--............A--N    124
06_cpx.ML.95ML84    -----------K--E--------------RM--L-----------L-------P---Q--ME---ST-R------K---------------KVT------H----I-----Q-V............A--N    124
06_cpx.SN.97SE1078  -----------K--E--------------R---L-------D---L-------A---Q--IE---SA-K---------F-------------KVQ-----V--V--I-----Q-T........--TT-A--N    124
10_CD.BFL061        -----------K--E------------Q-----L-------D---I-----------K--I-----AI------IK-------------E--KVA----------T------.............T----N    124
10_CD.BFL071        ----X------K--E--------------R---L-----------L-----------K--I-----AI------IK-------------E--QV----------G---T------..........--T---RN    124
10_CD.BFL110        -----------K--E--------------R---L-----------L-----------K--I------I------IK-------------E--KVT-------------T------...............N    124
11_cpx.CM.MP818     -----------K--S--------------R---L-----------L--S----G---Q-LM-----A-G--T---------L-------H----------------I-----Q-K...............N    124
11_cpx.FR.MP1298    ------I----K--A------------Q-R---L-----------L--S----A---Q--------A-G--T----------------H--DV------------I-----Q-K..........-G---N    124
11_cpx.FR.MP1307    -----------K--A--------------R---L-----------L--S----G---Q--ME---SA-R--T----------V---------V--IK----Q-K.........-E----N    124
11_cpx.GR.GR17      -----------K--A--------------R---L-----------L--S----G---K--M-----A-G--T----------A-----L---V---------V--I------...............-S-N    122
CPZ.CM.CAM3         --------T--R--A--------------MM--L-------D---C----M--A---E-L-R--E-A-K----G----F--L-V-W---K--SVE--QQ--T-LK-AVTA-REQE.....VA-PQQQQQD    125
CPZ.CD.CPZANT       ---G----R-EK--T--S-----------MI--L---RS--Q---LSSS--------EKAIH--S--IEIR-P-II--F--ICV-W---KGEK----EQ-VKTVKMKVMQTQAETGSSQT.ASRGMLLRL    129
CPZ.CM.CAM5         --------T--R--A------X-------MM--L-----------C----M--A---ENL-K--E-A-R----G----F--L-V-W---K--QVE--QQ--S-LK-TVAAQ.EQE.....VA-PQQQQQE    124
CPZ.GA.CPZGAB       --------T--K------V------R-R-MM--L-----------CD---M-SK---TKL-Q--E-A-K----G----F--L-V-W-I-SD-TVE--QK--EQLKRHHGEQQS-TESNSGSREGG-SQ-A    130
CPZ.US.CPZUS        --------T--R--A--------------MM--L-----------C----M--AD--L-L-K--E-A-K----G----F--L-V-W---S-VTVE--QQ--V-LK-VVQLQ-QQE......EKEQQQQEA    124
```

HIV-1/SIVcpz proteins

```
                                                                                                                cyclophilin  A binding domain (CyPA)
                                  p17 \/ p24                                                                    |-                                 -|
B.FR.HXB2           SN...........QVSQNYPIVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQ.AAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTST    242
CONSENSUS A         ss...........kv-q---iv--A---mv--sl-----------i-ekafs-----m----------q---m---i-----.---q---D-----------l--v----ip------------------    242
A.KE.Q23            -S...........N----------A----I--SL-----------I--------------------------M---I-----.-------D-----------L--------P------------------    242
A.SE.SE6594         -S...........K----------A----I--SL-----------I---G-N--------------------M---I-----.-------D-----------L--------P------------------    240
A.SE.SE7253         -S...........K----------A-------SL-----------I-----------V--------------M---I-----.-------D-----------L--A----V-------------------    240
A.SE.SE7535         -............K----------A-------SL-----------I--------------------------M---I-----.-------D-----------L--------P------------------    241
A.SE.SE8538         -S...........K--H---V---A----I--NL-----------I---G-N------------------------I-----.-------D-----------L--------P------------------    240
A.SE.SE8891         -S...........N--R-------A-------SL-----------I---G------------------G---M---I-----.-------D-----------L--------P------------------    240
A.SE.UGSE8131       -S...........K-------M--A-------VM-----------I--R-------------------H---M---I-----.---E---D-----------T--I----V-------------------    242
A.UG.92UG037        -S...........K----------A----I--SL-----------I----L---------------------M---I-----.-------D-----------L-------V-------------------    242
A.UG.U455           ..............----------A---P----L-------------D------------------------M---V-----.-------D-----------L--------P------------------    238
A2.CD.97CDKTB48     -SSQNYRG.....SS---------A--------V-------------------------T----------------------.-------D-----------L---Q----P--------------A--N    248
A2.CY.94CY017-41    .............SS---------A----------------------------------T----------------------.-------D--------------------P------------------    238
CONSENSUS B         ss???........qv-q-------lq------ai-------w---v----------------s---t---------------?---------------e---l--vh---ia-----e------------    242
B.AU.MBC18          --...........P----------M--------------------I------------------------------------.-------------------L--AQ---V------D------------    247
B.AU.SC49           -S...........-----------L--------------------I------------------------------------.-------------------L--A----N-------------------    242
B.AU.VH             -S...........K----------L-------P-------------------------------------------------A---------------D---L---------------------------    248
B.CN.RL42           NS...........-A---------L-------P-------------------------------------------------.-------------------L-------V-------------------    242
B.DE.D31            NS...........-----------L-------P-------------------------------------------------.-------------------L---------------------------    242
B.DE.HAN            R-...........-----------L---------------------------------------------------------.-------------------L---------------------------    242
B.ES.89SP061        -S...........---------------------------------------------------------------------.-------------------L-------HP------------------    242
B.GA.OYI            -S...........-----------L-------P-----------------------------A-------------------.-------------------L---------------------------    242
B.GB.CAM1           -S...........-----------L---------------------------------------------------------.-------------------L---------------------------    242
B.GB.GB8            --SN.........-----------L-------------------------------------A-------------------.---------------------------V-------------------    244
B.GB.MANC           -S...........---H-------L-------S---------------------------------I---------------.-------------------L-------V-------------------    242
B.JP.JH31           -S...........K--------------------------------------------------------------------.-------------------L--AQ-----------------------    242
B.KR.WK             -S...........-----------L-------P--------R---I------------------------------------.---------------D---L---------------------------    242
B.NL.3202A21        -S...........-----------L--------L------------------------------------------------.-------------------L---------------------------    242
B.TW.TWCYS          NS...........P----------L---------------------------------------------------------.-------------------L---------------------------    242
B.US.85WCIPR54      -S...........K--------------------------------------------------------------------.-------------------L---------------------------    242
B.US.AD8            -S...........-----------L---------------------------------------------------------.-------------------L---------------------------    242
B.US.DH123          -S...........--------------------L------------------------------------------------.-------------------L---------------------------    242
B.US.JRCSF          -S...........-----------L--------------------I------------------------------------.-------------------L---------------------------    242
B.US.JRFL           -S...........-----------M---------------------------------------------------------.-------------------L---------------------------    240
B.US.MN             RGNSS........------------E--------------------------------------------------------.-------------------L--------T------------------    245
B.US.NY5CG          -S...........-----------L---------------------------------------------------------.-------------------L---------------------------    242
B.US.RF             GS...........-----------L---------------------------------------------------------.-------------------L---------------------------    242
B.US.SF2            G-SS.........-----------L---------------------------------------------------------.-----------------------------------------------    244
CONSENSUS C         .............kv---------l-------ai-p--l------i--------v----T---------------n------.-------D-------e---l--vh---i----m-e--------t--t    239
C.BR.92BR025        .............K----------L-------P--A-----------------------T----------------------.-------D-----------L-------V-------------------    240
C.BW.96BW01B22      .............KI---------L-------PL-----------I-------------T----------------------.-------D-----------T-------V----L-G-----------N    239
C.BW.96BW0402       .............K----------L--------------------I-------------T----------------------.-------D-----------L---------------------------    240
C.BW.96BW0502       .............K----------L--------------------I-------------T----------------------.-------D-----------L---Q---V------D--------A---    239
C.BW.96BW1104       .............R----------L--------------------I-------------T---------------T------.-------D-----------L-------V-------------------    239
C.BW.96BW1210       .............K----------L--------------------I--------I----T----------------------.-------D-------G---L-------V------------------N    240
C.BW.96BW15B03      .............KI---------L--------------------I-------------T----------------------.-------D-----------L---------------------------    239
C.BW.96BW1626       .............K----------L-------------S------I-------------T----------------------.-------D---------------------------------------    239
C.BW.96BW17A09      .............K----------A----------------------------------T----------------------.-------D-----------L---------------------------    236
C.ET.ETH2220        .............KD---------M-------P--A-----------------------T----------------------.-------D-----------L-------V------D------------    240
C.IN.93IN101        .............K----------L--------------------I-------------T----------------------.-------D-----------L--I-----------------------S    239
C.IN.93IN904        .............K----------L--------L-----------I-------------T----------------------.-------D-----------L--I-----------------------S    239
C.IN.93IN905        .............K----------L--------------------I-------------T----------------------.-------D-----------L--------------------------S    239
C.IN.93IN999        .............K----------L-------P------------I-------------T----------------------.-------D-----------L------------I--------------    239
C.IN.94IN11246      ..............----------L--------------------I-------------T----------------------.-------D-----------I---------------------------    236
C.IN.95IN21068      .............K----------L--------------------I-------------T----------------------.-------D-----------L---P--------L--------------    239
C.IN.AF209990       .............K----------L--------------------I-------------T----------------------.-------D-----------L--------------------------S    239
D.CD.84ZR085        -S...........-----------L--------L-----------I-----N------------------------------.-------------------L---Q---V-------------------    242
D.CD.ELI            NS...........-----------L--------------------I------------------------------------.-------------------L---------------------------    242
D.CD.NDK            .............-----------L--------------------I------------------------------------.-----------D-------L-------V-------------------    239
D.CD.Z2Z6           NS...........-----------L--------------------I------------------------------------.-------------------L---------------------------    243
D.UG.94UG1141       -S...........-----------L------HPL-----------I-----N------------------------------.-------D-----------L-------V----L-------------N    242
CONSENSUS F1        ?.............----------L-------?i-----------I----------------------q-------------.-------D-----------L--vh---ip---m--------------    234
F1.BE.VI850         ..............----------L-------SL-----------I----------------------T-------------.-------D-----------L-------AP------------------    238
F1.BR.93BR020-1     ..............----------L-------SL-----------I------------------------------------.-------D-----------L--TQ----P---I--------------    238
F1.BR.BZ162         ..............----------L-------S------------I------------------------------------.-------D-----------L--AQ----P---I--------------    238
F1.CD.VI174         ..............----------L--------------------I------------------------------------.-------D-----------L---Q----P---I--------------    237
F1.FI.FIN9363       G.............----------L--------------------I------------------------------------.-------D-----------L--------P------------------    240
F1.FR.MP411         ..............----------L-------P------------I------------------------------------.-------D-----------L--A-----L------------------    238
F1.RW.VI69          ..............----------L--------------------I------------------------------------.-------D-----------L-------NP------------------    238
F2.CM.MP255         ..............-------------------L-----------I------------------------------------.-------D-----------L--------P------------------    238
F2.CM.MP257         -K...........G-------L--L-------SL-----------I------------------------------------.-------D-----------L--------P------------------    242
```

```
                                                                                                                                                cyclophilin  A binding domain (CyPA)
                                 p17 \/ p24                                                                                                     |-                                   -|
B.FR.HXB2           SN..........QVSQNYPIVQNIQGQMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQ.AAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTST    242
G.BE.DRCBL          -S..........-----------A--------------------------------T---------------------.-----------D-------L--QQ--------I-D-T------A---    235
G.FI.HH8793         -S..........-----------A-------------------------------------------------------.-------D-----------M--PQ----P---I-------------N    242
G.NG.92NG083        --..........P----------A----I-------------------------------------------------.-------D---D-------I--QQ----P---I---S----------    241
G.SE.SE6165         --..........-----------A---------T-----------------------------------L---------.-------D-----------M--QQ---FP---I-------------S    242
H.BE.VI991          EKDS........KI---------A-------------------------------------------A---------.-------D-----------L--------P------------------    244
H.BE.VI997          D-..........K----------A-------P--X-------------------------------------------A---------.-------D-----------L--------P------------------    242
H.CF.90CF056        D-..........K----------A-------------------------------------------A---------.-------D-------------------P------------------    242
J.SE.SE7022         D-S.........-----------L---P----L-----------I-----------------------------I----.-------D-------------------V----V-------------N    242
J.SE.SE7887         D-S.........-----------L---P----L-----------I-----------------------------I----.-------D-----------------------V--------------    242
K.BE.VI325          ............----------M-------PL-----------I----------------AD------------------.-------D-----------L--------P------------------    238
K.CD.EQTB11C        ............----------L--------L-----------I-------------------------------------.-------D-----------M---Q----P---I--------------    238
K.CM.MP535          ............----------L--------L-----------I-------------T---------------------.-------D---D-------L--------P------------------    238
N.CM.YBF106         -............I-R---L---A-------PLT----------I-----N--I----M-------S---S---------.--------V----------T--APV--LP-----D------------    244
N.CM.YBF30          ............NI-R---L--TA-------PLT----------I-------------M-------S-------------.--------V------D---T---PV--LP---L-D------------    244
O.CM.ANT70          QAG.................------S-A-------------------A------N--I----M------ISY-I-----AI----.G-L-V---V-----V----T--PPV--LP---I---T----------    240
O.CM.MVP5180        QTS.................------T-A-------------------A------N--I----M------V-Y-I-----AI----.G-L-V---V----------T--PAM--LP---I---T----------    240
O.SN.MP1299         QTG.................------T-A-------SL----------A------N--I----M------I-Y-T-----AI----.G-L-V---V----------T--PA---LPV--I---T----------    240
O.SN.MP1300         QTG.................----V-A-A-------SL----------A------N--I----M------I-Y-T-----AI----.G-L-V---V----------T--QA---LP---I---T----------    240
CONSENSUS 01        sS..........k--q-------Aq---v--pl----------v---G-n-------s-----a------M---I--g--.----m-ke-i---a-----v--------P---m-----s--a---st    241
01_AE.CF.90CF11697  NS..........K----------A-------PL---------------G-N--------------------M---I-----.-------D-----------L--------P------------X----N    242
01_AE.CF.90CF402    -S..........K----------A--------L---------------G-N--------------------M---I-----.----I--D--------------------P---------A--------    242
01_AE.CF.90CF4071   -S..........K----------A----I---L---------------G-N--------------------M---I-----.-------D-------------------P------------------    242
01_AE.TH.93TH057    -S..........K----------A-------PV---------------G-N--------------------M---I--E--.-----------------------------P------------------    242
01_AE.TH.93TH065    -S..........K----------A-------PL---------------G-N--------------------M---I-----.---------T------------------P------------------    242
01_AE.TH.93TH902    -S..........K----------AK------PV---------------G-N-------------S------M---I-----.------I--------------------P------------------    242
01_AE.TH.94TH7091   -S..........K----------A-------PV---------------G-N-------P-----S------M---I-----.------I--------------------P------------------    242
01_AE.TH.95TNIH047  -S..........T--H-------A-------P------------I---G----------------------M---I-----.-------------P--------------P---I------------T-    242
01_AE.TH.CM240      -S..........K----------A----A--PL---------------G-N--------------------M---I-----.-------------P--------------P------------------    242
01_AE.TH.TH022      -S..........K----------A-------PV---------------G-N--------------------M---I-----.-----------------------------P------------------    242
CONSENSUS 02        ss..........???--y-----Aq---t--s?----------i---a---------?------------M---I-----.-------D-------e---v--v----iP-----------------t    233
02_AG.FR.DJ263      -S.................--------A----T--PM----------I-------------T------------M---I-----.-------D-------------------P------------------    239
02_AG.FR.DJ264      -S.................--------A----T--PM----------I-------------T------------M---I-----.-------D-------------------P------------------    239
02_AG.GH.G829       ............----------A-------------------------------T------------M---I-----.-------D-----------T--I----NP-----------------N    238
02_AG.NG.IBNG       -S.................--------AK---T--SM----------I---G----------------------M---I-----.-------D-------------------P------------------    239
02_AG.SE.SE7812     -S.................--------A----T--S-------------------------------------M---I-----.-------D-------------------P------------------    239
02_AG.SN.MP1211     .............SS---F-----A-------S------------I-------------------------M---I-----.-------D-------D---T--------P------------------    239
03_AB.RU.KAL153-2   -S..........K----------A----T--SM----------I-------------------------M---I-----.-------D-----------L--AQ---FP---------------SS--    242
03_AB.RU.RU98001    -S..........K----------A----T--SM----------I-------------------------M---I-----.-------D-----------L--AQ---FP-----------------    242
04_cpx.CY.94CY032-3 -S..........N----------A-------S------------I-------------------------M---I-----.-------D-------D---T--------P------------------    242
04_cpx.GR.97PVCH    -S..........K----------A--------M----------I-------------------------M---I-----.-------D------S----A--------P------------------    242
04_cpx.GR.97PVMY    -S..........N-------A--A-------S---------------G----------------------M---I-----.-------D--------------A----NPA-----------------    240
05_DF.BE.VI1310     NSS.........-A------------------L----------I-------------------------------.-------------------L---Q---V------D------------    243
05_DF.BE.VI961      -G.........P----F-----L----------L----------I-------------T-------------------.-------------------AQ--------I--------------    243
06_cpx.AU.BFP90     -S..........NL---------A--------M----------I-D----------T------------M---I-----.-------D-------------------P---I--------------    241
06_cpx.ML.95ML127   -S..........NL---------A-------------------I-------------------------M---I-----.-------D-------------------P------------------    242
06_cpx.ML.95ML84    -S..........-L---------A----I--------------AI-------------------------M---I-----.-------D-----------M---Q---MP-----------------    242
06_cpx.SN.97SE1078  -S..........NL---------A-------------------AI-------------------------M---I-----.-------D--------------AQ----P---I-D------------    242
10_CD.BFL061        -S.........-----------L-------PLT-------------------------------------------.-------------------L---Q---V----I--------------    242
10_CD.BFL071        -S.........-----------T-----Y-S------------I-----------------------------------.G------------------L--------V------------------    242
10_CD.BFL110        -S.........-----------L-------PL----------I-------------T---------------------.-------------------L--AQ--------I--------------    242
11_cpx.CM.MP818     -S..........N----------A-------PV-------------------------------------M---I-----.-------D-----------L---Q----------D------------    242
11_cpx.FR.MP1298    NS.........N----------A-------TM-------------------------------------M---I-----.-------D-----------I--------P------------------    242
11_cpx.FR.MP1307    --.........K----F-----A-------------------------------------------M---I-----.-----------D-----------------P------------------    242
11_cpx.GR.GR17      --.........K----------A--------M----------I-------------------------M---I-----.-------D------------------AP----------------GN    240
CPZ.CM.CAM3         -A.............--R---V---A---L---PM----------I---N-N-------M----------V-----AI----.G---V---V----------T-------V----L--A-----------    242
CPZ.CD.CPZANT       LLLNKQWCQRHLSGEGR----IVDAG-IAR--PLT---------C----N-N----------------H-------A--D--.G---V---V----------L--T----VQA--L---T----------    258
CPZ.CM.CAM5         -S...........I-R-------A----I---M----------I---N-N-------M----------V-----A-----.G---V---V----------L------------L--A-----------    241
CPZ.GA.CPZGAB       -ASA........GI-G---L---A-------------------------------------L---V-----A-----.G---V---V----------L--T---------L--------------    250
CPZ.US.CPZUS        -GS........NIG-S---VI--A--------M----------A------N-------M----------V-----AI----.G---V---A------D$--T--L-----------------------    243
```

HIV-1/SIVcpz
proteins

HIV-1/SIVcpz proteins

```
                                                                 major homology region (MHR)
                                                                          |-                -|                                                                                                  p24 \/ p2
B.FR.HXB2          LQEQIGWMTN...NPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQ    369
CONSENSUS A        p---ig---g...---i---D--k--i-------------V-----k--------------F-tl-----t-e--g---e---v--------s--ra--aG-------t--------g------------    369
A.KE.Q23           P--------G...-------D------------------V-----K--------------F--F-----T-D------D------------A--R----G----------------------------    369
A.SE.SE6594        P--------S...-------D------------------V-----K--------------F-V------T----G-------I--------S--R---TG----------------------------    367
A.SE.SE7253        P--------G...-------D------------------V-----K--------------F--------T-D--------------------S--R---AG-------S--------S------------    367
A.SE.SE7535        P--------G...-------D------------------V-----K--------------F----------D----------I--------S--R---AG----------------------------    368
A.SE.SE8538        P---LQ---S...-------D------------------V--------------------F-A------T----G---------------S------TG----------------------------    367
A.SE.SE8891        P--------S...-------D------------------V-----K--------------F--------T----G---------------S--R---SG----------------------------    367
A.SE.UGSE8131      -----A---S...---V---D------------------V-----K--------------F-A------T----G---D------------S--RG--AG---------------S------------    369
A.UG.92UG037       P----A---G...-------D-----M-------------V-----K--------------F--------T----G-------I--------S--R---AG----------------------------    369
A.UG.U455          V--------G...-------D--R---------------V--------------------F--------T-D--------------------S--R----G----------------------------    365
A2.CD.97CDKTB48    ---------S...---------------------------V-----K--------------F--------T--------D------------S--R----G----------------------------    375
A2.CY.94CY017-41   ---------S...D--------------------------V------------------F--------T--------D-----------RS--R----G-S--------------S-----------H    365
CONSENSUS B        -----g---n...--pi-----------l-----------ts----r-----p---------k---a---s-e---wm---l----a----k--------a----------------g----v------q    369
B.AU.MBC18         ---------G...--A----------------------I-----K----------------------T-----------------------------------------------S------------    374
B.AU.SC49          ---------S...---V----------------------V-------------------------------D-----------------------------------------S----I-------    369
B.AU.VH            ----------...-------------------------------K---------------R---------------------------R-------------------------------------    375
B.CN.RL42          ----------...----------------------------C-----------------------------D-----------------------------------------S----I-------    369
B.DE.D31           ----------...-----------------------------------------------------------T-----------------------------------------------------    369
B.DE.HAN           ----------...-------------------------------K----------------------T-----------------------------------------------------    369
B.ES.89SP061       -----A----...-------------------------------K-----------------------------------------------------------------------------    369
B.GA.OYI           ----------...-----------------------------------------------------------D-----------------------------------------------------    369
B.GB.CAM1          ----------...-----------------------------------------------------------------------------------------------------------    369
B.GB.GB8           ----------...----------------------------V-------------------------------D-----------------------------------------------------    371
B.GB.MANC          ----------...-----------------------------------------------------------------------------------------------------S----I-------    369
B.JP.JH31          ---------S...---------------------------S-----------------------------------------------------------------------------------    369
B.KR.WK            ----------...----------------------------A-----------------------------------------------------------G----------------S------------    369
B.NL.3202A21       ---------H...-----------------------------------------S-----------------------------------------------G-----------------------    369
B.TW.TWCYS         ----------...-----------------------------------------------------T-----D-----------------------------------------------S----------R    369
B.US.85WCIPR54     ----------...-----------------------------------------------------------------------------S-----------------------------S------------    369
B.US.AD8           ----------...-----------------------------------------------------------------------------------------------------------    369
B.US.DH123         ----------...---------------M-----------------------------------------------------------S-------------G-----------------------    369
B.US.JRCSF         ----------...----------------------------V-------------------------------T-----------------------------------------------------    369
B.US.JRFL          ----------...-----------------------------------------------------------------------------------------------------------    367
B.US.MN            ----------...----------------------------S-----------------------------------RT-------------------------------------------    372
B.US.NY5CG         ---------H...-----------------------------------------------------------------------------------------------------------    369
B.US.RF            ----------...----------------------------I-------------------------------D--------F---------------------------------S----I-------    369
B.US.SF2           ----------...-----------------------------------------------------------------D-----------------------------------------------    371
CONSENSUS C        -----a---s...---v---d------il-----------V----ik-----p--------Fk------at-d------d-----------kt--r-l--g-sl-------------s---rv-----sq    366
C.BR.92BR025       -----T----...---V---D------------------V-----K--------------F--------T-D------D---------------R----G-S--------------------------K    367
C.BW.96BW01B22     -----A---A...---V---D------------------V-----K--------------F--------T-D------D---------------R----G-SI-------------S---------G-    366
C.BW.96BW0402      -----A---S...-------D------------------V-----K--------------F-------ST--------D---------------R----G-S--------------S-----------    367
C.BW.96BW0502      -----A---S...---V---D------------------V--------------------F--------T-D----------------------R----G-----------------------------    366
C.BW.96BW1104      -----A---S...-------D------------------V-----K-----S--------F-------S--------D-----------R--------G-S--------------S----I-------    366
C.BW.96BW1210      -----N---S...-------D------------------V-----K--------------F--------T-D------D---------------R----G-S--------------S-----------H    367
C.BW.96BW15B03     -----A---S...-------D------------------V------------------F--------T-D------D------------P--R----G-----------------------------    366
C.BW.96BW1626      -----A---S...---V----------------------V------------------F--------T-D------D---------------R-S--G-S--------------------------    366
C.BW.96BW17A09     -----A----...-------D------TM-----------V-----K--------------F--------T-D------D---------------R----G----------------S-----------H    363
C.ET.ETH2220       -----A---G...---V---D------------------V-----K--------------F--------T-D------D---------------R----G-S--------------A------------    367
C.IN.93IN101       -----A---G...---V---D------------------V--------------------F--------T-D------D---------------R----G-S-----------------------------    366
C.IN.93IN904       -----A---G...---V---D------------------V-----K--------------F--------T-D------D---------------R----G-S--------------S------------    366
C.IN.93IN905       -----A---G...---V----------------------V-----K--------------F--------T-D------D---------------R------S----------------------------    366
C.IN.93IN999       -----A---G...---V---D------------------V-----K--------------FR-------T-D------D---------------R----G-S-----------------------------    366
C.IN.94IN11246     -----A---S...---V----------------------V----TK--------------F--------T-D------D---------------R----G-S--------------S---K--------    363
C.IN.95IN21068     -----A----...---V---D------------------V--------------------F--------T-D----------------------R----G-S--------------S------------    366
C.IN.AF209990      -----A---G...---V---D------------------V-----K--------------F--------T-D------D---------------R----G-S--------------S------------    366
D.CD.84ZR085       ---------S...----------------------------V-------------------------------------------------------G----------------S------------    369
D.CD.ELI           -----A---S...---------------V-----------V-------------------------------D---------------------------Q----------------S------------    369
D.CD.NDK           -----A---S...----------------------------V-------------------------------D---------------------------Q----------------------------    366
D.CD.Z2Z6          -----A---S...----------------------------V-----------------------------G-------------------------Q----------------S------------    370
D.UG.94UG1141      ---------S...----------------------------V-------------------------------D-----------------------------------------S------------    369
CONSENSUS F1       -----Q---s...---v---di------------------Vs----r--------------F-t----q-t-e--g---dt--v--a-----t---al-pG----------q-----g----v---a---    361
F1.BE.VI850        -----Q---G...---V---D------------------V-----K--------------F-V--------D--G---D-------------------TG----------------S------------    365
F1.BR.93BR020-1    -----Q---G...---V----M------------------VG------------------F--------T----G---D--------------------G----------------S------------    365
F1.BR.BZ162        -----Q---S...---V----------------------V------------------F--------T----G---D--------------------G----------------------------    365
F1.CD.VI174        -----Q----...---V----------------------V------------------F--------T----G---D--------------------G----------------------------    364
F1.FI.FIN9363      -----Q---S...---V---D------------------V------------------F-A------T----G---D------------I---G--IG----------R----------I-------    367
F1.FR.MP411        -----Q---S...---V---D------------------V------------------F-----------------S-----S-------------G----------------------------    365
F1.RW.VI69         -----Q---S...-------D------------------V------------------F------E-T----G---D---I-------------M--G-------------------------T---    365
F2.CM.MP255        -----T---G...---V----------------------V-----K--------------F--------T---------------S-------------G--------------------I------K    365
F2.CM.MP257        -----A---S...---V----------------------V-----K--------------F--------T----G----------------------G----------------S----I------K    369
```

```
                                                             major homology region (MHR)
                                                                |-                 -|                                                      p24 \/ p2
B.FR.HXB2           LQEQIGWMTN...NPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQ    369
G.BE.DRCBL          -----R---S...---V-----------------------V--------------------F--------T----S---D---I--------I---G--QG----------------S------------    362
G.FI.HH8793         -----R---S...---------------------------V--------------------F--------T----G---D---------------R----G-----------------------------    369
G.NG.92NG083        -----R---S...---------------------------V--------------------F--------T----G---D---------------R----G----------------S------------    368
G.SE.SE6165         -----T---G...---------------------------V--------------------F-C--------D--G---D---------------R---QG-S--------------S------------    369
H.BE.VI991          ----VA---G...-------D-------------------V-----K--------------FRV------T-D------D-----------R-------RG-SI-------------S------------    371
H.BE.VI997          -----A---G...--S----D-------------------V-----K--------------F--------T--------D---------------R---QG-SI-------------S------------    369
H.CF.90CF056        -----A---G...--A----D-------------------V-----K--------------F--------T-D----------------------R---QG-SI-------------S------------    369
J.SE.SE7022         ---------G...---------------------------V--------------------F-A------T-D------D-------------------SG-----------------------------    369
J.SE.SE7887         ---------G...---------------------------V--------------------F-A------T-D------D-------------------SG-----------------------------    369
K.BE.VI325          -----T---S...---V-----------------------V-----K-----S--------F--------T--------D--------------------G-S-------------------I-------    365
K.CD.EQTB11C        -----T---S...---------------------------V--------------------FRV------T--------------------R-------SG-----------------------------    365
K.CM.MP535          -----A---S...---V-----------------------V--------------------F--------T--------D--------------------G-S--------------S----I-------    365
N.CM.YBF106         -A--VA---S...-------D--R---V----R-------V---E-K-----------------------T-D-------------------Q-------G----------------A----------A-    371
N.CM.YBF30          -A--VA---A...---V---D--R---V----R-------V---E-K-----------------------T---------------------QL------G----------------A------------    371
O.CM.ANT70          Q----H-T-RP..-Q-----D--RK--V-----M-K----V-----K-----------------------T---------------------Q---S---G-------V--------T----------AT    368
O.CM.MVP5180        Q----I-T-RGA.-S.----D--RK--V-----M-K----V-----------------------------T---------------S-----Q-------E-------V--------T---KI-----AS    368
O.SN.MP1299         Q---VH-I-RP..-Q-----D--RK--V-----V-K----V-----------------------------T---------------------Q---S---G-------I--------T----------AA    368
O.SN.MP1300         Q---VH-I-RP..-Q-----D--RK--V-----V-K----V-----------------------------T---------------------Q---S---G-------I--------T-----------A    368
CONSENSUS 01        -q-------n...---i---di--r-------------s-V----------e-f-d-----y--------T-e----------v--------S---a--TG-t--------------s-----------q    368
01_AE.CF.90CF11697  ---------S...-------D-------------------V--------------------F--------T---------------------S------TG-----------------------------    369
01_AE.CF.90CF402    -H-------S...-----------K---------------V--------------------F--------T---------------------S------TG-----------------------------    369
01_AE.CF.90CF4071   ---------S...---V-----------------------V--------------------F--------T---------------------S------TG-----------------------------    369
01_AE.TH.93TH057    ----------...-------D-----------------R-V-----------------------------T---------------------S---S--TG----------------S-----------H    369
01_AE.TH.93TH065    ----------...-------D-------------------V-----------------------------T---------------------S------TG----------------S------------    369
01_AE.TH.93TH902    ----------...-------D-------------------V----------G-S-E--------------T-G----------I--------S------TG-S--------------S------------    369
01_AE.TH.94TH7091   ----------...-------DV------------------V----------G-S-E--------------T-G----------I--------S------TG-S--------------S------------    369
01_AE.TH.95TNIH047  ----------...-------S-------------------V-----------------------------T---------------------S------TG----------------S-----------H    369
01_AE.TH.CM240      ----------...-------D-------------------V-----------------------------T---------------------S------TG----------------S-----------H    369
01_AE.TH.TH022      ---------S...-------D-------------------V-----------------------------T---------------------S------TG----------------S-----------H    369
CONSENSUS 02        l----g---s...-------?------?------------Vs-------------------F--------T-e---w---------a----ks--r---pG-t-----t--------g------------    358
02_AG.FR.DJ263      P--------S...-------D-------------------V--------------------F--------T-----R---------------S--R----G-----------------------------    366
02_AG.FR.DJ264      ---------S...-------D-------------------V--------------------F--------T---------------------S--R----G-S---------------------------    366
02_AG.GH.G829       -----A---G...--------------V------------VG-------------------F--------T------------------------R----G-----------------------------    365
02_AG.NG.IBNG       ---------S...--------------V------------V--------------------F--------T---------------------S--R---TG-----------------------------    366
02_AG.SE.SE7812     ---------S...--------------V------------V--------------------F--------T-D-------------------S--R----G----------------S------------    366
02_AG.SN.MP1211     ---------S...-------D-------------------V--------------------F--------T---------------S----R--------G-------S---------------------    366
03_AB.RU.KAL153-2   ---------S...-------D-------------------V--------------------F--------T-D----------------------R---SG--------V--------------------    369
03_AB.RU.RU98001    ---------S...-------D-------------------V--------------------F--------T-D----------------------R---SG-----------------------------    369
04_cpx.CY.94CY032-3 ---------S...---V-------------------T---I--------------------F-C------T---------------------S------TG----------------S------------    369
04_cpx.GR.97PVCH    ---------S...---------------------------V--------------------F-C------T---------------------S------TG----------------A------------    369
04_cpx.GR.97PVMY    ----V----S...-------D-------------------V--------T-----------F-C------T-D-------------------S------TG----------------S------------    367
05_DF.BE.VI1310     -----A----...---------------------------V-----------------------------T----------------------------QQ----------------S------------    370
05_DF.BE.VI961      -----T----...---------------------------V-------------------------------D---------------------------Q----------------S------------    370
06_cpx.AU.BFP90     ---------S...---------------------------VG----K--------------F--------T--------D---------------R----G-----------------------------    368
06_cpx.ML.95ML127   ---------G...---V-----------------------V--------------------F--------T--------D--------------------G-----------------------------    369
06_cpx.ML.95ML84    -L-------S...---------------------------V--------------------F--------T--------D--------------------G-----------------------------    369
06_cpx.SN.97SE1078  -----R---S...---------------------------V--------------------F--------T--------D--------------------G-----------------------------    369
10_CD.BFL061        -----R---S...---------------------------V-------------------------------D--------------------------------------------S------------    369
10_CD.BFL071        ---------G...--V------------------------I-------------------------------D---------------------------------------------------------    369
10_CD.BFL110        -----T---S...-------------V-------------V-------------------------------D----------------------------TS--------------S------------    369
11_cpx.CM.MP818     ---------G...-------D--R----------------V-----K--------------F-A------T----G---D---I--------S--R---TG----------------S------------    369
11_cpx.FR.MP1298    ---------S...---V------R----------------V--------------------F--------T-D------D---I--------S--R----G-----------------------------    369
11_cpx.FR.MP1307    ----V----S...---V---D--R----------------V-----K--------------F-A------T----G-------------------RG--QG----------------S------------    369
11_cpx.GR.GR17      ---------G...----------R----------------V--------------------F------------------------------S--R---VG----------------S------------    367
CPZ.CM.CAM3         -V---A---A...---V------R--VV-----V----C-V-----K-----------------------T--------D------------N--R----G----------------A------------    369
CPZ.CD.CPZANT       V---MQ--STPQQ-GGV---D-------M----V--X---V---E-K-----------------I-------P--A-------I--------H------TG-S----L---------A----------AS    388
CPZ.CM.CAM5         ----VA---A...----------R--VV-----V-K--C-V-----K-----------------------T--------D------------S--R----G----------------A------------    368
CPZ.GA.CPZGAB       -------T-A...-------DV-R--V------V----C-V--------------------------------------D------------Q-------G----------------S-----------M    377
CPZ.US.CPZUS        ----V----A...-------D--R--VV-----V-K--C-VG----K------------------------P---T-------------------R----G------V---------A----------C-    370
```

HIV-1/SIVcpz proteins

```
                                                                                  pol cds start |->                                                  vpr binding
                                p2 \/ p7          |-Zn - motif-|       |-Zn - motif-|     p7 \/ p1            p1 \/ p6            |-   -|
B.FR.HXB2          VTNSAT.....IMMQRGNFRNQRK.IVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCTE...RQANFLGKIWPSYKG....RPGNFLQSR...........PEPTAPPE.ESFRS.GVETT...   470
CONSENSUS A        vqqtn?.....i---rg--rGqkr..I----------L-r----p---------k----------...-------ki---sk-.....-----P-s-...........p--t---A?-ifgm.gEei....   467
A.KE.Q23           -QHTN..............KG-KR..I----------L------L----------------------...------------R--.....-----P-N-...........L------A.-TCGM.-E-------   467
A.SE.SE6594        -HQTN.........----K---G-KR..I----------L--------------------------...------------S--.....-----P---.....................A.-ILGI.-E-I....   465
A.SE.SE7253        AQQHT-.....V----A---G-KR..I----------L--------------------------...--------M---S--.....-----P---.....................A.-L-GM.-E-I....   466
A.SE.SE7535        -QHPN..............G-KR..I----------L--------------------------...-------R----S--.....-----P---...........L------A.-I-GM.RE-I....   466
A.SE.SE8538        -QQPN..............G--R..I----------L-K------------------------...------------S--.....-----P---.....................A.-I-GM.-EVI....   465
A.SE.SE8891        -QQAN......V-------KG-KR..I----------L--------------------------...------------NR-.....-----P---.....................A.-I-GM.-E-I....   465
A.SE.UGSE8131      AQQTN..............G-K-..I----------L--------------------------...------------H--.....-----P---...............S---A.-MM...-E-I....   465
A.UG.92UG037       -QHTN..............KG-KR..I----------L-K--------------R---------...------------S--.....-----P---.....................AA-I-GM.RE-I....   468
A.UG.U455          -QQTS..............GP-R..I----------L-K------------------------...------------N--.....-----P---.....................A.-I-GM.-EKM....   463
A2.CD.97CDKTB48    -Q-TN........I-----KG-KR..I----------L--------------------------...------------N--.....-----P---...........T------M.-......E-I....   468
A2.CY.94CY017-41   -QSTN-.............G-KR..I----------L--------------------------...------------N--.....-----P---...........T------A.-NL-M.-E-I....   465
CONSENSUS B        vtnsat?????i-m-rg-frn-rk.tv-------e--iakn-----kk---k--k---q--d-te?.?------------hkg....-pg--l-s-.??????????pe-t---e.e--rf.-eett...   470
B.AU.MBC18         A--A-......V----S-------.N----------I---------R------------------...------------H--.....----------.PEPTAPLQSR---S----.----F.-E------...   485
B.AU.SC49          --SP-N..............TI----------L--H-------------R----------...------------H--.....----------.......................----F.-E------...   470
B.AU.VH            ----------------------.T----------I---------R------------------...------------H--.....----------.......................----F.-E------...   476
B.CN.RL42          -----ITNSAT-------------.----------I-K-----------------L------...------------H--.....----------.......................----F.-E------...   475
B.DE.D31           ----------.V------------.T----------I-K------R------------------...-----------------------R-.......................----F.-E--A...   470
B.DE.HAN           --G--A.....------------.T----------I--------------------------...------------H--.....----------.......................----F.-E.A-...   469
B.ES.89SP061       ----------.----K--------.----------I-K-------R------------------...------------HR-.....-----P-N-........LEPTA-P----.----F.-E------...   473
B.GA.OYI           -.--V-.....V---K---K----.T----------I-K--------------R----------...------------H--.....-------N-....................A.---GF.-E------...   469
B.GB.CAM1          ----------.--------------.T--------V--I-K----------------------N-...------------H--.....----------.......................----F.-E-K-...   470
B.GB.GB8           -.----------.----K----S---.T----------I----------------------E---...------------H--.....----------.....PEPIAP--------.----F.-G------...   477
B.GB.MANC          ----------.--------------.T----------I-K------R------------------...------------H--.....----------.......................----F.-E------...   470
B.JP.JH31          ----T-.....---------------.-I----------L-----------------------N-...------------S--.....----------.......................----F.-E------...   470
B.KR.WK            A-----.....------------R.T----------I-K------------------------...------------H--.....----------...............S----.----F.-E------...   470
B.NL.3202A21       ----------.--------------.----------I--------------------------...------------H--.....----------.......................----F.-E------...   470
B.TW.TWCYS         -P--TN.............------Y-----.T----------I--------------------------...------------H-E...----------.......................----F.-EQ-----...   470
B.US.85WCIPR54     A-S---.....--------------.----------I--------------------------...------------H--.....----------.......................----F.-E------...   470
B.US.AD8           ----------.--------------R.T----------I-K------------------------...------------H--.....----------.......................----F.-E------...   470
B.US.DH123         I--TSAT.........-----G--------..I----------ISK----------------------...------------H-E...----------...............S----.----F.-E--A...   470
B.US.JRCSF         ---P--.....--------------.N----------I----------------------E---...----------------...----------.......................----F.-E--A...   470
B.US.JRFL          ---P--.....--------------.----------I--------------------------...----------------...----------.......................K---F.-E--A...   468
B.US.MN            ----------.--------------.-I----------I-K-------R------------------...------------C--......R--P---...........T-------.----F.-E------...   472
B.US.NY5CG         ---P--.......-I--------K-.T----------I-K------------------------...------------H--.....----------.......................----F.-E------...   470
B.US.RF            ------------L-K----D---.---------V--I-K----------------------NE.G------------H--.....----------.......................----F.-E------...   471
B.US.SF2           ---P-N.....--------------.T----------I-K----------R--R----------...----------------...----------.......................----F.-E-K-...   472
CONSENSUS C        anst?......imm-rs-fKgpkr.iv----------i-r--r---k-------k----m-d-t-...rq-n---ki---hk-?...----f--s-??????????pe-t---a.---rf.?e-tt...   464
C.BR.92BR025       -N-TN......-----S-CKGPKR.TI----------L---------------------V-----...------------HR-.....----L--N-...........T-------.----F.-E------...   467
C.BW.96BW01B22     ANSMN......-----S--K-PKR.----------I--------------------------...------------H--.....-------N-...........L--S---A.----F..E------...   465
C.BW.96BW0402      TNTN.......V----S--KGP-R.----------I--------------------------...------------H--.....----------.......................A.----F..E------...   465
C.BW.96BW0502      ANSVN......----KS--KGP-R.N----------I-K------------------------...------------H--.....-------N-.....SEPAAPTV-----A.----F..E------...   471
C.BW.96BW1104      AN--N......----KN--KGPTR.----------I----K-----------------------...--R---------H--.....----------.......................A.---....E------...   463
C.BW.96BW1210      AG-AG......--------KGP--.-P----------I----------------------S-...G----------H--.....----------...............S---A.----F..E------...   466
C.BW.96BW15B03     A-SAN......-L---S--KGPKR.-I----------I--------------------------...------------H--.....-------N-...........T------A.---KF..E------...   465
C.BW.96BW1626      AN-TN.......-I--S--KGP-R.S----------I----------------Q----------...---D--------H--.....----------.......................A.----F.-.------...   465
C.BW.96BW17A09     AN-TS......----K---KGP-R.TI----------L-K------X--------------E---...------------H--G...-------N-....................A.----F..E------...   463
C.ET.ETH2220       -N-TT......----KS--KGPKR.AI----------L--------------------------...-------RL---N--.....----------PEPTAPPESLR-------P.----F..E-A-...   477
C.IN.93IN101       ANST.......-L---S--KGSKR.----------I-K------------------------...------------H--.....----------.......................A.----F..E------...   464
C.IN.93IN904       TNSA.......-L--KS--KGSKR.-I----------L--------------------------...------------H--.....----------.......................A.----F..E--P...   464
C.IN.93IN905       AN-T.......-L---S--KGSKR.----------L-K------------------------...------------H--.....----------.......................A.----F..E------...   464
C.IN.93IN999       ANSN.......-L---S--KGSKR.T----------I--------------------------...------------H--.....-------N-..PEPTAPPAR-------A.----F..E------...   473
C.IN.94IN11246     TNSA.......-L---S--KGSKR.-I----------L--------------------------...------------H--.....----------.......................A.----F..E--P...   461
C.IN.95IN21068     TNSA.......-L---S--KGSKR.----------L--------------------------...------------H--.....----------.......................A.----F..E------...   464
C.IN.AF209990      TNST.......-L---S--KGPKR.----------I-K------------------------...------------H--.....----------.......................A.----F..E------...   464
D.CD.84ZR085       A-SASAA....-----KS--KGT--.----------I--------------------------...------------H--.....----------.......................A.-X-GF.-E-I-...   471
D.CD.ELI           A---V-TA...-------KGP--.-I----------I-K----------R-------L-----...-------R----H--.....----------.......................A.---GF.-E-I-...   471
D.CD.NDK           --G---A....V-------KGP--.SI----------K--------------R----------...------------H--.....----------.......................A.---GF.-E-I-...   468
D.CD.Z2Z6          A----AA....V-------KGP--.TI----------I-K------R------------L-----...------------H--.....----------.......................A.---GF.-E-I-...   472
D.UG.94UG1141      A--AN-A....--------KGPK-.-I----------K--------------R----------...------------HN-.....----------.....PPA------A.-I-GL.-E-I-...   473
CONSENSUS F1       Atnaa??....-m--KS-fkGq-R.iv-------e--i-k--------------r----------...r----------n--.....------l-s-...............-----t---A.--fgf.rE-it...   461
F1.BE.VI850        ANSA.......----KS--KG--R.V----------I----------------R----------...------------N--.....----------.......................A.---GF.RE-I-...   464
F1.BR.93BR020-1    A--T-......----KS--KG--R.----------I-K--------------R----------...------------N--.....-----I-N-...............S---A.----F.-E------...   465
F1.BR.BZ162        A--AT......-L--KS--KG--R.---------T--I-K--------------R----------...G----------H--.....-------N-....................A.---GF.RE-I....   464
F1.CD.VI174        A--T-AA....-----KS--KG--R.----------I----------------R----------...------------N--.....----------.......................A.---GF.-E-I-...   466
F1.FI.FIN9363      A.-TT......----KS---G--R.----------I----------------Q----------...------------N--.....----------.......................A.--LGI.RE-V-...   466
F1.FR.MP411        A--A-......----KS-YKGP-R.FI----------I-K------------------------...------------N--.....-------N-....................A.---GF.KE-I-...   465
F1.RW.VI69         A--A-......----KS--KG--R.T----------V-K--------------R----------...------------N--.....----------...............S---A.---GF.RE-I-...   465
F2.CM.MP255        A-ST-......----KS--KG-KR.----------I--------------------------...------------NR-.....-------N-....................A.-N-GF.-EGI-...   465
F2.CM.MP257        A-GA-......----KS--KG--R.----------I---------R------Q----------...--------M---N--.....-------N-....................A.---GF.-E-IA...   469
```

```
                                                                                  pol cds start |->                                   vpr binding
                                p2 \/ p7          |-Zn - motif-|       |-Zn - motif-|     p7 \/ p1            p1 \/ p6            |-   -|
B.FR.HXB2             VTNSAT.....IMMQRGNFRNQRK.IVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCTE...RQANFLGKIWPSYKG....RPGNFLQSR...........PEPTAPPE.ESFRS.GVETT...    470
G.BE.DRCBL            ASGA-AA...----KS--KGP-R.TI----------L-----------------------E---...------------N--...-------N-.................A.-N-GF.-E-IA...    464
G.FI.HH8793           A-GA-AA...----KS--KGPKR.NI----------L--------------------------...------------N--...-------N-.................A.---GF.-E-IA...    471
G.NG.92NG083          ASGA-AAA...----KS--KGP-R.-I----------L-----------------------E---...------------N--...-------N-...........T------A.---GF.-E-IA...    471
G.SE.SE6165           ASGA-AA...------S--KGP-R.TI----------L--------------------------...------------N--...-------N-...........T------A.--LGF.-E-IA...    471
H.BE.VI991            ---ASAA...----K---KGP-R.T---S-------I----------------Q---------G...------------S--...------P-K-...........L------A.---GF.-E-I-...    473
H.BE.VI997            ---ANAA...----KS--KGP--.------------I----------------R----------...------------S--...------------------------A.---GF.-E-M-...    471
H.CF.90CF056          ---TN-A...----K---KG---.F-----------I----------------R----------...------------S--...------------------------A.---GF.-E-M-...    471
J.SE.SE7022           ---TN...............DHKR.---------Q--I-K-------------------------...------------S--...------------------------A.--LGF.-E-I....    468
J.SE.SE7887           ---TN...........-----DHKR.---------Q--I-K-------------------------...------------S--...------------------------A.--LGL.-E-I....    468
K.BE.VI325            ---TT...............KG---.-I------R---I-KH-----------------------...------------N--...K-------------------------A.-N-GF.-E-I-...    465
K.CD.EQTB11C          ----......V-------KG--R.-I----------L------------------------S-...--------F--LN-E...------N-...................A.---GF.-EKI-...    465
K.CM.MP535            ---P......VV---K---KGH--.------------I--------------------------...------------H--...--------------------------A.---GF.-E-I-...    465
N.CM.YBF106           AQ.T--S....VFV-----KGI--.TI----------L----K---RR------Q--------KNE.GX-------G-SPF--...------P-TT.........TRR------L.--YG-.QE-.....    474
N.CM.YBF30            -QQPT-S....VFA-----KGI--.PI----------L----K---RG------Q--------KNE.G--------S-SPF--...------P-TT.........TRK------L.--YGF.QE-.....    475
O.CM.ANT70            AQQDLKGGYTAVF----QNP.I--GTI----------I----------------Q--------RN..GK-------Y--PGGT...-----YV-RP...........AH-S---M.-......E-.....    468
O.CM.MVP5180          AQQDLKGGYTAVF----QNP.N--GPI----------I-K-------R------Q--------KN..G--------Y--PGGT...-----YV-KQ...........VS-S---M.-......EA.....    468
O.SN.MP1299           A.QDLKGGYTAVF----QNP.S--GPI----------L----------------Q--------KN..G--------Y--PGGT...-----YA-RQ...........VS-S---M.T......E-.....    467
O.SN.MP1300           A.QDLKGGYSAVF----QNP.G--GPI----------L----------------Q--------RN..GK-------Y--PGGT...-----YA-RQ...........VS-S---M.T......E-.....    467
CONSENSUS 01          aqhan...........r---KGqkr..I------k---L------p--k------k----------...-------ki--sn--...------P---............p------a.-nwGm.-E-i????    466
01_AE.CF.90CF11697    -QHAN...............KG-KR..I----------L---------R------Q----------...-----------LN--...------P---............L------A.--LGM.-E-I....    467
01_AE.CF.90CF402      -HHTN...........----K---KG----.I----------L----------------R----------...------------S--...------P---....................M.--LGM.-E-I....    467
01_AE.CF.90CF4071     -QHAN...............KG-KR..I----------L--------------------------...--------R----S--...------P---....................A.--LGM.-E-I....    467
01_AE.TH.93TH057      AQHAN...............KG-KR..I----------L--------------------------...--------F---N--...------P---....................A.-NWGM.-E-.....    467
01_AE.TH.93TH065      AQHAN...............KGP-R..I----------L--------------------------...------------N--...------P---....................A.-.WGM.-E-I....    466
01_AE.TH.93TH902      AQHAN...............KG-KR..I------R---L--------------------------...------------N--...------P---....................A.-NWGM.-E-I....    467
01_AE.TH.94TH7091     AQHAN...............KG-KR..I------R---L--------------------------...------------N--...------P---....................A.-NWGI$-E-I....    467
01_AE.TH.95TNIH047    AQQAN...............KG-KR..I----------L------L-------------------...------------N--...------P---....................A.-NWGM.-E-.....    467
01_AE.TH.CM240        AQHAT...............KG-KR..I------R---L---------Q----------------...------------N--...------P---....................A.-NWGM.-E-I-GEE    471
01_AE.TH.TH022        -QQAN...............KG-KR..I----------L--------------------------...------------N--...------P---....................A.-NWGM.-E-I....    467
CONSENSUS 02          vQqsN......im------rG-rt..I----------L----k----k------------------...r-----------s--...------P---...............-t---A.--fGm.gE-I....    456
02_AG.FR.DJ263        -QQ-N...............G--T..I----------L----K----------------------...G----------S--...------P---....................A.--LGM.-E-I....    464
02_AG.FR.DJ264        -QQ-N...............G--T..I----------L----K----------------------...------------S--...------P---....................A.---GM.-E-I....    464
02_AG.GH.G829         -QSPN......-L------KG-KR..I----------L--------------------------...------------N--...------P---...............S---A.---GT.RE-I....    463
02_AG.NG.IBNG         -QQAN......V--------G--T..I----------L----K----R-----------------...------------S--...------P---....................A.---GM.-E-I....    464
02_AG.SE.SE7812       -QQPN...............G--P..I----------L----K----------------------...------------S--...------P---....................A.--LGI.-E-I....    464
02_AG.SN.MP1211       AQQ-N...............G--T..I----------L----K----R-----------------...------------S--...------P---....................A.---GM.-E-I....    464
03_AB.RU.KAL153-2     -Q-AN.........----KS---GPKR..I----------L--------------------------...-------R----S--...------P---...............S---A.-N-GM.-E-I....    467
03_AB.RU.RU98001      -Q-AN.........----KS---GPKR..I-------D--L-----------------------N-...------------S--...------P---...............S---A.-N-GM.-E-I....    467
04_cpx.CY.94CY032-3   AS-A-AA...----KSK-KG--R.TI----------L--------------------------...-------RM---S--...----------N-....................A.-CLER.KE-.....    470
04_cpx.GR.97PVCH      AS-A-AA...----KS---G-KR.-I----------L---------R-R----------------...-------RM---S--...--------------------------A.--LEM.KE-.....    470
04_cpx.GR.97PVMY      ASSA-AA...----KS--KGP-R.SI----------L-----------------------P-...----S--RM---S--...--------------------------A.---EM.KE-.....    468
05_DF.BE.VI1310       A----ATA....-------KGP--.-I----------I-K-----G--------R----------...--------V---H--...--------------------------S---A.----F.-E-I....    471
05_DF.BE.VI961        A-G-PA.....V-------KGP--.SI------------K------R------R--------I-..G------RV-L-H--...--------------------------S---A.---GF.-E-I....    470
06_cpx.AU.BFP90       ASVG-.......----KS--KGPKR.SI----------L--------------------------...------------N--...----------N-....................I.---GF.-E-I....    467
06_cpx.ML.95ML127     ASGTL-A.....----KS-YKGPKR.SI------R---L--------------------------...------------N--...------N-....................A.---GF.-E-.....    470
06_cpx.ML.95ML84      ASGTVAAA....----KS--KGP--..I----------L--------------------------...-------R----S--...------N-....................A.---GF.-E-.....    470
06_cpx.SN.97SE1078    -SGATAA.....----KS--KG---.NI----------------------------------...------------H--...-------N-......PEQNR---S---A.---GF.-E-I....    475
10_CD.BFL061          A-GGN-.........------KGPK-.SI------------K------RR------R----L---P-..R------------N--...--------------------------A.---GF.-E-I....    470
10_CD.BFL071          A-KGNA..............KGPKR.-I-------G--I-K--------------R----------...------------H--...--------------------------A.---GF.-E-M....    469
10_CD.BFL110          A-SGNA..............KGPK-.NI----------I--H----------E--R----------...-------R----N--...--------------------------A.---GF.-E-.....    468
11_cpx.CM.MP818       -QQTN......V----S--KG-KR..I----------L----------------R----------...------------S--...--------------------------A.---GF.-E-I....    467
11_cpx.FR.MP1298      -RQTN......V---K---KG--R..I----------L--------------------------...------------S--...------I---....................A.---GF.-E-I....    467
11_cpx.FR.MP1307      AQQTN.......------S--KG-K-..I------R---L----------------Q---------...------------H--...--------------------------A.---GF.-EN-GPSP    471
11_cpx.GR.GR17        AQHTN...............KG-KR..I----------L------------$---------------...------------S--...--------------------------A.---GF.-E-.....    464
CPZ.CM.CAM3           LQ-PTG.....VFL----NGKPTR.KI----------L--------R-------Q------N-PAGE.--------G-SPWS-G.SK-------G-...........K------L.--YGY.QEA-....    474
CPZ.CD.CPZANT         AN-AQGTA...VFL----GNRGG-RPL----------------K---R----R--Q----L-N-PATNTGKV-----PT-TWW-C...------V-KE.........EVV------I.-IY...QE-H....    495
CPZ.CM.CAM5           -Q--SN.....VF-----NGRVN-.KI--Y-------V-------R-------Q---N--E--TGG.--------D-SPWS-R.SK-----P-GG...........K------M.--YGF.QET-....    473
CPZ.GA.CPZGAB         -Q-.QGRAD..VFF-K-QGAGPKR.KI----------L----K---R----R--Q---------G...--V-----G---RS-...------V-N-...........T------I.--YGY.QE-E....    479
CPZ.US.CPZUS          MK-PSS.....VFL-K-.NAGKPGRKI----------L--------R-------Q---------A.GN--------H-SPSWSGGSK------EN-...........K------I.-D-GY.QE--VT..    478
```

HIV-1/SIVcpz proteins

```
                                                vpr binding
                                                |-     -|       \ p6 terminus
B.FR.HXB2           ............TPPQKQEP......IDKELY.PLTSLRSLFGNDPSSQ$          500
CONSENSUS A         ............aS-pk--q.....?kdreqnp-lv--K------lL---          498
A.KE.Q23            ............VS-L---Q......K-R-QAQ--V--K------LL---          498
A.SE.SE6594         ............AS-PR--Q...........TP-SV--K-------L---          491
A.SE.SE7253         ............AS-P---Q......K-R-QNS-SV--K------LL---          497
A.SE.SE7535         ............AS-P---Q........-GQDP--V--K------LL---          495
A.SE.SE8538         .............-S-P---Q......K---QVP--V--K-------L---          496
A.SE.SE8891         .............-S-P---K.....KDRESEAP--I--K-------L---          497
A.SE.UGSE8131       ............AS-P---Q..........NNP-SV--K------LL---          492
A.UG.92UG037        ............VS-P---Q......N-RDQNP-SV--K------LL---          499
A.UG.U455           .............-S-A---L......K-R-QT.--V--K-------L---          493
A2.CD.97CDKTB48     .............-SSL---N........R-PST-AI--K------LL---          497
A2.CY.94CY017-41    .............-SSL---L......ETR-P-N-AI--K-------LL--          496
CONSENSUS B         ....????????t-s----p??????idkel-.-la--rs---n--ssq-          500
B.AU.MBC18          .............--S----.....IDR-SS--.--A-------------          517
B.AU.SC49           ....TPPQKQEPL-S----T.......---D--.--A--K--------L--          508
B.AU.VH             .............--S----........------.--A-------------          506
B.CN.RL42           .............--S----........------.--A--K----------          505
B.DE.D31            .............--F----........------.--A-------------          500
B.DE.HAN            ............A-S----........------.--A--K----S------          499
B.ES.89SP061        .............--S----L......------.-SA--K-------L---          503
B.GA.OYI            ............--------......---G--.-----------------          499
B.GB.CAM1           .............--S----........------.--A-------------          500
B.GB.GB8            .............--S----.......-N--P-.--A-------------          507
B.GB.MANC           .............--A----........------.--A---P----------          500
B.JP.JH31           .............--S----.......R-----.--A-----------P-          500
B.KR.WK             .............--S----........------.--A-------------          500
B.NL.3202A21        .............--S----.......R-----.--A-------------          500
B.TW.TWCYS          .............--S----........---D--.--A--E----------          500
B.US.85WCIPR54      .............--S----........----Q-.--A-------------          500
B.US.AD8            .............--S----........------.-----K----------          500
B.US.DH123          .............--S----.............----.--A--K-------...-          495
B.US.JRCSF          .............--S----QKQEP..------.-----------------          504
B.US.JRFL           .............--S----........----M-.-----------------          498
B.US.MN             .............--Y----KKQET..---D--.--A--K-------L---          506
B.US.NY5CG          .............--S----........------.--A-------S------          500
B.US.RF             .............--S----K......------.--A--K----------          501
B.US.SF2            .............--S----........------.-----------------          502
CONSENSUS C         ............?-apk-ep.....?kdre???p-tslk----sdpLsq-          491
C.BR.92BR025        .............--SR---T......-----..-----K----S--L-T-          496
C.BW.96BW01B22      .............-AP----.......K-R-...----------S--L---          492
C.BW.96BW0402       .............-VP----.......K-R-...-----K----S--L---          492
C.BW.96BW0502       .............-AP--------.KDREPYRE---A------SG-L---          502
C.BW.96BW1104       .............-AP---T......K-R-...--I--K----S--L---          490
C.BW.96BW1210       .............-A----------K-R-..P--A--K-------L---          494
C.BW.96BW15B03      .............-AP----.......K-R-...--I--K----S--L---          492
C.BW.96BW1626       .............-AP----.......K-R-...-------------L---          492
C.BW.96BW17A09      .............-AP----.......K-R-...-----FK----S--L---          490
C.ET.ETH2220        .............-SP---L......K-R-...A----K------HLL--          504
C.IN.93IN101        .............-AP----.......K-R-...-----K----S-LL---          491
C.IN.93IN904        .............-AP----.......K-R-...----------S--L---          491
C.IN.93IN905        .............-AP----.......K-R-...----------S--L---          491
C.IN.93IN999        .............-AL----.......K-R-...-----K----S--L---          500
C.IN.94IN11246      .............-AP----.......KER-...----------S--L---          488
C.IN.95IN21068      .............-AP----.......K-R-...----------S--L---          491
C.IN.AF209990       .............-ALQ-G-......K-R-...----------S--L---          491
D.CD.84ZR085        .............-S----Q....KDK-----.--A--K-------L---          502
D.CD.ELI            .............-S----Q......K-----.-----K-------L---          500
D.CD.NDK            .............-S----Q......K-----.--A--K----------          497
D.CD.Z2Z6           .............-S----Q......K-----.-S-A-K-------LL--          501
D.UG.94UG1141       ............------Q......K-----.-----K-------L---          502
CONSENSUS F1        ............?-SPk--Q....??kdeglyP--A--K----n--...?          488
F1.BE.VI850         .............-SP---Q......K-G---P--A--K-------...-          491
F1.BR.93BR020-1     .............-SP---Q......K-EG--P--A--K-------...-          492
F1.BR.BZ162         ............N-SP---Q......KEEG--P--A--K-..........          486
F1.CD.VI174         .............-SP---Q......KGEG--P--A--K-..........          487
F1.FI.FIN9363       .............-SPR--Q......KEEGQ-P--A--K-------...-          493
F1.FR.MP411         .............-SP---Q....KDEGQG--P--A--K----S--...-          494
F1.RW.VI69          .............-SP---Q......K-EG-SP--A--K-..........          486
F2.CM.MP255         .............-SP---Q......KGE-QAP--V--K----S--...-          492
F2.CM.MP257         .............-SP---Q......K---QVP--I--K----S-Q...-          496
```

```
                                                vpr binding
                                                |-     -|        \ p6 terminus
B.FR.HXB2           ............TPPQKQEP......IDKELY.PLTSLRSLFGNDPSSQ$          500
G.BE.DRCBL          .............-SP---Q......KE----.--S--K------Q...-          490
G.FI.HH8793         .............-SP-P-.......KE--TH.--A--K----S--...-          497
G.NG.92NG083        .............-SP----......KE----.-----K----S--...-          497
G.SE.SE6165         .............-SP---M......KE----P...--K----S--...-          495
H.BE.VI991          .............-SPR--L......KEQ-..P------------Q...-          498
H.BE.VI997          .............SSP---L......K----..P-FA--K-------L---          499
H.CF.90CF056        .............-SP---Q.....LK----..P--A-------S--LL--          500
J.SE.SE7022         .............-SP----......K-----.----------S--L---          497
J.SE.SE7887         .............-SP----......K-----.-----K----S--L---          497
K.BE.VI325          .............-SPR--T......K---QGS-----K-..........          486
K.CD.EQTB11C        .............-SLR--M......K-Q-QGP-----K----S--L---          495
K.CM.MP535          .............-SPR--T......K---QSP-----K-------L---          495
N.CM.YBF106         ..........KS-QGKEMQENQEK..TETS--P------------LX---          511
N.CM.YBF30          ..........KS-QGKEMQENQER..TENS--P----------------          512
O.CM.ANT70          ...............V-GQENQEQKGGPN---.-FA--K----T-...--          498
O.CM.MVP5180        ...............V-EQENQSQKGDQE---.-FA--K----T-...--          498
O.SN.MP1299         ...............M-GQENQEQKEDQN---.-FA-------T-...--          497
O.SN.MP1300         ...............M-GQENQEQKGDQN---.-FA--K----T-...--          497
CONSENSUS 01        ?............-Sf?kqeq......K-kehpp-lv--K-------l---          496
01_AE.CF.90CF11697  .............-SFP---Q......K---HPS--V--K-------L---          498
01_AE.CF.90CF402    .............-SFP---Q......K--KQPP--V--K-------L---          498
01_AE.CF.90CF4071   .............-SFSR--Q......K-R-HPP--V--K-------L---          498
01_AE.TH.93TH057    .............-SLL---Q......K---HHP--V--K-----------          498
01_AE.TH.93TH065    .............-SLP---Q......K--D.PP--V--K-------L---          496
01_AE.TH.93TH902    .............-SSL---Q......K---YPP--V--K-------L---          498
01_AE.TH.94TH7091   .............-SAXNEDE......K---YPP--V--K-------L---          498
01_AE.TH.95TNIH047  .............-SSL---Q......K---PPP--I--K-------L---          498
01_AE.TH.CM240      I............-SLP---Q......K---HPP--V--K-------L---          503
01_AE.TH.TH022      .............-SFL---Q......K---HPP-SV--K-------L---          498
CONSENSUS 02        ............tsspk--p......r-kG--P--?--K-------???-          483
02_AG.FR.DJ263      .............-S-P---A......R-QG--P--A--K--------...-          492
02_AG.FR.DJ264      .............-S-P----......R-QG--P--A--K--------...-          492
02_AG.GH.G829       .............-SSPQ----.....R--G--P-----K--------...-          491
02_AG.NG.IBNG       ............P-SPQ----.......R--G--P-----K--------...-          492
02_AG.SE.SE7812     .............-SS-----.......G--G--P--A--K--------...-          492
02_AG.SN.MP1211     .............-SSP----.......G--G--P-----K-------Y---          495
03_AB.RU.KAL153-2   .............--SL---Q......K-R-QHP-SI--K----D--L---          498
03_AB.RU.RU98001    .............--SL---Q......K-RGQHP-SI--K-------L---          498
04_cpx.CY.94CY032-3 .............-SSL----.......R-----.-----K----S--L---          500
04_cpx.GR.97PVCH    .............-SSP----.......R-----.-----K-----H-L---          500
04_cpx.GR.97PVMY    .............-SSP---Q......R-----.-I---K----S--L-R-          498
05_DF.BE.VI1310     ............ASSP---Q......K-EG--P--A--K-------L---          502
05_DF.BE.VI961      .............--SP---Q......K-EGK-P--A--K-------L---          501
06_cpx.AU.BFP90     ............A-SP---S...KEKEE-G--.--A--K----S--...-          497
06_cpx.ML.95ML127   ............A-SPE-K-......KE----.--------------...-          497
06_cpx.ML.95ML84    ............A-SL-----....KEKE----.--A--K--------...-          499
06_cpx.SN.97SE1078  ............A-SP----......KE----.--A--K----S--...-          502
10_CD.BFL061        .............--S----Q.....K---PH.--A--K-------L---          500
10_CD.BFL071        .............--F----Q.....KDKE---.L-A--K-------L---          500
10_CD.BFL110        .............-S----Q......K-----.--A--K-V-----L---          497
11_cpx.CM.MP818     ............A-SP-----....KEKE----.-----K----S--...-          496
11_cpx.FR.MP1298    ............A-SP----......KE----.-----K-------L---          497
11_cpx.FR.MP1307    KQEQSFGIGENVA-SP----......KKE---.-----K-------LL--          513
11_cpx.GR.GR17      .............--SP----------KE-----.-I---K----S--L---          494
CPZ.CM.CAM3         ................GKRQEEEREKEKEP--P-----K----S----E-          507
CPZ.CD.CPZANT       ..................KRTQKGLKGEE--.P-SY--K----K-...--          522
CPZ.CM.CAM5         ....................MEKKKEEKEP--PQ----K----S---LE-          502
CPZ.GA.CPZGAB       ..................KSQEKKEGESS.--P-.---K----S------          508
CPZ.US.CPZUS        ...................QEKQGKEKEP..FQ----K----S----E-          505
```

HIV-1/SIVcpz proteins

```
                                                                                                          protease /        <-| gag cds end
B.FR.HXB2          FFREDLAFLQ....GKAREFSSEQT.....RANSP...........TRRELQVWGRD...............NNSPSEAGAD....RQGTV..SFNFPQVTLWQRPLVTIKIGGQLKEALLDTGADDTVL      89
CONSENSUS A        ----Nl--qq....gE-rk------.....?an-p..........?tS-dlwdggr-................sL-s-a--e....rqgt?p.tfS---I------l--Vkig---ke------------      87
A.KE.Q23           ----N---QK....-E---------.....GT--S...........-S-D-WDG---................SL------E....----GP.TLS---I---------VR------------------      89
A.SE.SE6594        ----N---Q-....RE--K------.....-----...........-S-DPWDRR--................SL---T---..........P.T-S---I---------V-V----R-------------      84
A.SE.SE7253        ----NV--Q-....-E--K------.....G----...........-S-AFWDG---................SL------E....----EL.T-S---I---------V------R-------------      89
A.SE.SE7535        ----N---Q-....-E-GK------.....G----...........-S-D-WNE---................SL------E.....--RP.T-S---I---------V--E----------------      87
A.SE.SE8538        ----N---Q-....RE--K------.....-----...........-S-D-WDG-S-................-L------E....----GP.TLS---I---------V-V-----------------      89
A.SE.UGSE8131      ----N---P-....-E-GK------.....G-I--...........-S-D..DG---................SL------K....-...P.T-S---I------I--VR------------------      83
A.UG.92UG037       ----N---Q-....RE--K------.....-T---..........TSS-D-WDE---................SL------E....---PEP.T-S---I---------V-------K-----------      90
A.UG.U455          ----N---Q-....-E---------.....-----...........-S-N-WDG-K-................DL-C-T--E....----D..--S---I---------V------I------------      88
A2.CD.97CDKFE4     ----N--SQ-....RE--K--P---.....-----...........-S-G-WNG---.........................K....K---AY.-C----I---------V--E---R------------      81
A2.CY.94CY017-41   ----N---Q-....RE--K-----N.....-----...........-S---ENG---................-LLP---TG....D---IQ.-C----I---------V--E----------------      89
CONSENSUS B        ----d---p-....-k-refs-e-t.....--n--???.???????trr--q---rd...............n--l----ad????r--tv??-fsf-qI------l-tiki-----------------      89
B.AU.MBC18         ----N---P-....-E---L-----.....-----SPE.QAGAISP-----------...............---L------.....---QQPV--S---I----------------------------     101
B.CN.RL42          --------P-....-----L-----.....-----...........--G--------...............---I------.....----I..--S---I-----------V----------------      89
B.DE.D31           ----------....--------A--.....-----...........-----------...............S--L------.....-------.--S---I----------------------------      89
B.GA.OYI           --------P-....-----------.....-----...........-S---R-----...............----------.....-------.---L--I-----I----------------------      89
B.GB.CAM1          ----N---P-....-E---------.....-----...........----------E...............---L------.....-------.--S---I----------------------------      89
B.KR.WK            --------P-....------P----.....--I--...........A----------...............---L-----N.....-------.--S---I--------A-------------------      89
B.NL.3202A21       --------P-....-----------.....-----...........-----------...............---L-----E....G-----..-LSL--I----------------------------      89
B.TW.TWCYS         ----N---P-....----K------.....-----...........--G-----T-................---L------.....---P-..--S--RI----------R-----------------      89
B.US.JRCSF         ----------....------P----.....-----...........-----------...............S--L-----EAGAD---I-..------I----------------------------      93
B.US.JRFL          ----------....-----------.....-----...........--K--------...............S--L------.....---N-..------I----------------------------      89
B.US.MN            ----------....---.------N.....-----...........-----------...............---L----EEAGDD---P-..--S---I------I---------------------      92
B.US.RF            ----N---P-....-----L-----.....-----...........-----------...............--L----E-.........--S---I------I--V-------------------      88
B.US.SF2           ----------....-----------.....-----...........---------GE...............---L------.....-------.------I----------R-----------------      89
CONSENSUS C        ----n---p-....-e--efps--t.....r-n-p???????????ts-e-qv?rgd...............?npr-eagae....rq-????tl-f--I--------s-kv---ik---------d---      85
C.BR.92BR025       ----N---P-....-E--KS----N.....-----...........-----------...............---L----D-.........AL----I--------N--V------------------      88
C.BW.96BW01B03     ----N---P-....------P----.....--I--...........-S-----.RG-................-PR-----E....G--....AL-L--I----------RV---I-------------      85
C.BW.96BW0402      ----N---P-....-E----P----.....-----...........-S-----.RG-................-PR--TR-E....G--....TL----I--------S--V---IR------------      85
C.BW.96BW0502      ----N---P-....-E----PP--I.....--S--.....NSTNSP-S-----.RG-................-PR-----E....G---LQGTL-C--I--------S--V---I----------N---      95
C.BW.96BW1104      ----N---P-....-E----P----.....G----...........-S-K-R..GN.................-PC----DE.........TL----I--------S--V---I-------------      83
C.ET.ETH2220       ----T---Q-....------P----.....-----TRESQTRANSP-T-----.RGS................-TF-----E.........-L----I----------------------------      96
C.IN.93IN999       ----N---P-....-E----P----.....-----SS..QTRANSPSS-----.RG-................-PR-----K.........TL----I--------S--V---IR------------      94
C.IN.94IN11246     ----N---P-....-E----PP---.....-----...........-S-----.RG-................TPS-K---E....-E-....TL----I--------S-RV---TR------------      85
C.IN.95IN21068     ----N---P-....-E----PP-$-.....----S...........-S-----.RG-................-PS-----E.........T-----I--------S-RV---I-------------      84
D.CD.84ZR085       ----N---P-....---G-L-----.....-----...........-S-X-R---G-................-PLP-T-TEGQ..-------..-LS---I------V---------------------      90
D.CD.ELI           ----N---P-....---G-L-PK--.....-----...........-S---R-----................-PL-KT--E....-------.------I--------A-------------------      88
D.CD.NDK           --------P-....---G-------.....-----...........-S---R---G-................-PL--T--E....-------.--S---I----------------------------      88
D.CD.Z2Z6          --------P-....---G-L-----.....-----...........-S---R-----................-PL--T--E....-------.---C--I----------------------------      88
D.UG.94UG1141      ----N---P-....W-----P----P...S-----...........-S-D-RIR-G-................-TS--T--E....-------.---L--I------V--V-----------------      90
F1.BE.VI850        ----N---Q-....-E--K-P----.....-----...........-S---R-QRG-................-PL-----E....-R---P.-LS---I---------------I------------      89
F1.BR.93BR020-1    ----N---Q-....-E--KLHP--A.....--V--...........AS-----R-G-................-PI-----E....-R---P.-LS---I----------RV-----------------      89
F1.FI.FIN9363      ----N---Q-....-E--K-P--$-.....-----...........AS--PRDQR-G................--L-----E....-R---P.-LS---I----------------------------      88
F1.FR.MP411        ----N---Q-....-E--K-----A.....-----...........ASG--R-QRGN................-PL-----E..GRGT---S.-LSL--I-F--------RV----R------------      91
F2.CM.MP255        ----N---Q-....-E--K------.....-----...........AS---R-RRG-................-PLP----E....-R--GS.-LS---I--------A-RV----R------------      89
F2.CM.MP257        ----NV--Q-....-E--K------.....-----...........AS---R-R-G-................S-LP----E....----GS.-LD---I------V----V----R------------      89
G.BE.DRCBL         ----N---Q-....-E----P---A.....-----...........------R-R-G-...............SPLP----E....GK--IS.-I.---I------I-KVR-----I------------      88
G.FI.HH8793-1-1    ----N---Q-....-E--------A.....-----...........------R-RRGN...............SPLP--R-E....GK-DT..-LS---I----------V-----I------------      88
G.NG.92NG083       ----N---Q-....-E--KL-P--D.....-----...........-S---RIRRG-................SPLP----K....GE-AI..-L----I----------V-----I------------      88
G.SE.SE6165        ----N---Q-....-E--------D.....-T---...........-C-KPR-RRG-................SPLP---DE....GK-AI..-L..--I----------V-----I------------      86
```

```
                                                                                                            protease /         <-| gag cds end
B.FR.HXB2           FFREDLAFLQ....GKAREFSSEQT.....RANSP...........TRRELQVWGRD...............NNSPSEAGAD....RQGTV..SFNFPQVTLWQRPLVTIKIGGQLKEALLDTGADDTVL    89
H.BE.VI991          ----N---Q-....------PP-EA.....-----...........-S---R-RRG-................HPL-----E....-T--...------I------I--V--E-----------------    87
H.BE.VI997          ----N---Q-....RE--K--P--A.....-----...........-S---R-R-G-................DLLP----E....G---...-LC---I---------V--E---R-------------    87
H.CF.90CF056        ----N---Q-....RE--K--P--A.....-T---...........-S---R-RRG-................DPL-----A...EG---...-LS---I---------V--E---R-------------    88
J.SE.SE7022         --------Q-....RE-----P---.....-----...........-S--PR-RRG-.................PLP-T--E....G----S.-.----I----------R-----R-------------    87
J.SE.SE7887         --------Q-....RE---L-P---.....-----...........-S--PRARRG-.................PLP-T--E....G----S.-.----I----------R-----R-------------    87
K.CD.EQTB11C        ----V--SQ-....RE--K------.....-----...........-S---W-R-E-................-PL--T-NE....-S--GS.------I------V--V-V----R-------------    89
K.CM.MP535          ----N---P-....-E---------.....-----...........-S---R-R-G-................-PL----DQ....----EP.------I------I----V----R-------------    89
N.CM.YBF106         ---KG-VS--....RET-KLPPDNNK...E--H--...........AT---W-S-GEEH...........TGKGDAG-P-E-....-DLS-P.TL----I------VXAV---KEIR------------I    96
N.CM.YBF30          ----E-VS--....RET-KLPPDNNK...E--H--...........AT---W-S-GEEH...........TGEGDAG-P-E-....-ELS-P.T-----I------VI-V---KEVR------------I    96
O.CM.ANT70          ---QI--SGG....HE--QLCA-TS.....TPI--...........-DGGGSEGT.................GE-GT-R-......PERAL..-VCL--IP--D--I--ARV--H-C-V-----------    85
O.CM.MVP5180        ----V--SGG....HE--QLCA-TS.....VPI--...........-NGGGSEGT-E.................-E--G-SG......RA-P..ICL--IP--D--I--A-V--H-C-------------    85
O.SN.MP1299         ----I--SGG....HE--QLCA-TS.....VPI--...........-DDGGNEGT-.................E-GT-G-......PER-L..-VCL--IP--D--V--ARV--H-C-V-----------    85
O.SN.MP1300         ----I--SGG....HE--QLCT-TS.....VPI--...........-DDGGNEGT-K.................-GT-G-PE....-..-L..-VCL--IP--D--I-PARV--H-C-V-----------    85
CONSENSUS 01        ----n--fQ-....gk-ge-s----.....ra--p...........tsrklgDG-rd??...........???nllteagaE....--gts.s-fs---I-l----l--vkig-q---------------    89
01_AE.CF.90CF11697  ----N--SQ-....-E---------.....-----...........-S---GDG---................-LLP----E....----P.F--S---I-F-------V--E-----------------    89
01_AE.CF.90CF402    ----N---Q-....-E--K-P----.....-----...........-NG--GDG---................-LLP----E....--E-A.S--S---I---------V-V------------------    89
01_AE.CF.90CF4071   ----N---Q-....-E---------.....-----...........-S---RDG---................-LLL----E....----S.S-LS---I------I--V--------------------    89
01_AE.TH.93TH057    ----I---Q-....---GK------.....-----...........AS-K-GDG---................-LLT----E....----S.S--S---I---------V--------------------    89
01_AE.TH.93TH065    ----N---Q-....R--G-------.....-----...........-S-M.GDG---................-LLT----E....---.S.S--S---I---------V--------------------    87
01_AE.TH.93TH253    ----N---Q-....-----------.....G---S...........AS-K-GDG-GA........................E....----S.S--S---I------------------------------    81
01_AE.TH.93TH902    ----N---Q-....---GK------.....-----...........-S-K-GDG---................-LLT----E....---IS.S-LS---I------I---R-------------------    89
01_AE.TH.94TH702    ----N---Q-....---G-------.....-----...........-S-K-GDG---................-LLT----E....---IS.S-LS---I------I---R-------------------    89
01_AE.TH.95TNIH047  ----N---Q-....---GK------.....-----...........-S---GDG---................-LLA-T--E....----P.S------I---------V----E---------------    89
01_AE.TH.CM240      ----N---Q-....---G-------.....-----...........-S-K-GDG---NG...........GRD-LLT----E....----S.S--S---I---------V--------------------    94
01_AE.TH.TH022      ----N---Q-....---GK------.....-T---...........-S-K-GDG---................-LLP----E....----P.S--S---I---------V--------------------    89
CONSENSUS 02        ----N---q-....gE--kf--e--.....Gtn-p...........ts--lwDg--d................nllse--te....Gq-tI.ssf?---I-------vtVr-gg--Iea--d---d----    88
02_AG.CM.MP807      ----N---Q-....-E--K------.....GT---...........-S---WDG---................TLC----TG....GE--I.S--S---I-------F-VR-----I-------------    89
02_AG.FR.DJ263      ----N---Q-....-E--K---K--.....GT---...........-S---GDG---................-L-----SE....GP--I.S-LS---I---------V-----I-------------    89
02_AG.FR.DJ264      ----N---Q-....-E--K---K--.....GT---...........-S---WDG---................-L-----TE....GP--I.SPSS---I---------VR-----I-------------    89
02_AG.GH.G829       ----N---Q-....RE---L-----.....G-I--...........-G---WDK--N................-LL-A--TE....G---I.S------I---------VR-----I-------------    89
02_AG.NG.IBNG       ----N---Q-....-E--K------.....GT--S...........-S---WDG---................T-L-T--TE....G--AI.S------I---------VR-E---I-------------    89
02_AG.SE.SE7812     ----N---Q-....-E--K------.....GT---...........AS--PWDR---................-LL----TG....G---I.S-LS---I---------V-----I-------------    89
02_AG.SN.MP1211     ----N---Q-....-E--KL-----.....GT---...........-S---WDG---................-LLP---TG....G---I.P------I-------F-VR-E---IK-------N----    89
02_AG.SN.MP1213     ----N---P-....-E--K------.....GT---...........-S--CWDG---................SLL----TE....GE--I.S------I-------FKVR--E--IKP--N---N----    89
03_AB.RU.KAL153-2   ----N---Q-....RE--K------.....--I--...........-S-K-WDG---................-PLP-T-TE....G---A.S------I---R-----VR-------------------    89
03_AB.RU.RU98001    ----N---Q-....RE--K------.....--I--...........-S-K-WDG---................-PLP-T-TE....----A.S---L--I---------VR-------------------    89
04_cpx.CY.94CY032-3 ----NV--Q-....RE--K-----A.....-----...........A-GM-REERG-................-LL----TE....G---I..------I-----------L---IR-------------    88
04_cpx.GR.97PVCH    ----NV--Q-....R--G------A.....-----...........-----RDERG-................-LL----TE....G---I..------I----P----------IR-------------    88
04_cpx.GR.97PVMY    ----NV--Q-....-E--K-----D.....-----...........A----RDERG-................-LL----TE....G---I..-H----I----------------R-------------    88
05_DF.BE.VI1310     ----S---P-....-E---LPP---.....G-L--...........AS-------G-................SLL-----E....GR---.P-LS---I------V---R-------------------    89
05_DF.BE.VI961      ----S---P-....------PP--A.....-TL--...........-S---R--RG-................-PLA----E....GR-E-.P-LS---I-F----V-----E-----------------    89
06_cpx.AU.BFP90     ----N---Q-....-E--------A.....-----...........-H---RFRRG-................SPLP-T-VE.GEGGK-AI..-LSL--I---------VR-----I-------------    91
06_cpx.ML.95ML127   --------Q-....-E--K-----A.....-----...........--G--R-RRG-................SPLPG-E-E....GK-AI..-L----I---------VR-----I-------------    88
06_cpx.ML.95ML84    ----N---Q-....-E-G-L----A.....-----...........-----R-RRG-................SPLP---TE..GKGK-AI..-LS---I---------VG-E---I-------------    90
06_cpx.SN.97SE1078  ----N---P-....-E-----P--ARTEQA-TL--...........-C---R-R-GN................SPLP----E....GE-AI..-LS---I---------V-V-E--I-------------    93
10_CD.BFL061        ----N---Q-....R----LP----.....----S...........-S---R---G-................-TL--T--E....---AA..-LS---I------------------------------    88
10_CD.BFL071        ----N---P-....-----LP----.....-----...........-S---R---G-................DTL--T-TE...RQG-A-..--S---I---------V-------------------V    89
10_CD.BFL110        ----N---Q-....R----LP----.....-----...........-S---R--R-A..................L--T--E....------.-LS---IS--------V--------------------    86
11_cpx.CM.MP818     ----N---Q-....R----L-P---.....-----...........-S---R-RRG-................SPLP-T--E..GEGE-AI..---L--I------I-K--VA-----------------    90
11_cpx.FR.MP1298    ----N---Q-....------YP--A.....-----...........-S---R-R-G-................SPLP----E....GE-AI..-L----I-----------VA-H---------------    88
11_cpx.FR.MP1307    ----N---P-....-E-----P--A.....-----...........-S---R-R-EHRSLPETGAELRDRGERSPLP-T--E....EG-AI..-S----I------I-----E---R-------------   104
11_cpx.GR.GR17      ----N---Q-....-E----PT--A.....-----...........-S---R-R-G-................-PLP-T--Q....GE--I..-Y----I------V-PV-VA-----------------    88
CPZ.CD.CPZANT       ----TDPHVV....-VQTRELCAEGG...SSGANSS..........-H-D-SGGAQE................D-EGSQ-GG....G..-T..-LV--EIP-----MMEVL-Q--KCQ-----------V    89
CPZ.CM.CAM3         ----R-VSVE....RGIK-TRELPAGQEGTHSS.............AP---W-S-SN................RKETGG-EGEG..KGTP-.S-V-L--I-------LGV--E---I------------I    94
CPZ.CM.CAM5         ----G-VSVE....REIKKTRELSAGREGTHSS.............-HG--W-S-NN................-GEK-G-......EGTP-.S-I-L--I-------LGV-VE---I------------I    90
CPZ.GA.CPZGAB       ----R---P-....RE--QLCA--N.....-T-G-...........-D---W-P-GR................EE-G-ERGR....E-.SI..-T-L--I-------IPV-VE---C------------I    87
CPZ.US.CPZUS        ----T-VPIVERGIKET--LPGK-E.....G-H-S...........-N-GFR-SR--................SDTG-T.....GKGK-AL..-A----I-------LR--VA--IV-------------    91
```

HIV-1/SIVcpz proteins

```
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
```

M41L

protease \/ p66, p51 RT

```
B.FR.HXB2           EEMSLPGRWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIK    219
CONSENSUS A         -DIN---K-k--m----------k----il-----K-a-----------------M--q------------e--------g----kv-------------T---te--k--k-------------i----    217
A.KE.Q23            -DIN---K---------------K-----------K------------------M-----------------------------------------------T------------------------I----    219
A.SE.SE6594         -DIN---K---------------K-----------K------------------M-----------------------------R-------------T------------------------I----    214
A.SE.SE7253         -DIN---K---------------K-----------K------------------M-----------------------------------------------T------------------------I----    219
A.SE.SE7535         -DIN---K---------------K----V------K------------------M-----------------------------------------------T------------------------I----    217
A.SE.SE8538         -DIN---K-R-------------K-----S-----K------------------M-----------------------------------------------T------------------------I----    219
A.SE.UGSE8131       -DIN---K----------------------------K------------------M-----------------------------------------------T------------------------I----    213
A.UG.92UG037        -DIN---K---------------K-----------K------------------M--L------------S-------------RI--------------T---AD--R--R-------------I----    220
A.UG.U455           -DIN---K----I---------------------K-T----------------M-----------------------E-------------------T---N-----------------------    218
A2.CD.97CDKFE4      -DIN---K---------------------A-----KR-----------------M-V-L-----------------------------------------T---G-------------------------    211
A2.CY.94CY017-41    --IN---K---------------------A-----KR-----------------M-V-L-----------------------------------------T---K-------------------------    219
CONSENSUS B         ee-n--gr---------------r---qipie------i----v---------------i-----------e----k---------------------a-v--------------i-------------k    219
B.AU.MBC18          -D-H---K---------------K---E--V-----------------------------------------------S-----------------------------------------------R    231
B.CN.RL42           -D-N------------------------P----------------------------L-----------K-----------------------------------------------------    219
B.DE.D31            ------------------------------------------------------------------------------------------------------------------------------    219
B.GA.OYI            ---N-------------------------------------------------------L-----------------------------------------V-I--------------V--------------    219
B.GB.CAM1           ---N------------------------P-------------------------------------------------------------------------------------------------    219
B.KR.WK             ----------------------------VA-------------I-------------------------------------------------------------------------------    219
B.NL.3202A21        ---N------------------------P-------------------------------------------------------------------------------------------------    219
B.TW.TWCYS          ---N---K---------------------P---------------------------------------D--------------------------V-I-----------------------------    219
B.US.JRCSF          -D-D------------------------P-D------V-----------------------------------------------------------------------------------------    223
B.US.JRFL           -D-D------------------------------------------------------------------------------------------------------------------------    219
B.US.MN             G--N--R----------------------T-G---------------------------L-----------------------------------------I-----------------------------    222
B.US.RF             ---N---K----------------------------------------------------------------------------------------------------------------------    218
B.US.SF2            ---N---K---------------------PV-------------------------------------------------------------------------------------------------    219
CONSENSUS C         -ein---k-k----------------dq-p-----K-------------------m---L-c----------------------------p--e---k--Ta--e---k----tk----------v----    215
C.BR.92BR025        --IK---N--------------------------K-------------------M---L-------------------------------L---------TA--D---R----T----------------    218
C.BW.96BW01B03      --IN---K--------------------------K-------------------M---L-----------------------------------------TA--E--------T----------------    215
C.BW.96BW0402       -DIN---K------------------E--------K-------------------M---L-----------------------------------------TA--E--------T----------------    215
C.BW.96BW0502       --IN---K---------------------V-----K-------------------M---L-----------------------------T---------T---E--------T----------------    225
C.BW.96BW1104       --IN---K-R-------------------P-----K-----------------------L-----------------------------------------TA--E--------T----------------    213
C.ET.ETH2220        --IN---K---------------------I-----K-------------------M---L-R---------------------------------------TA--E---Q-----R----------------    226
C.IN.93IN999        --VN---K---------------------P-----K-----------------------L-----------------------------------------TA--D---R----T-----------I----    224
C.IN.94IN11246      --VN---K---------------------P-----K-------------------M---L-------------------------------K---E--TI--N--------T-----------I----    215
C.IN.95IN21068      --V----K-R----------------EE-P-----K-------------------M---L-----------------------------------------TA--D--------T-----------I----    214
D.CD.84ZR085        -DIN---K-------------------H-------------------------------------------------------------------------T----D--------R----------I----    220
D.CD.ELI            ---N---K---------------------P-----Q-------------------------------------------------------------------T----D--------R----------I----    218
D.CD.NDK            --IN---K--------------------------Y--M-------------------------------------------------------------------T-------------R----------I----    218
D.CD.Z2Z6           ---N---K----------------------------------------------------------------------------------------------------T-------------RV---------I----    218
D.UG.94UG1141       --IN---K---------------------PL-------------------------------------------------------------------------I---S-L------------------I----    220
F1.BE.VI850         -DIN---K---------------K---N--------------------------M-----------V--------------------------------T---L-------------------------    219
F1.BR.93BR020-1     -DVN---K---------------K---S--------R------------------M--------H-----------------------------------T---M-------------------------    219
F1.FI.FIN9363       -DIN---K---------------K---H----------------------V---M------------------------------------------T----D--------R---------------    218
F1.FR.MP411         -DID---K----I----------K-----T-D----------------------M-----------------------------------------T----D-------------------I----    221
F2.CM.MP255         -DIN---K---------------------P-----Q-----------------------L-------------------------R-------------T-----------------------------    219
F2.CM.MP257         -DIN---K--------------------VS-----Q------------------M-----------------------------------------T-----------------------------    219
G.BE.DRCBL          --ID---K------------------------S-KR-----------I------M---------------------------R-------------T---N-------------------I----    218
G.FI.HH8793-1-1     -DIK---K------------------------S-K------------I------M------X------------X---------------------X-XT--------R--------------X-I---R    218
G.NG.92NG083        -GIN---K------------------------G-K------------I------M---------------------------R-------------T---KD------------------I----    218
G.SE.SE6165         --IN-----------------------VP---S-K-----I------I------M---------------------------R-------------T---K---E----------------I----    216
```

```
                                                                                                                             M41L
                                                                            protease \/ p66, p51 RT                           |
B.FR.HXB2           EEMSLPGRWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIK    219
H.BE.VI991          -DIN---K------------------E-VA---F-K-------------------I---M-----L----------T-----------------------T---L--------------------I----    217
H.BE.VI997          --IN-L----------------------VA-----K-------------------I-----------------------------R--------------T---M--------------------I----    217
H.CF.90CF056        --IN---K------------------E-VA-----K-------------------I--------------------------------------------T-------------R-------S--I----    218
J.SE.SE7022         --ID---K------------------NEVP---E-K-------I-----------M---L--------------------------I-------------TQ--A-L-E-----R---------------    217
J.SE.SE7887         -DID--RK------------------NEVP---E-K-------I-----------M---L--------------------------I-------------TQ--A---E-----RV--------------    217
K.CD.EQTB11C        --IN---K--------------------VCM----Q-------------------M--------------------------------------------------------------------------    219
K.CM.MP535          --IN---K--------------------V------Q----------------------------------------------------------------T-----------------------------    219
N.CM.YBF106         --IQ-E-K-------------------N-T-D-Q-R--V----------------F-----------------------------R-------A---E--R-------------R----------I----    226
N.CM.YBF30          --LQ-E-K-------------------N-TVD-Q-R--V------------------------------------------------------T---E--R-------------R----------I----    226
O.CM.ANT70          NNIQ-E-K---------------KE--NVTV--E-REVQ----------------I--GL-----------AP-------------------SK---E--TA--Q---Q-----R----------I----    215
O.CM.MVP5180        NNIQ-E---T-------------KE-NNVTV-VQ-KEVQ------------L---I--GL-----------AP-------------------SR---E--TA--Q---Q-----R----------I----    215
O.SN.MP1299         TNIQ-E-K-T-------------KE-N-VPV--E-REVL----------------I--GL-----------AP-------------I-----SK---E--TA--Q---Q-----R----------I----    215
O.SN.MP1300         NNIQ-E-K-T-------------KE-N-VPV--E-REVL----------------I--GL-----------AP-------------I-----SK---E--TA--Q---Q-----R----------I----    215
CONSENSUS 01        -DIN--gk---k-----------r-----l-----Kk------------------M--------------iD-v--t-----------------------T-i-Ke--E----------------v----    219
01_AE.CF.90CF11697  -DIN-------------------------------KR------------------M---------------D----T--------------$-----$--T---K---E---------------------    217
01_AE.CF.90CF402    -DIN---K---------------------------K-------------------M---------------D----T-----------------------T---K---E---------------------    219
01_AE.CF.90CF4071   -DIN---K---------------------I-----K-------------------M---------------D----T-----------------------T---K---E---------------------    219
01_AE.TH.93TH057    -DIN---K---------------------------K-------------------M---------------D----T-----------------------T---K---E---------------------    219
01_AE.TH.93TH065    -DIN---K---------------------------K-------------------M--------------XD----T-----------------------T-T-K---E---------------------    217
01_AE.TH.93TH253    -DIN---K---------------------------K-------------------M---------------D----------------------------T---K---E---------------------    211
01_AE.TH.93TH902    -DIN---K---------------------T-----K-------------------M---------------D-I--T-----------------------T---K---E---------------------    219
01_AE.TH.94TH702    -DIN--EK---------------------T-----K-------------------M---------------D-I--T-----------------------T---K---E---------------------    219
01_AE.TH.95TNIH047  -DIN---K---------------------------K-------------------M---------------D----------------------------T---K---E----------------I----    219
01_AE.TH.CM240      -DIN---K---------------K-----------K-------------------M---------------D----T-----------------------T---K---E---------------------    224
01_AE.TH.TH022      -DIN---K---R-----------------------K-------------------M---------------D----------------------------T---KK--E---------------------    219
CONSENSUS 02        -eIn---K--------------v------li----K--i-t------------r-m--q----------------v---------k----------i-a-td--t---------k----------v-a--    218
02_AG.CM.MP807      --ID---K---------------------------K-------------------M--------------------------------------------TD--L-------------------------    219
02_AG.FR.DJ263      --IN---K---------------------------K-------------------M--------------------------------------------T--------------------------G--    219
02_AG.FR.DJ264      --IN---K---------------------------K-------------------M--------------------------------------------TD--A-------------------------    219
02_AG.GH.G829       -DIN---K---------------------P-----K-------------------M--------------------------------------------TD--M--------------------I----    219
02_AG.NG.IBNG       -DIN---K---------------------------K-------------------M--------------------------------------------TD----------------------------    219
02_AG.SE.SE7812     --IN---K---------------------------K-------------------M--------------------------------------------TD------------R----------I----    219
02_AG.SN.MP1211     -KIN---K--------------L------------K--M-S------------K-I-------------------L--------------------L-H-TD--A-------------------------    219
02_AG.SN.MP1213     -KIN---K--------------L-------M----K----S------------Q-I--P--------------------------R--------------I-----------------------------    219
03_AB.RU.KAL153-2   -DIN---K---------------------------K-------------------M---L----------------T-----------------------TD--K-------------------------    219
03_AB.RU.RU98001    -DIN---K---------------------------K-------------------M---L----------------T-----------------------ID--K----------------------V--    219
04_cpx.CY.94CY032-3 --IN---K---------------------P-----K-------------------M---L----------------------------------------T----D-------------------I----    218
04_cpx.GR.97PVCH    -DIN---K---------------K-----------K----------------------------------------------------------------T------------------------I----    218
04_cpx.GR.97PVMY    --IN-------------------------T-----K--T---------A------M---L----------------------------------------R-------Q------V---------I----    218
05_DF.BE.VI1310     -DIN---K----------------------V----Q----------------------------------------------------------------T----D--------R----------I----    219
05_DF.BE.VI961      ---N---K------------------------------V-------------------------------------------------------------T----D--------R----------I----    219
06_cpx.AU.BFP90     -DIN---K---------------K-----------K------------I------M--------------------------------------------T---K-------------------------    221
06_cpx.ML.95ML127   -DIN---K---------------------------KR-M----------------M--------------------------------------------T-------------R----------I----    218
06_cpx.ML.95ML84    -DIN---K---------------------------K--M----------------M------------------------------I-------------T------------------------I----    220
06_cpx.SN.97SE1078  --IN-------------------------P-----KR------------------M--------------------------------------------T------------------------I----    223
10_CD.BFL061        ---N---K-----M---------------------Y----------------------------------------------------------------T---K---------R----------I----    218
10_CD.BFL071        -D-----K----------------E----P-----Y----------------------------------------------------------------T----D--------R----------I----    219
10_CD.BFL110        ---N---K---------------------------Y----------------------------------------------------------------T----D--------R----------I----    216
11_cpx.CM.MP818     --IN-------------------K--EE-I---E-K-------------------M---------------D----Q--A--------------------T-----------------------------    220
11_cpx.FR.MP1298    --IE----------------------EE-----E-K-------------------M---------------D--------------------S-------K-----------------------------    218
11_cpx.FR.MP1307    -DID----------------V-----ED-T---E-K------------A------MM--------------D----------------------------T-----L-----------------------    234
11_cpx.GR.GR17      -----------------------K--ED-T---E-K-------I-----------M--------------VD----------------------------T-----------------------------    218
CPZ.CD.CPZANT       -GIH-Q-N----T--------S-Q--NKVP-Q-GDRTVLA---L--N--------V-CLL---------KV--------E-----R------SK---E--K---DKL-A-N---R---D------I----    219
CPZ.CM.CAM3         DNVQ-T-K-R-------------K---N-A---E-R-TT----------------I-----------------------------R-------A------T---Q---------R----------I----    224
CPZ.CM.CAM5         DNVQ-T-K-R-R-----------K---N-V---E-R-TT----------------I--------------V--------------R-------A------T---Q--------------------I----    220
CPZ.GA.CPZGAB       -RIQ-Q-L---------------K-F-NVH---E-R-VV----------------I---L----V----S----------------------SA------T---Q--------------------I----    217
CPZ.US.CPZUS        DNIQIE-T-R---M---------K---HVN---E-R--Q-S--------------I-------------S-K----R--------R-------A------T---Q--------TRV---------I----    221
```

HIV-1/SIVcpz proteins

```
                    D67N K70R                                                                                                                  polymerase motif
                      |  |                                                                                                                        |-  -|
B.FR.HXB2           KKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLE   349
CONSENSUS A         --d-tk------------------------p------k------v---------d-s----------t--e---i---y----------a---s----------skn-eii--------------   347
A.KE.Q23            ------------------------------------------------------H-E----------T------V------------------------------SK--E----------------   349
A.SE.SE6594         --N--R----------------------------------N----------------H------------T--A-------X--------------------------SK----I---------------   344
A.SE.SE7253         ------------------------------------------------------S----------T------V------------------------------LK--E-I--------------   349
A.SE.SE7535         ---------------------------------------------L---------H-G----------T--A----------------------------------ER--EVI--------------   347
A.SE.SE8538         ----NR-----------------------------------------------------S----------T----------------------S-------------SK--E-I--------------   349
A.SE.UGSE8131       --N-DR-----------------------------------------------------S----------T------------------------------------SK--E-I--------------   343
A.UG.92UG037        ------------------------------------------------------------S----------T--------------------------A---------SK-------------------   350
A.UG.U455           ------------------------------T------------------------S------------------V--------------S--------------S-H-------------------   348
A2.CD.97CDKFE4      ------------------------------------------------------------------------S------------------S--------------SKY-EM----------------   341
A2.CY.94CY017-41    ---------------------------------------A--------------H------------T------V------------------------------SK-TELI--------------   349
CONSENSUS B         ---s-------------kr------v-------a----kk-v-v-------------kd----------i-----gi-------------------s---k-----rk---d-v--------------   349
B.AU.MBC18          ---G---------------------------S-----R------------------N------------------------------------R------R---------------------   361
B.CN.RL42           ------------------------A------------------------------K------------V--------------------------C-------------------------------   349
B.DE.D31            ------------------------------------------------------------------------V-------------------------------------------------   349
B.GA.OYI            ------------------------------------------------------K-------------------------------------------------------------------   349
B.GB.CAM1           ------------------------------------------------------K------------T------------------------------------------------------   349
B.KR.WK             ------------------------------------------------------------------T------------------------------------------------------   349
B.NL.3202A21        ----------------------------------------G-A-------------K------------V--------------------------C-------------------------------   349
B.TW.TWCYS          ------------------------------------------------------K-------------------------------------------R-----------------------   349
B.US.JRCSF          -----------------R------------------------------------K-------------------------------------------------------------I--------------   353
B.US.JRFL           -----------------K------------------R-----------------K-------------------------------------------------------------I--------------   349
B.US.MN             -----------------K------------------------------------K-------------------------------------------------------------------------   352
B.US.RF             ------------------------------------------------------KE----------------R------------------------------K----E-----------------   348
B.US.SF2            ------------------------------------------------------K-------------------------------------------------------------------------   349
CONSENSUS C         ------------------------------------------------v---------d-g----------i--e-------------------a---s-mtk------aq--eiv-----------s---   345
C.BR.92BR025        --------------------$-------------------------------------G-----------------------------------S-----T--------A---E-I--------------   347
C.BW.96BW01B03      ------------------------------------------------------G-------------------------------------------------AL--E-------------P---   345
C.BW.96BW0402       ------------------------------------------------------S-------------S--------------------------I-------TK-------------------   345
C.BW.96BW0502       ---------------------------------------------M----------G-----------------------------------------------L---E-----------------   355
C.BW.96BW1104       ------------------------------------------------------S-----------------------------------S--------------AK--EL-------------   343
C.ET.ETH2220        ------------------------------------------------------G----------T----------------------P------PQ------AP--E-----------------   356
C.IN.93IN999        ------------------------------------------------------Y------------R--------------------------A----------A---E-----------------   354
C.IN.94IN11246      ------------------------------------------------------Y-----------------------------------------------GR--E-D--------------   345
C.IN.95IN21068      ------------------------------------------------------Y-----------------------------------------N---R------A---E-----------------   344
D.CD.84ZR085        ----------------------------------------I----------I--C-----------------------------------------I----------EV-------------   350
D.CD.ELI            ------------------------------------------------------------------------S-----------------------------------------EM-------------   348
D.CD.NDK            ------------------------------------------------------------------------------------------------------------E-----------------   348
D.CD.Z2Z6           ------------------------------------------------------K-----------------------------------------------------E-----------------   348
D.UG.94UG1141       ------------------------------------------------------H------------T-----------------------------------------EMI--------------   350
F1.BE.VI850         ----S--------K---------------------------------------K--K---------V--------------------------C----------MK-------------------   349
F1.BR.93BR020-1     ------------------------------------------------------K-------S----T------V------------------Y------D---AK-------------------   349
F1.FI.FIN9363       ------------------------------------------------------K------------V--------------------------C----------TR-------------------   348
F1.FR.MP411         ------------------------------------------------------KE----------L-----------------------------------AK-------------------   351
F2.CM.MP255         ------------------------------------------------------KE-----------------------------------------C----------AK--E------------   349
F2.CM.MP257         ---------------------------------------R-----------------KE-----------------------------------------I-------E--E------------   349
G.BE.DRCBL          -----R--------------------P------G---Q-R------------------N----------T-----------X..------------------------T---E-----------------   346
G.FI.HH8793-1-1     ------------------------------------RENQSA---------------S----------T-------------------------------------IK--EM--------------   348
G.NG.92NG083        ---------------------------------------R-----------------K-----------------------------------------S-TK--EM----------------   348
G.SE.SE6165         ------------------------------------------------------------------------V----------------------R------AN--EM----------------   346
```

```
                    D67N K70R                                                                                                                  polymerase motif
                      |  |                                                                                                                        |-  -|
B.FR.HXB2           KKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLE    349
H.BE.VI991          --N--R-----------------------------------S-----G-------H-------------T-----------------------------------------EVI---------------    347
H.BE.VI997          ----------------------------------------S--------------K-------------------------------------------------E-I---------------    347
H.CF.90CF056        ----------------------------------------S--------------KE-----------------------------------------------A---E---EM----------------    348
J.SE.SE7022         -------------------------------------------------------Y---------------------------------------C------K---ER--E-----------------    347
J.SE.SE7887         -------------------------------------------------------Y---------------------------------------C------K---ER--E-----------------    347
K.CD.EQTB11C        -------I------------P-----------------------------------K-----------------------------------------C----------RK---M-L---------------    349
K.CM.MP535          -------------------------------------------------------K-------------------V-------------------H----------IK--EM----------------    349
N.CM.YBF106         --------------------E---------------Q------X--------C---K-------------------------------------------------KH-E-I---------------    356
N.CM.YBF30          ------------------------------------Q---------------C---K-----------------------------------------T---------EKH-E-I---------------    356
O.CM.ANT70          ---G-----------------E-----------G---Q-Q-------------C---P------------V-------------------------------D---RD--ELE-C------------P    345
O.CM.MVP5180        -------------------------------G---QRQ-------------C---P------------V------V--------------------------D----S--EVE----I---------P    345
O.SN.MP1299         -------------------------------G---Q-Q-------------C---P------------V-------------------------------D----D--ELE-C------------P    345
O.SN.MP1300         -------------------------------G---Q-Q-------------C---P------------V-------------------------------D----N--ELE-C------------P    345
CONSENSUS 01        k-d--k-----d------r-q--------i-h-----------------------d-sf------------n----i-----------k-------s-m------frik--Em-----m----------    349
01_AE.CF.90CF11697  E---------------------------V-------------------------Y-G-------------------------------------------------Y-AK--E-----------------    347
01_AE.CF.90CF402    -------------------------------------------------------G------------------V-------------------------------AR--E-----------------    349
01_AE.CF.90CF4071   -------------------------------------------------------H-SX---------------------------R-------C----------AK--E-----------------    349
01_AE.TH.93TH057    ------------------G-------------------------------------S-------------D-----------------------------------IK--EM----------------    349
01_AE.TH.93TH065    ------$----G-------------------R-------------------------S-----------------------------------------------IK--EM----------------    346
01_AE.TH.93TH253    -------------------------------------------------------S-----------------------------------------------IK--EM----------------    341
01_AE.TH.93TH902    ------------------G------------------------------------S-----------------------------------------------IK--EM----------------    349
01_AE.TH.94TH702    ------------------G-S----------------------------------S-----------------------------------------GIK--EM----------------    349
01_AE.TH.95TNIH047  --N--R-----------------------------------------------S-----------------------------------------------IK--EM----------------    349
01_AE.TH.CM240      -------------------------------------------------------S-----------------------------------------------IK--EM-----K----------    354
01_AE.TH.TH022      -----R------------------------R-------------------------S-------------------------------------C-T--------TK--E-----------------    349
CONSENSUS 02        kkd--k---------------df--v--g----------k-----------------kd-----------v-n--------------------a---a--tk---pf-tK--Eiv--------------e    348
02_AG.CM.MP807      ---------------------------------------R------------------G------------V---------------------------A----------TK--E---------------K    349
02_AG.FR.DJ263      -R--------------------S----------------------------------K------------T---------------------------A---N---HY-IK--E-M----------------    349
02_AG.FR.DJ264      --------------------G------R---------------------------K------------V---------------------------A----------IK--E-----------------    349
02_AG.GH.G829       -------------------------------------------------------K-------------------------------------------Y-IK--E-----------------    349
02_AG.NG.IBNG       -------------------------------------------------------K------------V---------------------------A----------TK--E-----------------    349
02_AG.SE.SE7812     --N--R-------------------I-----------------------------K-------------------------------------------A----------TK--E-----------------    349
02_AG.SN.MP1211     R------------------------------------------------------------------V---------------------------A----------TK--EL----------------    349
02_AG.SN.MP1213     ------------------------------------E-----------------K-------------S--------------------S------I-------AK--E-----------------    349
03_AB.RU.KAL153-2   ------------G------------------------------------------Q------------T-------------------------------------E-----------------    349
03_AB.RU.RU98001    ----------------------------L--------------------------Q------------T-------------------------------------E-----------------    349
04_cpx.CY.94CY032-3 -------------------------------------------------------PE-----------T---------------------------C----------FK--E-----------------    348
04_cpx.GR.97PVCH    --N--R-------------------------------------------------PA-----------T------V--------------------Y----------TR--E-----------------    348
04_cpx.GR.97PVMY    --N-NR-------------------------------------------------PE-------------------------------------C----------TK--EM----------------    348
05_DF.BE.VI1310     -----R-----------------------------------I---------------KE-------------------------------------C----------------EV----------------    349
05_DF.BE.VI961      -----R-----N------K-----------------------------------H-------------------F-------------------C----------------EM----------------    349
06_cpx.AU.BFP90     -------------------------------------------------I---K-----------T---------------------------I-------IK--E-----------------    351
06_cpx.ML.95ML127   -------------------------------------------------I---K-----------T---------------------------C------D---IK--EL----------------    348
06_cpx.ML.95ML84    ------------------------------------R-----------------G-N-----------L---------------------------I-------TK--E-----------------    350
06_cpx.SN.97SE1078  -------------------------------------------------------K-------------------------------------------I-------QK--E-----------------    353
10_CD.BFL061        -------------------------------------------------------Y-----------------------------------------------------QM---------------Q    348
10_CD.BFL071        -----------------R--------R-------------------------G-----------------------------------------------------EM----------------    349
10_CD.BFL110        -------------------------------------------------------Y-------------------T-----------------------------EM----------------    346
11_cpx.CM.MP818     -------------------------------------------------------N-------------------V-----------------------------T---E-----------------    350
11_cpx.FR.MP1298    -------------------------------------------------------KE-----------T---------------------------C----------------E-----------------    348
11_cpx.FR.MP1307    --N----------------------------------------------------S-----------------------------------------------T---------------M-----    364
11_cpx.GR.GR17      ------------------------I----------------------------L---S-----------L-----------------------------T---EV----------------    348
CPZ.CD.CPZANT       ---TS--------------------I-----------Q---------------I---Q-----------V---------C----------------A------A---DKY-AVE-------------M-    349
CPZ.CM.CAM3         -------------------------------------------------C----N-----------V----------------------------I---T---Q-H--LI----------------    354
CPZ.CM.CAM5         --------------------E-----------------------------C----N-----------V------V-------X------------------D-----H--LI----------------    350
CPZ.GA.CPZGAB       -------------------------------------------------C---K-------------------V---------------S--------------EK----T----------------    347
CPZ.US.CPZUS        ---------------------------------------R-------------C---KE-----------V-----------------------------D-----H--VI--------------N    351
```

HIV-1/SIVcpz
proteins

HIV-1/SIVcpz protein alignment: POL

```
                                   T215Y K219Q
                                      |    |
B.FR.HXB2          IGQHRTKIEELRQHLLRWGLTTPD.KKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAELELAENREILKEPVHGVYYDPSK    478
CONSENSUS A        -g---t------a---s--ft---.k--------------l---------e----e--------------------a----k---k----a----DiVt----------------d----v------    476
A.KE.Q23           -----A------A---S---I---.--------------------------E------------------------A----K--------A----D-VT------------------------------    478
A.SE.SE6594        -----A------S---S--F----.Q-----------------------------E------------------------K---R----A----DIVE----------------D-----------    473
A.SE.SE7253        ------------A---S--F----.--------------------------E----E--------------------A----K---R----A----DIVT----------------D-----------    478
A.SE.SE7535        ------------E---S--F----.---------------V---------------E--------------------A-------------A----DIVT----------------D----A------    476
A.SE.SE8538        ------------A---S-------.-------------------------N----E--------------------A----K--------A----DIVT----------------A-----------    478
A.SE.UGSE8131      -E----------A---S--FF---.Q-------------------------K----E--------------------A----K--------A----DIVT----------------D-----------    472
A.UG.92UG037       ------------E---K--F----.-------------------------E----E--------------------A----K-------------DIVT----------------D----A------    479
A.UG.U455          -----A------A---S--FI---.-------------------------Q--------------------------A----K--------A----DIVT----------------D-----------    477
A2.CD.97CDKFE4     -----A------A------F----.-------------------------K--------------------------A----K---R---------DIVT--K------E-------N-----------    470
A2.CY.94CY017-41   -S---V------A---K--FY---.-------------------------K--------------------------A----K--------A----DIVT--K------E-------T-----------    478
CONSENSUS B        -----t-i----q---r--f----?----------------p--------v--------------------------a----kq-------t------ip--e-------------kep--gv----s-    478
B.AU.MBC18         -----I----------K--F----.------------------------------------------------------------------------------------R------------    490
B.CN.RL42          -----A-------------F----.-----------------L--------M--------------------------A----KE-------------------------------------S----------    478
B.DE.D31           -------------------F----.-------------------------T--------------------------A-------------A----------K--------------------------    478
B.GA.OYI           -------------------F----.-------------------------M--------------------------A----KN-------------------------------------------    478
B.GB.CAM1          ------------------------.-------------------------M--------------------------A----K---------------V-----------------------------    478
B.KR.WK            -----I-------------F----.-------------------------------------------------------A----K---------------V----------------------------A-    478
B.NL.3202A21       -------------------F----.-------------------------------------------------------A-----H------------------------------------------    478
B.TW.TWCYS         -----I-V-----------F----X-------------------------------------------------------A-------------------V---K-------------Q----A------    479
B.US.JRCSF         ----------------K--F----.-------------------------------------------------------A----K-------------------K--------------------------    482
B.US.JRFL          -----A-------------F----.-------------------------------------------------------A----K-------------------------------------------    478
B.US.MN            -----A------R------F----.-------------------------------------------------------A----K-------------------------------------------    481
B.US.RF            -----I------E---K--F----.-------------------------------------------------------A----K---------------VQ--K-----------------------    477
B.US.SF2           -------------------F----.-------------------------M--------------------------A----K-------------------------------------E-------    478
CONSENSUS C        ---hrAk---lre---k--f---d.------------------------p-q--e-d----n------------------ik-rq--k----a-a--DiVpl------------------v---y-----    474
C.BR.92BR025       -----A------E---K--F----.-------------------------Q--------------------------------------A----DIV-----------------------------    476
C.BW.96BW01B03     ----TA----SKE---K--F----.-------------------------N----E-------------------------------A----DIV-----------------------------    474
C.BW.96BW0402      -----A------E---K------Y.-------------------------Q--D------H------------------R-KH-------A----DIV-----------------------------    474
C.BW.96BW0502      ---R-AQ-----E---K--F----.-------------------------Q--D-----------------------------------A----D-V-----------------------------    484
C.BW.96BW1104      -----AR-----K------F----.------------------------S-K----E-----------------------V-----------A----DIV-P---------------------------    472
C.ET.ETH2220       -----AP-----E---K--F----.-------------------------Q--------------------------------------A----DIVT----------------------F-----    485
C.IN.93IN999       -----A-------------F----.-------------------------Q------------------------------K---R----A-V--DIV-----------------------------    483
C.IN.94IN11246     -----A------E---K--F----.-------------------------Q--------------------------------------A----DIV-------------------I---------    474
C.IN.95IN21068     -----A------K------F----.-------------------------Q--------------------------------------DIV-----------------------------------    473
D.CD.84ZR085       -----A---K--E----------.------------------------S-T----E--------------------------K-----------------------------------M---------    479
D.CD.ELI           ---------K--E------F-R--.------------------------S-K----E--------N--ER-------------------------------------------------------    477
D.CD.NDK           ------------E------F----.-------------------------N----E--------------------A----K---------------V-----------------------------    477
D.CD.Z2Z6          ------------E------F----.------------------------S-K----E-------------------------------------------------------------------    477
D.UG.94UG1141      -----I------G---K--F----.--Y----------------------H---E-----------------------------C---A----------A-------------------A------    479
F1.BE.VI850        ------------E------F----.-------------H-----------Q--N----------------------------P-------A----DIV---A-------K-----R------------    478
F1.BR.93BR020-1    ------------E---K-------.-------------------------Q--D-----------------------------K--------A----DIV---T-------------------A------    478
F1.FI.FIN9363      ------------E---K--F----.-------------------------Q--D------D------------X----------------A----DMV---A--N----------------------    477
F1.FR.MP411        L----M------E---K--F----.-------------------------Q-----------------------------IK--------A----DIV-----------------------------    480
F2.CM.MP255        ------------E---K--F----.-------------------------Q----S-----------------------RIKH--R----A----D-V---A------------I-------------    478
F2.CM.MP257        -----A------E------F----.-----------------------A-Q--D-S-----------------------R-KH------------D-V---A-----------------------------    478
G.BE.DRCBL         -----A------E------F----.-------------------------Q--N-EN-----------------------K-----I--A----DIVSM-A---M----------------------    475
G.FI.HH8793-1-1    -----A--K---E------F----.-------------------------Q----E--------------------------------------A----DIV---A-----------------------    477
G.NG.92NG083       -----A------E---K-------.-------------------------Q----ED----------------------KH--R----A----DIV---A---M----------------H----    477
G.SE.SE6165        -----A------E---K--F----.-------------------------Q--D-E----------------------TH-------A----DIVS--A---M---------R------------    475
```

```
                                   T215Y K219Q
                                      |    |
B.FR.HXB2           IGQHRTKIEELRQHLLRWGLTTPD.KKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAELELAENREILKEPVHGVYYDPSK    478
H.BE.VI991          -----E------A------F----.Q-----------------------VK-----------------------------K---X----A-----IV---K-------------------A------    476
H.BE.VI997          -----A------A------F----.-----------------------VK-----------------------------------A----D-V---K-------------R------------    476
H.CF.90CF056        -----A------A---K--F----.-----------------------TVK--------------------------N---K--------A----DI----K-------------R--I----------    477
J.SE.SE7022         -E---R--K---E---K--FY---.-------------------------Q----ED------------------------IKE----I--A----DIV---R---------K--------------AR    476
J.SE.SE7887         -E---R--K---E---K--F----.-------------------------Q----ED------------------------K------K-A----DIV---R---------K-------------SA-    476
K.CD.EQTB11C        -----A------E------F----.-------------------------Q--D----------------------F-----K--------V----DIV---A--------------------------    478
K.CM.MP535          ---P--------E---K--F----.-------------------------Q--D----------------------------K--------V----DIV---A--------------------------    478
N.CM.YBF106         -A---ETV----G---K--F----.-------------------------K----EV----------------------K-----I---------VTF-Q---------------L-------G-    485
N.CM.YBF30          LA---EAV-D--D---K--F----.-------------------------K-----V----------------------R-K-----I---------VNF-----------------L-------G-    485
O.CM.ANT70          LTE--KR--L--E--YQ--F----.------------------------S-Q--N--V---------I----------Q--R--E----I----S----V--SR------E----R--Q-------Q-D-    474
O.CM.MVP5180        LAE--KRV-L--E--YQ--F----.-------------------------Q--D-EV--------------------Q--R-KE----I----S----V--SK------E----K----------Q-D-    474
O.SN.MP1299         LTE--KRV-L--E--YQ--F----.-------------------------Q--N-EE--------------------Q--R-KE----I--A-S---IV--SK------E----K----------Q-D-    474
O.SN.MP1300         LTE--KRV-L--E--YQ--F----.-------------------------Q--N-EE--------------------Q--R-KE----I----S----V--SK------E----K----------Q-D-    474
CONSENSUS 01        i-q-rt-i-e--a---s--F---d.k-----p-----------R------e--e-d----n----------------a----k----l---a----diVp---------a------kt-----------    478
01_AE.CF.90CF11697  ------------D---K--F----.------------------R------E----------------------------A----K--------A------V-----------------N-----------    476
01_AE.CF.90CF402    -------V-D--A---S--F----.------------------R------------------------------------A----DIVT-----------------T-----------    478
01_AE.CF.90CF4071   ----G-------A---S--F----.------S-----------R------Q--D-E----D---------------A----K--------A----DIV-----------------N-----------    478
01_AE.TH.93TH057    ------------A---S--F----.------------------R------E----------------------------A----K----P--------DIV-----------------T-----------    478
01_AE.TH.93TH065    -----I------A---S--F----.------------------R------E----------------------------K--------A----DIV-----------------RT----------    475
01_AE.TH.93TH253    ------------A---S--F----.------------------R------E----------------------------A----K-------------DIV----------E------RI----------    470
01_AE.TH.93TH902    T-P--A------A---S--F----.------------------R------E----------------------------A----K--------A----D-V-----------------RT----------    478
01_AE.TH.94TH702    T----A------A---S--F----.------------------R------E----------------------------A----K--------A----D-V-----------------RT----------    478
01_AE.TH.95TNIH047  ------------A---S--F----.Q-----------------R------E----------------------------A----K--------A----DIV-----------------T-----------    478
01_AE.TH.CM240      ------------A---S--F----.------------------R------E----------------------------A----K--------A----DIV-----------------T-----------    483
01_AE.TH.TH022      ------------A---S--F---G.------------------R------E----------------------------K--------A----DTV-----------------T-----------    478
CONSENSUS 02        -g---A------e---?--F----.k-----------------------iq--e---------------l--asq-ya---vKq--k----a-a--DIVt---------------------------t-    476
02_AG.CM.MP807      -----A------E------F----.-------------------------Q----------------------------A----K--------A----DIVT---------------------------A-    478
02_AG.FR.DJ263      -E---A------E---K--F----.-------------------------Q----------------------------A----K--------A----DIV----------------------------A-    478
02_AG.FR.DJ264      -----A------E---K--F----.-------------------------Q----------------------FA----K---R----A----DIV----------------------------A-    478
02_AG.GH.G829       -----A------E------F----.-------------------------M----------------------------A----KE-------A----DIVT---------------------------T-    478
02_AG.NG.IBNG       -----A------G---K--F----.-------------------------VE---------------------------A---IK---R----A----DIVA---------------------------T-    478
02_AG.SE.SE7812     -----A------E------F----.E-----------------------------------------------V--QVR--A----K--------A-T--DIVT---------------------------T-    478
02_AG.SN.MP1211     -----A------E---K--F----.-------------------------Q----------------------------K-------------DIV----------------------------T-    478
02_AG.SN.MP1213     -----A------E------F----.-------------------------Q--D----------------------A----K--------A----DIVT---------------------------T-    478
03_AB.RU.KAL153-2   ------E-----E------F----.-----------------------------------------------------A-------------A----------A--------------------    478
03_AB.RU.RU98001    -----I------D------F----.-----------------------------------------------------A-------------A-----I----A--------------V-----------    478
04_cpx.CY.94CY032-3 -----A------E------F----.-------------------------QPA----------------------------K--------A----DIV---T-------------------A------    477
04_cpx.GR.97PVCH    ------------E------FY---.-------------------------Q----------------------------K-------------DIV---T-------------------A------    477
04_cpx.GR.97PVMY    ----KA------E------FF---.Q------------------------Q-V--E--------------------A----K--------A----DIV---T-------------------A------    477
05_DF.BE.VI1310     -----A--K---E---A--F----.-------------------------Q--D----------------------------K-----------------------------------------    478
05_DF.BE.VI961      ------------E---K--F----.-------------------------Q--N----------------------------K---------------V----------------------------A-    478
06_cpx.AU.BFP90     -----A------E---K-------.-------------------------Q--D-E----------------------------------A----DIV---A--------------------------    480
06_cpx.ML.95ML127   -----A------E---K--F----.-------------------------Q--N-E--SI-----------------S-------------A----DIV---A-------------------A------    477
06_cpx.ML.95ML84    -----A------E---K--F----.-------------------------Q--D----------------------------KH------A----DIV---A-----------------------L-    479
06_cpx.SN.97SE1078  -----E-V----A---K--F----.-------------------------Q--D----------------------------K--------A----DIV---A--------------------------    482
10_CD.BFL061        --H-TI------G---K--F-A--.---------------------L---Q--K-N----------------------------------A-V--DIV----------------------------    477
10_CD.BFL071        -----A---D--G---K--F----.-----K-------------------Q----E----------------------------------DIV----------------------------    478
10_CD.BFL110        -----I------E---K--F----.-----------------------S-Q----------------------------R---------A----DIV----------------------------    475
11_cpx.CM.MP818     -----E-V----K-----------.-------------------------Q--D-EC---------------------------K-------------DIV---A--------------------------    479
11_cpx.FR.MP1298    -----E------K---K--F----.-----------------------T-Q--D-E--------------------S---IK-------------DIV---A-------------------A------    477
11_cpx.FR.MP1307    -----E-V----E--WK--FY---.-------------------------Q--D-EC---------------------------K-------------DIV---A-----------------------A-    493
11_cpx.GR.GR17      -----E-V----K---K--F----.-------------------------Q--D-E--------------------A----K--------A----DIV---A--------------------------    477
CPZ.CD.CPZANT       -TA--EM--K-----QV---E---.---------Q-------------K-K---P-D-----------------------TK---R-I--V-S--DRVQM-R------E--KQ--QQKIE-Y--Q-GL    478
CPZ.CM.CAM3         LRE--E-V-L------T--F----.-------------------------Q--Q-ET---------------------K-----IK------DIVTM-Q---M--E--X---XD---------E-    483
CPZ.CM.CAM5         --K--E-V-L------T--F----.-------------------------Q--Q-EI-------------------S----K-----IK--------VAM-Q---M--E-------D---------E-    479
CPZ.GA.CPZGAB       -D---K-V--------K--F----.-------------------------Q----EV---------I--------------IK-----I----K--D-V---P------------VST---------D-    476
CPZ.US.CPZUS        LEK--E-V-L---Y--T--F----.----E--------------------Q--Q-EI-----------------------IK-----IK-A------VNF-H---M--E-----------------E-    480
```

HIV-1/SIVcpz
proteins

```
                                                                                                                                                        p51 \/ Rnase H
B.FR.HXB2            DLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKT.PKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKL   607
CONSENSUS A          --ia------qD--t---y--p-----------KrSa-----k--a-v---Vvm--------t.-kfkl----e--et--md-------p-----------kl--q--kD--a-aet-y-d---nre-kl   605
A.KE.Q23             -----------D------F--------------K-S------R--A-V---VVM---------.---R--------D---MD-------L-------------------D--A-----------------   607
A.SE.SE6594          -----------D---------------------K-SV--------A-M---VV----------.----------------MD-------------------LW--L--ED-----DS-F-E---KEK-LX   602
A.SE.SE7253          ----------LD---------------------K-S------R--A-V---VVM---------.----------------MD---------------------------D--A-----------------   607
A.SE.SE7535          ---V-------D---------T-----------K-S---------A-V---VAM---------.-RS-------------ID---------------------------D--------------------   605
A.SE.SE8538          -----------D---------------------K-S---------A-V---VVM---------.----V-----------M----------------------------D--A-----------------   607
A.SE.UGSE8131        -----------D---------------------KKS------R--A-V---VVM---------.---------D--D---MD---------------------------D--A-V---------------   601
A.UG.92UG037         -----------D--I------------------K-S---------A-V---VVM---------.----------------MD-----$---------------------D--A-----------------   607
A.UG.U455            --V--------D---------------------K-S-----------V---VS---------I.---R---------A--M----------------------------D--A-----------------   606
A2.CD.97CDKFE4       -----------D--------------------KKKST---------------AI---------.---R----------------------------------------T--LA-----------------   599
A2.CY.94CY017-41     -----------D--------------------KR-ST----I-----------M---------.----------------A-------------------------------A-----------------   607
CONSENSUS B          d-ia-i----q-------y-epfkn-------rm--a--------------iat--------t.---k--------ea--te-w---------fv--p----------ke--vg--------------kl   607
B.AU.MBC18           -----V----E---------DQ-R----------------------------A----------.---R--------DA----------------------------------T-----------------   619
B.CN.RL42            ----------L-----------Y-----------------------------A----------.-------------A----------------I--S--------------EE----------------   607
B.DE.D31             ---------------------------------------------------------------.-------------A------------------------------T---------------------   607
B.GA.OYI             --V--L-----------------------------------------------Q---------.-------------A-------------------------------D--------------------   607
B.GB.CAM1            -----L--------------------------K---T---------------A----------.------------DA--ID----------------------------------------------R-   607
B.KR.WK              ---------------------------------------------------VAI---------.-------------A----------------------------------------------------   607
B.NL.3202A21         E--------------------------------------------------------------.-------------A----------------------------------------------------   607
B.TW.TWCYS           --V------------------T-----------T------------------A----------.-------------A----------------------------------I-----------------   608
B.US.JRCSF           ---V--------------F--------------T------------------AN--------I.------------------------------------------------------------------   611
B.US.JRFL            -----L------------------I--------T------------------AN--------I.------------------------------------------------------------------   607
B.US.MN              -----V----------------------------------------------A----------.---R---------------T$--------V-----------------------------------K   609
B.US.RF              ---------------------------------------------------VA----------.-------------A----------------------------------I-----------------   606
B.US.SF2             --V------------------------------------------------VS---------I.-------------A--M-------------------------------------------------   607
CONSENSUS C          -----i----qd---y-i------------y-KmrT------k--tea----am---v----t.---R-----et--t---D------------v----------q--ke-ia-a-----------et-i   603
C.BR.92BR025         -----------N--------------------K--T------R---------AL---I-----.---R---------A---D-----$------------------------A--------------I-M   604
C.BW.96BW01B03       ----------HD--------------------K--T----------------A---------X.---R-------------D------------A----------------------------------I   603
C.BW.96BW0402        -----V----HD--------------------K--T-----------V-----M---------.---R-----D-------D------------------------------A-----------------   603
C.BW.96BW0502        ----------HD--------------------K--T----------------AQ---------.---R-------------D------------------------------P-V---------------   613
C.BW.96BW1104        ----------GD--------------------K--T----------------SM--------I.---R------A--A---D--------------------------T--MA----------------I   601
C.ET.ETH2220         ----------ND---F-F------------F-KRGT----------AV----AL---------.---R---------A---D------------------------------A-V--------------I   614
C.IN.93IN999         --------------------------------K--T---------A------AM---------.---R---------A---D-----------------------R------A-V--------------I   612
C.IN.94IN11246       -----------D-----V--------------K--T----------------AM---------.---R-------------D---------------------------D--A-------------D--I   603
C.IN.95IN21068       -----------D--------------------K--T----------------A----------.---R-------------D------------------------------A----------------I   602
D.CD.84ZR085         -----L----------------------------------------------AI---------.---R------------ID------------------------------I------------K----   608
D.CD.ELI             ----------H----------------------------------A----R-S--------R-.---R------------A-------------------------------I-----------------   606
D.CD.NDK             -----L----D----------------------T------------------A----------.----------------I-------------------------------I-----------------   606
D.CD.Z2Z6            ----------H----------------------------------A-V----S----------.---R------------V-------------------------------I-----------------   606
D.UG.94UG1141        -----------D---------QY---------K---T---------------AQ-C-------.---R-------------------------Y-----------------------------------I   608
F1.BE.VI850          ----------D---------N-----------KV-S----------------AL--------RS-------L----D----D--------------------------T---A--D-------S-----K   608
F1.BR.93BR020-1      --------------------------------K--S----------------SL---------.---R---L----D-------------------------------T--------------S-----K   607
F1.FI.FIN9363        ---PKL-------------R------------K--S----------------AL---------.-------L----D-------------------------------T---A----------S-----K   606
F1.FR.MP411          -----L--------------------------KT-S-----I-E--D----VAR-C------S.---R---------A---D--------------------------T---I----------S-----K   609
F2.CM.MP255          ----------HD----------Y---------KRKS-----------V----A---------I.---R---------I------------------------------T---A-----------------   607
F2.CM.MP257          ----------HD----------H----------RKS-----------V---VA--G------V.---R---------I------------------------------T---I-----------------   607
G.BE.DRCBL           -----V----L------V----Y---------KGGS-----------V----A--G-I----I.-------K-----V------------------------------T---P-V--Y------------   604
G.FI.HH8793-1-1      E----V---E------------Y---------TRGT-----------V----A----------.-------R-----V------------D--------------R--T---P----Y------------   606
G.NG.92NG083         E----V----PD----------Y---------KRGS-----------V----A--G------I.-------R-----V--------A---------------------T---P----Y------------   606
G.SE.SE6165          E----V----LD----------Y---------KRGS-----------V----A----------.-------R-----I---D-----------------------R--T---P----Y------------   604
```

```
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
p51 \/ Rnase H
B.FR.HXB2          DLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKT.PKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKL    607
H.BE.VI991         E---------PD--------------------K--S-----------V----A---------I.---R--------------H---------------H---------T---E----Y-----------M    605
H.BE.VI997         ----------PD--------------------K--N---S------------A---------I.-----------------H--------$------H---------A---A----Y----------I    604
H.CF.90CF056       ------R-----------------------K--T-----I----------S---------I.---R-----------------------------H---------T---A----Y-I----------    606
J.SE.SE7022        E----V----LD--------------------KR-S---------SQV----AL-A-------.---R-------------D-----------------------------M----------S-----V    605
J.SE.SE7887        E----V----LD--------------------KR-S---------A-V----AL-A-------.---R----R--------D-----------------------------M----------S-----T    605
K.CD.EQTB11C       ----------H-----------Y----------I-S-----------V---VAM---------.---R--------G-----------------------------T--------------------Q    607
K.CM.MP535         ----------ND----------H------------S---------------A--G-------.---R-----------------------------------T---------------H----K    607
N.CM.YBF106        E---------------------Y---------KX-S-----I-E-AAV---VA----------.------V---V-------H-----------------------T---S----Y-------K----    614
N.CM.YBF30         E-V-----------------LH---------K--S-----I---V-V-R-VA----------.---R--V---V--A---DH-----------------------T---S-----------------    614
O.CM.ANT70         --WVN-----GE---------EH--------T-QKAS----IR--A-VI--VSQ---I----L.------VTR-------AD-------------S----I----R--S---M----Y------------    603
O.CM.MVP5180       --WVS---H-E------V--DEH----------QKAS----IR--A-V---VSQ-A------L.---R--VTR-------A--------------S----I-------T-----------------N---    603
O.SN.MP1299        --WVN-----W---------DEH--------T-QKAS----IR--A-VL-RVSQ-A-I----L.-------TR-------AD-------------S----I-------S---M----Y--------D---    603
O.SN.MP1300        --WVN-----K---------DEH--------TKQKAS----IR--A-VL--VSQ-A-I----L.-------TR-------AD-------------S-A--I-------S---I----Y--------D---    603
CONSENSUS 01       d-v--v----qd------------------a-krS------R--t-V---ia----------.---r----r-------m----------------------------d--v----f------S--t-l    607
01_AE.CF.90CF11697 E-V--------D--------------------K-S------R--A-V---VA----------.---R----R-------A----------------------------D--A----L------S--N--    605
01_AE.CF.90CF402   --V--------D--------------------K-S------R--A-V---VA----------.---R----R-------M----------------------------D--M----------S----Q    607
01_AE.CF.90CF4071  --V-------HN--------------------Q-S------R--V-V---VS----------.---R----R-------M----------------------------D--A----------S-----    607
01_AE.TH.93TH057   --V--V-----D--------------------K-S------R----V----A----------.---R----R-------M----------------------------D--------------S-----    607
01_AE.TH.93TH065   --V--V-----D--------------------K-S------R----V----A----------.---R----R-------M----------------------------D--------------S-----    604
01_AE.TH.93TH253   --V--V-----D-------------------S-K-S------R----V----A----------.---R----R-------M----------------------------D--------------S-----    599
01_AE.TH.93TH902   --V--V-----D--------------------K-S------R----V----A----------.---R----R-------M----------------------------D--------------S-----    607
01_AE.TH.94TH702   --V--V-----D--------------------K-S------R----V----A----------.---R----R-------M----------------------------D--------------S-----    607
01_AE.TH.95TNIH047 --A--V-----D--------------------N-S------R----V----A----------.---R------------M----------------------------I----------S-----    607
01_AE.TH.CM240     --V--V-----D--------------------RGS------R----V---VA----------.---R----R-------M----------------------------D--------------S-----    612
01_AE.TH.TH022     --V--V-----D--------------------K-S------R----V----A----------.---------------M----------------------------D--------------S-----    607
CONSENSUS 02       --i--i-----D---y---------------ak?rsahtndvkq-t-V---vat---vi----.---r----r----a--M---------e------p-----------d--v------------r---l    604
02_AG.CM.MP807     --V--------D--------------------KR-S-----------V---VA----I-----.---R---------A--M--------------------------------------------    607
02_AG.FR.DJ263     -----------D--------------------KR-S---------A-V---VV----------.---R----R----A--M---------D------------------D--------------------    607
02_AG.FR.DJ264     -----------D--------------------KR-S-----------V---VA----------.---S----R----A--M---------D------------------D---------------K----    607
02_AG.GH.G829      -----------D---F---------------QERVAPYDLSITE---V---V-----I-----.---R----R----A--M----------------L-----------D--------------------    607
02_AG.NG.IBNG      --V--L-----D--------------------KK-S-----------V---VAM---------.---R----R-------M----------------------------D-------------------I    607
02_AG.SE.SE7812    --V--------D--------------------KK-S-----------V---VA----------.---R----R----A--M----------------------------D-------------------I    607
02_AG.SN.MP1211    -----------D--------------------KK-S-----------V---VA-----V----.---R----R----A--M----------------------------D--I----------------    607
02_AG.SN.MP1213    -----------D--------------------KK-S------R--A-V---AAM---------.---R----R-------M----------------------------D--A----------------    607
03_AB.RU.KAL153-2  --V-----------------------------L------------------A----------.-----------------------------------------------------------S    607
03_AB.RU.RU98001   --V-----------------------------L-S----------------A----------.----------------A----------------------------------------S    607
04_cpx.CY.94CY032-3 ----------------------H---------T-S------R---------AM-C-------.---R--------D-----------------------------TD--A----------------Q    606
04_cpx.GR.97PVCH   ----K-----------------Y---------KT-S------R----V----AM-C-------.---R--------D--------------------------$----PD--A----------S---RR    605
04_cpx.GR.97PVMY   ----------L--X--------Y---------KT-S------R---------AM-C-------.---R--------D---M----------------------T---A----------S---NQ    606
05_DF.BE.VI1310    -----V----L----------------------A--V--------A------A----------.---R---------I--------------------------------L-----------------    607
05_DF.BE.VI961     -----------E---------------------T-N---------A------A----------.---R-----------------------------------------M-----------------    607
06_cpx.AU.BFP90    -----V----Y-----------H----------IKS----------------AL---------.---R-----------------------------------------T--------------------Q    609
06_cpx.ML.95ML127  ----------------------H---------KIKST---------------AR---------.---R-----------------------------------------T-------------------K    606
06_cpx.ML.95ML84   -----L----------------H------------S-----I----------AL-A------I.---R-----------------------------------------T----------------K---K    608
06_cpx.SN.97SE1078 --------------------DTH----------TKS------R---------AL---------.---R---H-------------------------------------T---I----------------I    611
10_CD.BFL061       -----------D----------H---------KR-T----------------AQ-C-------.--------------I---D--------D-----------------------------------    606
10_CD.BFL071       ---V------HD----------H---------KR-T----------------AM---------.---R--------------------------------R------------V-----------D---    607
10_CD.BFL110       -----------D----------H---------K--T-----------V----AQ---------.---R-----D-------D-----------------------------A-----------------    604
11_cpx.CM.MP818    -----V----LD--------------------KR-A------R----VI---CL-A-------.---R----R--------D---------------------------QG--L----------------I    608
11_cpx.FR.MP1298   EI---V----LD--------------------KR-T------R--A-I----AL---------.---R----------D-----------------------------I----------------V    606
11_cpx.FR.MP1307   -----V----LD------F-------------KR-T------R--A-V----A----------.---R----R--------D----------------------H------I----------------M    622
11_cpx.GR.GR17     -----V----LD--------------------KR-T------R--A-V----AL-G------I.---R----------D-----------------------------I-----------------    606
CPZ.CD.CPZANT      P-K-T-----S---------NEG-L--A-----PT-T---E-R--AGV----GL-------EV.---Q---TR---DA--SD------------------IR---N-LAD--PE------------NSQ-    607
CPZ.CM.CAM3        E----X----NS------F--QH----------Q-S-----IR--A-V----A----------.---R--V---V-----S---------D-X----------------T---P----------------K    612
CPZ.CM.CAM5        E----V----NS------F--QH----------Q-S-----IR--A-V----A----------.---R--V---V-----S---------D-----------------P---P----------------K    608
CPZ.GA.CPZGAB      E---------NC------F---H----------Q-S-----IR--A------A----------.---R--V---S--A--A-------------I----------S--T---PTTD-Y----------T    605
CPZ.US.CPZUS       E-V--V----RS------F--RH----------Q-S-----IR--V-V----A---------V.-----LV---V-----S---------D-------------K--T-A-E----------Q----K    609
```

HIV-1/SIVcpz proteins

```
                                                                                                                      p66, Rnase H \/ Integrase
B.FR.HXB2            GKAGYVTNRGRQKVVTLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQD.EHEKYHSNWRA  736
CONSENSUS A          -kagy-tdr---k--s-te--n-k--lh-ihla-----Se-n--t---y-l--------r----l------K-i?-d----s-------------------S---------------E.e--r--s---?  732
A.KE.Q23             -------DK------S--E--------H--H-------S--------------------K----I------K--E-D----S-------------------S---------------E.---R------T  736
A.SE.SE6594          -TGWF-AEG---V--P-AK--H-T---P-LHPS-----SV-I--N---S-G--------R-----------K-VG-D------------------------S---------------E.---R------T  731
A.SE.SE7253          -------D-------S--E--------H--H-------S--------------------S----I------K--G-D----S-------------------S---------------E.---R-------  736
A.SE.SE7535          ---------------S--E--------H--L-------S--------------------R-----------K--G-D----S-------------------S---------------E.---R------T  734
A.SE.SE8538          -------D-------S--E-------VH--H-------S--------------------S-----------K--G-D----S-------------------S---------------E.---R-------  736
A.SE.UGSE8131        -------D-------S--E--------H--H-------S--------------------K-----------K--E-D----S-------------------S---------------E.---R------T  730
A.UG.92UG037         -------D-------S--E--------H--H-------S--------------------R-----------K--E-D----S-------------------S---------------E.---R-------  736
A.UG.U455            -------D-------S--E--------H--H-------S--------------------R----I------K--E------S-------------------S---------------E.D-----C----  735
A2.CD.97CDKFE4       -------D-----I-S--E--------H-------------------------------R-----------K--E--R-----------------------S---------------E.---R-------  728
A2.CY.94CY017-41     -------D-----I-S--E--------H------------------------------ER----I------K--E--R---S-------------------S---------------E.---R-------  736
CONSENSUS B          -------dr------slt------------h-------l-----t-----l--------K----l-sq---q---------a------------qv-k---a--------d---k--e.e-----s--r-  736
B.AU.MBC18           -------D-------S-A------------H-------S--------------------K----I--N--------------------------K------T------------R--E.D----------  748
B.CN.RL42            --------K-----------------------------V--------------------K------S----------------------------I---------------------E.------C----  736
B.DE.D31             -------D-------S--------------H-------------S-----I--------K------S--------------T------------------------------------.-----------  736
B.GA.OYI             -------D-------S--------------H----------------------------K------S--------------------------------------------------E.-----------  736
B.GB.CAM1            -------D-------P-------------------------------------------K------S----E---------------------------------------------E.-----------  736
B.KR.WK              -------D-------P--------------H----------------------------K------S--------------------------------------------------E.-----------  736
B.NL.3202A21         -------D-------S-N------------N----------------------------K------S--------------------------------------------------E.-----------  736
B.TW.TWCYS           -------DK------S--------------H----------------------------K------S----E---------T--------------------------------R--E.-----------  737
B.US.JRCSF           -------S-------S--------------H----------------------------K------S--------------------------------------------------E.D----------  740
B.US.JRFL            ---------------S--------------H----------------------------K------S--------------------------------------------------E.D--------K-  736
B.US.MN              ---------------S--------------H----------------------------K------S--------------------------------------------------E.D----------  738
B.US.RF              -------D-------S--------------H----------------------------K------S------------------------------R---T----------------.-----------  735
B.US.SF2             -------D-------SIA------------H----------------------------K------S-------------------------------------------N------E.-----------  736
CONSENSUS C          -------Dr--qki-sl-E----kt-----q-------s---------------q---dK----l-n-------k--r---S------------q------s---kv------d---E.---k--s----  732
C.BR.92BR025         -------D-----I-SI-E-----------Q-------S--------------------K-----------------R---S-------------------S-----------N---E.-----------  733
C.BW.96BW01B03       -------D----RI-S--E-----------Q-------S--------------------K-----------------R---S-------------------S---------------E.------N----  732
C.BW.96BW0402        -------DK----I-P--E----RA-----Q-------S--------------------K------T----------R---S-------------------S---------------E.------C----  732
C.BW.96BW0502        -------D-----I-S--E-----------Q-------S--------------------K--------------Q--W---S-------------------Q----M----------E.------N----  742
C.BW.96BW1104        -------DK---E-----E-----A-----Q-------P--------------------K---------------------S------------K------S---E-----------E.-----------  730
C.ET.ETH2220         -------D-----I-S--E-----------Q-------S---------------L----K----I---------S--R---S-------------------S---------------E.-----------  743
C.IN.93IN999         -------D-----I-P--E-------------------S--------------------K-----------------R---S-------------------S---------------E.-----------  741
C.IN.94IN11246       -------D-----I-S--E-----------C-------S--------------------K--------------N--R---S-------------------S---R-----------E.-----------  732
C.IN.95IN21068       -------DK--N-I-S--E-----------C-------S-------------------NK-----------------R---S-------------------S---------------E.---R-------  731
D.CD.84ZR085         -------D-------PF-------------N----------------------------K------S----------------------------------S---------------E.-----------  737
D.CD.ELI             -------D-------P--------------N----------------------------K-----------------------------------------Q---------------E.------N----  735
D.CD.NDK             -------D-------PF-------------N----------------------------K------S----------------------------------Q---------------E.------N----  735
D.CD.Z2Z6            -------D-------PF-------------N----------------------------K------S----------------------------------Q---------------E.------N----  735
D.UG.94UG1141        -------D-------S--------------N----------------------------K----V-S--------------S-------------------N----I----------E.------N----  737
F1.BE.VI850          -------DK-K----S--E-----A-------------S--------------------K----I---------Q--R---S---------------------V--I----------E.------N----  737
F1.BR.93BR020-1      -------D-----A-S--E-----A-----Q-------S--------------------K---------------------S-----------------------------------E.------N----  736
F1.FI.FIN9363        -------D-------S--E-----A-----H-------S--------------------K--------------Q---I--S------------------------I----------E.------N----  735
F1.FR.MP411          -------D-----A-I--E-----A-----H-------S----------V---------K----I--K---K--Q--R---S---------------R--------I----------E.------N----  738
F2.CM.MP255          -------D-------P--E-----------H-------S--------------------K--------------Q------S-------------------S---------------E.------N----  736
F2.CM.MP257          -----I-D-------S--E-----------Q-------S-----------------H--K----I---------Q--R---S-------------------T---------------E.-----------  736
G.BE.DRCBL           -------DK---III---E-----A--H--Q------QS--------------------R----I----------------S-------------------S---------------E.---R-----K-  733
G.FI.HH8793-1-1      -------DK-K--II---E-----------Q-------S--------------------R---------------------S-------------------S---------------E.D----------  735
G.NG.92NG083         ----H--DK-K--II---E-----A--H--Q------RP--------------------R-G-------------------S-------------------S---------------E.---R-------  735
G.SE.SE6165          -------DK-K--II---E-----A-----Q------RS--------------------R--A------------------S-------------------S---------------E.---R--N----  733
```

```
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                
p66, Rnase H \/ Integrase
B.FR.HXB2 GKAGYVTNRGRQKVVTLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQD.EHEKYHSNWRA 736
H.BE.VI991 -------D--K--I-S--E----------E--P----------K----------E------F--S----------S----------V.Q---------- 734
H.BE.VI997 -------DK-K----A--E----------K----------E---------S----------S----------E.A--R--N---- 733
H.CF.90CF056 -------D--K----S--E----------K----------E---------S----------S-V----------E.---R--N---- 735
J.SE.SE7022 -------DK-----I----------I----------K----------E---------S----------S----------E.---------- 734
J.SE.SE7887 -------DK----------H------R----------K----------E---------S----------S----------E.D---------- 734
K.CD.EQTB11C -------DK-----ISI-E----------H-------S----------K----------DR---S----------S----------EX------N---- 737
K.CM.MP535 -R-----D-------SI-E-----A-----C-------S----------K---D----------RI--S----------E.------N---- 736
N.CM.YBF106 ----F--D-------SIEN-----A-----L----E--Q-A---------M---HS---K---D--G----E-----R---S----------X---S---X-------E---E.---R-----K- 743
N.CM.YBF30 ----F--D-------SIA------A-----LM---E--RD----------M---HS---K------S----E-----R---S----------S----I------E---E.D-DR-----K- 743
O.CM.ANT70 -------EQ-K--IIK-DE-----A--M--L------KET----------V-SS--T----PI-Q----E-T---Q---T----------KI-----KD--R----E---Q--E.D--------K- 732
O.CM.MVP5180 -------EQ-K-NIIK-EE-----A--M-VLI-----KEQ---------V----SS--T--D-PI-Q----E-T---R---T----------KI-----KD--R----E---Q--E.D---------- 732
O.SN.MP1299 -------EK----IIK-EE----RA--M-VL------KET----Q----V----SS--T----S--Q----E-T---Q---T----------KI-----KD--R----E---Q--E.D---------- 732
O.SN.MP1300 -------EK-K--I-K-EE-----A--M-VL------KET---------V----SS--T----P--Q----E-TQ--Q-F-T----------KI-----KD--R----E---Q--E.D---------- 732
CONSENSUS 01 -----v-d---q--vS--Ett-q-t--Ha-h---q---se--------y----------R---ev-n-i--e--k--kv--Sw---------n-------sS----------d---E.e--R--s---t 736
01_AE.CF.90CF11697 -------D-------S--E--------H--H-------SD--------H----------R----I------K--E--R---S----------S----------E.---R------T 734
01_AE.CF.90CF402 -------D-------S--E--------H--H-------P----------R---DI------K--E------S----------S----------E.D--R------- 736
01_AE.CF.90CF4071 -----A-D---H---S--E--------H--R---R---S----------R----I------K--E------SR----------S----------E.---R------T 736
01_AE.TH.93TH057 -------D-------S--E--------H--H-------S----------R----V------E---------S----------D--------S----------E.---R------T 736
01_AE.TH.93TH065 -------D-------S--E--------H--H-------S----------R----V-S-M--E------D--S----------S----------E.---R------T 733
01_AE.TH.93TH253 -------D------IS--E--------H--H-------S----------R----V-S----E---------S----------IS----------N---E.---R------T 728
01_AE.TH.93TH902 -------D-------S--ESN-R----HR-H-------S----------R----V------E---------S----------S----------E.---R--R---T 736
01_AE.TH.94TH702 -------D---A---S--E--------H--H-------S----------R----V------E---------S----------S----------E.---R--R---T 736
01_AE.TH.95TNIH047 -------D-------S--E--------H--H-------S----------R----V---L--E---------S----------S----------E.---R------T 736
01_AE.TH.CM240 -------D-------S--E--------H--H-------S----------R----V------E---------S----------S----------E.---R------T 741
01_AE.TH.TH022 ----------S--E-----S--H--H-------S----------R----V------E---------S----------S----------E.---R------T 736
CONSENSUS 02 -k-----D----kvvS-tE--------H--h----d--S----------r------nq---k--e-Dk---S----------k---n-i--v-----i----E.--ery----r- 733
02_AG.CM.MP807 -------D-----AIS--E--------H--H-------S----------K----------K--G-D----S----------N----------E.---R------- 736
02_AG.FR.DJ263 -------D-------S--E--------H----------S----------R----------K--E-D----S----------N----------E.--GR------- 736
02_AG.FR.DJ264 -------D-------S--E--------H----------S----------R----------K--E-D----S----------N----------E.---R------- 736
02_AG.GH.G829 -------D-------S--E--------H--H-------S----------R----------K--G-D----S----------N---S-F--I-----L----E.----F------ 736
02_AG.NG.IBNG -------D-------S--E--------H--H-------S----------R----------K----D----S----------N----------E.---R-----K- 736
02_AG.SE.SE7812 -------D-------S-AE--------H--H-------S----------R------T----E----DR---S----------S----------E.---R------- 736
02_AG.SN.MP1211 -R-----D----N--S--E--------H--L----E--S----------R----------K--E-D----S----------K----------E.---R------- 736
02_AG.SN.MP1213 -------D-------S--E--------H----------S----------R-------L---K--E-D----S----------N----------E.---------- 736
03_AB.RU.KAL153-2 -------D-------S----------H----------K------S----------E----------E.------G---- 736
03_AB.RU.RU98001 -------D-------S---I-------H--H----------K------S----------F----------E.A---------- 736
04_cpx.CY.94CY032-3 -------D-------S-SE----------S----------I--------R---D----------R-D----S----------N----------E.------N---- 735
04_cpx.GR.97PVCH -------D-------S-SE----R----------K---S----------I--------R---D-------L--Q-D----S----------N----------E.D--------- 734
04_cpx.GR.97PVMY -------D------I--PE----------S----------M-----K--K---D----------Q-D----S----------N----------E.------N--K- 735
05_DF.BE.VI1310 -----I-DK----A-S--E-----A-----H----------K----------E-K----S----------.---------- 736
05_DF.BE.VI961 -------DK----A-S--E--------H----------S----------K-------K-----VQ--R---S----------N----I----------E.------N---- 736
06_cpx.AU.BFP90 -------D----------E----------N-------P----------K----------I--S----------N----------E.D--R------- 738
06_cpx.ML.95ML127 -------D-------S-AE----------N-------S----------K----------S----------N----------E.D--R------- 735
06_cpx.ML.95ML84 -------DK---R-IS--E----------N-------S----------K------S----------S----------E.D--R------- 737
06_cpx.SN.97SE1078 -------D-------S--E----------N-------S----------R----------N------S----------S----------E.---R------- 740
10_CD.BFL061 -------D----RAISI----------N-------S----------K----------P----------R----------E.------N---- 735
10_CD.BFL071 -------D------ISI----------N-T-----S----------K----------N------S----------H----S----------E.---------- 736
10_CD.BFL110 -------D-------P--E----------K-------S----------K----------Q--R---S----------N----------E.------N---- 733
11_cpx.CM.MP818 -------DK-----I---E--------E--H----------K------S----------T----------I-----S----------E.D-DR------T 737
11_cpx.FR.MP1298 -------DK---------E--------E--H----------K------S----------S----------N----------E.---R------- 735
11_cpx.FR.MP1307 -----I-DK---------E-----A--E--H---------A----------K------S-------M------S----------S-V----------E.D--R------T 751
11_cpx.GR.GR17 --------K-----I-PIE--------E--H---K----------K------S----------S----------S----------PE.G--R------- 735
CPZ.CD.CPZANT -------D---SR-KH-QK----QA-----LM--E--TGP----------VL-GT------P--EE--QK---R-QI--S----------Q---Q----E------E.D-D------S 736
CPZ.CM.CAM3 -------D-----I-N-EN-----A--T-V------AEHI---------V----HS----------L---E------S--S-------V----------S-NQ-SSLFRWYRLSTRKNMR---N---- 742
CPZ.CM.CAM5 -----I-D-----IIN-EN-----A--T-V------AEHI---------V----HS----------GS---E------S--S----------S----------E.------N--K- 737
CPZ.GA.CPZGAB -------DK-K--IIS-EN----QA--K-LL------DQQ---------V-----S---H----------E------I--S----------R--E.---R-----K- 734
CPZ.US.CPZUS -------D-----II--EN-----A--T-V----K--ENT--V--G---V----HS----------E------S-IS----------I-----S----------E.D-D------T- 738
```

HIV-1/SIVcpz proteins

HIV-1/SIVcpz proteins

```
B.FR.HXB2          MASDFNLPPVV.AKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLA.GRWPVKTIHTDNGSNFTGATVRAACWW.AGIKQEFGIPYNPQSQ    863
CONSENSUS A        ---------?v.--------d--------m---------i-------------l---------i------------t---l----.------vv---------S-AvK-a---.-n?q--f---------    857
A.KE.Q23           ---------I-.----------------------------------------V-------------------------------.------IV---------S-A-K-----.-N-Q------------    863
A.SE.SE6594        ----------I.-----------------I-------------------------------------------------------.------VV---------S-A-K-----.-N-Q------------    858
A.SE.SE7253        ----------I.-----------------------------------------------------------------------.------IV---------S-AFK-----.-S-Q------------    863
A.SE.SE7535        ---------I-.-----------------------------------------------------------------I----.------V----------S-A-K-----.-NV-------------    861
A.SE.SE8538        ----------I.--------N------------------M------------------------------------------.------VV---------S-AFK-----.-NVQ--Y---------    863
A.SE.UGSE8131      ---------I-.-------------------------------------------------------------A--------.------VV---------S-AFK-----.-NVQ------------    857
A.UG.92UG037       ---------I-.-------------------------------------------------V--------------------.------VV---------S-A-K-----.-NV-------------    863
A.UG.U455          ------------.--------N-------------------------------------------------------I----.------V----------S-A-K-V---.-N-Q------------    862
A2.CD.97CDKFE4     --Q------I-.-------------------------------------------------------------------I----.------V------P---S---K-----.--VQ-----------    855
A2.CY.94CY017-41   --H--------.-------------------------------------------------------T----D----I----.------V------P--IS---K-----.---Q------------    863
CONSENSUS B        --s------v-.------s--------------------------------i----------------------------l----.-----kt------s--tStt-K-----.agi--e--i-------    863
B.AU.MBC18         ---------I-.-------------------------------------I----------------------------I----.-----T----------IST--K-----.--------T-------    875
B.CN.RL42          ------------.-------------------------------------I-------------------------------.-------------R---SNS-K-----.----------------    863
B.DE.D31           ------------.-------------------------------------------------------------------I----.-----------------ST--K-----.--V-------------    863
B.GA.OYI           ------------.-------------------------------------I----------------------------I----.-----------------ST--K-----.----------------    863
B.GB.CAM1          ------------.-------------------------------------------------------------------.-------------G--IST--K-----.----------------    863
B.KR.WK            --G--------.------C-----------------------------------------------------------.-----------------SN--K----$.-R--------------    862
B.NL.3202A21       ------------.-------------------------------------I----------------------------I----.-----T-----------S---K-----.----------------    863
B.TW.TWCYS         ------------.-------------------------------------I-------------------------------.-----------------S-A-K-----.-----K----------    864
B.US.JRCSF         ---------I-.-------------------------------------I-------------------------------.-----T-----------ST--K-----.----------------    867
B.US.JRFL          ------------.-------------------------------------I-------------------------------.-----------------ST--K-----.----------------    863
B.US.MN            ---------I-.-------------------------------------------------------------------.-------------P---ST--K-----.T---------------    865
B.US.RF            ------------.-------------------------------------I----------------------------I----.------V----------ST--K-----.----------------    862
B.US.SF2           ------------.-------------------------------------I-------------------------------.-----------------ST--K-----.----------------    863
CONSENSUS C        --se--l--i-.a-----s-dk--l----i----d-s--------------i-----------i---------------fi----.-----kv------s--tSaa-K-----.---Q------------    859
C.BR.92BR025       ---E-----I-.-----------------T--------------------I----------------------------I----.------V---------ISN--K-----.---Q------------    860
C.BW.96BW01B03     ---E-----I-.------------------------G--------------------------M---------------YI----.------V------T---S-A-K-----.---Q------------    859
C.BW.96BW0402      ---E-----I-.-----------------I--------------------I----------------------------YI----.------V----------STA-K-----.---Q------------    859
C.BW.96BW0502      --DE-----I-.-----------------I-----------------------------------M---------------I----.-----RV------T---S-A-K-----.---Q------------    869
C.BW.96BW1104      ---E-----I-.-----------------I-------------------------------------------------YI----.------I----------S-A-K-----.---Q------------    857
C.ET.ETH2220       --NE--I----.P-----C----------I----N----------------I-------------------------------.-----RV----------SNA-K-----.---Q------------    870
C.IN.93IN999       ---------I-.--------GQ--Q--------------------------I----------------------------YI----.------V----------S-A-K-----.---Q------------    868
C.IN.94IN11246     ------------.---------Q----------------------------I-----------M----------------I----.------V----------S-A-K-----.---Q------------    859
C.IN.95IN21068     ------------.---------Q-------T--------------------I-----------M----------------I----.------V----------S-A-K-----.---Q------------    858
D.CD.84ZR085       ------------.-------------------------------------------------------------------A--------.------VV---------S---K-----.----------------    864
D.CD.ELI           ------------.-------------------------------------------------------------------.------VV---------S-A-K-----.----------------    862
D.CD.NDK           ------------.-------------------------------------------------------------------.------VV---------S---K-----.----------------    862
D.CD.Z2Z6          ------------.-------------------------------------------------------------------I----.------IV----------S-A-K-----.----------------    862
D.UG.94UG1141      ---E-------.------------V----L--------------------G--------------------------------.------VV---------S-A-K-----.----------------    864
F1.BE.VI850        ---------I-.-----------------------------------------------E------------------I----.------I----------S-A-K-S---.---Q------------    864
F1.BR.93BR020-1    ------I-A--.-------------------------------------I-----------L----------------------.-------------T---S---K-----.---Q------------    863
F1.FI.FIN9363      ------------.-------------------------------------I-----------------------D----I----.------M----------S-A-K-----.---Q------------    862
F1.FR.MP411        ------------.-------------------------------------I----------------------------I----.------I----------SSA-K-----.---Q------------    865
F2.CM.MP255        ------------.------N--------------------------------I----------------------------I----.------I----------STV-K-----.---Q----V-------    863
F2.CM.MP257        ------------.-------------------------------------I-----------------------A---I----.------I----------S-V-K-----.---Q------------    863
G.BE.DRCBL         ---------I-.-------------------------------------I----------------------------I----.------I----------S-A-K-----.-S-T------------    860
G.FI.HH8793-1-1    --------HI-.--------H------------------------------I----------------------------I----.------V----------S-A-K-----.-D-T------------    862
G.NG.92NG083       ------------.-------------------------------------I-I--------------------------I----.------V------P--IS-A-K-----.-N-T------------    862
G.SE.SE6165        ---------I-.-------------------------------------I-I--------------------------I----.-----TV----------S-A-K-----.-N-T------------    860
```

```
B.FR.HXB2           MASDFNLPPVV.AKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVIPAETGQETAYFLLKLA.GRWPVKTIHTDNGSNFTGATVRAACWW.AGIKQEFGIPYNPQSQ    863
H.BE.VI991          ---------I-.-----------------------------------------I----------------------------I----.------M----------S-A-K-----.-D-H------------    861
H.BE.VI997          ---E-----I-.-------------------------------------------------------XP--------------I----.------M------T---STA-K-----.-D-Q-D----------    860
H.CF.90CF056        V--------I-.-------------------------------------Q-----------------------K----------.S-----V----------S-A-K-----.-D-Q------------    862
J.SE.SE7022         -----------.-------------------------------------------------------------A-F-I-----.-G----A----------SGA-K-----.-D--------------    861
J.SE.SE7887         -----------.-------------------------------------------------------------A-F-I----.------V----------SGA-K-----.-D--------------    861
K.CD.EQTB11C        -----------.-----------------I---------------------I----------------------------I----.-----RV----------S-V-K-----.-D--------------    864
K.CM.MP535          ---------I-.-----------------I---------------------I----------------------------I----.------V------T---STV-K-----.--V-------------    863
N.CM.YBF106         ---------I-.----------------------IN----V-----------I-----------L----------------I----.------V------P--IS---K-----.---------------    870
N.CM.YBF30          ---------I-.----------------------N----V-----------I-----------L----------------I----.------V----------S---K-----.-N--------------    870
O.CM.ANT70          L--E-G-----.----I---P--HI----I--------EV--I----M---I-I--------F----------------------.A-----V------P---ST-MK-----.-N-QH-----------    859
O.CM.MVP5180        L----G---I-.----I---P--HI----T-----Y--E---M--------I-I-------DF----------------------.A-----V------P---S-AMK-----.T--QH-----------    859
O.SN.MP1299         L----G-----.----I-N-PQ-HI----I--------EV--M--------I-I--------F----------------------.A-----V------P---S-AMK-----.-N--H-----------    859
O.SN.MP1300         L----G-----.----I-N-P--HI----I-----Y--EV--I--------I-I--------F----------------------.A-----I------P---S--MK-----.T---H-----------    859
CONSENSUS 01        ---------i-.-----an---------am----------------------------v--g-----------------------.-----kVi----------S-avK-----.-NVr----------s-    863
01_AE.CF.90CF11697  ---------I-.------N-------------------------------------------------------------.--$---V----------S-AMK-----.-NVQ------------    860
01_AE.CF.90CF402    ---------I-.-------------------------------------------------------------------.-----RV----------S-A-K-----.-NVQ------------    863
01_AE.CF.90CF4071   ---------I-.------N-------------------------------------------------------------.------V----------S---K-----.-NVQ------------    863
01_AE.TH.93TH057    ---------I-.------N-------------------------------------------------------------.------V----------S-A-K-----.-NVR----------G-    863
01_AE.TH.93TH065    ---------I-.------N-------------------------------------------------------------.------V----------S-A-K-----.-NVR------------    860
01_AE.TH.93TH253    ---------I-.------N-------------------------------------------------------------.------V----------S-A-K-----.-NVR------------    855
01_AE.TH.93TH902    ---------I-.------N---------P----------------------------L----------------------.------V----------S-A-K-----.-NVR------------    863
01_AE.TH.94TH702    ---------I-.------N---------PX-----------------------------X----------------------.------V----------S-A-K-----.-NVR------------    863
01_AE.TH.95TNIH047  -----------.------N-------------------------------------------------------------.------V----------S-A-K-----.-NVR------------    863
01_AE.TH.CM240      ---------I-.-----TN-------------------------------------------------------------.------V----------S-A-K-----.-NVQ------------    868
01_AE.TH.TH022      ---------I-.------N-------------------------------------------------------------.------VV---------S-A-K-----.-NVR------------    863
CONSENSUS 02        ---d-----iv.------s--k--l-----------s-----l--------i-l--------------------------I----?----v-vi----------S-A-k-----?anvt--f--p------    860
02_AG.CM.MP807      ---------I-.-----------------------------------------I-M--------------------------I----.------V----------S-A-K-----.-NVT------------    863
02_AG.FR.DJ263      ---------II.------C--Q--------------G--------------I----------------------------I----.------V----------S-A-K-----.-NVT------------    863
02_AG.FR.DJ264      ---------I-.------C-----------------------------------I----------------------------I----.------V----------S-A-K-----.-NVT------------    863
02_AG.GH.G829       ---------I-.-----------------------------------------I----------------------------I----.------I----------S-A-------.-NVT------------    863
02_AG.NG.IBNG       ---------I-.------------M-----------G--------------I----------------------------I----.------V----------S-A-K-----.-NVT------------    863
02_AG.SE.SE7812     -----------.-----------------------------------------I----------------------------I----.------V----------S-A-K-----.-NVT------------    863
02_AG.SN.MP1211     ---E-----I-.------------------------------I--------I-I--------------------------I----X----X-VV---------S-A-K-----XPKLH--L--X------    865
02_AG.SN.MP1213     ---------I-.-------------------------------------------------------------------I----.------V----------S-A-K-----.-NV-------------    863
03_AB.RU.KAL153-2   -----------.-----------------------------------------I----------------------------V----.------I----------STA-K-----.----------------    863
03_AB.RU.RU98001    -----------.-----------------------------------------I----------------------------V----.------V---------ISTA-K-----.----------------    863
04_cpx.CY.94CY032-3 --------S--.--------N-------------------------------M----------------------------I----.------M--A---P---S-A-K-----.-D-N------------    862
04_cpx.GR.97PVCH    -----------.--------N------------X-----------------I---P------------------------I----.------I------P---S-A-K-----.-DVQ------------    861
04_cpx.GR.97PVMY    -----------.--------N-----------------------------I----------------------------I----.------I----------S-A-K-----.-N-Q----V-------    862
05_DF.BE.VI1310     -----------.-------------------------------------------------------------D----I----.------M------P---SGA-K-----.---Q------------    863
05_DF.BE.VI961      -----------X-------------------------------------------------------------------I----.------MV---------S-A-K-----.----------------    864
06_cpx.AU.BFP90     --N------I-.-----------------I---------------------I----------------------------I----.------V----------S-A-K-----.-N-T------------    865
06_cpx.ML.95ML127   ---------I-.-----------------------------------------I----------------------------I----.------V----------S-A-K-----.-N-T------------    862
06_cpx.ML.95ML84    ---------IL.------C-----------------------------I----------------------------I----.------V----------STA-K-----.-NVT------------    864
06_cpx.SN.97SE1078  --N------I-.----------------P---------------------I----------------------------I----.------V----------S-A-K-----.-NVT------------    867
10_CD.BFL061        -----------.----------------L-------------------------------------------------.------VV---------S-A-K-----.----------------    862
10_CD.BFL071        -----------.----------------L-------------------------------------------------.------VV---------S-A-K-----.----------------    863
10_CD.BFL110        -----------.----------------L-------------------------------------------------.------VV---------S-A-K-----.----------------    860
11_cpx.CM.MP818     ---------I-.-----------------------------------------I-M----------------T---------I----.------V----------SS--K-----.-N-Q------------    864
11_cpx.FR.MP1298    ----------I.-------------------------------------------------------------------I-R--.------M----------S-A-K-----.-N-Q------------    862
11_cpx.FR.MP1307    ---------I-.-----------------------------------------I----------------------------I----.------V----------S-A-K-----.-N-Q------------    878
11_cpx.GR.GR17      ----------I.------N------------------R------------I----------------------------I----.------V----------SNA-K-----.---Q------------    862
CPZ.CD.CPZANT       L-DEY----I-.----I-Q----HV----R------------V----------I-----S--F-----MAD---KS----------.S------------A---S-A-K-----.-N-Q------------    863
CPZ.CM.CAM3         ------I-HI-.---L--R-------------------------------------------L----------------I----.------V------P---SN--K-----.---Q------------    869
CPZ.CM.CAM5         -----------.-------------------------------------------I-----------L----------------I----.------V------P---SN--K-----.---Q------------    864
CPZ.GA.CPZGAB       ---------I-.------H-----V----------------V----------I--------------------------------.-------------P---S-A-K-----.-D--------------    861
CPZ.US.CPZUS        -----------.------------P----I--------------------I-I---------L----------------I----.------V----------SS--------.---Q------------    865
```

HIV-1/SIVcpz proteins

```
                                                                                                                                            vif cds |->
B.FR.HXB2           GVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGG.IGGYSAGERIVDIIATDIQTKELQKQIT.KIQNFRVYYRDSRNPLWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMA    991
CONSENSUS A         ------------------E-----k--------------k--.----------I----t-i--k----q-t.k--n---y-----D-i-------l-----------ns---------a--ir-------    985
A.KE.Q23            ------------------E-----------------------.----------I--------------H--.-------------D---------F----------------------------------    991
A.SE.SE6594         ------------------E-----------------------.----------I---------R------I.-------------D-I------------------------------------------    986
A.SE.SE7253         ------------------E-----------------------.----------I-----------------.---K---------D-I--------------------N------------L--------    991
A.SE.SE7535         ------------------E--------------------R--.----------I-----------------.-------------D-I------------------------------V-----------    989
A.SE.SE8538         ------------------E-----------------------.----------I-----------------.-------------D-I------------------------------------------    991
A.SE.UGSE8131       ------------------E-----R-----------------.----------I-----------------.X------------D-I-------------------Q----------------------    985
A.UG.92UG037        ------------------E-----------------------.----------I----S-L----------.---K---C-----D-I------------------------------V---K-------    991
A.UG.U455           ------------------E-----------------------.----------I----------------S.-------------D-I------------------------------------------    990
A2.CD.97CDKFE4      -----------R----I-------------------------.----------I----------------I.-------------D-I------------------------------------------    983
A2.CY.94CY017-41    ------------------------------------------.----------I-------------R---.-------------D-I------------------------------------------    991
CONSENSUS B         --v---nke--k-------------t----------------?----------vd---t----k-----i-.----------dsrd-l---p----------------sd--------a-i-------m-    991
B.AU.MBC18          ------------------------------------------.---------------S------------.-----------N-D--------------------------------V-----------   1003
B.CN.RL42           ------D-----------------------------------.--------------------R-------.----------G--D--------------------------------V-----------    991
B.DE.D31            ------------------------------------------.----------------------------.-------------D-----------------------E--------------------    991
B.GA.OYI            -------N----------------------------------.----------------------------.-------------E--------------------------------------------    991
B.GB.CAM1           -------N----------------------------------.----------I-----------------.-------------D--------------------------------------------    991
B.KR.WK             ------------------------------------------.----------I---------------V-.-------------D--------------------------------------------    990
B.NL.3202A21        ------------------------------------------.----------------------------.-------------D--------------------------------------------    991
B.TW.TWCYS          --I---------------------------------------X----------------------------.-----------N-D--------------------------------------------    993
B.US.JRCSF          ------------------------------------------.----------I-----------------.-----------N-D-I------------------------------V-----------    995
B.US.JRFL           ------------------------------------------.----------------------------.-------------E--------------------------------V-----------    991
B.US.MN             --I----------------------R----------------.-----------G----------------.-------------D----------------------N-----------V-------T-    993
B.US.RF             --------Q--Q------------------------------.----------------------------.-------------D-----H--------------------------------------    990
B.US.SF2            -------N----------------------------------.----------------------------.-----------NKD--------------------------------------------    991
CONSENSUS C         ----s-------------d----------m---------k--.----------I-i-at----k---k--i.-iq----------D-i----------------i-------------a---k-------    987
C.BR.92BR025        ---------------------------------------X--.----------I----------------M.-------------D-I----------------L-------------V---K-------    988
C.BW.96BW01B03      ------------------E--------------------R--.----------I----------------M.-------------D-I------------------------------V---K-------    987
C.BW.96BW0402       ------------------E-----------------------.----------I----------------I.-------------D-V------------------------------V-----------    987
C.BW.96BW0502       ------------------------------------------.----------I----------------I.-------------D-I----------------------------------K-------    997
C.BW.96BW1104       ------------------------------------------.----------I-M--------------I.-------------D-I------------------------------V-----------    985
C.ET.ETH2220        ------------------E--------------------R--.----------I----S--------N--L.-------------D-I------------------------------------------    998
C.IN.93IN999        -----------------------------V---------R--.----------I----------------I.--H----------D-I------------------------------------------    996
C.IN.94IN11246      ----A-------------------------------------.----------I---S-----R------I.-------------D-I----------------------------------K-------    987
C.IN.95IN21068      ------------------------------------------.----------I-----------------.-V-----------D-I----------------------------------K-------    986
D.CD.84ZR085        ------------------------------------------.----------I----S----R-------.-------------D-I------------------------------------------    992
D.CD.ELI            ---------------------------------------RR-.----------I----------------I.-------------D-I-------------------K----------V-----------    990
D.CD.NDK            ------------------------------------------.----------I---------R------I.-------------D-I------------------------------V-----------    990
D.CD.Z2Z6           ------------------------------------------.----------I-----------------.-------------D-I------------------------------V-----------    990
D.UG.94UG1141       ------------------E-----------------------.----------I----------------I.-------------D-V---------------------E--------------------    992
F1.BE.VI850         -----I------------------------------------.----------I---S-----R-------.---------------V---------------------E--I-----------------    992
F1.BR.93BR020-1     ----------------I-------------------------.---------TI---------R------I.-------------D-V---------------------E--------------------    991
F1.FI.FIN9363       ------------------------------------------.----------I---------------V-.-------------D-V---------------------E--------------------    990
F1.FR.MP411         ------------------------------------------.----------I---S-----R------I.-------------D-V---------------------E--------------------    993
F2.CM.MP255         ----------------I-------------------------.----------I-----------------.--------F----D-V--------------------NE--------------------    991
F2.CM.MP257         ------------------------------------------.----------I----------------S.--------F----D-V--------------------NE---I----------------    991
G.BE.DRCBL          ------------------------------------------.----------I----S------------.-------------D-I--------------------NE--------------------    988
G.FI.HH8793-1-1     ------------------------------------------.----------I----S--------T---.-L------F----D-V--------------------NE--------------------    990
G.NG.92NG083        -----------------G------------------------.----------I----S-----------I.-------------D-I--------------------NE-----------LK-------    990
G.SE.SE6165         ------------------------------------------.----------I----S------------.-------------D-V--------------------NE--------------------    988
```

```
                                                                                                                                   vif cds |->
B.FR.HXB2           GVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGG.IGGYSAGERIVDIIATDIQTKELQKQIT.KIQNFRVYYRDSRNPLWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQMA    991
H.BE.VI991          ------------------------R-----------------.------R---I-------P--------S.Q--K---------D-I---------------------E--------------------    989
H.BE.VI997          ------------------------R-----------------.----------I----------------S.N--K---------D-I---------------------E--------------------    988
H.CF.90CF056        ------------------------------------------.----------I----------------S.N--K---------D-I---------------------E-------E------------    990
J.SE.SE7022         ------------------E-----------------------.----------I-----------------.-------------D-I---------------------E--------------------    989
J.SE.SE7887         ------------------E-----------------------.----------I---------R-------.-------------D-I-------P-------------E--------------------    989
K.CD.EQTB11C        ------------------E-----------------------.----------I-----------------.-------------E-I-----------------.X--E--------------------    991
K.CM.MP535          ------------------------------------------.---------------------------L.N--K---------E-I---------------------E--------------------    991
N.CM.YBF106         -A--------------I-------------------------.---XT-----I---------TK--T--L.-V---X-------D-I--------------------G---------------------    998
N.CM.YBF30          -A--------------I-------------------------.----T-----I---------TN--T--L.-V-----------D-I--------------------G---------------------    998
O.CM.ANT70          ----A------S--Q---------R--------V--------.----T-----I--L-SQ---T------L.-XHK---------D-I-----Q-------------KG--------------E------    987
O.CM.MVP5180        ----A------S--Q------------------V--------.----T----LI--L-SQ---T------L.--N----------D-I-----Q-------------KG---------------------    987
O.SN.MP1299         ----A------S--Q------------------V--Y-----.----T-----I--L-SQ---T------F.---K-Q-------D-I-----Q-------------KG--------------H------    987
O.SN.MP1300         ----A------S--Q------------------V--Y-----.----T-----I--L-SQ---T------F.---K-Q-------D-I-----Q-------------KG--------------H------    987
CONSENSUS 01        -----------ki---v-e--------vq-------------.-gg--a----I----td----el--q--.-------------D-I-------------------------v----------------    991
01_AE.CF.90CF11697  ------------V-----E-----------------------.----------I-----E-----X-----.-------------D-I------------------------------------------    988
01_AE.CF.90CF402    ------------------E-----------------------.--E-------I----------A------.-------------D-I------------------------------------------    991
01_AE.CF.90CF4071   ------------------E-----------------------.----------I--------------H--.-------------D-I------------------------------------------    991
01_AE.TH.93TH057    -----------E------E-----------------------.----------I-----------------.-------------D-I------------------------------------------    991
01_AE.TH.93TH065    ------------------E--------FX-------------.-X---V----I----A------------.-------------D-I------------------------------------------    988
01_AE.TH.93TH253    ------------------E-----------------------.----------I-----------------.-------------D-I------------------------------------------    983
01_AE.TH.93TH902    ----------------I-E-----------------------.----------I-----------------.-------------D-I-------------------------I----------------    991
01_AE.TH.94TH702    ----------------I-E-----------------------.----------I-----------------.-------------D-I-------------------------I----------------    991
01_AE.TH.95TNIH047  ------------------------------------------.----------I-----------------.-------------D-I------------------------------------------    991
01_AE.TH.CM240      ------------------E-----------------------.----------I-----------------.-------------D-I------------------------------------------    996
01_AE.TH.TH022      ------------------E-----------------------.----------I-----------------.-------------D-I------------------------------------------    991
CONSENSUS 02        ----s---------g-----------------f------k--.----------I----S-----------t.k--n---------D-i---p---------------n----------a--ir-------    988
02_AG.CM.MP807      ------------------------------------------.----------I----S------------.-------------D-I------------------------------V-----------    991
02_AG.FR.DJ263      ----A-------------------------------------.----------I----S------------.-------------D-I-------------------K--------------K-------    991
02_AG.FR.DJ264      ------------------------------------------.----------I----S-----------I.-------------D-I-------------------K--------------K-------    991
02_AG.GH.G829       ------------------------------------------.----------I----S------------.-------------D-I---------------------------------L--------    991
02_AG.NG.IBNG       ------------------------------------------.----------I----S-----------I.-------------D-I------------------------------------------    991
02_AG.SE.SE7812     --------------------------------L------R--.----------I----S------------.-------------D-I------------------------------V--V--------    991
02_AG.SN.MP1211     ---------------------------------------R--.----------I----S------------.N--K---------D-I------------------------------T--L--------    993
02_AG.SN.MP1213     --------------T---------------------------.----------I----S------------.-------------D-X---X------------------------------K-------    991
03_AB.RU.KAL153-2   --------Q--QT-----------------------------.----------I----------------I.-------------D-I--------------------N---------------------    991
03_AB.RU.RU98001    --------Q--Q------------------------------.----------I----------------I.-------------D-I--------------------N---------------------    991
04_cpx.CY.94CY032-3 ------------------------------------------.----------I----S------------.-------------E-I------------------------------------------    990
04_cpx.GR.97PVCH    --------------K---------------------------.----------I----S------------.-------------D-I------------------------------------------    989
04_cpx.GR.97PVMY    ------------------------------------------.-E--------I----S-----------I.------------KD-I----------------------------K-------------    990
05_DF.BE.VI1310     ------------------------------T-----------.----G-----I---T------------I.-------------D-V---------------------E--------------------    991
05_DF.BE.VI961      ------------------------------------------.--------G-I---S-------------X-------------D-V---------------------E--------------------    993
06_cpx.AU.BFP90     ------------------E-----------------------.----------I----S-----------I.-------------D-I---------------------E--------------------    993
06_cpx.ML.95ML127   ------------------E-----------------------.----------I----S------------.-------------D-I---------------------E------------K-------    990
06_cpx.ML.95ML84    ------------------------------------------.----------I----S------------.-------------D-I------------------------------------------    992
06_cpx.SN.97SE1078  ---K------------I---------------Y---------.----------I----S------------.--R----------D-I---------------------E--------------------    995
10_CD.BFL061        ----------------------Q-------------------.----------I----------------I.-------------G-I------------------E-----------V---K-------    990
10_CD.BFL071        ------------------------------------------.----------I--------S-------I.-------------D-I------------------------------V---K-------    991
10_CD.BFL110        ------------------------------------------.----------I----------------I.-------------D-I---------------------E--------V---K-------    988
11_cpx.CM.MP818     ----------------I-------------------------.-----------------L----------.-------------D-I------------------------------------------    992
11_cpx.FR.MP1298    ------------------E-----------------------.----------I------L---------I.-------------D-I------------------------------------------    990
11_cpx.FR.MP1307    ------------------E-----------------------.----------------------------.-------------D-I------------------------------------------   1006
11_cpx.GR.GR17      --------G--E------E--------------------G--.----------I------L----------.---K---------D-I--------------I---------------------------    990
CPZ.CD.CPZANT       --------Q--Q----I-----Q-----V---H---------.----TP-Q--L--L------TQ--N--L.---Q---H-----D-V-----Q-----------K-QEE------------KE-R-KIE    991
CPZ.CM.CAM3         ----------------I-E-----R-----------------.----T-----I---------TN-----L.-V-----------D-I-----R-----------KEREEV--I----------------    997
CPZ.CM.CAM5         ----------------I-----Q-------------------.----T-----I---------TN-----L.-V-----------D-I-----R-----------KE-EEV-------------------    992
CPZ.GA.CPZGAB       -----L------------------------------------.----T-----I---------S------L.-V-K---------D-I-----T-------------QGEL-------------------    989
CPZ.US.CPZUS        ----------------I-------------------------.----T-----I----SEL--DL-----L.-V-----------D-I-----------------KE-EEV-----------K-------    993
```

```
B.FR.HXB2           GDDCVASRQDED$          1003
CONSENSUS A         ----v-g------           997
A.KE.Q23            ------G------          1003
A.SE.SE6594         ------G------           998
A.SE.SE7253         ------G------          1003
A.SE.SE7535         ------G------          1001
A.SE.SE8538         ------G------          1003
A.SE.UGSE8131       -------------           997
A.UG.92UG037        ------G------          1003
A.UG.U455           ----M-G------          1002
A2.CD.97CDKFE4      ------C------           995
A2.CY.94CY017-41    ------G------          1003
CONSENSUS B         ------s------          1003
B.AU.MBC18          -------------          1015
B.CN.RL42           -------------          1003
B.DE.D31            ------G------          1003
B.GA.OYI            -------------          1003
B.GB.CAM1           -------------          1003
B.KR.WK             -------------          1002
B.NL.3202A21        -------------          1003
B.TW.TWCYS          ------G------          1005
B.US.JRCSF          -------------          1007
B.US.JRFL           -------------          1003
B.US.MN             -------------          1005
B.US.RF             -------------          1002
B.US.SF2            -------------          1003
CONSENSUS C         -A--v-gr---d-           999
C.BR.92BR025        -A--M-------           1000
C.BW.96BW01B03      -A----G------           999
C.BW.96BW0402       -A----G------           999
C.BW.96BW0502       -A----GG---N-          1009
C.BW.96BW1104       -A----G------           997
C.ET.ETH2220        -A----G------          1010
C.IN.93IN999        -A----G------          1008
C.IN.94IN11246      -A----G------           999
C.IN.95IN21068      -A----G------           998
D.CD.84ZR085        -------------          1004
D.CD.ELI            -------------          1002
D.CD.NDK            -------------          1002
D.CD.Z2Z6           -------------          1002
D.UG.94UG1141       -------------          1004
F1.BE.VI850         V-----G------          1004
F1.BR.93BR020-1     ------G------          1003
F1.FI.FIN9363       ------G------          1002
F1.FR.MP411         ------G------          1005
F2.CM.MP255         ------G------          1003
F2.CM.MP257         ------G------          1003
G.BE.DRCBL          ------G------          1000
G.FI.HH8793-1-1     ------G------          1002
G.NG.92NG083        -G----G------          1002
G.SE.SE6165         ------G------          1000
```

```
B.FR.HXB2            GDDCVASRQDED$          1003
H.BE.VI991           ------G------          1001
H.BE.VI997           ------G------          1000
H.CF.90CF056         ----------             1002
J.SE.SE7022          ------G------          1001
J.SE.SE7887          ------G------          1001
K.CD.EQTB11C         ------G------          1003
K.CM.MP535           ------G------          1003
N.CM.YBF106          --G----GX--NQ          1011
N.CM.YBF30           --G----G---NQ          1011
O.CM.ANT70           -T-SM--G-T-SE          1000
O.CM.MVP5180         -T-SM-N--T-SE          1000
O.SN.MP1299          -T-SM--G-T-SE          1000
O.SN.MP1300          -T-SM--G-T-SE          1000
CONSENSUS 01         ------g--d-d-          1003
01_AE.CF.90CF11697   ------G----N-          1000
01_AE.CF.90CF402     ------G--N---          1003
01_AE.CF.90CF4071    ------G------          1003
01_AE.TH.93TH057     ------G------          1003
01_AE.TH.93TH065     ----------             1000
01_AE.TH.93TH253     ------G------           995
01_AE.TH.93TH902     ------G------          1003
01_AE.TH.94TH702     ------G------          1003
01_AE.TH.95TNIH047   ------G------          1003
01_AE.TH.CM240       ------G------          1008
01_AE.TH.TH022       ------G------          1003
CONSENSUS 02         ----v-gr-d---          1000
02_AG.CM.MP807       ----------             1003
02_AG.FR.DJ263       ------G------          1003
02_AG.FR.DJ264       ------G------          1003
02_AG.GH.G829        ------G------          1003
02_AG.NG.IBNG        ------G------          1003
02_AG.SE.SE7812      ------G------          1003
02_AG.SN.MP1211      ------GG-N---          1005
02_AG.SN.MP1213      ----M-G------          1003
03_AB.RU.KAL153-2    ----------             1003
03_AB.RU.RU98001     ----------             1003
04_cpx.CY.94CY032-3  -N----G------          1002
04_cpx.GR.97PVCH     ------G------          1001
04_cpx.GR.97PVMY     ------G------          1002
05_DF.BE.VI1310      ------G------          1003
05_DF.BE.VI961       ------G------          1005
06_cpx.AU.BFP90      ------G------          1005
06_cpx.ML.95ML127    ------G------          1002
06_cpx.ML.95ML84     ------G------          1004
06_cpx.SN.97SE1078   ------G------          1007
10_CD.BFL061         -A----------Q          1003
10_CD.BFL071         -A----------Q          1004
10_CD.BFL110         -A----------Q          1001
11_cpx.CM.MP818      ------G------          1004
11_cpx.FR.MP1298     ------G-----Q          1003
11_cpx.FR.MP1307     ------G----N-          1018
11_cpx.GR.GR17       ------GS-----          1002
CPZ.CD.CPZANT        DR-DL-G--N---          1003
CPZ.CM.CAM3          ---SM-GG---SQ          1010
CPZ.CM.CAM5          --GSM-G--N--Q          1005
CPZ.GA.CPZGAB        ---------N---          1001
CPZ.US.CPZUS         -A-SM-G----S-          1005
```

HIV-1/SIVcpz proteins

```
                                                                                                                Hydrophilic region
                                <- pol cds end                                                                  |-    -|
B.FR.HXB2          MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHYESPHPRISSEVHIPLGD..ARLVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPELADQLIHLYYFDCFSDSAIRKALLGHI    128
CONSENSUS A        m-----vmi------------N-------y?-kk-kg-f----y-sRh-kv-------l-d..-rlvvrt-w--htg-kd-h--hg----w-l?r----id-Dl---l--lh-fd---d------il-qv    126
A.BY.97BL006       X------X-------------N----------K---E$V-------RQ--V----------..----VK--X---AX-K--Q--HX----X-QEX----I--D---------------E------IV---    127
A.DE.AF200476      ---------------------N----------K--K-------F-CR--KA----------..----VR----------------------Q---------D-------------------I--E-    128
A.KE.AF233689      ---------------------N--------I-K--KN---------R--KV---------E..----VR-----Q---K-----H-------Q------I--D--------H------E------I--QV    128
A.KE.Q23           ------A--------------N-------H--K--KR---------R--KV----------..-T--VRA--------K-----H-------L------IT-D-------MH-------------IV-QV    128
A.SE.SE6594        ---------------------N----------K--KN-V----F--R--KV----------..-K--VR---------KE----H-------LN-----I--DQ-------H---------------QV    128
A.SE.SE7253        ---------------------N----------R--KD---------RN--T----------..----VR---------K--Q--H-------LR-----I--D----Q---H-------------I---V    128
A.SE.SE7535        ---------------------N-------C--K---N---------R--KV----------..----VR-----Q---K-----H-------L------I--D----------L-------------QV    128
A.SE.SE8538        ---------------------N--------I-K--KN------F--R--KV---------E..-K--VR-----Q---K--Q--H-------LR-----I--D-------MH-XN----------I--QV    128
A.SE.SE8891        --------V------------N--------I-K--K--L----F--R--KV---------E..----VR-----Q---K-----H-------LR-----I--D--------H--K----------I--E-    128
A.SE.UGSE8131      ---------------------N--------I-K--A---------R--KV----------..---IVR------P-----Q--H-------LR-----I--D--------H--N----------I--RV    128
A.UG.92UG037       ---------------------N--------I-RR-K---------R--KV-------I--..--I-VR-----Q---K-----H-------L------I--D--------H--N----------I--QV    128
A.UG.U13-2         ---------------------N--------I-K--K---------R---V----------..--I---------K--K--Q--H-------LR-----I--D--------H--N-------------QV    128
A2.CD.97CDKFE4     -----------H---------N------I---R--KD-S-------R--KV---------E..----VR-----QP--QA----H-------Q------I--D---H-----------E------I--E-    128
A2.CD.97CDKS10     ---------------------N--------I-K--KK---------R---V---------E..-T-IVR------P--KE-Q--H-------Q------I--D---H-----------E------I--E-    128
A2.CD.97CDKTB48    ---------------------N--------K---E---------R---V--------KE..---IVR------P--K-----H-------QG-----I--D---H-----------E----R-I--E-    128
A2.CY.94CY017-41   ---------------------N--------I-R--K--V-K-----RN-----------E..--IIVR------I--K-----H-------QN--H--I--D---H-----------E------II-E-    128
CONSENSUS B        ---rw--m----v-----rt-k---k--m-isk--kgwf-r---esth-ri--e-h--l-d..ar--it------t---d-hl-q-vsi-w-kkr-s--v--dl--q---ly-----se---rnail-hi    128
B.AU.MBC18         ---------------------N---------K--K--I------NT--K-----------..-----------------------------ERT-R-------------IH-----------A-I---R    128
B.CN.RL42          ------------------K-----------I-R--K----K---D-T--K-----------..-----------------------------------G---------------E----N-I--RV    128
B.DE.D31           ------------------------------------EK---K-----TN-------------..-----------------Q--H-----------------G---------------E----N-I--R-    128
B.ES.89SP061       ------------------------------R--K----------T-------------N..-K-------------------------L--R-----L--N--------H------E----N-I----    128
B.FR.vi02011A1H    -------------------A----------KT----K-G-----D----K-----------..-K---------------------------R--------N---H-----------E----N-I--YR    128
B.GA.OYI           ------------------------------K--K----------T---------------..-T--V----------E------A---------------G-------T-------E----N-I--N-    128
B.GB.CAM1          ----------------------------I----KK-S-------T-------------E..----V-----------------------T-G-N-----D---------------E------IV-RL    128
B.JP.D70887        ----------------------------I-R--K----------T---------------..-K---I---------------------------------G--------------------N-I----    128
B.JP.PT7-6         ----------------------------I----K--C-----------V--A--------..S---------------N-------------R--------N---------------E----N-I--RR    128
B.KR.WK            ------------------K-----------I-K--KE-V-------T---------------..-K--------------E---------------N-----D---K----H-----------H-I---R    128
B.NL.3202A21       -------------------A----------K-K--ER---------T-----------V-E..-----------------------------------G---------------E----N-I---V    128
B.TW.TWCYS         -------------------A----------I-K--K--L-K-----T---------------..-T----------------------------R--------DQ-------------E------IV-CR    128
B.US.1-2           -------------------------R------K--KE-S-------T--K-----Q------..----------------E-------------E-----I--D--------H-----T-----N-I--RL    128
B.US.AF019528      ---S--------------K--H------I-I-E--K--I-K-----NN-K----------E..-K---A-------------F----------RK-------D-------Q-------E----N-I--Q-    128
B.US.BC            ---------------------I--------I-R--K----------T--K-----------..----------------------H---V-------------D---------------E----N-I----    128
B.US.JRCSF         ---------------------N--------I----K--I-K-----TN--V----Q------..-------------------------M---TR--------D---------------E----N-I----    128
B.US.JRFL          ----------------------------T----K--I-------T---------------..-------------------------V-------N-----D---------------E----N-I----    128
B.US.MN            ----R-------A----------------I-K--K-R--------T---------------..-----------------------------------D---H----H-------------I---R    128
B.US.RF            ----------------------------I-R--K----------T-----------P--..E-----------------------------R--------D---------------E-----PS----    128
B.US.SF2           ----------------------------I-K--K----------T---V-----------..-K--------------E-------A------K-------G--------H------E---KN-I--YR    128
B.US.WC001         -------K----------------------------KK---------T---------------..-N----------S--------------------------G---------------E----N-I---R    128
CONSENSUS C        -------L--------k----n-------yv-rr-n--f----ye-Rh-kv----------e..-r-v-k-----q----d----H------rlRr-s-qvd-G---q---MH-----ad------i--hi    128
C.BR.92BR025       -------L--------K----N----------RR-S--Y-------R--K----------E..-----I-----Q---------H-------LR----R---G-------MH-----A-------I---R    128
C.BW.96BW01B21     -------L-------------N----------RR-S-------F--R--KV---------E..-----K-----Q----E----H-------LR--------G-------MH-----A-------I----    128
C.BW.96BW0402      -------L--------K----N----------R--N----------R---V----------..-----K-----Q---------H-------LR--------G-------MH-----A-------I--Q-    128
C.BW.96BW0504      ----$--L--------K----N-------HI-KR-K----------R--KV---------E..-----K-----Q---------H-------LRK-----E-G-------MH-----A-------I--E-    127
C.BW.96BW1104      -------L-------------T----------RR-N--S----F--R--KV---------E..-----K----------------H------TVRG---$---G-------MH-----T----------QV    127
C.ET.ETH2220       -------L--------K----N-------HI-RR-N--V-----D-R--KV---------E..---I-K-----Q---------H-------LRS-N-----G---H---MH-----AE------I--YR    128
C.IN.94IN11246     -------L--------K----N----------RR-N--------D-RN-KV---------E..-I---K-----Q---------H-------LR--N--IE-G-------MH-----A-------I--R-    128
C.IN.95IN21068     -------L--------K----N----------RR-N----------R--KV---------E..----------Q---------H-------LRK-----E-G-------MH-----A-------I----    128
CONSENSUS D        ------------------rt----v-h-myv-K--?g-f-----dsphpk----------e..-r--vk-----h---re-------------rr-s--v--G---q---m-y----aes-----i--hi    127
D.CD.84ZR085       ------------------N-------Y--HI-K--K--------D----K----------..----V----------E-------------R--------G-------M------A-------I----    128
D.CD.ELI           ------------------K-------------K--NR------------K----------E..-----K----------E-------------R--------G-------M------E------I--D-    128
D.CD.NDK           ------------------N-------Y-----K--NR-------D-H--K----------E..----V----------KE------------R--------G-------M------AE------I----    128
D.CD.Z2Z6          ------------------------------K--SR-------D----K----------E..-M--VK----------------------R--------G-------M-------EA-----I----    128
D.UG.94UG1141      ----------------------------I-K--K--L-----DC-N-K----------E..----VK----------E-------------G--N--I--G-------I-H----AE------I--QV    128
D.UG.U18-0         ------------------------------K--Q--W-----D----K----------E..----VK----------E-------------E-------G-------M------AE------I----    128
D.UG.U25-6         -------------------I--------IH--K--T--L-----DCL--K----------E..----VK----------E-------------R--------G-------I------AE------I----    128
D.UG.U36-0         -------------------------Y---I-K--Q--V-----DC-NSK----------E..----VK----------E-------------R--------G-------M------AE---------Q-    128
D.ZR.AF233690      ------------------N-----I-Y-----K--S----------------------E..-K--VI-----N----------------------G-------G---R---M-------E------I--Q-    128
F1.BE.VI850        ------L----------N-------Y-----K--K--S----FQ-R---V--------EE..VK----------P---E------------QGK-R--I--G-------I-------E------I---R    128
F1.BR.93BR020-1    ------------------N-------Y--HI-K--K-------F--R--K---------ET..-E---------LP---E------------QG--R--I--G-------I-------E------I---K    128
F1.DE.AF200475     ---I--------------NQ------Y--H--K--K-------F--R--KV--------KE..-K-----------------------QG--R-----G-------I-----------------R    128
F1.FI.FIN9363      -------T----------N-------Y--H--K--KR------F--R--K---------EE..-E---------N----------------QG--R--I--G-------I-------E------I---R    128
F1.FR.MP411        ------------------S-------Y--H--K--KN------FQ-R--K---------EE..-K--V----------------------QG--R--I--D-------I-------A------I---R    128
F2.CM.MP255        ---------------------N----------RR-K----------R--KV--------E-.DSK---I----------------H-------Q---R-----D--------R------E----N-I---R    129
F2.CM.MP257        ----------------K----N--------I-KR-A----------R---V--------EE.DSK---I----------------------Q---R-----G--------H-------------I--QR    129
```

```
                                                                                                                                Hydrophilic region
                                   <- pol cds end                                                                                |-    -|
B.FR.HXB2          MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHYESPHPRISSEVHIPLGD..ARLVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPELADQLIHLYYFDCFSDSAIRKALLGHI    128
G.BE.DRCBL         ---------------------N--------I-K--K----------R--KV-----------..-K--VR--------DKE----H-------QG--------D---H----H--N---E------I--QT    128
G.CG.CNG30         ---------------------N--------I-K--K----------R---V---------N..-M-IVR---------K-----H-------Q---G--I--D--------H--K-------------Q-    128
G.FI.HH8793-1-1    ---------------------N------L---K----$-----F--R--KV-----------..-K--VS-----N---K--Q--H-------QR-----I--D--------H------E-----N-I--R-    127
G.NG.92NG083       -------V-------------N----------K--K----------R---V--------R-..-T--VR------A--K--Q--H-------Q------I--NT--H-----------E------I--E-    128
G.SE.SE6165        ---------------------H----------K--------P--A-R---V-----------..-T--V----------K--Q--H-------QR--R---E-D---H----H-------------I--Q-    128
H.BE.VI991         ------------------K--N----------K--KK-V-------TN-KT-------V--..-------------------H-------QE-----I--D--------H-------------I---R    128
H.BE.VI997         ------------------NK------Y---I-K--NR-V-----D----------------..-----------------------------Q---------G-------TH------E----G-I--RV    128
H.CF.90CF056       ------------------N-------Y--HI-R----------F--T-------------E..-----------N----E------------L-------E-G-------MH------E------I--RV    128
J.SE.SE7022        ------------------N-------Y--N--K---Q-L-----D-R--K----------E..----V------Q----------------R---R-----G-------MH-------------I--Q-    128
J.SE.SE7887        ------------------N-------Y--N--K---K-L-----D-N--K----------E..-I---------Q----------------QR--R-----G-------MC-------------I--Q-    128
K.CD.EQTB11C       ------------------N-------Y-----K--NR-R-----D-N--K-----------..-E--V----------E------------L-K-R-----G-------I-------E----------R    128
K.CM.MP535         ------------------N-------Y--HI-K--NR-Y-------R--K-----------..-E--V------L-----------------L---R---E-D-------I-------E--V---I---R    128
N.CM.YBF106        --------V-----X-K----N----------K--K--Y------TK--KT---X---V-X..-X---V-----T---QP----HX------QGK-K------M--K---C---N--TA----Q-V--RP    128
N.CM.YBF30         --------V-------K--K-N----------K--K--Y------TH--K--------V-Q..----TV-----T---QS----H-------LRK-K------M--K----H-----TA----Q-V--RP    128
CONSENSUS O        ---r--vl-v--v--QkVKa-N----Y-K-rsrktr?-y----?-srN-rv--g-Y--vgg..a?ivv------MP--rdEhl-H----e-qykk-k--I---T--Rm--lh--t--Te--i-k-I--QR    125
O.CM.ANT70         -------L-------QKVKA-N----Y-K-R-R-TEN-W-------RN--V--S-Y--V-V..-HV-V------MP----E---H-------Y-K-K--I---T--RM---H--T--TA--V---I--QR    128
O.CM.MVP5180       -------L----I--QKVKA-N----Y-K-M-K--AN-R-------RN-KV--A-Y--VAE..-DI-V------MP---EE---H------QY-E-K--I---T--RM---H--T--TE------I--QR    128
O.FR.HIVY16019     ------A--------QKVKV-N----Y-K-R-R-TKE-C----F-YKN--V--G-Y--V-M..-WI-L------MP----E---H------QY---K--I---T--RM------T--TE------I--QR    128
O.FR.HIVY16020     ---K---L-L-----QKVKV-N----Y-K-R-R--KD-Y-------TN--V--S-Y--V-P..-HVLV------MP--K-E-M-H------QC-K-T--I---T--RM---H--T--TE------I--QR    128
O.FR.HIVY16021     -------L-------QKVKV-N----Y-K-R-R--KE-H-------RN-----S-Y--T-P..-HI-V------MP----E-M-H------QY-E-T--I---T--RM---H--A--TE------I--QR    128
O.FR.HIVY16022     -------L-------QKVKA-N----Y-K-R-RQT-E-H-------RN--V--G-Y--V-G..-NV-V------MP----E---H------QY-K-K--I---T--RI--QH--T--T-------I--QR    128
O.FR.HIVY16023     -------L-------QKVKA-N----Y-K-RT--T-N-Y----F--RN--V--G-Y----M..-CV-V------MP----EQ--H------QY-K-T--I---T--RI---H--T--T-------I--QR    128
O.FR.HIVY16024     -------L-------QKVK--N----Y-K-R--RT-Q-Y----F-HRN-----S-Y--V-A..-Y--V------MP----E---H------QY-K----I---T--RI------T--TE----Q-I--QR    128
O.SN.99SE-MP1299   -------L-------Q-VK--N----Y-K-R---T-D-Y----F--RN--V--G-Y--V-G..PWI-V------MP----E---H------Y-K-K--I---T--RM--I---A--TE------I--QR    128
O.SN.99SE-MP1300   -------L-------Q-VK--N----Y-K-R---T-D-Y----F--KN--V--G-Y--V-G..PWI-V------MP----E---H----D--YRK-T--I---T--RM--T---A--TE------I--QR    128
CONSENSUS 01       ----------q------rt-n--------i-K--kk-f-------Qh-KV-------lge..-r-v-r-----q---K--h--H-------Qrk-s--i--D---ql---q--d---d----k-i--Qv    128
01_AE.CF.90CF11697 ---------------------Y--------I-K--KN---------Q--KV-------I--..-----R---------K-----H-------QRK-------D---RQ---Q-------------M--QV    128
01_AE.CF.90CF402   -------------------A-N--------S-K--AK---------Q--KV-----------..---I-R---------K-----H-------QRK$---I--D--------Q---------------Q-    127
01_AE.CF.90CF4071  ------------------K--N--------I-K--KK-V-------Q--KV---------E..-----R---------K-----H-------QEK----I--D---R----Q------E------I--QV    128
01_AE.DE.K08DE     -----------R----------N--------I-K--KK---------Q--KV----------..-----K-----Q---K-----H-------QR--R--I--D--------Q-------------I--QV    128
01_AE.TH.93TH057   ---------------------N--------I-K--KK---------Q--KV---------E..-----R-----Q---K--Q--H-------QRN----I--D---R----Q-----------R-I--QV    128
01_AE.TH.95TNIH047 ---------------------N----------K--KK---------QN-KV--------EE..-K-I-R-----Q---K--Q--H-------QRT----I--D--------H-----------R-I--QV    128
01_AE.TH.CM240     ---------------------N--------I-K--KK---------Q--KV---------E..-----R-----Q---K-----H-------QRK----I--D---K----Q--G----------I--QV    128
01_AE.TH.TH022     -----------R---------N--------I-K--KN---------Q--KV---------E..-K---R-----Q---K--Q--H-------QRK----I--D--------Q-------------I--QV    128
CONSENSUS 02       ---------------------N----h---v-?--k?------y--r--kVs----------..-rlvvr---g-ht-e-d-h--H----e-kQrr----I--D--------h--d--s?----k-i--qv    125
02_AG.CM.97CMMP807 ---------------------N----------K--NK---------R--KV-----------..-T---K----------E----H-------QR-----I--D--------H--N---E----Q-I--QV    128
02_AG.FR.DJ263     ---------------------N--------I-R--KN------F--R--KV-----------..---IVR------A--------H------KQRK----I--D--------H-----AE------I--EV    128
02_AG.FR.DJ264     ---------------------N--------I-R--KN------F--R--KV-----------..---IVR---R--A-----Y--H------KQRK----I--D--------H-----T-------I--QV    128
02_AG.GH.G829      ---------------------N----------R--KD---------R--KV-----------..--I-VR---------------H-------Q------I--D--------H--T---E------I--EV    128
02_AG.NG.IBNG      ---------------------N----------K--K-------F--R--KVC----------..-----VR---------------H------KQ------I--D----------N----------I--E-    128
02_AG.SE.SE7812    ---------------------N----Y--K-R--KD---------S---V----------..-----VR---------------H------KQR-----I--D--------H-------------I--Q-    128
02_AG.SN.98SEMP1211 --------------------N----------K--HR---------R--KV-----------..-T-IVR-----QP--------H----V-QQ------I--D--------H------E--------QV    128
02_AG.SN.98SEMP1213 --------------------N--------I-K--K----------Q--KV-----------..-----VR--------D---Q--H------KQRK----I--D--------H-------------I--QV    128
03_AB.RU.KAL153-2  ---------------------N------I-I-K-----V-K-----RN--------------..-K---K-----------------A------E--------N---------------E----N-I---R    128
03_AB.RU.RU98001   ---------------------N------I-I-K-----V-K-----RN--------------..-K---K-----------------A------E--------N----------------N-I---R    128
04_cpx.CY.94CY032-3 -A--------------K----N----------K--K-$--------R--KV---------E..----VR-----QP--Q-----H-------LR--------D-------MH------E------I---R    127
04_cpx.GR.97PVCH   ----------------K----N--------I-K--K--S-------R---V----Q-----..-----R-----QP--K-----H---M---LR--------D-------MH------E----Q-I--YR    128
04_cpx.GR.97PVMY   --------T-------K----N-------HI-K--K--V-K-----RN--A---------E..-K--VR-----QP-KK-----H----X--LRS-------D-------VH------E------I---R    128
05_DF.BE.VI1310    ------------------N-------Y--H--K--NR-C----F--RN--------------..-K--V-------------------------R--------S-------V-------E----N-I--R-    128
05_DF.BE.VI961     ------------------N-------Y--H--K-TKA----------R-------------EE..-K---I----------E-------------G-----I--G-------M-------E------I--YR    128
06_cpx.AU.BFP90    ------------------N-------Y--HI-K--KR-N-----D-N--K----------S..-E---------N----K------------LR--R-----SM------IH------E----------R    128
06_cpx.ML.95ML127  ------------------N-------Y--HI-K--KR-T-----D-N--KV---------C..-E---------N----K------------LR--R-----G-------MH------E------I--Q-    128
06_cpx.ML.95ML84   ------------------NA------Y--N--K--K--L-----D-N--K----------C..-E--V------N----E------------L-K-R-----S-------MH------E------I---V    128
06_cpx.SN.97SE1078 ------------------N-------Y--HI-K---K-A-----D-H--K----------K..-E--V------N----K------------L---R-----S-------MH-----TE----E-I----    128
10_CD.BFL061       -------L-------------N----------R--N--S-------R--K----------E..-S-I-K----------E----H-------L------I--G-------MH-----AE------I--QV    128
10_CD.BFL071       -------L-------------N--------I-RR-K----------R--K----------..---IVK----------E----N-------LR-S------G-------MH-----A-------I--Q-    128
10_CD.BFL110       -------L-------------N----------R-----Y-------R--K----------E..----VK------P--KE----H-A-----LR--------G-------MH-----AE------IV-Q-    128
11_cpx.CM.97CM-MP81 --------------------H----------K---R-M-------R--K-----------..-D--V----------E------------M---R-----G-------IH------E----E-I---R    128
11_cpx.FR.99FR-MP12 --------------------H--------I-K--KR-L-------R--K----I----R-..-E--V-------------------------L---R-----G-------MH-----T-----T-I--YR    128
11_cpx.FR.99FR-MP13 --------------------H-------HI-K-TT--L-------R--K----I-------..-K--V------Q----E----H-------L---R-----G---H---MH-----AE------I--QR    128
11_cpx.GR.GR17     -------------A-------N----------K--K--L-------R---V---------EE..DM--V----------E------------R---R-----------MH-----AE------I---R    128
CPZ.CD.CPZANT      I-K-$KIEMI--------NI--------IWETKVLKP-K------ND--KKGE------PTLDKK--V-VF---QC---P----H-------CGK-I------T---M--Q---P----Q-V-Q-I--ER    129
CPZ.CM.CAM3        -----L-----------VKA-N----C---KTK--KD---------S--KV---I---IEK..-K--V------MP-QHA----H-------CGS--------M--K----Q-----AA------V--KQ    128
CPZ.CM.CAM5        --------V---------K--N----T---RTK--KK---------DN-KV--------DT..-K--V------T--QHA----H-------CGS-K------T--R----H-----AA------V--KQ    128
CPZ.GA.CPZGAB      ------------------K--N----Y-I-R-K-----------DH-N-KVA--I---FR-.YSK-IV----A-SP---A----H----Q--LGS-V-----FT--R---SQ-----AET---R-I--QL    129
CPZ.US.CPZUS       -------L---------VKA-N----Y---R-K--K----------TN--VN------IDK..-KI--I-----K---QA----H-------QGT-K--I---M--K----------TA------V--RP    128
```

HIV-1/SIVcpz proteins

```
                                                   vpr cds ->
B.FR.HXB2          VSPRCEYQAGHNKVGSLQYLALAALITP.KKIKPP...LPSVTKLTEDRWNKPQKTKGHRGSHTMNGH$   192
CONSENSUS A        -s-r-e-qa--nk---l-----k--Vt-.k?tk--...---vrk-a---wn-pq--rgh-gsht-ngc-   189
A.BY.97BL006       -----N-P--------------K--V--.TRER--...----RX-----X------R-R-XN----XC-   191
A.DE.AF200476      -R--------------------K--VA-.--T---...---LK--A----------R----NRS-----   192
A.KE.AF233689      ----------------A-----K--V--.TRK---...----R--A----------R----HRS---C-   192
A.KE.Q23           ---K------------------K--V--.--T---...----RI------------R-L-E------C-   192
A.SE.SE6594        -----------K----------R--VA-.R-T---...----RI-A----------RDP-E------C-   192
A.SE.SE7253        ---I---HT-------------K--V--.TRT---...----R--A----S-----R----------C-   192
A.SE.SE7535        ---S---HT---Q---------K--V--.--TR--...----RI-A------SR--R-P--------C-   192
A.SE.SE8538        ---S------------------K--V--.-------...----K-------------R----N---H-Y-   192
A.SE.SE8891        ----------------------K--V--.---R---...----R--A----------------------   192
A.SE.UGSE8131      ---S----T-------------K--V--.-RT---...----R-----------------------C-   192
A.UG.92UG037       -----D--T-------------K--V--.SRM---...----K--A----------R-R-E------C-   192
A.UG.U13-2         --------T-------------K--V--.TRT---...----R-------S-----R-P------S-C-   192
A2.CD.97CDKFE4     ----------------------K--VA-.TRA---...----K--V----------R----------C-   192
A2.CD.97CDKS10     ----------------------K--VV-.N-AR--...----R--V-------R--R----------C-   192
A2.CD.97CDKTB48    -R--------------------R--VAS.TRT---...----R--V----------R----------C-   192
A2.CY.94CY017-41   ----------------------K-VVAS.TRT---...----R--V----------------------C-   192
CONSENSUS B        vs-r-eyq---n---s------ta-itp?-kik--...----?----------qk----rgs--mn---   191
B.AU.MBC18         -R-S------------------S-----.------...----A------------------------   192
B.CN.RL42          ---S-D----------------T--T--.-NR---...----------------R----------S---   192
B.DE.D31           G--S---R--------------------.------...----A----------R---------------   192
B.ES.89SP061       ----------------------T--VR-R------...----------------------K-N-------   193
B.FR.vi02011A1H    -----------S----------------.------...----V----------------KE--------   192
B.GA.OYI           -------P------------------K-.------...-----------------------------   192
B.GB.CAM1          ----------------------T---A-.------...----R--------------------------   192
B.JP.D70887        ----------------------T-----.--T---...----K-----------------N-------   192
B.JP.PT7-6         -------R-------F------------.---R---...---------------------E--------   192
B.KR.WK            -R-K------------------T-----.------...----R--------------------------   192
B.NL.3202A21       --------------------------K-.------...-----------------------------   192
B.TW.TWCYS         ----------------------T--VQ-.------...----V----------------------H---   192
B.US.1-2           ----------------------T-----.-R----...----R----------------------H---   192
B.US.AF019528      -----------------------V-V--.-RR---...----K--------------------------   192
B.US.BC            ----------------------------.-R----...-----------------------------   192
B.US.JRCSF         -----------S----------T---K-.------...----K--------------------------   192
B.US.JRFL          ----------------------T---KS.------...----K--------------------------   192
B.US.MN            ---I--F--------P------T-----.------...----K--------------------I----   192
B.US.RF            -------------------------T--.------...----K--------------------------   192
B.US.SF2           ----------------------------.--T---...----K--------------------------   192
B.US.WC001         ---S------------------V--VA-.------...-----------------------------   192
CONSENSUS C        -ip--d------kv--------T---K-.-ki---...---vk--V--r--k---tRgRr-n-t-n---   192
C.BR.92BR025       --S--D----------------T---K-.------...----K--V----------RDR--N-------   192
C.BW.96BW01B21     -I---D----------------T---K-.--R---...----K--V-----D----R-R----------   192
C.BW.96BW0402      ----------------------T---K-.--R---...----R--V-----E----R-R--N-------   192
C.BW.96BW0504      -I---D------Q---------T---K-.-QR---...----R--V----------R-R--N-------   191
C.BW.96BW1104      -I---D----------------T---K-.------...----R--V----------R-R--N-------   191
C.ET.ETH2220       -----D----------------T---K-.--A---...----S--V--K-------R-R--N-------   192
C.IN.94IN11246     -I---D-------I--------T---K-.------...---IK--V-----N---IR-R--N-------   192
C.IN.95IN21068     -I---D----------------T---K-.------...---IK--V-----N---IR-RK-N-I-H---   192
CONSENSUS D        -spr-?-q---n----l-----t--ia-.kki---...-----r--t-----k-q?-kghr--h------   189
D.CD.84ZR085       ----------------------T---A-.--R---...----K----------RQ---R----------   192
D.CD.ELI           ----------------------T---A-.-Q----...----R-----------Q-R------------   192
D.CD.NDK           ---S--------------------A-.------...----R----------------R----------   192
D.CD.Z2Z6          --H--------S----------T--VA-.------...----R----------------K---------   192
D.UG.94UG1141      -Y---N----------------T--V--.R-----...----G-----------R--------------   192
D.UG.U18-0         -----D----------------T---K-.------...----R-----------Q--------------   192
D.UG.U25-6         -T---N----------------T---A-.------...----R-----------R--------------   192
D.UG.U36-0         -T---N----------------T---K-.------...----R----------R--A-----------   192
D.ZR.AF233690      -----D-P--------F-----T--LE-.--R---...----Q--V-----N----RA----R------   192
F1.BE.VI850        I----N----------------T---A-.E-T---...----Q--V--------E-R------------   192
F1.BR.93BR020-1    I----N----------------T---A-.--T---...----Q--V---------R---E--------   192
F1.DE.AF200475     -----N----------------T--VA-.--T---...----Q--V--------E-R------------   192
F1.FI.FIN9363      I----D----------------T--VS-.--A---...----K--V--------EIR------------   192
F1.FR.MP411        I----N----------------T---A-.--T---...----K--V--------E-R------------   192
F2.CM.MP255        -----N----------------T-----.------...----R--V-----N----------------   193
F2.CM.MP257        -----N----------------T-----.------...----R--V-----N----R------------   193
```

```
                                                        vpr cds ->
B.FR.HXB2           VSPRCEYQAGHNKVGSLQYLALAALITP.KKIKPP...LPSVTKLTEDRWNKPQKTKGHRGSHTMNGH$  192
G.BE.DRCBL          -R-S---P--------------KV-VA-.TRRR--...----R--------E----R---ENP------  192
G.CG.CNG30          -R-S---PT---Q---------K--V--.PR----...----R-------------R-----R-R---W  193
G.FI.HH8793-1-1     -R--------------------K--V--.T-TR--...----R---------S---R---E-P------  191
G.NG.92NG083        -------P-------------SK--V--.TRKR--...----G--A----------RD--ENP------  192
G.SE.SE6165         ------------Q---------KV-V-S.-RSR--...-----E-A----------R---ENP------  192
H.BE.VI991          ---I-D-----R----------T---S-.-RT---...----R--V----------R------------  192
H.BE.VI997          ------------Q---------T--VA-.--T---...----K--V--G-------R---------R--  192
H.CF.90CF056        -R---N-P---KQ--T------T--VA-.-------...----R--V----------R------------  192
J.SE.SE7022         -----D----------------T---K-.-RR---...----Q--V----------RD--E--------  192
J.SE.SE7887         -----D----------------T---R-.-RR---...----Q--V----------T---E--------  192
K.CD.EQTB11C        -----------TQ---------T---A-.--T---...V---Q--V----------R--------S-Q-  192
K.CM.MP535          ------C---------------T--VA-.RRP---...V---K--V----------R-----Q------  192
N.CM.YBF106         -L---X-P---XQ--T-----XT-XVGV.--RR---...-------------ER---Q----NPI---P-  192
N.CM.YBF30          -L-----P---KQ--T------T-WVGA.--R---...-------------EH--MQ----NPI-----  192
CONSENSUS O         vLtk---pa--SQ--T--l---ravvKa.krnk--...----Qk----r-n?hlrIrDQle-hs-n---  188
O.CM.ANT70          -LTK---PT--SQ--T--L---R-VVKA.RSR---...----Q---------HLRIRDQLK-PS-----  192
O.CM.MVP5180        -LTK---L---SQ--T--F---K-VVKV.-RN---...----QR---------W-IRDQL---S-----  192
O.FR.HIVY16019      ILVK---P---SQ--T--L---R-IEKA.-RNR--...----Q-----K-S-HLMIRDQL---P-----  192
O.FR.HIVY16020      ILTK---S---SQ--T--L---R-VVKE.-RN---...----Q---------HLRIRDQLE--S-----  192
O.FR.HIVY16021      -LTK---S---SQ--T--L---R-VVKE.-RN---...----Q---------HLRI-DQLE--S-H---  192
O.FR.HIVY16022      -LTK---P---SQ--T--L---R-VVKA.-RS---...----Q--------RHLRIRDQPE--P-----  192
O.FR.HIVY16023      -LTK---T---SQ--T--L---R-VVKA.-RN---...----Q--------RHLRIRDQLE--S-----  192
O.FR.HIVY16024      ILSK---P---SQ--T--L---KVVVKA.RRN---...----Q-----K-SRRLRIRDQLE--S-----  192
O.SN.99SE-MP1299    -LT----P---SQ--T--L---RVVVKE.-RN---...----Q-------SRHLRIRDQLE--S-----  192
O.SN.99SE-MP1300    -LT----S---SQ--T--L---RVVVKE.-RH---...----Q-------SRHLRIRDQL---S-----  192
CONSENSUS 01        -Rr--e-Ps-------------K--ttp.-rir--...----rkl-----nkp--?rg--EnP--n---  191
01_AE.CF.90CF11697  -R-----PT------------K--A--.--TR--...----R-------------R---ENP------  192
01_AE.CF.90CF402    -R-----P--------------K--TKT.--T---...----R-----------------E-P------  191
01_AE.CF.90CF4071   -R---D-PE------------K--A--.---R---...----R-------------R---ENP--S---  192
01_AE.DE.K08DE      -RH--D-PS------------K--TA-.-R-R---...----K-------S----IRD--ENP------  192
01_AE.TH.93TH057    -RR----PS------------K--T--.-R-R---...----K------------IW---ENP------  192
01_AE.TH.95TNIH047  -RR----PS------------K--T--.-R-R---...----K---------H--.GD--ENP------  191
01_AE.TH.CM240      -RR----PS------------K--T--.-R-R---...-----.EI----------.R---ENP------  190
01_AE.TH.TH022      -RR----PS------------K--T--.-R-----...----R--------E---IR---EYP------  192
CONSENSUS 02        -rpr---q---nk---------K--Vt-.trtk--...----k--a----ne--k-r--r-sRs---hW  190
02_AG.CM.97CMMP807  -R--------------------K--VI-.TRTR--...----K-----------R----K--RS----W  193
02_AG.FR.DJ263      -R---------KQ---------K--V--.A-T---...----K--A----------R----NRS-----  192
02_AG.FR.DJ264      ------------Q---------K--V--.V-T---...----K--A----------R----NRS---Q-  192
02_AG.GH.G829       -R--------------------K--V--.TRK---...----R--A-----E----R-----RP---R-  192
02_AG.NG.IBNG       -R--------------------N--VA-.T-T---...----R--A----KE----R-----RP-----  192
02_AG.SE.SE7812     -R---------T----------K--V--.TRT---...----K--A-----E----R-----RS-----  192
02_AG.SN.98SEMP1211 -R-K------------------K--V--.TRR---...----K--A-----E----R----NRS-----  192
02_AG.SN.98SEMP1213 -RS----P--------------K--V--.IRR---...----K--------E----R-----RS---R-  192
03_AB.RU.KAL153-2   ---S---R-----------------R--.-------...----------------R--D-------S---  192
03_AB.RU.RU98001    ---S---R-----------------R--.-------...----------------R--D-------S---  192
04_cpx.CY.94CY032-3 -------------------------S-.--T---...----K--V----------R-R-ENQI-----  191
04_cpx.GR.97PVCH    ------------E-------------S-.--T---...----K--V----------R-R-ENQI-----  192
04_cpx.GR.97PVMY    -------------P------------S-.--T---...----K--V------S-----R-E--I-----  192
05_DF.BE.VI1310     ------F---------------T-----.--T---...----R---------------R--N-----Y-  192
05_DF.BE.VI961      ---------------------T------.Q-----...----R-----------R------C-----Y-  192
06_cpx.AU.BFP90     -----D----------------T---K-.E-R---...----Q--V----------R---E--------  192
06_cpx.ML.95ML127   ---Q-D----------------T---K-.--R---...----Q--V----------RD--E--------  192
06_cpx.ML.95ML84    ---K-D----------------T---K-.R-R---...----Q--V----------RD--EC-------  192
06_cpx.SN.97SE1078  -----DFE--------------K--VKT.-RR---...----Q--V-----------D--E--I-D---  192
10_CD.BFL061        -T---N-----KQ---------T-----.--T---...----R-----------R--------------  192
10_CD.BFL071        -T---NF-----Q---------T-----.--TR--...----R--------E--R--------------  192
10_CD.BFL110        -T---N------Q---------T-----.--T---...----R-----------R-----E--I-----  192
11_cpx.CM.97CM-MP81 ------------Q---------K--V--.TRA---...----R--A----------R----N--A--C-  192
11_cpx.FR.99FR-MP12 ----------------------K--VA-.TRT---...----R--A----------R----N--L--C-  192
11_cpx.FR.99FR-MP13 -------P--------------R--V--.TRAR--...----R-----------------..AI--C-  190
11_cpx.GR.GR17      -------P--R-----------K--V--.-RT---...-----R-------------R-R--N--V--C-  192
CPZ.CD.CPZANT       ILTY-H-KK--SQ--T-----FCKILEFRGYP-G-RRQF--LSI---------RRMR---ENQ$-S-C-  196
CPZ.CM.CAM3         -Y-K---P---QQ---------R-WVRV.G-K---...--------------RH--IRV-QE-------  192
CPZ.CM.CAM5         -Y-K---PT--QQX--------R-WVGV.Q-KT--...---IX--A------Y---R--QEN----X--  192
CPZ.GA.CPZGAB       -A-----KE--RQ-----F---K---SE.RRHR--...----A---------H-R--V-QENL-R----  193
CPZ.US.CPZUS        -I-K---PT---------F---K--VGQ.S-RR--...--------------C-----QPEN---S---  192
```

HIV-1/SIVcpz
proteins

HIV-1/SIVcpz proteins

```
                                 <- vif cds
                              amphipathic alpha helix
                                |-          -|                                H(S/F) RIG motifs
                   |-         oligomerization      -|                        |-  -|  |- -|      tat cds ->
B.FR.HXB2          MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIWLHGLGQHIYETYGDTWAGVEAIIRILQQLLFIHXFRIGCRHSRIGV...TRQRRAR...NGASRS$   97
CONSENSUS A        ---a----------yne-m-e-led--h-a------p--h----h--n-y----E-----i-t------v-.---gcqhsri-I..?Irg--v-...d-sg---   96
A.KE.Q23           ------------------M----DD--H--------P----------N------E-----T--------V-.-----Q-----I...IPG--G-...--PG---   96
A.SE.SE6594        --------------Y---M-----D--H--------P----------N------E-------T------V-.-----Q-----I...I-G--V-...--S----   96
A.SE.SE7253        --------------Y---M-D---D--H--------P--Q-------N-H----E-----V-T------V-.-----------I...I-G--V-...D-SG---   96
A.SE.SE7535        ---V----------Y---M-----D--Q--------P----------D------E-------T------V-.-----Q-----I...I-G--V-...--PG---   96
A.SE.SE8538        --------------Y-A-I-D------H--------P----------D------E-----T-T------V-.-----H-----I...I-G--V-...D--G---   96
A.SE.SE8891        --------------Y---A---V----H--------P----------N------E-------T------V-.-----Q-----I...I-R--V-...D-S----   96
A.SE.UGSE8131      --------------Y---M-----D--H--------P-------Y--Y------E-----V-T------V-.---X......-I...I-G------...D-S----   90
A.UG.92UG037       --------------Y---M-D---D--H--------P----------H------E----------------V-.-----Q-----I..NI-G--V-...D-SG---   97
A.UG.U455          --------------YA-$A--------H-D------Q----------N------E-----------------.-----Q-----I...IPG--G-...-------   95
A2.CD.97CDKFE4     --------------Y---V--------Q-------GQ----------N------E-----T-T------V-.--M--H------...----------...--S--P-   96
A2.CD.97CDKS10     ----------P-------V--------Q--------Q-------L--S------E-------T------V-.-----------I...X-R--V-...X------   96
A2.CD.97CDKTB48    ----------P-------V--------Q--------P----------N-----RE-----V--------V-.-----Q-----I...IPR--V-...D----P-   96
A2.CY.94CY017-41   --------------Y---M--------Q--------H-------Y--N------E---V---Y------V-.-----Q-----I...I-R--V-...D----P-   96
CONSENSUS B        ---a--d---q-e-Y-e-t---l----s--------i--hs-g-h--e------a----i--i--------.-----q---igI...irqr-a-?..--a--s-   96
B.AU.MBC18         ------------Q-Y---A-----------------V-------Y---------E-----------------.-----Q-----I...F----------...--V----   96
B.AU.VH            --------------Y------------S-------------------N-----------L--S--------.-----------I...NP----------...--S----   96
B.CN.RL42          --------------Y------------R--------------------------------T--------.-----Q-----I...IQ----------...-------   96
B.DE.D31           --------------Y------------S------------S-------------T----L--M--------.-----Q-----I...----------...-------   96
B.ES.89SP061       --------------Y-------------------------S-----------------------------.-----Q----DI...L-R----------...-------   96
B.GA.OYI           --------------Y------------S------------S-E-Y---------E-----------------.-----Q-----I...----------...-------   96
B.GB.CAM1          ---T--------Q-Y------------S--------P--LA-------------P----L--S--------.-----------I...IQ---S-...-------   96
B.GB.I4663         --------------Y------------S------------S-----------------L-----------.-----Q-----I...I-RS----------...-------   96
B.JP.PT1-4         --------------Y---------------------T---S-----------------------------.-----------I...IQR----------...-----P-   96
B.KR.WK            --------------Y-------S----S--------P---S-------------T-----------------.-----H---L-I...I----------...-------   96
B.NL.3202A21       ------N-------Y------------R--------P-----------------------------------.-----Q-----I...IQ-----R..-------   97
B.TW.TWCYS         ----------P---Y-A----------R--------P---S---Y--------------------------.-----Q-----I...I----------...--S----   96
B.US.JRCSF         --------------Y-------------------------S---Y--------------------------.-----------I...----------...-------   96
B.US.MN            --------------Y-Q-A-----------------------------------------------------.-----------I...I----------...-------   96
B.US.RF            --------------Y------------S--------L---S-----------------------------.-----Q-----I...----------...-------   96
B.US.SF2           --------------Y------------R--------P---S---Y--------------------------.-----Q-----I...IQ-----R..-------   97
B.US.WC001         --------------Y------------S--------L---S-----------------------------.-----------I...I----------...-------   96
CONSENSUS C        --qa----------Y-e-t--------q-------rPw--g---y?--------t---ai--i--q-l-i-.-----Q-----I...Lrq--a-...--as---   95
C.BR.92BR025       --------------Y------------R--------P-----------------T----------R---V-.-----Q-----I...L----------...-------   96
C.BW.96BW01B21     --RP----------Y------------Q--------P-------Y---------T---NM---------V-.-----Q-----I...LQ----------...-------   96
C.BW.96BW0402      --R-----------Y------------R--------P---S---Y---------T----L--T--------.-----Q-----I...L----V-...--TN---   96
C.BW.96BW0504      --------------Y------------Q-------GP-------YV--------T---TL----------.-----Q-----I...L----T-...-------   96
C.ET.ETH2220       --------------Y---A--------Q--------P---N---Y---------S----L--T----M---.-----Q-----I...L----------...-------   96
C.IN.93IN101       ---S----------Y------------Q--------P-------YV--------T-------L--------.-----Q-----I...L----------...-------   96
C.IN.94IN11246     ---S----------Y------------Q--------P-------YV--------T---T----------.-----Q-----I...LPR----------...---N---   96
C.IN.95IN21068     ---S----------Y-A-A--------Q--------PC------YV--------T---M----------.-----Q-----I...LP----------...-------   96
CONSENSUS D        -----ed-------Y---t--------S-------ri---s---?---------a----i----------.-----q----gI...tr---a-...--s---Q   96
D.CD.84ZR085       -----A--------Y------------S------------S---Y--------------------------.-----Q-----I...-P----------...--S----   96
D.CD.ELI           -----A--------Y---A--------S------------S------------V----------------.-----Q-----I...I----------...--S----   96
D.CD.NDK           --------------Y------------S------------S-----------------------------.-----Q----SI...----------...--S----   96
D.CD.Z2Z6          --------------Y------------S------------S---Y--------------L----------.-----Q-----I...----------...--S----   96
D.JP.PT14-4        --------------Y------------S-----------------------------------------.-----------I...-Q----------...-------   96
D.UG.94UG1141      ------N-------Y------------S-------GP---S---Y---------T----------------.-----------I...N----I-...-------Q   97
F1.BE.VI850        -----G--------Y---A--I--------------P----------N------E----------------.-----------I...VP---V-...-------   96
F1.BR.93BR020-1    --------------Y-----D--------------P---S------N------E----------------.-----------I...-----V-...--T----   96
F1.FI.FIN9363      -----G--------Y---A----------------PF---I--Y--N------E------K----------.-----H-----I...I----V-...-------   96
F1.FR.MP411        -----G--------Y---------------I-----P-------Y--N------E----------------.-----Q-----I...I----V-...--T----   96
F2.CM.MP255        ---P----------Y--------------------EY------Y--N------R---TM----------.-----H-----I...I----L-...--S----   96
F2.CM.MP257        ---P----------Y------------H--------E----------N------E----------------.-----H-----I...I----I-...--SG---   96
G.BE.DRCBL         --R---------------A--------H--------P-------L--N------E----------------.-----Q-----I..IPQR--V-...D-SG---   97
G.CG.CNG30         --------------YQ-------------------WT------Y--N----S-E--I-------------.Y----Q-----I...NV--GG-...-------   96
G.FI.HH8793-1-1    ---V----------Y---A-----G----------P----------N------E----------------.-----Q-----I...-LR--V-...--P----   96
G.NG.92NG083       --------------Y--------------------P-------Y--N------E----------------.-----Q-----I...-P---V-...D-PG-P-   96
G.SE.SE6165        --------------Y---A----------------L----------N------E----------------.-----Q-----I...-PR--V-...D-PG---   96
H.BE.VI991         --------------------------S-----Y--P---D------N------E-------T------A-.-----Q-----I...-----V-...--P----   96
H.BE.VI997         -----------------$--------S--------D---Q------N------G----L--T--------.-----H-----I...I----V-...--P----   95
H.CF.90CF056       ------------------------I----------V---Q------N------V----L--T--------.-----Q-----I...-----V-...--P----   96
```

```
                                 <- vif cds
                            amphipathic alpha helix
                              |-              -|                                H(S/F) RIG motifs
                     |-      oligomerization      -|                          |-  -|  |- -|      tat cds ->
B.FR.HXB2            MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIWLHGLGQHIYETYGDTWAGVEAIIRILQQLLFIH#FRIGCRHSRIGV...TRQRRAR...NGASRS$   97
J.SE.SE7022          --------------YH--------------------P-------YV-S------E----V-----------.-----H-----I...IP---G-...-------   96
J.SE.SE7887          -------H------YH--------------------P---S---Y--S------E----------------.-----H-----I...IP---G-...-------   96
K.CD.EQTB11C         --------------Y----I-I-----R--------P---D------NN-----E----------------.-----H-----I...L----GG...-------   96
K.CM.MP535           --------------N------I-----R--------P---N------T------E-L--------------.-----H-----I...IP---G-...--S----   96
N.CM.YBF106          --R----A------Y---A----------------------------N------E----------------.Y----Q-----I...-P---.-...-------   95
N.CM.YBF30           --R----A------Y---A----------------------------N------E----------------.Y----Q-----I...-P---.-...--T----   95
CONSENSUS O          --q-ped---Ar--f-e-a-el---l-a--v----rp--qal-qY--et-----v--M--i-i--qll-a-.f----Q-----InpsntRG-G-R..--S--s-  100
O.CM.ANT70           ------N---AK--F---A--------A--------P---A---Y---------V--M-----------T-.Y----Q-----I...NPRG-G-R..--S----   97
O.CM.MVP5180         ----L-N---A---F------------E--------P--QAC--Y---------E--M-----------T-.Y----Q-----ILPSNTRG-G-R..--S----  100
O.FR.HIVY16019       -----D----A---S---A--------A-------KP--QA---Y---------E--M-----------A-.-----Q-----INPSNTRG-G-R..--S----  100
O.FR.HIVY16020       ----------A---F------------T--------P--MS---Y--D------V--M-------L---A-.-----Q-----INPSNTRG-G-R..--S--P-  100
O.FR.HIVY16021       ----------A---F-A----------A-------GH---S---Y---------V--M-------LI--A-.-----Q-----INPSNTRG-G-R..--S--P-  100
O.FR.HIVY16022       ----------A---S---A-K------A--------P---A---Y---N-----L--M--------------.Y----Q-----INP...RG-G-R..--S----   97
O.FR.HIVY16023       ----------A---F---A--V-----A--------H--QA---Y---------V--M-----------A-.-----Q-----INPSNTRG-G-R..--S----  100
O.FR.HIVY16024       ----------A---F------I-----A--------Q--QA---Y---------V--M--T-----I--A-.-----Q-----INPSNTRG-G-R..--S----  100
O.FR.HIVY16025       --R-------A---F---A--I-----A--------Q--QA---Y---------V--M--T-----I--A-.-----Q-----INPTNTRG-G-R..--S----  100
O.FR.HIVY16026       ------N---A---F---A--------E--------P--QAC--Y---------E--M---------M---.Y----Q-----IIPSNTRG-G-R..--S----  100
O.FR.HIVY16027       ------N---A---F---A--T---I-A--------P--QS---Y---------E--M---------I-A-.-----Q-----ITPSNARG-G-R..--S----  100
O.FR.HIVY16028       ----------A---F------------A--------P--QA---Y--D------V--M----L--LMI-A-.-----Q-----INPSNTRG-G-R..--S--P-  100
O.FR.HIVY16029       ----------A---F------------A--I-----P--MS---Y---------V--M-------L-I-A-.-----Q-----INPSNTRG-G-R..--S--P-  100
O.FR.HIVY16030       -----Q----A---F---A--T-----A--------H--QA---Y---------V--M--T-----I--A-.-----Q-----INPSNTRG-G-R..--S----  100
O.FR.HIVY16031       --R-------A---F---A--T-----A-------QP--QA--NY---------V--M-----------T-.-----Q-----INPSNTRG-G-R..--S----  100
O.SN.99SE-MP1299     ----------A---F---S--------A--------P--QA---Y--D------V--M-----------A-.-----Q-----ITPSNTRG-G-R..--S----  100
O.SN.99SE-MP1300     ----------A---F---S--------A--------P--QA---Y--D------V--M-------H---A-.-----Q-----INPSNTRG-G-R..--S----  100
CONSENSUS 01         --q---dq--q-e-y---t--------n--------p--hg--qh--Nn-----E--e--i-i------V-.-----Q-----I...ipgr-g-...--aG---   96
01_AE.CF.90CF11697   --------------Y------------I--------P-------Y--N------E-------T------V-.-----Q-----I...I-G--G-...---G---   96
01_AE.CF.90CF402     --------------Y------------T--------P------K---N------E--------------V-.-----Q-----I...I-GT-G-...---G---   96
01_AE.CF.90CF4071    ------------------------------------P--Y-------N------E--------------V-.-----Q-----I...I-X--X-...---G---   96
01_AE.TH.93TH902     ----------P------$I-----------------T---D------NN-----E--A--T--------V-.-----Q-----I...IPG--G-...---G---   95
01_AE.TH.94TH702     ----------P------$I-----------------T---D------NN-----E--A--T--------V-.-----Q-----I...IPG--G-...---G---   95
01_AE.TH.95TNIH047   ------G.------Y--------------------S-------Y--NN-----E-------F------V-.-----Q-----I...IPG--G-...---G---   95
01_AE.TH.CM240       ------E.------Y--------------------P--Q----Y--NN-----E-------M------V-.-----Q-----I...MPG--G-...--TG---   95
01_AE.TH.TH022       --R---------V-Y--------------------T----------NN-----E--------------V-.-----Q-----I...IPG--G-...---G---   96
CONSENSUS 02         --r-----------f----l----e--h---r----P-------h--e------E--?-----------v-.f----q-----I...irG--G-...---g---   95
02_AG.CM.97CMMP807   --------------F------------H--------P----------------E--R-----------V-.-----Q-----I...-LG--G-...---G---   96
02_AG.FR.DJ263       --------------F------------H--------P----------D-----E--Q-----------V-.-----Q-----I...I-G--G-...---G---   96
02_AG.FR.DJ264       --------------F----V-------H--------P----------D-----E--Q-----------V-.-----Q-----I...I-G--G-...--------   96
02_AG.GH.G829        --R-----------S------------H--------P----------------E--------------V-.-----H-----I...I-G--G-...---N---   96
02_AG.NG.IBNG        --R-----------S------------H--------P----------------E--K-----------V-.-----Q-----I...IQG--G-...---G---   96
02_AG.SE.SE7812      --R-----------F------------H---T----P----------------E--G-----------V-.-----Q-----I...I-G--G-...---G---   96
02_AG.SN.98SEMP1211  --R-----------F------------H--------P----------------E--G-----------V-.-----Q-----I...IPG--G-...---G---   96
02_AG.SN.98SEMP1213  --R-----------F---------A--Q--------P-------Y---------E--V--------------.Y----Q-----I...-PG--G-...---G---   96
03_AB.RU.KAL153-2    --------------Y------------S--------V---S---Y--------------------------.-----Q-----I...IQR-----..-------   96
03_AB.RU.RU98001     --------------Y------------S--------V---S---Y--------------------------.-----Q-----I...IQR-----..-------   96
04_cpx.CY.94CY032-3  --------------N--------------------P----------N------E----------------.-----Q-----I..TPQR--G-..XD------   98
04_cpx.GR.97PVCH     --------------N--------------------P-------Y--N------E-----T-M------V-.-----Q-----I..ILQR--G-..XD-G--FS   99
04_cpx.GR.97PVMY     ---V----------------T------I--------P---S---Y--Y-----RE-------T--------.-----Q-----I..LPQR--G-GXGD------  100
05_DF.BE.VI1310      --------------Y---I--------H--------P---S---Y--N------E-----V----------.------------I...L----I-....--S----   96
05_DF.BE.VI961       ------------V-Y---V--------S--------P---D------N------E---V-T--------T-.-----H-----I...------------S--P-   96
06_cpx.AU.BFP90      --------------Y------------S--------P---NI--Y--S------E----LL-T--L---V-.-----H-----I...IP---G-...-------   96
06_cpx.ML.95ML127    --------------Y------------S--------P--LN---Y--D------E----------------.-----H---V-I...I----G-...-------   96
06_cpx.ML.95ML84     ------------V-Y------------S--------P---N---Y--------E----------------.-----Q-----I...M----G-...-------   96
06_cpx.SN.97SE1078   ----------------G----------S--------P--LD------N------E----L-----------.-----H-----I...IL---G-...-------   96
10_CD.BFL061         --------------Y--------------------N------------T----------------.-----P-----I...V---------...-T-T--   96
10_CD.BFL071         --R-----------Y---A--------T--------M---S------------T----------------.-----Q-----I...I-------..-------   96
10_CD.BFL110         --------------Y---A--------S-----Y------N------------T----L-----H-----.-----Q-----I...M-------..-------   96
11_cpx.CM.MP818      --------------YS--M--------H--------L----------D-----E--------------V-.-----Q-----I...I-G--G-...-------   96
11_cpx.FR.MP1298     --------------YT--M--------H--------P----------N------E---V----------V-.-----Q-----I...I-G--G-...-----P-   96
11_cpx.FR.MP1307     -------------..G-Q-M--------H--------P------L---N------E-----T----------.-----K-----I...IQG--G-...------P-   94
11_cpx.GR.GR17       --------------YS--M--------H--------P----------N------E--------------V-.-----Q-----I...I-G--G-...---R-Y-   96
CPZ.CD.CPZANT        -----Q-E------M---L--T---I---------QPT-QH--NWV-AN---S-R--QVL-IL--KA--T-.Y-H--A-----...............H-PR-R-   88
CPZ.CM.CAM3          ----------P---Y---A----------------A-------F--D----------------------..--L--Q-----I..IPGR--..-------   95
CPZ.CM.CAM5          ---I------P---Y---X--I-----T--L-----P---R---Y---------I----------H---V-.--L--Q-----I..LXER--.....--T--X-   95
CPZ.GA.CPZGAB        ----------P---YQ--A--T--------------P---Q---F--D------V----------H------.--L--Q-----I..LPQR--S.....--SN---   96
CPZ.US.CPZUS         ---V------A---Y---A-DV-----E--L-----P--Q----Y--S------E----------------.Y----N-----ISLAR-TPQG-R..--S----  100
```

HIV-1/SIVcpz proteins

HIV-1/SIVcpz proteins

```
                                                 intramolecular
                                                disulfide bonding
                                                |-                -|  rev cds ->|-  nls  -|         exon \ / exon
B.FR.HXB2          MEPVD...PRLEPWKHPGSQPKTACT.NCYCKKCCFHCQVCFI.TKALGISYGRKK..RRQRRRAHQNSQTHQASLSKQ.PTS.QPR.GD.PTGPKE$KKKVERETETDPFD$   100
CONSENSUS A        -d---...-nl---n----q-t-p-n.kcy--k--y--?v--l.n-G------r--..-kq--gtP-snKD--npipk-.piP.qtQ.g?.st-pe-Sk--v-SKaet-Rf?-    98
A.KE.93.Q23        -D---...-N----N------T---N.K-------Y------L.N-G---------..-----G-P-S-KD--N-IQ--.-IP.-TQ.-V.S-----S---M-SKA-A-R---   101
A.SE.SE6594        -D---...-N----N------S---N.K----I--Y--P---L.N-G----------..-TR--G-P-SNKD--NPI---.-IP.-TQ.---.S---E-S-----SKA---R---   101
A.SE.SE7253        -D---...-N----N------T-P-S.K-------Y--I---Q.N-G----------..------TP-S-KD--NPIP--.-IP.-TQ.-V.S---E-S-----SKA-A-R---   101
A.SE.SE7535        -D---...-N----N------T---N.K-------Y---A--L.N-G----------..-K---G-P-SNKD--NPIQQ-.SIP.-TQ.-I.----E-S-----SKA---R---   101
A.SE.SE8538        -------..-NI---N------T-P--.K-------Y--PA--L.N-G----------..-K---GTP-GNKD--NPIP--.S-P.-TQ.-I.S---E-S-----SK--A-R---   101
A.SE.SE8891        -D---...-N----N------A---S.A-------Y--P---L.N-G----------..------TP-S-KD--NPVP--.SIP.-AQ.---.S---E-S-----SK----R-A-   101
A.SE.UGSE8131      -D---...-N----N------T-P-N.KYF--R--Y--L---Q.H-G----------..-----GTP-S-KD--DPIP--.SIP.PAQ.-I.----E-S-----SKA---R-A-   101
A.UG.U455          -------..-N-----------T---S.-----V--W---L--L.K-G------K--..-KP--GPP-G-KD--TLIP--.-LP.-SQ.RV.SA-QE-S-----SKAK--R-A-   101
A.UG.UG273A        -D---...-N----N----K-T-P-S.K----V--Y--P---L.K-G----------..-S---GTP-SNKD--NPIP--.-IP.-TQ.-I.S-----SQ----SK----R-A-   101
A.UG.UG275A        -D---...-N----N------T-P-N.K-------Y--T---L.N-G----------..-K----TP-CNKD--NPIP--.-IP.-TQ.-V.S---E-S-----SK----R-T-   101
A.UG.92UG037       -D---...-S----N--------P-N.K----V--Y---C--L.N-G----------..-KP--GTP-SNKD--NPIP--.-IP.RTQ.---.S---E-S-----SK--A-RYA-   101
A2.CD.97CDKS10     -------..-K----N----------N.K-------Y---L--L.N-G----------..--P--GPS-S-KD--NPIP--.SLP.-TQ.RV.S--TE-S--E--SK--P-R---   101
A2.CD.97CDKTB48    -------..-N----N----------N.--------Y---R--L.N-G----------..--P--GPD-GNTD--NPVP--.SLP.-TQ.R-.S---E-S-----SKA---R---   101
A2.CY.94CY017.41   -------..-K----N-----------.K----R--Y---L----.N-G----------..--P--KPSPSNKD--NPIP--.SLP.-AQ.RV.----E-P--E--SKA---R---   101
CONSENSUS B        -e---...-r----k------k-act.n----k--l------t.t-g---sy----..-----raPqdsqt--vs-s--.pas.qpr.---.pt-pk-skk---r--etd-vdQ   102
B.AU.MBC200        -------..------------------.T---------------.--G----------..-------P-D-E---V------.---.-H-.---.------P------------V--   101
B.CN.RL42          -------..-----------------N.--------L------M.K-G---F------..-------P-D-----------.-A-.---.---.-A----S-----S------R--   101
B.DE.D31           -------..------E-----------.---------------T.K-G----------..-------PED-----V------.-A-.---.---.------S-----T------I--   101
B.ES.89SP061       -------..-------------R----.----------------.--G----------..-------P-D-----V------.RA-.--G.---.------S------------VN-   101
B.GA.OYI           -------..----------------SN.-----R--L------T.K-G----------..-------P-D-K---V------.-A-.---.---.------S------------E--   101
B.GB.CAM1          -------..-------------R----.-----Q----------.--G----------..------TP-S-K----------.-A-.-FQ.---.------S-----G----H-G--   101
B.KR.WK            -------..----------------P--.K-------L------M.--G----------..-------P-DNKN--V------.----.RA-.---.---QE-S-E---K--VV--VT-   101
B.NL.3202A21       -------..-K--------------N.---------------T.K-G----------..------SP-D-E---V------.-A-.---.---.------S------------V--   101
B.TW.TWCYS         -D---...------------------.--------L------M.--------N----..------TP--------D----.---.---.---.Q--Q--ST-----------N-Q   102
B.US.WC001         -------..-S----------------.-----Y--L------T.R-G----------..-------P-D-----V------.-A-.---.---.------SQ----S------SA-   101
B.US.RF            -------..-----------------N.--------Y------L.--G----------..-----GPP-G-----V------.---.---.---.------S-E----------AVQ   102
B.US.MN            -------..------------------.T---------------T.K-----------..-------PED-----V--P--.-AP.-F-.---.------S----------H-V--   101
B.US.JRCSF         -------..-S----------------.--------L------T.--G----------..-------PP-D-----V--P--.-S-.-Q-.---.------S------------DN-   101
B.US.JRFL          -------..-S----------------.T-------L-------.--G----------..-------P-D-----------.-S-.---.---.------S------------D--   101
CONSENSUS C        me-v-...-n----n------k---n.?-y--k-sY--L---q.tkG-g-s------..-----saPpS-ed--nlisk-.-Lp.rTq.-d.p--sE-S-----sKtet-pfd-   100
C.BR.92BR025       -------..-N----N----------N.-----R-SY--L---Q.--G----------..-----S-PPS-ED--NPIP--.-LP.-T-.---.Q--SE-S-----SK--------   101
C.BW.96BW01B22     -------..-N----N----------N.-----H-SY--L---Q.--G---Y------..-----S-PPS-ED--NLI---.-LP.RTQ.---.---SE-S-----SK-K------   101
C.BW.96BW0402      ---I-...-N----N----------N.K-F--H-SY--L---Q.--G-S--------..-----S-PPS-ED--NPV---.-LP.-T-.---.---LE-S-----SK--------   101
C.BW.96BW0504      -------..-N-----------R----.K----Y--Y--L---Q.--G----------..-----STPPS-ES--NLI-E-.-LP.RTQ.-N.---SE-S-----SK--A---A-   101
C.DJ.DJ259A        SV$I-...------N----------N.K------SY--L---Q.--G----------..-----S-PPS-ED--TPI---.-LP.-T-.---.----E-S-----SK--A-R---   100
C.DJ.DJ373A        -------..-N----N-----------.K------SY--L---Q.--G----------..-----S-PSS-ED--NLI---.-L-.RTQ.---.----E-S-----SK-KA--L--   101
C.ET.ETH2220       -------..-N----N----------N.Q------SY--L---L.-XG----------..-------P-S-KD--NLI---.-L-.HT-.---.---SE-S-----SKA----YA-   101
C.IN.93IN101       -------..-N----N----------N.Q----R-SY--L---Q.K-G----------..-----S-PPS-ED--NLI---.-LP.RTQ.---.---SE-S-----SK-K------   101
C.IN.94IN11246     ---I-...-N----N----------N.-----H-SY--L---Q.R-G----------..-----S-P-S-ED--DLI---.-LP.RTQ.---.---SE-S-----GK-TS-----   101
C.IN.95IN21068     -------..-N----N----------N.-------SY--L---Q.K-G----------..-----S-P-S-ED--NPI---.-LP.RTP.---.---SE-S-----SK-K------   101
CONSENSUS D        md-v-...-nlep-N------r-p-N.k-y--k--Y--qv--i.t-G----------..--q--rppqgg-ah-dpiPk-.-S-.-p-.-d.---p-eQ-----Sea-tdp--W   102
D.CD.ELI           -D---...-N----N------R-P-N.K-H-----Y--P---L.N-G----------..-----GPP-GG-A--VPIP--.-S-.---.---.------Q-----S-A----$    99
D.CD.NDK           -D---...-N--S-N------R---N.K-H-----Y--------.--G----------..-----KPP-GD-A--VPIPE-.-S-.-S-.---.-----K------S-A------W   101
D.CD.84ZR085       -D---...-NID--N------R---N.--------Y--------.--G----------..------PPHS-----DPIP--.-S-.---.---.---Q--------SK--VH---W   101
D.CD.Z2Z6          -D---...-NI---N----------N.R-H-----Y--------.--G----------..------PS-GG----DPIP--.-S-.---.---.------------S-A------W   101
D.SN.SE365A2       -D---...-N----N--------P-N.--------Y---N----.--G----------..------TP-GN----V-VP--.-S-.-H-.---.------------SKA---Q--W   101
D.UG.UG274A2       SV$I-...------N------S-P-N.K-------Y---L----.--G----------..------PP-S--A--DPIP--.-S-.----.-N.-----K------SKA-A----W   100
D.UG.94UG1141      -------..-N----N------R-P-N.K----R--Y--Y---V.--G----------..--P---TPPGG-AN-DPVP--.-S-.----.-N.------------S---A----W   101
CONSENSUS F1       --l-...-n-D--Nh-----t-P-t.k----r-----?w--t.--G----------..-r--h-TP-S--i-qd?vPk-.-?-.-a-.gn.-----eS--E--SKakt--c?-    97
F1.BE.VI850        -------..-S-D--N------T-P--.K----R------W--T.--G----------..----H-TP-S--V--N--P--.-L-.-A-.---.------S--E--SKAK---CA-   101
F1.BR.BZ126        --L-...-N-D--N------T-P--.K----R------W--T.--G----------..-K--H-TP-S--I--DLVP--.-I-.-A-.-N.------S--E--SKATR--CA-   101
F1.BR.BZ163        --L-...-N-D--N------T-P--.K----R------W--T.--G----------..-K--H-TP-S--I--DLVP--.-I-.-A-.-N.-----KS--E--SKA----$    99
F1.BR.93BR020.1    --L-...-N-D--N------T-P--.R----W-----YW--T.--G----------..----P-TP-S--I--DFVP--.-I-.-A-.-N.------S--E--SKAK---$    99
F1.FI.FIN9363      --L-...-N-D--N------P-P-N.K----R-----YW--A.--G----------..----H-TP-S--I--DPVP--.-L-.----.RN.------S--E--SKAK---C--   101
F1.FR.MP411        --L-...-N-D--NQ-----T-P--.K----------YC--A.--G----------..------TP-S--S-KNPIPE-.-L-.-A-.---.------S--E--SK--------   101
F2.CM.MP255        --V-...-KID--N------E-P-N.-----------PL--M.K-G----------..------TP-G-KI--DPVP--.-L-.-T-.---.----E-S-----SQ-----$     99
F2.CM.MP257        --V-...-N-D---------E-P-N.K-----------L--T.R-G----------..------TP-SGEV--DPV---.-L-.-T-.---.-K--E-S-----SK-K---S--   101
```

```
                                          intramolecular
                                         disulfide bonding
                                         |-                -|  rev cds ->|-  nls  -|         exon \ / exon
B.FR.HXB2           MEPVD...PRLEPWKHPGSQPKTACT.NCYCKKCCFHCQVCFI.TKALGISYGRKK..RRQRRRAHQNSQTHQASLSKQ.PTS.QPR.GD.PTGPKE$KKKVERETETDPFD$   100
G.BE.DRCBL          -D---...-K-------------P-N.-----SVAAL.----L.N-G----------..-KH--GTPHS-KD--TPVP--.-F-.TT-.-N.----Q-S--E--SK--------   100
G.FI.HH8793-12.1    -D---...-N----N----------N.K----I--W------L.N-G----------..-KH---PP-DNKD--NPVP--.-LP.TT-.-N.------S-----SK-K------   101
G.NG.92NG083        -D---...-K----N------T-P-N.K----V--W------L.N-G----------..--P--GTP-G-KD--NPVP--.-LP.ITS.-N.---SEKP--E-ASK-----L--   101
G.SE.SE6165         -D---...-N----N--------P-N.K-F--V--W------L.N-G----------..-KH--GTP-S-KG--DPVP--.-LP.TT-.-N.------S--E-ASKA-A-QC--   101
H.BE.VI991          -D---...-NQ---N------R---N.--------Y---L--L.K-G---Y------..-----GTPKSL-D--TLIP--.-L-.RTS.--.----EK-----ASK-------W   101
H.BE.VI997          -D---...-NQ---N---N----P-N.-----Q-S----L--L.K-G----------..-SR--ATPASV-D--NHIP--.-L-.RT-.--.-----------ASK--A--C--   100
H.CF.90CF056        -D---...-K----N------Q---N.--------Y---M--L.K-G----------..-S--H-TPASL-D--N-I----.-L-.RTH.--.------Q--E-ASK-----$    99
J.SE.SE7887         -------...-NR---N-----------.--------Y------L.Q-G----------..-----S-PPG-K---DLIP--.-L-.-TQ.RK.----E-S--E--SKA-P-R---   101
J.SE.SE7022         -------...-NR---N-----------.--------Y------L.Q-G----------..-----S-PPG-KN--DLIPE-.-LF.-TQ.RK.----E-S--E--SKA-P-R---   101
K.CD.EQTB11C        -------...-NI---NQ---------N.Q-------Y---L--L.Q-G---C---E-..----TTTPYA-KN-KDPIP--.-LP.-A-.--.------S--E--SK-K---$    99
K.CM.MP535          -D---...-NI---NQ---------N.Q----R--Y---I--L.K-G----N----..--P--TTPY--EN--DP-R--.-L-.---.-E.Q-D---S-----SK-K--Q---   101
N.CM.YBF106         -------...------N----------N.K-------Y--MC--T.K-G----------..-S----PPKS-KD--DPIPE-.-L-R-QP.---.Q--Q-KQ--AL-GK--A--C--   102
N.CM.YBF30          -------...------N----------N.-----R--Y--LY--T.K-G----------..-S----TP-S-KS--DLIPE-.-L-.-QQ.---.Q--Q-KQ-EAL-SK--A--C--   101
O.CM.ANT70          -D---...-EVP--H------QIP-N.-----R--Y--Y---V.R-G----------..-GRP..-AASHPD-KDPVP--.SPT.ITK.RK.QERQE-QEEE--KKAGPGGYPR   100
O.CM.MVP5180        -D---...-EMP--H----K-Q-P-N.-----R--Y--Y---T.K-G----H-----..--RP.A-AASYPDNKDPVPE-.SL-.HTG.RK.QKRQE-QE----K--GPSGQPC   101
O.SN.MP1299         -D---...-EMP--H------Q-P-N.K----A--Y--Y---A.S-G----------..-.R-PA-AARHPDN-DIVPE-.L-Y.ITN.RK.QKRQE-QE-E--N-ACPR$    97
O.SN.MP1300         -D---...-EMP--H------Q-P-N.K----A--Y--Y---A.S-G--L-------..-.R-PA-AARNPDN-DIVPE-.-PPITNN.RK.HKRQE-QE-E--K--GS-RYPR   102
10_CD.BFL061        ----H...-N----N----------N.K-H--A-SY--L---Q.---X--------..-----S-PPS-ED--NLI----.-S-.---.--.--------------A-A----W   101
10_CD.BFL071        -------...-N----N--------P-N.K-F--A-SY--L---Q.--G----------..-----S-PPS-ED--NLI----.-S-.---.--.------------SQA-------   100
10_CD.BFL110        -------...-K----N--------P--.E----H-SY--L---Q.K-G----------..-----S-PPS-ED--DLI----.-S-.---.--.------------SKA-A----W   101
CONSENSUS 01        --l--...-n----n------t---s.K----k--W---L--L.K-G----y-----..-KH--gtpqssKd--npIpk-.-lp.ii-?gn.--d---S--e-asKae--pcdW   102
01_AE.CF.90CF11697  -------...-S----N------T---S.K-------W---L--L.K-G----------..-KH--GPPPG-KD--NPIP--.-LP.TT-.-N.------S--E-AKKAK----A-   101
01_AE.CF.90CF402    -------...-N----------T---S.K----M--W---L--L.K-G----------..-KH--GPS-D-KD--N-IP--.-LP.TS-.-N.------S-----SKA----$    101
01_AE.CF.90CF4071   --L-----...-K----N------T---N.K----I--W---L--L.K-G----------..-KH--GTP-G-KG--DPI----.-LP.II-.-N.------S--E--SK-K---CA-   101
01_AE.TH.CM240      --L-----...-N----N------T---S.K-------W---L--L.K-G----------..-KH--GTP-S-KD--NPIP--.-LP.II-.RN.--D---S--E-ASKA---QC--   101
01_AE.TH.93TH057    --L-----...-N----------S---S.K----I--W---L--L.K-G----------..-KH--GTP-SRKD--HPIPE-.-L-.II-.-N.--D---S--E-ASKA----C--   101
01_AE.TH.93TH065    -------...-N----N------T---S.K----I--W---L--L.K-G----H-----..-KH--GTP-S-KD--SPIPE-.-LP.II-.-N.--D---S--E-ASKA----C--   101
01_AE.TH.93TH253    --L-----...-N----N------T---S.K-------W---L--L.K-G----H-----..-KH--GTP-SRKD--YPIPE-.-LP.II-G-N.--D---S--E-ASKA----C--   102
01_AE.TH.95TNIH047  --L-----...-N----N------T---S.K-------W---L--L.K-G----H-----..-KH--GTP-S-KD--TPIR--.-P-.II-.-N.--D---S-----SKA----$   100
01_AE.TH.TH022      --L-----...-N----N------T---S.K-------W---L--L.K-G----H-----..-KH---TP-S-KD--YPIPE-.-SP.II-.-I.--D---S--E-ASKA---QC--   101
02_AG.CM.MP807      -------...-S----N------T---S.K----M--W---L--L.N-G----------..--L--GTP-SR-D--N-V----.-LP.TT-.-N.----E-S--E-ASK-----$    99
CONSENSUS 02        --lv-...-s----N----q-T---S.k----i--w---L--L.N-G----------..-rR--GTP-SrqDh-nPVp--.pLP.tT-.-n.ptg---S-ke-asKte-dpcd-   101
02_AG.FR.DJ263      -------...-S----N------T---S.K----I--W---L--L.N-G----------..--R--GTP-SR-D--NPVP--.-LP.TT-.--.--D---S--E-ASK----QG--   101
02_AG.FR.DJ264      -------...-S----N------T---S.-----I--W---L--L.N-G----------..-XR--GTP-SR-DR-NPV----.-LP.TT-.-N.------S-RE--SK-K---C--   101
02_AG.GH.G829       --LI-...-S----N------T---S.K----I--W---L--L.N-G----------..--R--GTP-SH-D--NPV---.SLP.-T-.--.------S--E--SK------A-   101
02_AG.NG.IBNG       --L-----...-S----N------T---S.K----M--X---L--L.N-G----------..--R--GTP-SR-D--NPVP--.-LP.TT-.-N.--D---S--E--SK-K---C--   101
02_AG.SE.SE7812     --L-----...-S----N------T---S.K----I--W---L--L.N-G----------..-KR--GTP-SR-DN-DPVP--.-LP.TT-.-N.-A----S--E-AGK-----C--   101
02_AG.SN.MP1211     --L-----...-K----N------T---S.K----L--W---L--L.N-G----------..--R--GTP-SR-D--NPVP--.-LP.TT-.-E.Q-----S--E-ASK---G-C--   101
02_AG.SN.MP1213     --L-----...-S----N----K-T---S.K-------W---L--L.N-G----------..--R--GTP-SRPD--NPVP--.-LP.NT-.-N.------S----ASKAK---$    99
03_AB.RU.KAL153-2   -------...------------------.--------L------M.K-G----------..-------P-DN--D-V--P--.-A-.---.--.----------M------H----   100
03_AB.RU.RU98001    -------...------------------.--------L------M.K-G----------..------PP-DN--D-V--P--.-A-.---.--.------------------H----   100
04_cpx.CY.94CY032-3 -------...-D----N------T-D-N.K-F-----W------L.K-G----------..-KH--GSL-G-KG--NLIP--.-L-Q--N.---.S---E-Q----ASK--A---A-   102
04_cpx.GR.97PVCH    --A--LVY-N----N------T---N.K-F-----W---A--L.K-G----------..-KRNEDLLGF-RDR-NPIP--X-L-.--N.-N.-E----Q----ASK--A-----   105
04_cpx.GR.97PVMY    -------...-N----N------T----.K-F--V--Y------L.R-G----------..-KH---PP-G-RDR-NPIP--.-L-Q-HS.--.------Q-EA-ASK--SN----   102
05_DF.BE.VI1310     -D---...-N----N------R---N.Q----M--Y---N---.--G----------..--P---PP-G--A--DPVPE-.-P-.---.--.-----KQ--E--SK--A-Q--W   102
05_DF.BE.VI961      -D---...-N----N------R---N.Q-H-----Y-------.--G----------..--S---PP-GG-A--IPVPE-.-S-.-A-.--.---Q--Q-----SKA------C   102
06_cpx.AU.BFP90     -------...-KI---NQ---R------.K-------Y--P---L.N-G----------..-----Q-PPG-KN--DPV----.-L-.-TQ.RE.Q---EKS--E--SKA-P-R---   101
06_cpx.ML.95ML127   -------...-NI---NQ---R------.P------SY---L--L.N-G----------..-----T-PPG-KN--DPVP--.-L-.-TQ.RG.----EKS-----SKA-P-R---   101
06_cpx.ML.95ML84    -------...-KI---NQ---R------.S----R--Y--PL--L.N-G----------..-----T-PPG-KN--DPVP--.-L-.-TQ.RK.S---E-S--E--SKA---R---   101
06_cpx.SN.97SE1078  -------...-NI---NQ---R------.S----Q--Y--P---L.N-G----------..-----T-PLG-KS--DPVP--.-L-.-TQ.RE.S---EKS-EE--SK--P-R---   101
11_cpx.CM.MP818     -------...-NI---NQ---R-----N.Q----R--Y--PH--L.K-G---Y------..-----A-SHS-EN--DPIP--.-ST.--N.RK.----E-S--E--SKA-P-R---   101
11_cpx.FR.MP1298    -------...-KI---N----------N.Q-F--R--Y---I--L.K-G----------..-----T-YHD-KN--DP-PE-.-LP.-T-.RK.----E-S--E--SK--A-R---   101
11_cpx.FR.MP1307    -------...-NI---N------Q---N.Q----R-SY---I--L.K-G----------..-----ATPYS-KN--DPVPE-.-LP.-TT.RK.--D-E-S--EM-SKA-PTR---   101
11_cpx.GR.GR17      ---G-...INID--NQ---------N.Q----R--Y---H--L.K-G----------.W----T-SRS-KN--DPIPE-.-LP.-AS.RN.----E-P--E--SKA-PA----   101
CPZ.CM.CAM3         -------...-D----L------R---N.-----M-----I---E.K-G---------RRG-R-LS-ASSTH.N-G--QQ-.-L-.RQQ.-N.Q--QE-Q--TL-SQAKA---A-   102
CPZ.CD.CPZANT       -D-----...AETP--L--PAT-A-P-N.-----C--Y--PL--T.K-G--------RRA--N--TTAES-ENN-DPV--$.SLP.KTS.RI.QSSQ-K-E----EK-GSGGGPC   102
CPZ.CM.CAM5         -------...XN----L----------N.-----C-----M---E.K-G---------RRG-RHSL-SSSTH.N-D-IRQ-.-L-.RQQ.-N.Q--EE-Q--T--SQA-A--R-W   103
CPZ.GA.CPZGAB       -D-I-...-D-----------R-V-N.-----A--Y--IY--T.K-G---------.RTTR--T-PAG-KNN-D-IP--.-L-.-S-.-N.KE-SEKST-E-ASK--A-Q$    101
CPZ.US.CPZUS        -D---...-NI----Q---------N.-----------V---T.K-G---------.RG-R--TTAKS-ADN-DP-R--.-F-.-QL.-N.Q--Q--Q--AL-KQ-TS-QR--   102
```

HIV-1/SIVcpz
proteins

HIV-1/SIVcpz proteins

```
                                                  high affinity binding site          Leu-rich effector domain
                                        exon \/ exon    |-    nls     -|                          |-     -|
B.FR.HXB2          MAGRS.GDS..DEE.LIR.TVRLIKLLYQSNPPP.NPEG.TRQARRNRRRRWRERQRQIHSISERILGTYLGRSAEPVPLQLPPLERLTLDCNEDCGTSGTQ.......GVGSPQILVESPTVLESGTKE$   116
CONSENSUS A        --g--.-ns..-ee.-Lk.Ai-i--il-q---y-.k-k-.s---rK-r----rA--r--dsl-e-i-stC--rsa---p-q--pl-r-hl-csed-g----qqsqg?etgvgrp-vs?-s?vv-g---kn-   120
A.KE.Q23           -----.---..----.-L-.A--I--I--K---Y-.K-K-.S----K-------A-----D-------S-C---PT-------------H---C---------QSQGAET---RH-VS----VI-G----N-   123
A.SE.SE6594        -----.-PG.----.-LK.AI-I--I------Y-.K-R-.-----K-------A-----D-------SAC----T-------------N---R---------QSQGVET---R--VSG---VI-G----N-   123
A.SE.SE7253        -----.---..----.-L-.A--I--I------Y-.R-K-.S---QK-------A-----D-------S-C----E--------S--T-H---HD--------QSQGVET---R--VPG-PS---G----T-   123
A.SE.SE7535        -----.-N-.----.-L-.AI-I--I--N---Y-.K-K-.S----K-------A-----D-------NAC--G-T-------------N---R---------QSQGIET---R--VS----VI-G------   123
A.SE.SE8538        -----.-N-.----.-L-.AI-I--I------H-.K-R-.S----K-------A-----D-------S-C----------------G-H---C--------E.......---R--VS----G--D----N-   116
A.SE.SE8891        -----.---..----.-LK.A--I--I------Y-.K-R-.-----K-------A--K--D-L-----S-C---PQ-----L---I---H---S---------QSQGAET---G--VSE--SII-G----T-   123
A.SE.UGSE8131      -----.---..----.-LK.A--T--I------Y-.Q-K-.S----K-------A-----D-L-Q---SAC----E-------------HI--H-------G...........T-VSG--SV--D----D-   112
A.UA.ukr970063     -----.-NT..-A-.-LT.A--I--FV-----Y-.T-R-.-----K-Q-----A-----N-L-----SAC--------S-----I---H--YS---------QSQGTET---G---SG--SGI-G------   123
A.UG.92UG037       -----.-NP..----.-L-.AI-I--I------Y-.E-K-.-----K-------A-----DTL---V-S-C---P----------I---H---S---------QSQGVET---RT-VSG---V--G----N-   123
A.UG.U455          --R--.-NP..--D-.LK.A--I---------C-.--R-.S----K-------A-----D-L-----SDC---P----------I---R---S-S-------QPQGTET---G---S---SA--G----N-   123
A.UG.UG273A        -----.-V-..----.-LK.AI-I--I------Y-.K-K-.S----K-------A-----D-L-Q---S-C---P----------I---R---E-------.---RN-VS--P-V--G---QN-   116
A.UG.UG275A        -----.-S-..----.-LN.AI-I--I------Y-.K-K-.S----K------KA-----D-L-----S-C---P----------I---H---S--------G...........--VS---TSI-G----N-   112
A2.CD.97CDKS10     -----.--P..--D-.-L-.A-KI--I------Y-.K-R-.S----K-------A--N--D---------C---PT---------I---H---S---------QSQGTET---G---FG-PS---G--IEK-   123
A2.CD.97CDKTB48    -----.--P..--D-.----.AI-I--I------Y-.K-R-.-----K-------A-----D-------S-C---PT-------------H---S-----C---QSQGAET---R--TS---SGI-G--IED-   123
A2.CY.94CY017-41   T----.D-P..--S-.-LQ.AI-T--I------Y-.K-R-.S---Q--------A-----D-----V-R-C---PT-------------HI--S--------LQSQGTET---RS-ES---SVI-G---E--   123
CONSENSUS B        ----s.g--..dee.-Lk.tv-l--fl---pp-.spe-.-rq----------e--rqirsi-gw--sty--rpaep-p-q---le--t--csed--t--tq.......-v-s-qil----av-es-tkee   117
B.AU.MBC18         -----.----..----.-LK.------SI-------.----.------------------L----A---S-F---PE-S-----------------------R........-M----------P---T--T-E   117
B.AU.VH            -----.----..-D-..-LK.------F--------.S---.------------A-----RQ--GW--S-----P------------------SK--------.........---G--V-----A-----AA--   116
B.CN.RL42          -----.E--..----.-LK.-------------L-.S---.------------A-----RE--D---V----G-T----------------SK--------.........------------A--D------   116
B.DE.D31           -----.---..----.-LK.------F--------.S---.------------Q-----Q-------S-----PEQ----P----------S---------.........------------A---------   116
B.DE.HAN           -----.---..----.-LK.------F--------.SN--.-PT----------------R-------S-F---P------------------S----N---.........------V-----A---P-----   116
B.ES.89SP061       -----.---..----.-LK.------F-----A--.SQ---.------------------R---GW--SN----P-----------------------.........-------------I-D------   116
B.GA.OYI           -----.---..----.-LK.------F--------.----.------------------RK--GW--S-------------------N---S---------.........-----E------A---P-----   116
B.GB.CAM1          -----.---..----.-LK.A-------------L-.SSK-.-----------------H-RA--NW--S-H---P------------------SK--------.........------------A---------   116
B.KR.WK            -----.---..----.-L-.-I-I--F------L-.E---.------------R--WW-Q-L-GW--N-H---P-------------------E--------.........---N--------A---------   116
B.NL.3202A21       -----.---..--D-.-LK.------F--------.S---.------------------R-------S-----------------------D---------.........------------A---------   116
B.TW.TWCYS         -----.---..----.-L-.-------I-------.----.------------------RT--GW--SN----P------------------D---------.........-------F------D------   116
B.UA.AF193278      -----.---..----.-LK.-I----F--------.----.-----------------HL--------S-----PE-----P----------S---------.........---------------D-----X   117
B.US.JRCSF         -----.---..--D-.-LK.------F--------.SN--.------------------RT------S-----P-----------------------------.........---N-E----------------   116
B.US.JRFL          -----.---..----.-LK.----------------.S---.------------------RT------S-----P-----------------------------.........-----E----------------   116
B.US.MN            -----.---..----.-LK.------F--------.SS--.-----------------H-R---AW--SN----P----------.Q------S---------.........--------------------.E   115
B.US.RF            ----R.---..--D-.LK.A-----S--------.S---.------------------RRC--W--D------VD--Q------------SS---------.........------V-----A-----A---   116
B.US.SC            -----.----..E--..-LK.------F--------.K---.------------------R---GW--SN----L-----------------------------.........--------------------   116
B.US.SF2           -----.---..----.-L-.----------------.S---.------------------R---GW--S-----------------------S----N--A-.........------------A--D------   116
B.US.WC001         -----.---..----.-LK.------F--------.----.------------A-----RAL-GW--SN----P----------I-------S---------.........------------A---------   116
CONSENSUS C        -----.g-s..-ea.-Lq.Avri-ki---s--Y-.k-e-.--q-rk-------A--r--hsise---stcl--pa---plq---i---hidcSesggts--qqsqgttegv-s-$isGkscav-gsgakKE   123
C.BR.92BR025       -----.---..--A.-LQ.A--I--I------Y-.K---.------------A-------------S-CV--P-----F----I---NIN-S-SG------QPQGNTER--N-$VFGRPCA----RV-KE   123
C.BW.96BW01B21     -----.---..--A.-LQ.A--I-RI------Y-.E-R-.--R--K-------A-------------T-C---FT----F----I---HIGDS-SG------QSQGTTE---N-$VSGK-FA--G--A-KE   123
C.BW.96BW0402      -----.---..--A.-LQ.A--I--I------Y-.K---.----W--------A-------------S-C---PT---------I---HI--S-SS-A----QSQGTTE---N-$VSGK-CAI-G-RA-KE   123
C.BW.96BW0504      -----.---..--A.-LQ.A-KA--I------Y-.E-K-.-----K-------A-------L-----SAC---P-----F----I---HI--SKSV-----E.......-----$-SGKPCA--G-RA-KE   116
C.BW.96BW1104      -----.--N..--A.-LQ.A--I--I------Y-.K---.-----K-------A--K--N-------S-C---------FL---I---HISDS-SG------QSQGTPE---N-$-SGK-CA--GARA-KE   123
C.BW.96BW1210      -----.---..--A.-LQ.A--I--I---N--Y-.K---.-----K-------A-------------S-C---P----------I---HIGGS-NS--T---QSQGTTE-----$-SGK-CA--G----KE   123
C.BW.96BW15B03     -----.E--..--A.-LH.A--I--I------Y-.E-K-.-----K-------A-----D---T---S-C---PE----F----I---NIGDS-SG------QSQGTTE-----$VSGK-CA--G----KE   123
C.BW.96BW16B01     -----.-----..-AA.-LQ.A--I--I------Y-.K-K-.-----K-------A-----RA--A---SAC---P-------------HIGSS-SG------.........-----$-SGKPCA--G-RA-KE   116
C.BW.96BW17A09     -----.--N..--A.-LQ.AMGI--I--------Y-.K---.--R--K-------A-------------S-C---PT---H-P-------H---S---EPT--EQSQGATE-----$-SGKPWA--G--A-KE   123
C.DJ.DJ259A        -----.---..--A.-LQ.A--I-R-------Y-.K---.------------A-----D-------SAC---P----------I---N---S-SG--A--E.........---N-$VSGKPCA--G--A-KE   116
C.DJ.DJ373A        -----.---..--A.-LQ.A--I--I------Y-.E-K-.----Q--------A-------------S-C---P-----F----I---N---S-SG-----E.........---N-$VSGKPCAI-G--A-KE   116
C.ET.ETH2220       -----.---..----.-LK.A--I--I------Y-.T---.------------A------TL-----SNF---P----------------N---S--S------QSQGTTE---N-$-SGKPCA--G--A-KE   123
C.IN.93IN101       -----.---..--A.-LQ.A--I--I------Y-.E-K-.-----K-------A-------------S-C----T---------I---HIGGS-SG-------.......QL---$-SGK-CA--G--A-KE   116
C.IN.94IN11246     -----.---..--A.-LK.A--I--I------Y-.E-R-.-----K-------A-------------SAC---P----------I---HISGS-SG------QSQGTTER----$-SGK-CA--G--A-KE   123
C.IN.95IN21068     -----.---..--A.-LK.A--I--I------Y-.E-R-.-----K-------A-------------S-F---P----------I---HISGS-SA-----PQSQGTTE-----$-SGK-C---G--AEKE   123
C.SO.SO145A        -----.---..--A.-L-.A--I--I------Y-.K---.-----K-------A-------------S-C---PT---S-----I---HIGGS-SG------QSQGTTER----$VSGK-WP--G--A-KE   123
C.UG.UG268A2       -----.---..----T.-LQ.A--I--I------Y-.----.------------A-----RE-NQ---S-C-------------------I--S-SS-------QSQGTTE-----$VSGKPCA--G-R--KE   123
CONSENSUS D        -----.--s..-ed.-Lk.Avr---?---s----.Sp--.------------A---q-hs-g---ls-y---se---------l--ln-N-s-d-gts---.......g-gsp-is--s-av-?sg-eeQ   115
D.CD.84ZR085       -----.---..--D-.-LT.A-----I--------.S---.------------A---Y----G----S------E------------IN-N-S---------.........-------S----AI---R-E--   116
D.CD.ELI           -----.---..--D-.LK.A-----F--------.S---.------------A-----RE-A----------P--------------N-N-S---R---------.---H---S-----------E-Q   117
D.CD.NDK           -----.---..--N-.LK.AI----F--------.S---.------------A--------G---IC-F---PE-------------N-N-S---------.........-------P--P-A------E--   116
D.CD.Z2Z6          -----.--R..--D-.LK.A-----I--------.S---.------------A--------G----S------E------------N-N-S----A---------.------S-----------E-Q   117
D.SN.SE365A2       -----.---..----.-LK.AI----F---N----.ST--.------------A-----N--G----S----------------V---N-N-S---------.------R-----VS----A--DTE-EK-   116
D.UG.94UG1141      -----.--R..----.-LQ.A-----I--------.S---.------------A--------G---IS-----FE-------------N-N-S---------.------.----H--S----A--D------   116
D.UG.UG269A        -----.----..--D-.LK.A-----F--------.S---.------------A-------------S-----PE-------------N-N---N-------.------.----H--S----KI-D------   116
D.UG.UG274A2       -----.-----..-AD.-LK.A-K---I--------.S---.------------A--------G----S------E---------------N-S----NP---.........-E-H--S----A--D------   116
CONSENSUS F1       -----.-ds..dt-.-L?.Avry--I------Y-.k-e-.------------A--rq-Rai-e---ssC---pe---------l---hINcSed-g.......QGaeE---sp-iSg-sh?--?------   113
F1.BE.VI850        -----.----..-T-.-LK.A-KC--I------Y-.K---.------------A-----RAL-D---SSC----E-------------HIN-S---------.QGPEE----S--SG--HA---------   116
F1.BR.93BR020-1    -----.-----..-Q-.-LK.A--Y--I------Y-.K---.------------A-----RE--D---SSC---P-------------HIN-S---------.QGAEE----S--SG--H---G------   116
F1.BR.BZ126        -----.-N-..--T-.-L-.A--Y--I------Y-.K---.------------A---E-RAL-----SSC---PE------------NINSS---E.......QGAKE------TSG--HA--G------   116
F1.BR.BZ163        -----.-N-..GT-.-L-.AA-Y--I------Y-.K---.------------A-----RE--R---SSC---PE---------P----INNS-N-E.......QGAEE------TSG--HA--G------   116
F1.FI.FIN9363      -----.-----..-T-.-LK.A-KY--I------Y-.S-D-.------------A-----RA------SSC---LE-------------HIN-S---------.QGTEE-------SG-HH----------   116
F1.FR.MP411        -----.--N..-----.-L-.A--A--I------Y-.K---.------------A--K--R-------VAC---PE-------------HIN-SK--------.QGTNE---N---SM-PR----------   116
F2.CM.MP255        -----.---..----.-LK.A--Y--I------Y-.K---.------------A------E------T-----PT---LF---------IN-S---------.QGSER-------S---HAA-G--I---   116
F2.CM.MP257        -----.--R..----.-LK.A--Y--I------Y-.KL--.--K----------A------Q------S-C---LQ---R----L--K-HIN-S---------.QGTEK-------S---RA--G------   116
```

```
                                                   high affinity binding site          Leu-rich effector domain
                                   exon \/ exon    |-    nls     -|                    |-     -|
B.FR.HXB2           MAGRS.GDS..DEE.LIR.TVRLIKLLYQSNPPP.NPEG.TRQARRNRRRRWRERQRQIHSISERILGTYLGRSAEPVPLQLPPLERLTLDCNEDCGTSGTQ.......GVGSPQILVESPTVLESGTKE$     116
G.BE.DRCBL          ------.-ST..----.-LT.A--I---------S-.P---.-------------A-------------S-C---PE-------------H---S--G------QSQGTEI-------F---SV--G-----     123
G.FI.HH8793-1-1     ------.-ST..--A.-L-.-I-I--I------Y-.P---.-----K-------A-------------SSC------------------H--SS--GE---P-QSQGTET---G---F---SV--G--A---     123
G.NG.92NG083        ------.--P..----.-L-.A--I--T------Y-.S-A-.-----K-------A-------------SAC---P-----F------G-S---SK-G------QPQGTET---R--V---P-V--G------     123
G.SE.SE6165         ------.-ST..----.-L-.A-KA--I------Y-.P---.-------------A-----SA------TA----P--------------H---S--S------QPQGTET---RS$-F---SV--GQ-----     122
H.BE.VI991          ------.--N..--G.-L-.AC-I-R-------Y-.E-A-.----Q--------A--------G--V-A-C--GP------------------S-------EK.......------TSG---A--GT-A---     116
H.BE.VI997          ------.-AG..--Q.-PQ.VCKI--II-----Y-.E-A-.-------------A-----RA------TDC---PP------------------K------EK........-K-G---P---S---GT-----     116
H.CF.90CF056        ------.-A-..-T-.-LQ.VCKI--I------C-.E-T-.-------------A-----RE------TSC---PP---T-----------N-S-------EK........-E-----SL--S-I-GT-----     116
J.SE.SE7022         ------.--N..-DQ.-LL.A--I--I------YS.K-N-.S------------A--N--D-------SSC---P----------I---R---S----N----........---D---SG-PCM--GA-----     116
J.SE.SE7887         ------.----..-DQ.-LL.A-----I------Y-.K-N-.S------------A--N--D------PSSC---P----------I---R---S----N----........---D---SG-PCM--GA-----     116
K.CD.EQTB11C        ----R.----.EQQ.-LT.P--I--I------Y-.K---.-------------A-----RE--Q-V-SSC----T-------------S-N-D--S-........QGTEGEL-----P--PD---G--D---     116
K.CM.MP535          ----R.--P..--Q.-LT.---T--I--E---Y-.-L--.S--T---------A--K--S-----L-SAC--------------I-K-N-N-D--P-........KGTEG-L-----S--PC----------     116
N.CM.YBF106         ------.-VN..--D.-LK.A-KI--I------Y-DSSQ-.-------------A-----RA------A---GPQ---D-P----G-----HK--S-DP--ESQQGTATTE$TQNT--GNNSI-GKRV-R-     123
N.CM.YBF30          ------.-VN..----.-L-.A--V--I------Y-.-SK-.-------------A-----RA------SSC--GPP---D-P----D-----TE--S--P--ESQQGTATTE$TQNT--GNTCI-GKRV-G-     122
O.CM.ANT70          -----.E-...-Q..-LQ.AIQI--I------Q-.S-R-.S-N--K-------R--A-VDTLAA-V-A-VVHGPQNNNIVD-----Q-SIRDP-GDQL-EAWTVDPRAEDN$CL-N-CSCN-I-ATRIA--     120
O.CM.MVP5180        -----.EE...-QQ.-LQ.AIQI--I------C-.T-A-.S-N--K-------R--A-VD-LAT---A-VVHG-QDNNLVD-----Q-NIRDP-ADRLP--GTVDPGTKDN$-L$T-WSCNAI-ATRIEK-     120
O.SN.MP1299         -----.DG...-QP.-L-.AIQI--I-----SHT.S-T-.S-S----------T-HAHVDTLAA---A-VVHGPQDNNLVE-----Q-SIRDPDGDQP---WTVDSGTEDN$CL-T-HSCN-I-ATRVA--     121
O.SN.MP1300         -----.DG...-QP.-L-.AIQI--I------H-SPTT-.STS----------K--A--DTLAA---A-VVHGPQDNNLVE-----Q-SIRDPDGDQP---WTVDPGTEDN$CL-N-HSCN-I-ATRVA--     122
CONSENSUS 01        -----.-sT..--e.-Lr.avri-ki-yq---y-.ss--?t--t-k----r--a-----rais----s?Cl---t---p-q---l---hl--s--c------QSqGTETg--r---Sg--svI-Gp--kN-     122
01_AE.CF.90CF11697  ------.-NT..--D.-LQ.A--I--I------Y-.P---.-----K-------R-------L-----VACV---T-------------HI--S---------QSQGTET---G---SG--SVI-G----N-     123
01_AE.CF.90CF402    ------.-ST..--D.-L-.---I--I------Y-.PA--.-----K----K--A------K-G----S-C----P-------------H---S---------QSTGTETE--R---SG--SVI-G----N-     123
01_AE.CF.90CF4071   ------.-ST..----.-L-.AA-A--I-F----Y-.SS--.-----K-------A-----RAL-----SAC-----------------H---S---------QSRGTET---R---SG--SVI-G---EN-     123
01_AE.TH.93TH057    ------.-ST..----.-L-.A--I-NI------Y-.SS--.---T-K-------A-----RA------I-C----T-------------H-------------QSQGTET---R---SG--SVI-GP---N-     123
01_AE.TH.93TH065    ------.-ST..----.-L-.A--I--V------Y-.SS--.---T-K-------A-----RA------NACV---T-------------H---S---------QSQGTET---R---SG--SVI-G----N-     123
01_AE.TH.93TH253    ------.-ST..----.-L-.A--I-NI------Y-.SS--G---T-K-------A-----RA------S-C----T-------------H---S---------QSQGTET---R---SG--SVI-GP---N-     124
01_AE.TH.95TNIH047  ------.-ST..----.-L-.A--I-----E------.SS--.---T-K-------A-----RE------SSCV---T-----P-------H---S---------QSQGTET---R---SG---VI-GP---N-     123
01_AE.TH.CM240      ------.-ST..----.-L-.A--I--I------F-.SS--.---T---------A-----SA------S-C----T---------V---N---S--G------QSQGTET---R---SG--SVI-GP---N-     123
01_AE.TH.TH022      ------.-ST..----.-L-.A-KI-NI------L-.SS--.S--T-K-------A-----SA------SAC----T---S-P-------H---S---------QSQGTET---R---S---SGI-GP---N-     123
CONSENSUS 02        -----.gda..-ee.-L-.?v-iikI------Y-.p---.t--a-k---r---A----irai-e-i-S-C---sa---plq---i---c---s-dc-t----qsqgtetgvgspqis?--sii-Gs-----     121
02_AG.CM.97CMMP807  -----.----..-A-.-L-.A--I--I------Y-.P---.-----K-------A-----RE------S-C---PT---------I---C---C-G-------QSQGTET-------SG--SII-GP-----     123
02_AG.FR.DJ263      ------.--A..----.-L-.V--I--I------Y-.P---.---T-K-------A-----RA----F-S-C--------------I---C---S-G-------QPQGTET-----P-SG--S-I-G------     123
02_AG.FR.DJ264      -----.X-A..----.-L-.V--IV-I------Y-.P---.-----K---G---A-----RA--Q---S-C----------------C---S-G-------QSQGTET-L-----SG--SDI-GA-----     123
02_AG.GH.G829       ------.--A..----.-L-.VI-I--I------Y-.K---.-------------A-------L-----S-C---PE---SF--------N---S----N----........S-AD---PG--RAI-G------     116
02_AG.NG.IBNG       ------.--A..----.-L-.A--I--I------Y-.P---.---T-K-------A-----RA------S-C--------------I---N---S---------LSQGTET-------S---YII-G------     123
02_AG.SE.SE7812     ------.-NA..----.-L-.A--T--I------Y-.P---.-----K-------A-----RA------S-C--------------I---R---S----------------------S---SI--G------     116
02_AG.SN.MP1211     ------.--A..--G.-L-.A--I-RI------Y-.P---.S------------A----VRA------S-C---P------P---I---C---S--S-------QSQGTET----S-TS---SVI-G------     123
02_AG.SN.MP1213     ------.--A..----.-L-.V--T--I------Y-.T---.-----K-------A------K------S-C--------------I-------S----------------------S---SII-G------     116
03_AB.RU.KAL153-2   ------.----..----.-LK.-I----F---------.S---.-----------------H------Q--S-----PE---L-H-----------S----------------------D------     116
03_AB.RU.RU98001    ------.----..---D.-LK.-I----F---------.----.-----------------H-----Q---S-----PE-----H-----------S----------------------D------     116
04_cpx.CY.94CY032-3 ------.-NI..--D.-FK.AA-A--I------Y-N--T-.-------------A--K----L-----A-----P------------K---N-S-------DK.......------VS--L-A--GT-A---     117
04_cpx.GR.97PVCH    ------.-SX..N-D.-LGXS-GIV-I------Y-.--T-.--K----------A--K--------V-A-----P------------K---N-S-------EK.......------VS--P-A--GT-A---     117
04_cpx.GR.97PVMY    ------.-ST..--D.-LK.A-GIV-I------Y-N-TA-.-------------A--NR---------AAC---P----------I-K-----S-------DK.........---N--VP--P-A--GT-D---     117
05_DF.BE.VI1310     ------.--R..--D.-LK.A-----I------L-.S---.-------------A-----N--G--L-S------E-------------N-N-S----------------------S--P-AI--------     116
05_DF.BE.VI961      ------.--R..--D.-LK.A-----F---------.R---.-------------A-----R--AD--VD-----PE-------------N-N-S----------------------S----------A---     116
06_cpx.AU.BFP90     ------.----..-DR.-LL.A--I--I------Y-.K-N-.S------------A--N--D-----V-S-C--------------I---R---S----N----........---N---SG-PDML-GT--T--     116
06_cpx.ML.95ML127   ------.--N..--Q.-LL.A--I--I---N--Y-.K-N-.G------------A--N--D-------S-C---PT----F----I---R---S----N----........---D---PG-PGV--GT-----     116
06_cpx.ML.95ML84    ------.--N..--Q.-LL.A--T--I------Y-.K-S-.S------------A--K--D-------S-C--------------I---R---S----N----........---D---P--PGVL-GT-----     116
06_cpx.SN.97SE1078  ------.----..--Q.-LW.A--V--I------Y-.KLS-.S------------A--N--D-------SSC----E---------I---R---T----N----........---D---PG-PGVL-GT-----     116
10_CD.BFL061        ------.----..--A.-LQ.A--I--I---------.S---.----------------------GG---DA----PE-------------N-N----------H.........---GH--S-----I-D------     116
10_CD.BFL071        ------.----..--A.-LQ.A--I--I---------.S---.-------------A-------------CDC---PE-------------N-N----------------------S----AI-DP----Y     117
10_CD.BFL110        ------.----..--A.-LQ.AA-I--I------L-.S---.-------------A--------G---IS------E-------------N-N----------------------N--S--P-EI-D------     116
11_cpx.CM.MP818     ------.--N..--Q.-LT.A-KI--I------Q-.-T-.S------------A--N--D---Q---SDC----E---------I---H---S----NP---........---DS--SG--D---GPR-E--     116
11_cpx.FR.MP1298    ------.--N..--Q.--T.---I--I------Y-.K---.N---------K--A--K--D---Q---S-C----E---S-P---I---H-H-S--S-D----........---NS-TSG--S---G-R-E--     116
11_cpx.FR.MP1307    ------.----..-AQ.-LT.A--I--I------Y-.K-Q-.S--TQ--------A--N-LD---Q---S-C---PE----F----I---H--SS----D----........---DS--P---N-I-GPR-E--     116
11_cpx.GR.GR17      ------.----..-AQ.-LA.AA-I--I----S-Y-.K-A-.-------Q-----A--N-L----Q---S-C----E-----P---I---H---S----N----........---ES--PG--S---GPR-E--     116
CPZ.CM.CAM3         -----E--D..-AC.-LQ.A--I--V--S.--Y-.DNK-.-------------A-------L-----SS-----P--PA-L---I----I-TG--I-KP--EPQQGTATAEELQ--P-G-DNL-GK-ATN-     123
CPZ.CD.CPZANT       ----EELEGT.--Q.-LK.A-KI--I------Y-.K-A-.S-A----K----KK--D-VEGLAA-V-R-LV-GPD-HN-VD--D-QN-S-EPLGSPP-PLQN.......YFQTV-PSG-QSQHPLDNNQN-     118
CPZ.CM.CAM5         -----E--E..-TR.-LQ.A--I--I--D.--Y-.DNK-.-------------A------E-GQ-V-SSC----S--PD-L---ID----NPV--SAN---EHQQGTATVE-L$Y-PKG-TNI-GKRDTN-     122
CPZ.GA.CPZGAB       -----.EPQ..-DAR-LQ.A-KI--I------Y-.S----..--K----------A--K--SE--G-V-A-----PPK-GD-E--E-DK-S-Q-V-TTQDV--SNTSQPQTAT-ETVPAGGNYSI-GK-A-N-     124
CPZ.US.CPZUS        -----G--A..--Q.-L-.A--I--I--A---F-.-NS-.------------KN-----NE--Q---S-C----S--PD-P---IGG--INSE--STNP--EHQQGTATVE$TQST-LGGNRL-GK-A-S-     123
```

HIV-1/SIVcpz proteins

HIV-1/SIVcpz proteins

```
                                                                                                                  |- alpha helix -|
                                                                                  |-    alpha helix    -|    -> env cds
                                transmembrane domain \/ cytoplasmic domain     phos |   | phos
B.FR.HXB2          TQPI......PIVAIVALVVAIIIAIV.VWSIVII.EY.RKILRQRK..IDR.LIDRLIERAEDSGNESEG...EISALVEMGVEMGHHAP.WDVDDL$...  82
CONSENSUS A        MspL?...??e-?-ivg-i-alil--v.--T--gi.ey?kkllk-r-..i-r.-ie-Ir------g---d-Dte-Ls?....L?em-nYdL.gddNnl-...  78
A.KE.Q23           MS-L......E-S---G-I--L-L--A.--T--G-.---.K--RK---..---.--E-IR----------D-DTE-L-T....LID--DYDL.G-NNI-....  81
A.SE.SE6594        M-SL......V-L-------VL-----.--T--F-.-W.GRLR-----..X--.--E-IR----------D-DTE-L--....LI---NYNL.GFDNN-....  81
A.SE.SE7253        ML-L.....ST-W--IG-I--L-L---.--T--G-.---.K-L-K---..--K.-VE-IR----------D-DTN-LA-....LI---NYNL.G-DN-----...  82
A.SE.UGSE8131      MS-L......E-W---G-I--L-L---.--T--GL.---.K-LK----..---..--E-IR----------D-DTE-L-T....LI-L-DYDL.G-DNN-....  81
A.UA.ukr970063     MT-L......E-Y-------VFV----.--T--G-.--.--L-K---..---..--K-IR------------DAE-L-T....LM-V-NY-L.L-DNNV-...  81
A.UG.92UG037       M-LL......E-C-V-G----L-----.--T--G-.---.K-L-K---..---.-V--IR----------D-DRE-L-L....L-D--DYDL.G-DNN-....  81
A.UG.U455          MT-L......E-W--TG-I--L-L---.--T--G-..$XK-L-K-K-..---.-LN-IR----------D-DTE-L-L....L----NYDL.GVDNN-....  81
A.UG.UG273A        ML-L......N-C---G-I--L-L---.--T--G-.D-.K-L-K---..---..--E-IR------A-----DTE-LA-....L----NYNL.GFDNN-....  81
A.UG.UG275A        MS-LP...PLH-C---G----L-L---.--T--G-.---.KRL-K---..---..--K-IS----------D-DTD-L-K....L----NYDL.G-ANN-....  84
A2.CD.97CDKS10     MLSL......E-A---G----L-A---.--T--F-.--.---RE---..--W.--Q-IS----------D-DTD-L--....L----NLNF.G-ANN-....  81
A2.CD.97CDKTB48    MS-L......A-LS--G----S-L---.--TV-F-.--.---KK---..--W.-LE-IS----------D-DTE-L-K....M-G--NLGF.--DN-V-...  81
A2.CY.94CY017-41   ML-L......V-L---G-I--L-L---.--T--F-.--.K--KK---..--W.--K-IS----------D-DTE-L--....L--R--LDF.G--NNV-...  81
CONSENSUS B        mqpl......qil-i-a---?a-i--v.--t--fi.ey.rkilr-rk?.idr.-i-rIr-ra-----e-e-Dqeelsa....lvemghha-.-dvd-l-...  80
B.AU.MBC18         M--L......Q---------A-----.--T--L-.--.---------..---.----IR--------D---DQE-L--....L-D----D-.--IN-----...  81
B.AU.VH            M-SL......Q---------VT----.--T--F-.--.-R------..---.----IA------------DQE-LA-....L--------.--------...  81
B.CN.RL42          M-AL......T-L-------VA-----.--T--LL.--.---------..---.----IR----------D-DQE-L--....FM-------.--------...  81
B.DE.D31           M-SL......QVA-------VA-----.--T--F-.--.---------..---.----I-------------DQE-L--....L--------.--------...  81
B.ES.89SP061       M-SL......Q-L-V------A-----.-----F-.--.---R--N-..-E-.----I-----------D-DQE-L--....L-----D--.---N------...  81
B.GA.OYI           IYSL......Q-L-------VT----.--T--LL.--.---------..---.----IR--------D---DQE-L--....L--------.-----M-...  81
B.GB.CAM1          ML-L......Q-$-------VA-----.--T--F-.--.---R--KQ..V--.----I-------------DQE-L--....LM----N--.---N------...  80
B.GB.I4663         MI-L......Q-A-------VA-----.--T--L-.--.---------..----.-L--IX--E-------D-DQE-L-K....L--------.---N------...  81
B.JP.JH32          M-SL......$-A-------VA-----.-----F-.--.---------..---.----IR------------DQE-L--....L-----D--.--I------...  80
B.JP.PT1-01        M-SL......Q-L--------------.-----F-.--.---------..--K.----IR----------G-DQE-L--....L--IR----.$-IN------...  80
B.KR.WK            M-LL......A-L---G----A-L---.--F--F-.--.K---K-K-..---.----IR------------DQE-L--....L--------.--------...  81
B.NL.3202A21       M--L......VVA-------VA-----.-----F-.--.---------..---.----IA------------DQE-L--....L--------.--------...  81
B.TW.TWB101        M--P......L-Y--------A-----.-----F-.K-.-R--K---..---.----IT------------DQE-L--....L--R--L--.--I------...  81
B.TW.TWCYS         M--L......H-L--------A-----.--T--LL.--.---------..---.----IR------------DQD-L--....L----Y---.---N-Q-...  81
B.UA.UKR1216       M--L......A-A-------VG-----.-----F-.--.---------..---.----IR------------DQE...-....LM----LV-.--A------...  78
B.US.JRCSF         M--L......Q-L--------G----I.-----L-.--.---------..---.---KIR------------DQE-L--....L--R--L--.--IN------...  81
B.US.MN            M--L......V-A--------G-----.-$---F-.--.---R----..---.----IS------------DQE-L--....L-G---D--.-VIN------...  80
B.US.RF            M-SL......E-L--------A-L---.--T--G-.-I.--T---X-K.---.----IR----------D-DEE-L--....L--------.--------...  82
B.US.SF2           M-SL......Q-L---S---VA-----.--T--L-.--.-------X..---.----IR-K---------DQE-L--....L-----L--.--------...  81
B.US.WC001         M--L......E-L--------A-----.--T--F-.--.---------..---.----IR------------DQE-L--....L--------.-N------...  81
CONSENSUS C        m?dlla?kvdYRlgvg-liv-L--a-v?--t-vY-.--.--Llr-rk..idw.l?k-ir----------d-DtE-LsT....m-d--hLRL.Ldvndl-...  84
C.BR.92BR025       MLELIG.RIDYRLGVG--I--L--V-I.--T-AY-.--.--LV---R..--W.-VK-IK----------G-DTE-LET....M-D---LRL.L-GN-----...  86
C.BW.96BW0402      MISLAA.RVDYRIGVG-FI--L----I.--T--Y-.--.--LV----..--W.-VK-IR----------D-DTE-L-T....M-D---LRL.L-..-N-...  84
C.DJ.DJ259A        MIDLPA.KVDYRL-VG--I--L-----.--T--Y-.--.--L----Q..--W.--E-IR----------D-DTE-LAT....V-G--NLRL.L--V-----...  86
C.ET.ETH2220       MVDLLA.KVDYRIV---FI--L-----.--T-AY-.--.--L----R..----.--K-TR----------D-DTE-L-T....M-D--NLRL.L--N------...  86
C.IN.93IN101       MIDLX....XYRLGVG--I--L-----.--T--Y-.--.--LV--S-..-NW.--K-IR------------DTE-L-T....M-D--RLRL.L--N------...  83
C.SN.SE364A        MVDLLAXKVDYRLGVG--I--L----IX--I-AY-.--.--L---K-..L-W.-VK-IR----------D-DVE-L-T....M-D---LRL.L-AI-----...  88
C.SO.SO145A        MLNLLA.GVDYRI--G-FS--L-----.--T--Y-.--.--L-----..--W.-VK-IR------------DTE-LAT....M-D---LRL.L--N------...  86
C.TW.TWC2          ARVD......YRLGVG--II-L--T--.--I--Y-.--.--L-----..----.IRE-IR------------DTE-LAT....M-D---LRL.LYDNN-----...  81
C.UG.UG268A2       MLNLLA.GVDYRIG-G--LI-L-----.--I--Y-.--.--L-K--R..M--..--K-IR----------D-DTE-L-T....M-D--NLRL.L-IN------...  86
CONSENSUS D        mqpL......e-l--v-lvv-l-l---.--t--f-.-?.rrikr--k..--w.-i--Ir--a---------dkeeL-?....Lv---hha-.wdvdd?L...  79
D.CD.84ZR085       M-SL......Q-L------L-L-----.--T--F-.--.-R-K-----..--W.----IR--E---------DKE-L-T....L--------.-----DL...  82
D.CD.ELI           M--L......G-I--A-------L---.--T--F-.--.-R-KK--R..--C.-L--IT----------DREKL-K....L---------.--I------...  81
D.CD.NDK           M--L......V-I--A-----------.--T--Y-.--.-R-K----..--C.----IR------------ERE-L-K....L--------.---------...  81
D.SN.SE365A2       M-SL......E-L--A-I---L-L---.--T--F-.-V.---R----..--W.----IK------------DEK-L--....L---------.-N---M-...  81
D.TZ.TZ005         M-SL......V-L----V-A-L-----.--T--F-.-C.-RLS---Q..--W.----IR------------DKE-L--....L---------.-NI--M-...  81
D.UG.94UG1141      M--L......E-L--------L-L---.--T--F-.-C.K-LR----..--W.----IR------------DKE-L--....L-----D--.--A--M-...  81
D.UG.UG266A2       IN-L......E-W-----I--L-L---.--A--S-.-C.--LK----..--W.----IR--E---------DTE-L-K....LI---NYN-.GNDNN-....  81
CONSENSUS F1       MS?l......laisi?-lIv-L--a--.--T-vy-.--.k-L?-----..-Nr.-YerIre-----------DAE-LAa....LGem-pFIp.G-inN-...  78
F1.BE.VI850        MSYL......LAIG-A--I--L-----.--T--Y-.--.K-LV----..-NK.-YK-IR------------DAE-LA-....LG---PFI-.G-INN-....  81
F1.BR.93BR020-1    MSNL......LAIG-A--I--L--T--.--T-AY-.--.K-LV----..-N-.-YK-IS------------DAE-LA-....LG-V-PFI-.G-INN-....  81
F1.BR.BZ126        MSYF......LAIS-T--I--L-----.--T-AY-.--.--L-----..-N-.-YE-IRD-----------DAE-LAT....LG---PFI-.G-I-N-....  81
F1.BR.BZ163        MSDL......LAISVT--I--L-----.--T--Y-.--.--L-----..-N-.-YEGIR------------DAE-LA-....LG---PFI-.G-IVN-....  81
F1.FI.FIN9363      MSDL......LAIT---FI--L--V--.--T--F-.--.K-LV----..-N-.-YI-IR------------DAE-LA-....LGK--PFI-.G--NN-....  81
F1.FR.MP411        MSNL......YVLS---FII-L-----.--T--F-.--.K-L-----..-N-.-YE-IR------------DAE-LA-....LG---SFIS.G-INN-....  81
F2.CM.MP255        MPSL......LA-G-S--I--L--T-I.--T--YL.--.K-L---KR..-N-.-YE-IR------------DAE-LA-....LG-V-PFI-.G-INN-....  81
F2.CM.MP257        MSLS......L--V-A-YI-VL-L--I.--T--Y-.--.K-----KR..-N-.-YE-I-------------DAE-LA-....LG-V-PLI-.G-INN-....  81
CONSENSUS G        M-sL......eisaivg-I--f-A-i?.--t--fi.e-.-?-Rk-k?..iek.-l--IR--a-------e-Dte-Lat....L?emgdFD-.-VGdN-...  77
G.BE.DRCBL         M--L......E-S---G-I--S-A---.--T--F-.--.---RK-KR..-EK.-L--IR------------DTE-LAT....LM-L-DFD-.-VG-N-...  81
G.CG.CNG30         M--L......EVAS--G-I--A-A--I.--I--F-.$-.-E-R--K-..--K.-L--IR--E---------DTD-LAK....L----NFD-.-VG-N-...  80
G.FI.HH8793-12-1   M-SL......E-S---G-I--F-A--I.--T--LV.Q-.-E-RK----..VE-.-----IR------------DRE-LAT....LM---DFD-.-VG-N-...  81
G.NG.92NG083       M-AL......E-S...X-I--F-A-TI.-----F-.--.---RK-K-..-EK.-L--IR------------DTE-LAT....LM---DFD-.-VGNN-...  78
G.SE.SE6165        M-SL......V-L---G-I--F-A---.--T--F-.--.-E-RK-KR..-GK.-L--IR----------D-DTE-LVT....L----DFD-.-VG-N-...  81
G.TW.TWG1          M-SL......E-A---G-I--A-A---.--T--F-.--.---RK-KR..----.-L--IR------------DTE-LAT....L-D-VDFD-.-VG-N-...  81
```

```
                                                                                                                |- alpha helix -|
                                                                |-    alpha helix    -|    -> env cds
                                transmembrane domain \/ cytoplasmic domain     phos |   | phos
B.FR.HXB2           TQPI......PIVAIVALVVAIIIAIV.VWSIVII.EY.RKILRQRK..IDR.LIDRLIERAEDSGNESEG...EISALVEMGVEMGHHAP.WDVDDL$...  82
H.BE.VI991          MNIL......GIG-G-----F-----.--T-AY-.--.--LXK---..---.--E-IR----------D-DTE-L-K....L-----LNL.GY-A-----...  80
H.BE.VI997          M.Y-......IGIG-G--I--F-----.--T--Y-.--.--LVK-K-..---.--Q-I--G--------DE...-L-T....M--R--LTF.GY-A-----...  77
H.CF.90CF056        MYIL......GLG-G----TF---VI.--T--Y-.--.K-LV--K-..---.--E-IG----------D-DTE-L-K....LM----LNL.GY-A-----...  80
J.SE.SE7022         MVSL......Q-------I--FFL--C.--T--Y-.--.K-L-----..--K.--N-IR----------D-DTD-LAE....L----P-DL.-N-N-----...  81
J.SE.SE7887         MI-L......Q-A----FI---FL--G.M-T--Y-.--.K-L-----..--K.----IR----------D-DTE-LAD....L--R-P-DL.-N-N-----...  81
K.CD.EQTB11C        MV-L......T-G-I---A-L-L--I.--T-AYL.--.--VV--KR..-NW.-F--IR------------DTE-LA-....LG-T--LIL.G-INN-----...  80
K.CM.MP535          MVSL......A-S.-------L-L--I.--T--Y-.--.--LVK-KR..-NW.----IR------------DAE-LAD....IG-L--LIL.GNI-N-----...  80
N.CM.YBF30          MLSL.........GFI--GA-VS--VI.--ALLYR.--.K--KL-E-..-KH.IRQ-IR--E-------D-DAEWLDGDE-WL-TLLSSSKLDQGNWV-...  83
CONSENSUS O         Mhhr......dLltliiisAlllinvi.l-mf?Lr.?yLeqkkQdrrErEIleRLr-ireirD--DYe-ngeeeQEVr....?dLvhs?GFdNPMFEl-...  82
O.CM.ANT70          MHHR......DLL--IIISALLF-NVI.L-GFILR.K-LEQKEQD--EREILERLR-IR-IRD--DY--N-EEEQEVM.....DLVLS-GFDNPMFEP-...  85
O.CM.MVP5180        MHQE......NLL-LI--SALCL-NVL.I-LFNLR.I-LVQRKQD-REQEILERLR-IK-IRD--DY--NEEEQQEVM.....ELIHS-GFANPMFE-...  85
O.FR.HIVY16019      MHHR......DLLTLIIISALLFLNVI.L-TF-LN.K-LEQKKQD-REREILERLR-IR-IKD--DY--D-EEEQEVM.....DLVHDYGFDNPMFE-...  85
O.FR.HIVY16020      MHHK......DLLLLIVILALLL-N-I.L-RFNLR.K-LEHKEQD-REREILERLR-IR-IRD--DY--NEEEEQEVR.....DLVHTFGFDNPMFE-...  85
O.FR.HIVY16021      MLHR......NLLLSIIISALLLTN-I.L-MYILR.R-LEQRKQDSREREILERLR-IR-IRD--DY--NEEEEQEVR....-QLVHNFGFDNPMFE-...  86
O.FR.HIVY16022      M-HK......DLLIVIISSALLF-NVI.L-TYNLK.T-LEQRKQD-REREILERLE-IRKIRD--DY--N-EEEQEVM.....DLVHSYGFANPMFE-...  85
O.FR.HIVY16023      M-HR......DLLTLIITSALLL-NVI.I-MF-LR.LCLEQRRQD-REREIIRRLR-IR-IED--DY--N-EEEQEVR.....DLVHS-GFDNPMFE-...  85
O.FR.HIVY16024      M-HR......DLLTLIIISALLL-NVI.L-MF-LR.QCLEQKKQN-REREIIERLR--R-IKD--DY--N-EEGQEVR.....DLVHSYGFDNPMFE-...  85
O.FR.HIVY16026      MHQR......DLL--I--SALCL-S-I.--I-NLR.N-LEQRKQD-REQEILERLR-VR-IKD--DYD-N-EEEQEVM.....DL-HS-GFDNPMFEW-...  85
O.FR.HIVY16029      MHHR......DLLLL-IISALLL-N-I.I-MFNLS.K-LEPKKQD-REREILERLR-IR-IRD--DY--N-GGGQEVR....-HLVHTFGFANPMFEI-...  86
O.FR.HIVY16031      MHHS......DLPTLII.SALLL-NVI.L-MFILR.QCLKNKKQD-REREILERLR-IR-TQD--DY--NEDEEQEVR.....DLVHS-GFDNPMFE-...  84
O.SN.MP1299         MHHR......DLL-LITTSALLLTNV-.L-TFILR.Q-LKQKKQDKREREILERLR-IRQIED--DY--D-TEEQEVR.....DLVHSYGFDNPMFE-...  85
O.SN.MP1300         MHHR......DLLTLIIISALLLTNVI.L-AFILR.Q-L-QKKQD-REREILERLR-IRQIED--DY--D-KEEQEVR.....DLVHGYGFDNPMFEP-...  85
CPZ.CD.CPZANT       MTN-......FEY-FL-FSIVL.WI-C.IPILYKLYKIYKQQQIDN-.RNQ-IIEVLSRRLSI--AI-EDEEADTYYL.........GSGFANP-YREGDE  83
CONSENSUS 01        mtpL......?is-i-g-Ival-la--.--Ti-a-.-y.kkilr--k..----.-vk-Ir--------e---DTd-LAk....L--mgdFD-.-Vg-Nl-...  80
01_AE.CF.90CF11697  MSAL......Q-$---G-I--L-L---.--T--F-.--.K-------..---.--E-IR--------G---DTN-LAT....L--V-DFD-.-VG-N-----...  80
01_AE.CF.90CF402    RL-L......H-C---G-I--L-----.--T--A-.--.K-LR----..----.-VQ-IS------------DTE-LAK....L----DFD-.-VG-N-----...  81
01_AE.CF.90CF4071   ML-L......Q-S---R-I--L-L---.--TL-A-.--.XR------..---.--E-IR------------DTD-LAK....L----DFD-.-VG-NV----...  81
01_AE.TH.93TH057    MS-L......E-S---G-I----LV--.--T--A-.-F.K---G---..----.-VK-IR------------DTD-LAK....L----DFD-.-VG-N-----...  81
01_AE.TH.95TNIH047  MT-L......E-S-V-G-I-VL-L---.--T--G-.$L.K-------..----.-VK-IR------------DTD-LAQ....L---EDFD-.-VG-N-----...  80
01_AE.TH.CM240      MT-L......E-S---G-I--L-L---.-$T--A-.-V.K-------..----.-VK-IR------------DTD-LAK....L----DFD-.-VG-N-----...  80
01_AE.TH.TH022      MT-L......Q-S---G-I--L-L---.-$T----.--.K---K---..----.--K-IG------------DTD-LAK....L----DFD-.-VG-N-----...  80
01_AE.TW.TWE13      MS-L......E-S---G-II-L-L----.--T--L-.--.K---S--I..----.-LK-IG------------DTD-LAK....L----NFD-.-V--N-----...  81
01_AE.TW.TWE6       MT-L......QLS---G-I-VL-L---.--T--A-.-L.K---K---..----.-VK-IR------------DTE-LAK....L----DFD-.-VG-N-----...  81
CONSENSUS 02        MqsL??..??eia---G----fiaa--.--t--y-.--.---RK-k-..--k.-L--IR--a-------d-DtE-L-t....Lm---ydni?ldndN-----...  81
02_AG.CM.MP807      MT-LTX..XLQ-A---G----F-V---.--I--L-.--.---RK-K-..----.-L--IR----------D-DTE-L-T....LL---NDDF.--NNN-----...  85
02_AG.DJ.DJ258A     MNSL......E-A---G----FTA---.--T--L-.--.---RK-K-..--K.-L--IR----------D-DTE-L-T....LM---YDNI.LNN-N-----...  81
02_AG.FR.DJ263      M-SL......E-T---G----F-A---.--A--F-.--.---RK-K-..--K.-L--IR----------D-DAE-L-T....LM---YDHI.L-N-N-----...  81
02_AG.FR.DJ264      MKSL......E-A---G------AG--.--T--V-.--.---RK-K-..--K.-L--IR----------D-DTE-L-T....LL---YDNI.L-N-N-----...  81
02_AG.GH.G829       M-SL......E-A---G----F-A---.--I--Y-.--.---RK-K-..--K.-L--IR--E---------DAE-L-K....LM---X-DF.-XI-N-----...  81
02_AG.NG.IBNG       M--L......T-T---G----F-A----.-----Y-.--.---RK-K-..----.-L--IR----------D-DTE-L-T....LM---YEYI.L-N-N-----...  81
02_AG.SE.SE7812     MNSL......D-----G----F-A---.--T--Y-.--.---RK-K-..----.-L--IR----------D-DTE-L-T....LM---YDNI.L-N-N-----...  81
02_AG.SN.MP1211     M--L......AVA---G----F-L---.--T--F-.--.---RK-G-..--K.-L--IR----------D-DTE-L-T....LL---YDNAAL-I-N-----...  82
03_AB.RU.KAL153-2   M-SL......A-A-------VG------.-G---F-.--.---------..----.----IR------------DQE...-....LM----LV-.--A------...  78
03_AB.RU.RU98001    M-FL......V-A-------GG-----.-----F-.--.---------..----.----IR------------DQE...-....LM----L--.--A------...  78
04_cpx.CY.94CY032-3 MLFW......E-W---G----L--V--.--TL-F-.--.K-LR---R..--S.-YN-IR----------D-DAE-L-T....L-G--NFD-.-VG-N-----...  81
04_cpx.GR.97PVCH    MSSLS...SLE-WT----I--L-----.--T--Y-.--.K-LQ----..----.-YK-IR----------D-DTE-L-T....L-G--DFD-.-VGNNV----...  84
04_cpx.GR.97PVMY    MSSMS...SLQVW--T--I--L----I.--T-AL-.--.--L-----..----.-CK-I---------D-D-DTE-L-T....L-D--DFH-.LVGNN-----...  84
05_DF.BE.VI1310     MSDL......LTI-V---I--L-----.--I--Y-.--.K-LV----..-N-.-YK-IR------------DAE-LA-....LG-V-PFI-.G-INN-----...  81
05_DF.BE.VI961      MSDL......LA---A-FI--L-----.--T--YL.--.--LV-----..-N-.-YK-IR------------DAE-LA-....LG---PFI-.GNINN-----...  81
06_cpx.AU.BFP90     M-AL......E-A---G----FLA---.--T--F-.Q-.-E-RK-K-..-EK.-L--IR------------DTD-LAT....LM---DFD-.-VG-N-----...  81
06_cpx.ML.95ML127   M-SL......E-S---G----FLA---.--T--F-.Q-.-E-RK----..-EK.-LN-IR------------DTE-LA-....FM---NFD-.-VG-N-----...  81
06_cpx.ML.95ML84    M--L......E-S-V-G-I--FLA---.--T--Y-.Q-.-E-RK-K-..-EK.-L--IR------------DTD-LAT....LM---NFD-.-VG-N-----...  81
06_cpx.SN.97SE1078  M--L......E-S---G-I--FLA---.--T--F-.LH.-Q-KK-K-..-EK.-L--IR--E-----D---DTE-LAT....LM---NFD-.-VG-N-----...  81
10_CD.BFL061        MLSLLA.RVYYKLRVG--I--L-----.--IR-Y-.Q-.--LVEPT-..--W.--K-IR-----R-HD-D-DTE-L-T....M-D--NLRL.L--N-----...  86
10_CD.BFL071        MLDLIA.RIDYRLGVG--A--L-----.--T--Y-.--.--LV----..----.--K-IR----------D-DTE-L-T....M-D---LR-.L--N-----...  86
10_CD.BFL110        MLNLTA.RIDYRLGVG--II-L----I.--T--Y-.-$.--LV-----..--W.--K-IR----------D-DTE-L-T....M-D---LRL.L--N-----...  85
11_cpx.CM.MP818     MI-L......Q-AG----I--A--V-I.--T--F-.--.--L-K-K-..----.----IR------------DTE-LAR....L----P-DQ.-N-N-----...  81
11_cpx.FR.MP1298    MVSL......Q--S-A--I--F---V-.--T-AF-.--.--W-K-K-..-EK.----IR------------DTE-LAQ....L----P-DL.-N-N-----...  81
11_cpx.FR.MP1307    ITTL......Q--G----I--F---V-.--T-AG-.--.--W---KR..----.----IRD-----------DTD-LAK....L----P-DL.-H-------...  81
11_cpx.GR.GR17      MITL.....RTVKHSSIS$QHLL---I.--T-AG-.--.-RL--H--..X--.-XX-IRX-----------DTE-LAQ....L----P-DL.-N-N-----...  81
CPZ.CM.CAM3         MLTW......EQIGLIGIGIE------.A-G-AFKEWR.KGKEEENRAQKI-T--E-I-S-Q-------D-EDQRQLD.....R-IHVYGFDNPMF-W-...  85
CPZ.CM.CAM5         MLIW......EQIGLI--GIEL--V--.--G-AYKKWR.EIKEERNRTQRIYN-AE-IRI-Q-------N-EEEAQLD.....NLLHANGFDNPMF-WP$   86
CPZ.GA.CPZGAB       MTLL......VGL.VLI--GL-AWN-C.I-GYI-K.WGY-RYK-H-LETEIER-NLI-R----------N-EEE-RLE.QLIHNYNHNNHFANPMF-----...  88
CPZ.US.CPZUS        MLNW......FEIGLI--GIEG-LVVI.I-GL-ARLW.RQIKIKENTQQEIQN-LE-IRI-E-------D-EEE-TL.....AKLLSSLELDNPRIV....  84
```

HIV-1/SIVcpz proteins

HIV-1/SIVcpz proteins

```
                                   signal peptide \/ gp120
                                              <- vpu cds end
                                                 *                    *                    *            ^^^                          *
B.FR.HXB2          MRVKE....KYQ.HL.WRWG..WRWGTMLLGMLMICSAT..EKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLVNVTENFNMWKNDMVEQMHEDIISLWDQSLKPCV   120
CONSENSUS A        -rvmGiqr.Ncq.hl.wr--......i-ilgmiIi-saa..en----------v-kd-e------------dt-v--------------n---ihle----e-n----n--e--ht----------k---   119
A.GB.MA246         --AMGIQR.NC-.S-.----------.I-I---II----A..GN-------------D-E------------T--------------------IN-E----K------N------T----------Q---   119
A.KE.K89           ---MGTQR.NC-.--.L----------.--I---II----A..-N-------------N-E------------E--K------------------IH-E----E------N------T----------   119
A.KE.Q23           ---MGIQR.NC-.--.LT---------.I-I--TIIF---V..-N-------------RD-D------------E--K------------------IH-D----K------N------T----------   119
A.RW.PVPI          ---MGIQM.NC-.N-.----------.-I----I---VA..GN-------------D-E---------------A------------------IL-K----E------G------T----------   118
A.SE.UGSE8131      ---MGTQR.NC-.--.LN---------.I-I---II---TA..-N-------------D-E------------EK--------------------LY-E----D------N------T----------   119
A.UA.ukr970063     -KARGMQR.N--.--.----------.---FW-IIM-K-A..-N------------RD-E-------------K--------------------IS-E----K-D----N-----QT----------   119
A.UG.92UG037       ---MGIER.N-P.CW.-T--------.I-I---II--NTA..-N----------I--D-N-----------------------------S---LKME----E------N------T----------   119
A.UG.U455          ---MGIQR.N-P.C-.----------.--I--LII--N-...QQ-------------D-V-------------A--------------------ID------E------N--D----------------   118
A.CD.97CDKS10      ---MGTQT.S--.--.----------.ILI----I--K---.D.W------------D-E--------D---E--A-------------------N-K----D------N--------------------   118
A.CD.97CDKTB48     T--MGTQR.NC-.KW.-E--------.ILVF--I-M-K-A...D------------RD-D------------A--K-------------------N-A----D------N------A----------   118
A.CY.94CY017-41    ---MGTQR.N--.--.--G--------.ILI----IM-K---...D-------------D-D-I-----------------------------IN-E----------N-----Q----------------   118
CONSENSUS B        -rvkgiRk.nyq.hl.-rw-??????tml--mlm-csaa..enl-----------k-at-------d-kaydt-v-----th----t--n---vvle----n-n-w--n-ve--h-di-s---q------   119
B.AU.MBC18         -T--GTRK.N-R.--.-T--------.--I--I------A..NNF-----------R----------------E---------------------E-I-----------N--------------------   119
B.CN.RL42          ---TGIRK.N--.--.----------.----------N-A..-N-------------------------------------------------G----------N--------------------   119
B.DE.D31           -K---IRK.N--.--.----------.----------..------------------------------K---------------D------E----D------N--------------------   119
B.DE.HAN           -K---TRK.N--.R-.--G--------.L-------S-VA..GN-------------------------------------------------MG----------------Q----------------   119
B.ES.89SP061       ---T-IRK.N--.--.-K--------.----------K-A..-N-------------------------------------------------E-K----------N--D----------------   119
B.GA.OYI           -TARGTRK.N--.R-.----------.------------A..-N-------------------------R--A-----------------------G------D----N-----Q----------------   119
B.GB.CAM1          --A-GIRK.NC-.R-.----------.------------A..D-------------------------------------------------E----------N--------------------   119
B.GB.WB            ----GIRN.N-....-.-K--------.----------K-A..-Q-----------------------------------------D----A-E----------N--------------------   117
B.JP.ETR           -----IRK.N--.--.----IML---------------A..-Q-------------------------------A-----------------------------N-------N-------------   125
B.JP.JH32          ----GIRK.N--.--.----------.-----I------A..-Q---------------A-------------------------------------E----K------N--------------------   119
B.KR.WK            ----GIRK.N--.-W.----------.I-----W-----A..-----------------------------------------------I-------F-E----------N--------------------   119
B.NL.3202A21       -K---TRK.N--.--.----------.------------A..-Q-----------------------------S-----------------------A-E----D-D----N----------N----------   119
B.TH.93TH067       -K-TGIRK.N--.--.----------.------------A..-N---------------N-----------------------------------G----------N--------------------   119
B.TT.QZ4589        ---RGIRR.NC-.--.-T--------.-------I------..-Q-------------------------------K--------------S------G----------N--------------------   119
B.TW.TWCYS         ---RGTRM.NC-.--.----------.---------S--A..-N---------------------------T-----------------------T-G----------N-AD--Q--------E------   119
B.UA.UKR1216       -----IRK....--.----------.LF-----------..-NS---------------------------SK--------Y-------S---IP-K------S-G--N--------------------   116
B.US.DH123         ---MGIRK.N--.--.-KG-------.L---I------A..-Q---------------N-------------------------------IL-E----D------N--------------------   119
B.US.JRCSF         ----GIRK.N--.--.-KG-------.IL---T------V..----------------T-------------------------------------E----D------N-----Q--V-N----------   119
B.US.JRFL          ----GIRK.S--.Y-.-KG-------.-L---I------V..-------------------------------------------------E----H------N-----Q----------------   119
B.US.MNCG          ----GIRR.N--.-W.-G--------.-----L-------..-------------------------------------Q------------E----------N--------------------   119
B.US.RF            ---M-MRK.NC-.--.-K--------.------------A..-D-----------------------E----K--------K-------------L-E----------N--------------------   119
B.US.SF2           -K--GTRR.N--.--.----------.L-----------..--------------------------R----------------------------G----------N-----Q----------------   119
CONSENSUS C        -rvmGiqr.Ncq.qw.-iw-------.ilgfwmlmicngm..gnl-----------ke-kt--f----ak--et-v----a-h-------n---mvle----n----k-d--d---q-i-----qs-----   119
C.BI.BU910112      --ATGIQR.NC-.QW.-I--------.ILGFW-----NVM..GN-------------------------------A------------------M--K----D---------D---Q----------   119
C.BR.92BR025       ---EGIQR.NWK.QW.-I--------.ILGFW-V--YNVR..GN---------------K-------------A--------------------M--E---------E--------Q----------   119
C.BW.96BW0402      ---MGIQR.NC-.QW.-I--------.ILGFW-IINGMGS.....---------------KA--------R---R--------------------N-E----------D--------------------   116
C.CN.AF268277      ---TGIRK.N--.--.----------.------------A..-N---------------K------------EK--------------------M--E----------N-----V----------   119
C.DJ.DJ373A        ---MGIQM.NC-.QW.-IG-------.ILGFW--LM-NGM..GN------------Q--NP-----------E-----------------S---L-ME----------G--D---Q-------EG-----   119
C.ET.ETH2220       -K-MGIQR.NC-.QW.-I--------.ILGFW-----NGM..GN------------D-SP------------------G-F-------S---LG-E----------Q--------G-----   119
C.IN.93IN101       ---RGTLR.N--.QW.-I--------.VLGFW-----NGG..GN---------------K---L--------ER--------------------I--G----------D-----V----------   119
C.SO.SO145A        ---RGILR.NW-.QW.-I--------.ILGFW------GG..GN------------R--K--------S---ER----------------S---ML-G----------D--------------------   119
C.UG.UG268A2       ---MGIQR.NC-.QW.-I--------.ILGFWI----NVM..GN-------------------------E--------------------I--E----S---------D---Q-V----------   119
CONSENSUS D        -rvre??r.Ny-.hl.-k--------.i----m--tcs?a??enl----------v----------s-et-a-nI----------q-Iele----------N--e-----i------s-----   116
D.CD.84ZR085       ----GIKR.N--.P-.-K--------.I-------MTYSAA.DN-------------------------S-K--A--I----------IE-K----------N--D----------------   120
D.CD.ELI           --ARGIER.NC-.NW.-K--------.I----I--T---A..DN-------------------------S-E--A--I----------IA-E----------N--------------------   119
D.CD.JY1           ---MGIRM.N--.--.-K--------.I----I--T--VA..-D-------------------------S-EP-A--I----------R-IEME----------N----------N-----   119
D.CD.NDK           --AR-KER.NC-.N-.-K--------.I-------T---A..-D----------I-----------------KK-A--I----------IE-E----------N--------------------   119
D.CI.CI13          -KAR-MNR.N--.--.----------.I-------T--VA..-N-------------------------S-E--K--I----------IK-E----------N--------------------   119
D.SN.SE365A2       --AR-MKR.N--.--.----------.--------T--VA..-------------------------S-E--K--I----------IE-A----------N--------------------   119
D.TZ.87TZ4622      ---R-MRR.N--.--.----------.----------..-Q-------------------------S-E--A--I----------IK-E----------N--------------------   119
D.UG.92UG024-D     ---R-MQR.N--.--.-K--------.I----------G-.E-QS-------------------------$KA-A--I----------R-IQ-E----------N--------V----------   119
D.UG.WHO15-474     ---RGIQR.N--.--.-G--------.--------T--TA..--K-------------------------S-K--A-DI----------R-IE-E----------N--------------------   119
F1.BE.VI850        ---RGMQR.NW-.--.GK--------.LLF--I-I--N-A..DN-----------------------ER-A-------------------F-K------D----N------T----------   119
F1.BR.93BR020-1    ---RGMQR.NW-.--.GK--------.LLF--T-I--N-A..-N-------------------------S-EK-A---------------------E----R----E-N------T----------   119
F1.BR.BZ126        ---RGMQR.NW-.--.GK--------.LLF--I-I--N-...-N-------------------------S-EK----------------------E------D----N------T----------   118
F1.FI.FIN9363      ---RGMQR.NW-.--.GK--------.LLF----I--K-A..DD----I----------N---------S-EK--------------------A-.----------E--------K----------   118
F1.FR.MP411        ---RVMQR.NW-.--.GK--------.LLF--I-I----A..DN-------------------------G-ER--------------------IW-K------D----N----------------   119
F2.CM.CA20         ---R-MQR.NW-.--.G---------.LLF--I-I----A..D-----------------------ER--------Y-------S------G----------N--D----------E------   119
F2.CM.HIM277819    ---MGIER.N--.--.-K--------.-------LMTYSAA.DN--------------S--------------I------Y-------S---LF-E----------N------A----------   120
F2.CM.MP255        ---R-MQR.NW-.--.GK--------.LLF--I-I--N-NATDD--------------P--------------P-------Y----------L--G---------E-N--D---L----------   121
F2.CM.MP257        ---R-MQR.NW-.--.G---------.LLF--I-I----A..D-----------------------ER--------Y-------S---L--G--S-K------N--D----------E------   119
```

```
                                        signal peptide \/ gp120
                                                <- vpu cds end
                                                *                   *                   *           ^^^                         *
B.FR.HXB2          MRVKE....KYQ.HL.WRWG..WRWGTMLLGMLMICSAT..EKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVLVNVTENFNMWKNDMVEQMHEDIISLWDQSLKPCV   120
CONSENSUS G        -rvkGIqR.NW-.h-.-k-g......tLi--LVI--sas..nn----------v-ED-dtt----s---aySt-s--v-----------n---I?le-vt-n--m---n-----h------dE------   118
G.BE.DRCBL         ----GIQR.NW-.--.-N-......ILI--LVI----...--------------ED-NAP---------HS--S--I-----------S---INMR-----------N-------------E------   118
G.FI.HH8793-12-1   ----GIER.NW-.--.-K-......-LI--LVI----S..NN------------ED-K-----------S--R--I-----------D---IP-G-------V-----------------E------   119
G.GA.LBV217        --A-GILR.NW-.--.-K-......ILI--LVI--N-S..GN----------A-ED-D-----------S--S-----------------IS-N-----------N-------------E------   119
G.NG.92NG083       ----GIQR.NW-.--.-K-......-LI--LVI----S..DN------------ED-D-P--------S-SS-K------------------IAIE-----------N-----Q-------EE------   119
G.NG.NG1937        ----GIQR.NW-.Y-.-T-W......-LI--LVI---TS..NN------------ED-D------P-----S--R------------------IT-E-I--I------N-------------E------   119
G.NG.NG1939        -K--GIQR.NW-.--.-T-W......-LT--LVI----S..NN------------ED-D-P----------S--K------------------IE-K--K--------N-------------E------   119
G.SE.SE6165        ---TGIQR.NW$.--.-K-......-LI--LVI----S..NN------------ED-D----------S-SA-S------------------IIME----Y------N-------------E------   118
H.BE.VI991         T--M-TQR.N-P.S-.---------LI----L---VV..GN---------------K--------------R------------------M--E----T----V--------T--------------   119
H.BE.VI997         T--M...R.N-P.QW.--G-......IL-----L-Y--A..GN---------------K------------EP-K--------------S---M--A---------D-------QT--------------   116
H.CF.90CF056       T--M-TQR.N-P.S-.---------LI----L----A..QN---------------K------------E--K------------------M-ME----S----E-N------T--------------   119
J.SE.SE7022        T--M-TQT.SWL.S-.---------L-IF--------R..-N------------RD-K------------S--K------------------MS-P--------------D--Q-----V--E------   119
J.SE.SE7887        T--M-TQK.NW-.T-.--G......L-IF------K-K..-D-------------D-K------------S--K--------------S---MN-P--------------D--Q-----V--E------   119
K.CD.EQTB11C       --AR-IQR.NW-.--.GKR-......ILF--I-I----A..NN----------------------------E------------------------E-----------N------T-------E------   119
K.CM.MP535         ---RGMQR.NW-.T-.GN-......ILF--I-I---NA..D-----------------P-----------EK---------------------EME-----------N------T-------E------   119
N.CM.YBF106        ...MGMQS.GWP............FFCL-ISLTI...GS..DPH-----------RD-E-V----------HS--A--I---Q-----------L-T----Y----E-K-A---Q-------E-------   109
N.CM.YBF30         -K-MGMQS.GWM.GMKSG-.....LLFY--VS-IKVIGS..-QH-----------R--E----------HS--A--I---Q-----------L-P----K----E-K-AD--Q-------E-------   121
CONSENSUS O        ????MkvMekkNkKl.w.......ilyiaMAliiP-Lss?.?qlYA---a-----ed-tPV---a--aNLTS--k--I--SQ-------t-?-yP-h---Dn-nI-k-Y--e--q-------e--l---v   114
O.CM.ANT70         ....MKAME-RNKK-.-.......TLYLAMALITP-LSL..RQ-YA---A-----ED--PV-------NLTS--K--I--SQ-------T-Y-YP-H---DD--I---Y-----Q---------------   116
O.CM.CM4974        ....MKVME-KNKKS.-.......ILCIAMAL-IP-LNCX.NQ-YA---S------D--PV-------NLTS--K--I--SQ-------S-H-YP-E---DK-DI---Y--D-----------------   117
O.CM.HIV1CA9EN     ....MKVME-KNRK-.-.......ILYIVMALIIP-LSS..KQTYA---A-----GD-APV---T--VNLTS--K--I--SQ-------T-H-YP-H---DD--I-N-Y-----Q----------S----   116
O.CM.MVP5180       ....MKVMK-NNRKS.-.......SLYIAMAL-IP-LSYS.KQ-YA---S-----E--APV-------NLTS--Q--I--SQ---------H-FP-G---D--DI---Y--D---------E------E   117
O.GA.VI686         ....MRVME-KNKK-.-.......ILC-IMALIIP-LSS..NH-YA---A-----ED-NPV-------NLTS--Q--I--SQ-------T-L-YP-K---D---I---Y-----Q-------E-------   116
O.GQ.193HA         ....MKAMERKNKKS.-.......IWYL-MALITP-LSS..KP-YA---A-----ED--PV-------NLTS--K--I--SQ-------T-Y-YP-H---DK--I---Y-----Q-------E--F----   116
O.SN.MP1300        -T-TMKVME-RNRK-.G.......ILC-VMALITP-LSH..NQHYA---A-----E---PV------VNLTS--Q--I--SQ-------S-Y-YP-K---D---I-E-Y-----Q-------E-------   120
CPZ.US.CPZUS       -K-....ME-K.KR-.-.......LSYCL-SSLIIP.GLS...S--A---------RDVE------------KQ-A--I---Q------------H-P----K-D--E-N-A---Q--------------I   114
CONSENSUS 01       -rvk?Tqm.NWP.n-.-K-g......-LIl-LVIi---S..dn-----y------rD-d-t-------k-he-----v----a----------IhlE------n----n-v---Q--V-----q--k---   118
01_AE.CF.90CF11697 -G--GTQM.NWP.--.-K-......-LI--LVI----S..DT------------RD-D---------N-QE----------V----------IH-E------D----N-----Q--V------------   119
01_AE.CF.90CF402   ----GTRR.NWP.N-.-K-......-LI--LVI----S..DN------------RD-D-I---------HV--------------------IY-E------D----N-----Q--V--------Q---   119
01_AE.CF.90CF4071  ----GTQM.NWP.N-.-K-......-LI--LVIM---S..DN------$-----RD-D-----------HE-----I---------------ID-E-----------N-----Q--V---$--------   117
01_AE.CM.CA10      --ARGTQM.NWP.N-.-K-E.....-LIF-LVI----S..D------H-------D-D-----------HE--------------------IH-E-----------N-----Q--V-----E------   119
01_AE.TH.93TH065   -----TQM.N$P.N-.-K-......-LI--LVI----S..DD------------RD-D-----------HE--------------------IH-E-----------N-----Q--V------------   118
01_AE.TH.CM240     -----TQM.NWP.N-.-K-......-LI--LVI----S..DN------------RD-D-----------HE--------------------IH-E-----------N-----Q--V------------   119
01_AE.TH.KH03      -----TQM.NWP.N-.-K-......-LI--LVI----S..DN------------RD-E-----------HE--------------------IH-E-----------N-----Q--V------------   119
01_AE.TH.TH022     -----TQM.NWP.N-.-K-......-LI--LVI----S..-N------------RD-D-----------HE--------------------IHME----------K-A---Q--V------------   119
CONSENSUS 02       ---MGiq?.nyp.ll.-r-......?tiifWimi--n-...e?-----------wrd-e------------dt----------------n-q-ih-k----k------n------e--------------   116
02_AG.FR.DJ263     ---MGMQR.N-P.L-.-K-......-IIFWIMI--N-...--------------RN-E-------------A--------------------IH-K----K------N--------------------   118
02_AG.FR.DJ264     ---MGIQR.N-P.LF.-K-......MIIFWIMI--N-...--------------RD-E-------------V--------------------IH-K----K------N--------------------   118
02_AG.NG.IBNG      ---MGIQK.N-P.L-.---------NIFWIMI--N-...-Q-------------T-E----------------------------------IH-E----K------N--------------------   118
02_AG.NG.NG1921    ---MGTQR.NCS.L-.---------II-WLII--N-...KD------------RD-D------------EK----------------D-H-ML-K-----------N--------------------   118
02_AG.SE.SE7812    ---MGIQK.N-P.L-.---------MIIFWIMT------.GN------------RD-E----------------------------------IH-G----D------S--------------------   118
02_AG.SN.MP1211    ---MGILK.SCP.PF.--------MI---WI-I--N-...-N-----------RRD-E------------E-----------------S---MH-E-----------N------V--------------   119
03_AB.RU.KAL153-2  -----IRK.....--.----------LF-----------..-N----------------------------SK--------Y-------S---IP-K--------G--N--------------------   116
03_AB.RU.KAL68-1   -----IRK.....--.----------LF-----------..-N----------------------------SK--------Y-------S-$-IP-K--------G--N--------------------   115
03_AB.RU.RU98001   -----IRK.....--.----------L-----------..-N----------------------------SK--------Y-------S---IP-E--------G--N--------------------   116
04_cpx.CY.94CY032-3 ---MGMQR.N-P.--.-E-......-LI--LVI----S..NN------------RD-E-------E----EK----I----------------A-I-----------------------NEG----A   119
04_cpx.GR.97PVCH   ---MGIQR.N-P.--.-E-......-LI--LVIM---S..KDM-----------RD-E-KPX---------K-I--I---------------LA-T---------E-ST-------------EG-----   119
04_cpx.GR.97PVMY   -T-MGTQR.NCP.R-.-T-......-FI-WLVI----S..NN------------RD-E-SP----------K----I---------------IF-K----D------N-------------EEG-----   119
05_DF.BE.VI1310    ---RGMQR.NWP.--.GK-......LLF--I-I------.D-F-----------R--K----------G--K-----------------G------------------T--------------   119
05_DF.BE.VI961     ---RGMQR.NW-.--.GK-......LLF--I-I----A..DN---------------N----------G-EK-A------------------LA-E-----------N------A-----------S--   119
06_cpx.AU.BFP90    ----GIQT.SW-.--.-K-......-LI--LVI----S..KNM---------A-ED-D-I----------SPDK------------------IS-K-----------N--D-----------E------   119
06_cpx.ML.95ML127  ---RGIQK.NW-.--.-K-......-LI--LVI----S..NN----------A-ED-D-I----------SA-K------------------IP-K---------E-H-------------E------   119
06_cpx.ML.95ML84   ----GIQM.NW-.--.-K-......-LI--LAI------..SN----------ED-D-I----------SA-K------------------IP-E-----------H--Q----------E------   119
06_cpx.NG.NG3670a  -T--GIQK.NW-.--.-K-......-LI--LVI----S..NNM---------A-E--D-I----------S--K------------------IT-E-----------H-------------E------   119
06_cpx.SN.97SE1078 -T--GIQR.NW-.--.-K-......-LI--LVI----S..QNM---------A-ED-D-I----------SA-K------------------IK-E-----------N-------------EE-----   119
10_CD.BFL061       -T-TGIQR.NCR.QW.-I-......ILGFW-----N-R..----------------T-------------KA-A--I-------------R-I--E-----------G--D-----------G----I   119
10_CD.BFL071       -K-MGIQR.N-P.QW.-I-......ILG-W---.ICN...GN--------------------------Y-KE-A--I--A------------L--E-I-----------D-----------------   117
10_CD.BFL110       ---MGIQK.NC-.PW.-I-......ILGFW-----NV-..G---------------T----------G-K-----I---------------I--E-----------G--D-----------G-----   119
CONSENSUS 11_cpx   -rvk?tQr.NWh.n-.-r-......-L-If------N-t..dnl-----------rD-?-t-------kaySt-K------ha------n-q-ip-e----n--m---n--e--h--i-----E-----v   117
11_cpx.CM.MP818    --ARGTQK.NWH.D-.---------L-IS------N---.DN------------D-N--------Q--SP-K-------Y----------IL-G-----------N-------------E------   119
11_cpx.CM.CA1      -K-RGSQM.NWH.N-.-K-......L-IF------N---.D------------RD-D----------S--K--------------D---IY-E-----------N-------------E------   119
11_cpx.FR.MP1298   -----TQR.NWH.N-.---------L-IF------N---.DN------------RD-N----------VS--K-------S----------IF-E-----------N-------------E------   119
11_cpx.FR.MP1307   -----TQM.NW-.N-.---------L-IF------N-A..GN------------ED-N----------S--K------Y----------MP-K----E------N-------------E------   119
11_cpx.GR.GR17     -----TQR.NWH.N-.---------L-IF------N-...--M----------RD-D-I-------RT-S--K-------S--------R-LS-E-----------N--------V-----E------   118
CPZ.CM.CAM3        .......ME-I.KGS.......LIEKY-FM-LIIP-LTGS.DQ--A----------D-------E-S-LNK-A--I--SQ----------QIP---------N-T-A---NT-------------   114
CPZ.CD.CPZANT      -------KPIH..II.-.....GLALLIQFIE....KG-..NEDY---F------RN--P-----TN-SMTS---------TS---I--D-IV-R-.-TSVW--AY--Y---S-T--MXQ-FQ--H----   110
CPZ.CM.CAM5        .......MERK.KLN.-......IICY--M-LITP-LTG..HEW-A----------E-------E-S-LNK-A--I--SQ-----V-------YMP--S-S------N-A---Q---S----------   113
CPZ.GA.CPZGAB      -K-....ME-K.KRD.-..NSLSIITIITIIL-TP-LTS...E----------HD-DPV---------HS--A--I---Q-------S----F-P--I-S------N--D-----------------   119
```

HIV-1/SIVcpz proteins

HIV-1/SIVcpz proteins

```
                                 |-              V1                -|-                          V2                    -|
                          *    *    ^^^            ^^^            ^*^ ^^^                      ^               ^^       *^^^      *            *
B.FR.HXB2          KLTPLCVSLKCTDLKNDTN..........TNSSSGRMIMEKGEIKNCSFNISTSIRGKVQKEYAFFYKLDIIPIDN.............DTTSYKLTSCNTSVI.TQACPKVSFEPIPIHYCAPAGFAI    225
CONSENSUS A        k------T-n-sn??tnnt.......??????ntn?tdgmreemk---f-mT-EL-DkkqkVySL--?--iVqin?.???????????nssnq-R-IN----ai.---c---s------h--a---f--    217
A.GB.MA246         -------T-N-S-VNVNAT...........-YNVNDT-NM-E--R-----MT-EL-DRK--V-SL--R--VVQMN-............SNNSNQ-R-IN----A-.----------------------Y--    225
A.KE.K89           -------T-N-SNVNVTAT..............--TTHGGR-RM----Y--T-EL-DRR--V-SL--R--VV--NE...........NNSSNNE-R-IN----A-.-------T----------------    223
A.KE.Q23           -------T-H--NVTSVNT..................TGDREGL------MT-EL-D-R--V-SL--R---V--NE.............NQGSE-R-IN----A-.-----------------T------    217
A.RW.PVPI          Q------T-D-NYNITINS.......TKTPSPVTNISEEMRE--------MT-EL-D-K--VHSL------VHLND.............SKPKQ-R-IN----A-.------------------------    227
A.SE.UGSE8131      Q------T-N-SNNVTANT..........NST-ANLTDSV---MR------T-EL-D-KK-V-SL------VK-NK.NKSFRGKNSSGNSSSDR-R-IN----A-.------------------------    237
A.UA.ukr970063     -------T-N-AEP-STSS........NNSSVN-NSSDSLFE-M------MT-EL-D-RKTVHSL------VSTGX.............NGSGQ-R-IN----AM.-------T------Y---------    227
A.UG.92UG037       Q------T-D-SYNITNNI.............---ITNSSVNMRE--------MT-EL-D-NR-V-SL-----VVQ-N-...........GNNSSNL-R-IN----AL.---R---T-------------Y--    227
A.UG.U455          -------T-D-HNITINNT...........-NNTNITDGVRE-M------MT-EL-D-K--V-SL--R---VQ-NK.............T-NN--R-IN----T-.------------------------    223
A2.CD.97CDKS10     -------T-N-SNANTNST............N-TSAPS-GP---------VT-EV-D-EK-V--L-----VVQ-NE.......SDSNSTK-S-Q-R-IN----A-.------------------------    228
A2.CD.97CDKTB48    -------T-N-SNANTTNT...................NSTE------Y-MP-ELKD-T--V-SL--E--VVLLNR.........SKNSSYST-R-I-------.----------------Y--    218
A2.CY.94CY017-41   -------I-N-SNANTS-H...........SN--STQSPINE------Y-TT-IL-D-T--V-SL--R--VVQL-E......SENKNTSGSN-L-R-IN----T-.-------T----------------    230
CONSENSUS B        kl-----t-n-tdlnnttn???????????tsssggtmekg-ik----nittsirdkvqkey-lfykl-vvpidn.?????????????nntsyr-is-nt-vi.t-----v-----pi-y-?-----i    221
B.AU.MBC18         --------N---A--T--...SNTNSSSSTN--SLEQGKA----------T-NM-D----Q--L--S---V---D............KGND---R-I-------.------I------------------    233
B.CN.RL42          -------T-N--N---A--...............TSST--G----------T---KT--.-D--L---V-VV--G-..............-S---R-IN------.-----------------T------    218
B.DE.D31           -------T-N------A--..........TN---WTMTGEM----------T----D-------H-----VV----...............-N---R-I-------.-----------------T-----M    225
B.DE.HAN           -------T-----YN-A--..............-SWGR------Q----KVT-N--D-----S-L---T-VV----........NKTSNRDN----M-IH--R---.------------------------    227
B.ES.89SP061       -------T-N---YG-----...TNNSSATNPT--SGGMEGR----------TR---D--K----L--S--V---KD..............-N---R-R-------.------------------------    231
B.GA.OYI           -------T-D---VNTTSS....SLRNATNTT--SWET-----L------TT----D-M-EQ--L-----VL---K..............ND-KFR-IH----T-.------I------M---T------    230
B.GB.CAM1          -------T-I--NVN-TRT........NSSDWDRREGEKM----------VT----N--R----L-----VV---K..............AN---T-IH------.-----------------T------    226
B.GB.WB            -I-----T-----VTFNCT....GTSNCTDTN--SPLK-----M------VT-N--NQ-R----L-----VV--NK..............ND-NFI-IH------.------------L-----------    228
B.JP.ETR           -------T-H----R-T--.............NN-SIEEKM----------VT-N--D-------L-----VV----------.....DNNSTN-C-R-I--D----.-----------------T-----L    232
B.JP.JH32          -------T-N-I-WG---S......PNATNTT---GEK-----M-------T----D-----H-L---H-VV--N-.....STKDNIKNDNS-R-R-I-------.------I-----------------    237
B.KR.WK            -------T------NDTNTNNSSTSENNTNPTI--GEG-GE--M------VT-N--D-------L-----------------------------.--A-RH------.------------------------    232
B.NL.3202A21       -------T-N---FG-A--............TT--SGV-I----------K-N-NMKD-A-I---L----VV----.........NNTNTSY---R-I-------.---------------F-T-----L    228
B.TH.93TH067       -------T-N----R-A--...............TSSP-GG----------T-.SIRTKV-D--L--TV-VV---D..............-N---R-LN------.-----------------T------    218
B.TT.QZ4589        -------T-N---KLRNNI..............-SGEK--I----------TSG--N-M-----L----------D....DNNRNNSSNNTVRN-R-IN------.-----------------T------    230
B.TW.TWCYS         E------T---N-TTMSK-............D-NTLT------------VT--L-N-------S--R--LVQT-E..............NS---R-I---A---.K-----------------T------    221
B.UA.UKR1216       -------T-N-----KNAT.......STN-S-ITNTSSI-RT-M-------T-DL-D--K----L--N--VVQ-----------------.D--R-I------V.------I-----------------    222
B.US.DH123         -------T-H------G--..........LK-GTKIIGKSMR---------VTKN-ID--K----L--RH-VV---R.............NI---R-I-----TL.-----------------------    225
B.US.JRCSF         -------T-N-K-VNATNT............T--SEGM--R----------TK---D-------L-----VV---------------.KNN-K-R-I-------.------------------------    223
B.US.JRFL          -------T-N-K-VNATNT...........TNDSEGT--R----------T----DE------L-----VV----------------NN---R-I--D----.------I-----------------    222
B.US.MNCG          -------T-N----R-T--....TNNSTAN-N-NSEGTIKG--M-------T----D-M-----LL-----VS----------------S---R-I-------.------I-----------------    230
B.US.RF            -------T-N---ANLNGT.........NVT----GTM--N--------QVT--R-D-T--K--L-----VV--EK.GNISPKNNTSNNTSYGN-T-IH--S---.-----------------T------    238
B.US.SF2           -------T-N----GKA--...........TN--NWKEEI-----------T----D-I---N-L-RN--VV----------.ASTTTNY-N-R-IH--R---.-----------------T------    228
CONSENSUS C        -------T-n-tnvnin?t....?????????g?ntynsm??eik-----iTtelrDkkqk?y-l--r--iVpLnn........??????ns?e-R-In----a-.-------sfd-------a---Y--    215
C.BI.BU910112      -------T-D-KNANATH-............GTIDNRTMG----------T-ELKD-K-RAH-L--S---VQL-G..............GG--R-I-----A-.---------D-----------Y--    220
C.BR.92BR025       -------T-H-SNRTI-Y-..............N-TDNMG---------MT-EV-D-RE-VH-L--R---V-LK-..........ESSNTSGD-R-IN----A-.---------D-----------Y--    223
C.BW.96BW0402      -------T-N--EVNGTSD......SSVPA-GT-NSSVSM-E-MR------T-EL-D-NKQ-R-R--R---VQLN-............NSNSNE-R-IN----T-.-------T-D-----------Y--    227
C.CN.AF268277      -------T-E-RNVSSNGT................YNETYN--------AT-VL-DRK-TV--L--R---V-LNK........KNSSENSSEY-R-IN----A-.-------T-D-------T---Y--    223
C.DJ.DJ373A        -------T-N-SNKVTA-G............NG-VTYNSTEKDMR------T-EL-D-K--K--L------V-LG-..............GNEE-R-IN----S-.---------------------Y--    222
C.ET.ETH2220       -------T-N-NAI--N-K............VTNNSINSAND-M-------T-EL-D-KR-A--L------V-LN-..............GS-D-R-IN----T-.--------LD-----------Y--    222
C.IN.93IN101       -------T-E-RNVSRNVS............-YNTYNGSVE--------ATPEV-DRK-RM--L--G---V-LNK........KNSSENSSE-R-IN----A-.-------T-D-----------Y--    226
C.SO.SO145A        -------T-N--EANINIT...........SVGNVTSNDN-E--------AT-ELSD-K--V--L--R---V-LNE..........GENNSRK-R-IN----A-.---------D-----------Y--    226
C.UG.UG268A2       -------T-N--NVNVNIT....NNANATNSPYENGKL--Q---------VT-E--D-K-TAH-L-----VVSLEG..............NSNT-R-IN----A-.-------TLD-----------Y--    230
CONSENSUS D        ----l--T-N-tdasrn?t....?????????ntngp?mekgemkn----ittv?rD-kkqv?aL--r--?vpidn....???????nntnyts-R-IN-n--a-.t-a---vt-------y--------    216
D.CD.84ZR085       ----R--T-N---ASRNST...........DNN-TLPTVKP--M-------T-VVTD-RKQVH-L--R--VVQ---------.EGKNEINDTYGT-R-IN----A-.------------------------    232
D.CD.ELI           -------T-N-S-ELRNNG............TMGNNVTT-EKGM------VT-VLKD-K-QV--L--R---V----------.DSSTNS-N-R-IN----A-.------------------------    226
D.CD.JY1           -------T-N--NAGGNKT......TNGNNTTNQEEQM-----M-------T-V-SD-KKQVH-L--R--VV---D....DNSANTSNTN.Y-N-R-IN----A-.-------T----------------    237
D.CD.NDK           -------T-N---ELRNSK..............GNGKVE-EEKR------V.....-D-REQV--L------V----------.NNRTNS-N-R-IN-D--T-.------I--------F--------    220
D.CI.CI13          -------T-N-S-VQGN-M..............ETGQ-----M-S-----T-VL-D-QE-VA-L--R---V----------.NS--R-IN----T-.-------T----------------    218
D.SN.SE365A2       -------T-N-R-ISS-AT.......SNTISNVT-IPM-G---MQ------T-E--D-R-NV-SL--R--LV--G-........SNKNSTNY---R-IN------.K------N----------------    233
D.TZ.87TZ4622      -------T-N--NII-G-S........ANNGTNTNGSVT-DPGM-------T-E--D-KKQVH-L------M--GG...........NNSTSS--R-IN----A-.------IT----------------    229
D.UG.92UG024-D     -------T-N-IEW--ENE...........-KGNKTVDPDKDIGM------VT-EV-D-KKQ---L-----VVQMND...........NNTN---R-IN----A-.--E----T----------------    226
D.UG.WHO15-474     -------T-N---ASWNET....KSGNDTSRNETSDIR-----M-------T-VV-D-QKQVH-L--R--VVQMN-........DNSTTANYS--R-IN----A-.-------T----------------    236
F1.BE.VI850        -------T-N--NAT-NSQ...................EKP-AMQ-----MT-EV-D-KL-LS-L--R---V--G-.............NNSSE-R-IN----T-.--------WD-----------Y--    216
F1.BR.93BR020-1    -------T-D-RNIATNGT........ND-IAINDTLKEDPEA-Q-----TT-E--D-QL-VH-L------VQ-NK.............D-NRT-R-IN-DA-T-.--------WD-----------Y--    227
F1.BR.BZ126        -------T-H--SATPNAT..........ATQNGTTLKE-P-A-Q-----M--EV-N-Q--VH-L------V--S-...........SSNNSRE-R-IH-D--T-.--------WD-----------Y--    226
F1.FI.FIN9363      -------T-N--NATTTND.........TLSDQ-STLKE-P-A-Q-----MT-EVED-K--VH-L--R---E--S-............NNSREE-R-IT----T-.--------WD-----------Y--    226
F1.FR.MP411        -------T-H-S-VNITS-.........A-TTNDTSTPE-S-A-Q-----MT-EVKD-KLRVN-L---------N-............SSSSD-R-IN----T-.K-------WD-----------Y--    226
F2.CM.CA20         -------T-N-A-VNVTLE...................E-R----------T-E-KD-KK----L--R--VV--NS.............NNSD-R-IN----A-.---------------F--------    215
F2.CM.HIM277819    -------T-N-DNATINDT.........NGTNVTATLKE-P----------T-E-KDRKK-QR-L--R--VV--NA...........SDSNSSE-R-I-----TV.---------------------Y--    229
F2.CM.MP255        -S-----T-N---VNITMS..........DINGTSLKEDQ---------VT-ELKD-KR-QQ-L--R--VE--K-..............SSNI---I---M-TV.---------D-----------Y--    225
F2.CM.MP257        -------T-N--KAIINVT........SSN-TTLAPNVTISE-M-------T-E--D-QK----L-----VVQ-N-..............SN---R-IN----TL.---------D--------------    226
```

```
                                 |-              V1              -|-                          V2                          -|
                          *    *    ^^^          ^^^           ^*^ ^^^                         ^               ^^       *^^^      *            *
B.FR.HXB2          KLTPLCVSLKCTDLKNDTN..........TNSSSGRMIMEKGEIKNCSFNISTSIRGKVQKEYAFFYKLDIIPIDN.............DTTSYKLTSCNTSVI.TQACPKVSFEPIPIHYCAPAGFAI    225
CONSENSUS G        k-----vT-N--nvnn???..?????t????t?nn?tvt?eee?k----niT-Eirdkkkk---l--rl-Vvpind......??????n?n?t?-R-in--v-t-.K------t-d---------a----    214
G.BE.DRCBL         -------T-N--EIN-NST................-N-T-EYRMT-----MT-EL-D-KKA---L--RT-VV--NE......MNNENNGTNS-W-R--N--V-T-.K------T----------------    225
G.FI.HH8793-12-1   -------T-N---ANVT-V...............ANESVSAQ---------T-E--DRKR----L--R--V---ND.......DSSNSTGNYSN-R-IN--V-T-.K------D-D---------G----    226
G.GA.LBV217        -------T-N--NAIANVT...NSNKTSNNS-LGNSTVSSIE-M-------T-EV-D-SK----L--R--VV--NG.........NENSS--D-R-IN--A-T-.K------T-D----------------    235
G.NG.92NG083       ------IT-N--NVNSANH............TEANNTVEN-E-------K-T-ERG--KKE---L-----VV--S-............GNK---R-IH--V-T-.K------N-D----------------    223
G.NG.NG1937        -------T-N--SDN-VNC.......TNANCTNNSTRSL-EANLR------T-E--DRQK----L---P-VV---D.......QSNSSNSSYS--R-IN--V-T-.K--------D----------------    234
G.NG.NG1939        R------T-N--NVNTTSC.......TNN-TTKNTTVDTMCDSMT-----VT-E-KDRQK----L--R--VVQFKD............NDNSSK-R-IN--V-A-.K--------D----------------    229
G.SE.SE6165        -------T-N---VT-KG-..KRNNSTDNSTETNNSTVDNP---------VT-E--D-KK-------R--VV--N-...............A-N-R--H--V-T-.K------T-D---------T----    230
H.BE.VI991         -------T-D-SSVNATNV........TKS-N-TDIN-G-IQ-QR-----VT-A--D-N--VH-L--RA--VQ--E.........GERNKSDNH-R-IN------.K----------------------    231
H.BE.VI997         -------T-D-SNITRNDT........NSSSTVNATSSPSAN-LT-----VT-V--D-Q-RVH-L--R--VV---E.......TSNNNNSNS-K-R-IN------.---------D----------------    230
H.CF.90CF056       -------T-N--NVR-N-S.............N-TSSMEAG--LT-----VT-VL-D-Q--VH-L--R--VV----------------NS-Q-R-IN----------------------------    221
J.SE.SE7022        -I-----T-N-S-VNSNNS...........-D-N-SASNNSPEIM------VT-E--N-RKQ---L--RQ-VV--NS.............-NK--I-IN------.K--------Q----------------    224
J.SE.SE7887        -I-----T-N-SNITSNS-............TT-NSSVSSPDIMT------T-E--N-RKQ---L--RQ-VV---S.............NNKN-I-IN------.K--------Q----------------    222
K.CD.EQTB11C       -------T-T--NVT-NRT......NANKNDTNINATVTSTD---------T-ELKD-KKRVS-L------VQ-KQ.........SEINQSESEDR-IN----TV.-----------------------    233
K.CM.MP535         E------T-N---Y-GTNS...........TNNATSTVVSPA---------T-E-KD-KK--S-L--R--VL-LNG..........EGNNSS-E-R-IN----T-.--T----T----------------    227
N.CM.YBF106        -------TML-NNSNGNSA.........GNSTTNRTEDL-DRQM-------T-E--DRKKQV-SL--VE-VV--KD............GT-NNT-R-IN---TAV.------TT-----------P----    217
N.CM.YBF30         -------TML-N-SYGEER........NN--MTTREPDIGYKQM------AT-ELTD-KKQV-SL--VE-VV--NA.............YNKT-R-IN---TAV.------T-----------P----    228
CONSENSUS O        qM-F---QMn-tn??gtt?....??????????????s?enlmkq-e--vT-VlkD-KEkkQ-L--vs-L?klnn.......????n??nstm-t-iN--STt-?k-a-------p--------a-y--    206
O.CM.ANT70         QM-F---QME--NIAGT-......................ENLM-K-E--VT-V-KD-KE-KQ-L--VS-LMELNE.......TSSTNKTNSKM-T--N--STT-.---------------------Y--    217
O.CM.CM4974        QM-F---QMN--NAQLRII...................QYK-HXM-R----TT-VFKD-KEQKQ-L--VS-LTA-HD...........KNNNK-I-T-IN--STT-XK----------------------    218
O.CM.HIV1CA9EN     QM-F---QMN--TPPQPNS...........-S--ENVTSIPENPM-Q-E---T-VLKD-KE-KQ-L--VS-LVKLN-..........TGNGND-M-T-IN--STT-.K--------------------Y--    225
O.CM.MVP5180       -M-F---QMN-V--QTNKT...............-LLNETIN-MR-----VT-VLTD-KEQKQ-L--VS-LSKVND.........SNAVNG-T-M--N--STI-.K------------------T-Y--    221
O.GA.VI686         QM-F---QMN--NIS-A-S....SNISNA-S-DISKDTSSENSV-Q-E---T-VLKD-KE-KQ-L--VS-LVKT--...............S-M-T-IN--STT-.K--------------------Y--    226
O.GQ.193HA         QM-F---QMN--NPNGTAA...................NSTENHL-Q-E--VT-VVKD-KERKQ-L--RA-L-KLE-...............S-M-T-IN--STT-.K--------------------Y--    212
O.SN.MP1300        QM-F---QMN--NVNDE-............S-VKNDTSSSENLM-K-E--VT-VLKD-KE-QQ-L--VS-LM.............KVNENND-M-T-IN--STT-.K-T-------A----------Y--    225
CPZ.US.CPZUS       -------TMT-LNPDSNSS...................AVNTTDIMR------T-EL-D-KKQV-SL--VD-LAH-N-................NT-R-IN---TA-.------T-----------P----    210
CONSENSUS 01       k------T-n--?anltn?....??????tnnv?kiignitdeVr--s-nmT-ElrD-k-?vh-l--kl-iVqi?g?????????????nnsseyr-In--t---.K-----vs-d-------t---y--    219
01_AE.CF.90CF11697 -------T-H--KA-LNDT..........Y-GTAKLNDTIGD-VR-----VT-EL-D-K-EVH-L--VP--VR-GE.......KNKNSSGNSSE-I-IH------.K------T-D-------T---Y--    231
01_AE.CF.90CF402   -------T-H--KASFTNA.............TSDR-KMEDAVR-----MT-ELQD-K-EVH-L--TS-VVQ-SSSVQNNNNSNTSGQNNSHKFR-IH------.K--------D-----------Y--    234
01_AE.CF.90CF4071  -------T-D--KADFY-T......KFNT-EKPEIEISEMQK-VS--P---T-EL-D-E-EV--L--RS-LV--ER........NSGENNGS--R-IH------.K--------D----------Y--    231
01_AE.CM.CA10      -------T-N--KANFNNT....STGNSTDEDKFKSIGNITESV-----DMT-E--D-K--VH-I------VQMGG.......QSSNNSNGYNK-R-IN--S---.K----------------------    237
01_AE.TH.93TH065   -------T-N--NANLTNV..........N-NV-NIIGNITD-VR-----MT-EL-D-K-QVH-L------VQMNK.............NSSE-R-IN------.K-----I--D-------T---Y--    223
01_AE.TH.CM240     -------T-N--NANLTNG........SSKTNV-NIIGNITD-VR--T--MT-ELTD-K--VH-L------VQ-ED............KKTSSE-R-IN------.K-----I--D-------T---Y--    228
01_AE.TH.KH03      T------T-H--NANLTDV........TE-TNVPKIIGNVTD-VR----KMT-EL-D-Q-QD--L------V--EH............KNNSSE-R-IT------.K-----I--D-------T---Y--    228
01_AE.TH.TH022     -------T-N--NVNATNV..........S-TTEAPN-VGTD-V------VT-EL-D-T-QVQ-L------VQMGG.............N-SGE-R-IN------.K--------D-------T---Y--    225
CONSENSUS 02       k------T-d-hnvnss?s.?????????????nissdmqg-i------mT-elrD-Kqkvs-L--rl-vvq-ne....?????????s?nsq-R-IN-n--a-.tq---kvsf------------f--    215
02_AG.FR.DJ263     -------T-D-HNVNSSNS.STSNSSNSS-PINRTIDSDMQE--------MT-EL-D-K--VS-L--R--VVQ-NE.............SGNSQ-R-IN-H--A-.-K---R-TL------------L--    233
02_AG.FR.DJ264     E------T-D-YNVSSNSS...........I-VKSISRDMQ---------MT-EL-D-K-QVS-L--R--VVQ-NE.............SSNSQ-R-IN----A-.K------T----------------    223
02_AG.NG.IBNG      -------T-D-HNFN-SYS............N--NLTSDMN----------T-EV-D-KK-MH-L--R--VVQ-NE.............NNGSQ-R-IN----A-.----------------------    222
02_AG.NG.NG1921    -------T-D-YNVNSSYS.......NSSSHNNRSISSEMQR--------MT-LL-D-KE-VS-L--RQ---K-NK....NENGSEETETSSNQ-R-IN----T-.------I----------------    236
02_AG.SE.SE7812    Q------T-H-Q-NLTSSG...............NISENMQ---------MT-EL-D-K--V--L--RY-VVQ-NE............TG-NIQ-R-IN----A-.----------------------    220
02_AG.SN.MP1211    -------T-E-HNYNYTR-...............NSK-N-VQ-M------MT-VLKD-KK-MA-L---I--V---K.............NA-Y-R-IN----A-.----------------------    219
03_AB.RU.KAL153-2  -------T-N-----KEVT.............-TNTSSIKMM-M-------T-DL-D--K----L-----VVQ---------------D--R-I------V.------I----------------    216
03_AB.RU.KAL68-1   -------T-N-----KNVT.............-TNTSSIKMM-M-------T-DL-D--K----L-----VVQ---------------D--R-I------V.------IF---------------    215
03_AB.RU.RU98001   -------T-N--EV-TNDT..............STNASGIEMM-------T-DL-D--K--H-L-----VVQ---------------D--R-I------V.------I----------------    214
04_cpx.CY.94CY032-3 ---S---TFT-INATTTNS..............TNGTVI-EG------D-T-E--D-KK----L--RI--V--NA...RVPINGSNRNNS-EE-M-IN--A-T-.K----------------------    230
04_cpx.GR.97PVCH   -------A-N-SNATIII-....STNNST---TGNSTVKSTA---------T-EV-D-QK--H-L--R--VV--N-...NVPINNTSNTSEYRE-R-M-----NH.K------TL--NSHTLLC-GW-CD 241
04_cpx.GR.97PVMY   -------A-N-G-ATIKST.............N-TPNVTTTN-M-------T-E--D-KK-A--L--R---V--ND....NNSTNSRRSSN-SD-M-IN--V-S-.------IK---------------    231
05_DF.BE.VI1310    -------T-N---F-ANST........ANSTTN-TTLKE-T-AVQ-----MT-EVND-KL-VH-L--R---V--SS...........DDSSNS--R-IN----T-.K-------WD-----------Y--    229
05_DF.BE.VI961     -------T-N---ATSA-T.TSKNISATP-SNPNDTLKE-Q-A-Q--T---T-EVKD-NKRVH-L--R---VS-NS.............SRKE-R-IN----T-.K------CWD-----------Y--    233
06_cpx.AU.BFP90    -------T--NATLGNKTLGNNSTNSTLGNN-TIVDDISK---------T-E--D-TK----L--RP--V--GD...........DSNNSD-R-IN--V-T-.K----------------------    236
06_cpx.ML.95ML127  -------T-I--NINITST............--TKNITVESGE--------VT-E--D-QKE---L--R--VV--ND...........GSNNN--R-IN--A-T-.K------T----------------    225
06_cpx.ML.95ML84   -------T-N--NVTDHGI............NNNNTVEG-E--------VT-E-KD-KK--R-L--T--VV--ND...........NGNNST-R-IN----T-.K----------------------    223
06_cpx.NG.NG3670a  -------T-Y--NVTKNNS...........TGNNSSTVEDRE-MT-----VT-E--D-QK----L-----VVQ--G............SNKT-R-IN--V-T-.K------T-G-----------G-    223
06_cpx.SN.97SE1078 -------T-N--NVNATKY........SNETVGKSLTVKDRE---------T-EV-DQKKT---L--RP-VMQV-G.............KNSST-R-IN--A-T-.K----------------------    227
10_CD.BFL061       -------T-N-NGINATND..................TVGDPRMT-----TT-EVKD-KKQ---L-----VVQ-GD.............SN---R-IN----A-.-------T----------------    216
10_CD.BFL071       -------T-N-S-VASNA-...........VT-GTNSTDSDTGM-------T-E--D-QK--H-L-----VVQ--G............SN---R-IN----A-.K------T----------------    221
10_CD.BFL110       -------T-N-SEWNATNH...........EGRTVNATNNVGM-------T-E--D-KKQ---L--RS-VVQL-G.............NTN---R-IN------.-------T----------------    223
CONSENSUS 11_cpx   -------t-n-tdv?snnt.......??????????v???a?mk-csfnvttel?dkkkr-yalf-kldvVp-nd.......?????nsn??d-rlin--vstv?K-acp--t-------------f-i    208
11_cpx.CM.MP818    -------T-N-AEVTSYNT.....................T-M------VT-ELIDRRKQ---L------V--ND........NNNSSNVSD-R-IN--V-T-.K------T----------------    217
11_cpx.CM.CA1      -------T-N---V-TKNR................TVEDQA-L------VT-ERNN-KKR---L-----VV--NA.........NNNSTAGD-M-IN--V-TV.K------T-------------Y--    221
11_cpx.FR.MP1298   -------T-N--N-NVTNS.......TAVGADVTNS--E-----E----RVT-EVTY-NEV---L---H-VV--NV..............N-T-R-IR--V-AV.K------T----------------    226
11_cpx.FR.MP1307   -------T---ANVTSNST............ADTTEVDRQDDML-GTLS....LPERKLE-NTHE---FKVL-KV............NIKVND-NINK--IVT-NK-GS---T---------------X    220
11_cpx.GR.GR17     -------A-N---ARDNAT.......................D----T---T-ELED-KKN-R-L--R--VV--ND...........SSSNIGQ-R-IN--V-AV.K---S--T----------------    213
CPZ.CM.CAM3        ----I--TME-RKVTFNST...SNRNKTS-MTTNSPNEKXDSTV---T--MT-EV-D-EK-V-SL--VD-LVL----------------DT-R-IN---TA-.------T-----------TP-Y--    226
CPZ.CD.CPZANT      ----M-IKMN--GYNGTPT..TPSTTTSTVTPKTTTP-VDGMKLQE-N--Q--GFKD-K--MK-I---G-LMKCQD............NNE-NC-Y-WH---TT-.--S-E-ST-------------Y--    225
CPZ.CM.CAM5        ----I--TMQ--T---IST...........--TN-TELKNI-EIM---T--MT-EV-D-KKEV-SL--VD--V--GG...........NKSNDT-R-IN---TTV.------T-------------P----    220
CPZ.GA.CPZGAB      -------T-Q-SKANFSQA...........KNLTNQTSSPPL-M------VT-EL-D-KKQV-SL--VE-VVNLG-............ENNT-RIIN---TA-.------T----------------    223
```

HIV-1/SIVcpz
proteins

HIV-1/SIVcpz proteins

```
                                                                                                                  |-        V3 neutralizing loop          -|
                    * ^^^        * ^^^    *                ^^^            ^^^           ^^^   ^*^   ^^^                                        *^^^     ^^
B.FR.HXB2           LKCNNKTFNGT..GPCTNVST.VQCTHGIRPVVSTQLLL.NGSLAE.EEVVIRSVNFT.DNAKTIIVQLNTSVEINCTRPNNNT..RKRIRIQ.RGPGRAFVTIGK.....IGNMRQAHCNISRAKWNNT    341
CONSENSUS A         ---kdke-n-t..-p-k---t.-------k--v------.-----e?kevm---E-it.n-a-?-i--ltep-k-n-t--nnn-..-ksi-i???-p-QafYat-d....i--dI-q-h--Vsrse--?t    330
A.GB.MA246          ---KD-E----..-L-K----.-------K---------.------.SK-M---E-I-.N---N-L---TSP-N-S-I------..--SV--...---Q--YAT-E....I---I---Y--VN-SE--EA    340
A.KE.K89            ---KD-K----..--------.-------K---------.------.G--M---E-I-.N---N----FAEP-K----------..-MS---...---Q--YAT-D....I---I------V---E--T-    338
A.KE.Q23            ---KDEG----..-L-K----.-------K---------.------.KNIT---E-I-.N---I-----VQP-T-K-I-------..--S---...---Q--YAT-D....I--DI------VT-SR--K-    332
A.RW.PVPI           ---KRAN-T--..---K---S.----------I------.------EKD-Q---E-I-.N-------FTKA-K-------KT-..--GV--...---Q-WYAR-N....M--DI-K-Y--V--TE--Q-    343
A.SE.UGSE8131       ----EDE----..-T-R----.-----------------.-----K.---R---E-IS.---------FTKP---I--------..--S---...---Q--YGM-D....I--DI-K----V--S---E-    352
A.UA.ukr970063      ---KDTN-T--..---K----.-------K---------.------.K--M---E-I-.--S-I-----TEP-N-T-I--G---..-TS---...---QT-YAT-D...V--DI-K-Y--V---A--S-    342
A.UG.92UG037        ----D-E----..-L-K----.-----------------.------.GK-M---E-I-.N-V-N------E--T----------..-RSV--...---QT-YAT-D....I--DI------V-GSQ--K-    342
A.UG.U455           ---KDPE---K..---R----.-------K---------.------.R-IR---E---.N---------VNP-K---S--Y.--...--N--RYSI-S-Q--YVT--....I--DI------V--RD--R-    340
A2.CD.97CDKS10      ---EDPR----..---N---S.-------M--A------.------.K--M---E-I-.N---N----F-E--P-T-I-------..--G-P-...---QV-Y-.SD....I--DI---Y-S-NKT--DAS    342
A2.CD.97CDKTB48     ---KD-E---K..-S-S---S.---A------A------.------.GK-M---E-I-.----N----F-KP-P----------..--S--F...---Q--Y-NNN....I--DI--------ITE--A-    333
A2.CY.94CY017-41    ---KDPR----..-S-K---S.-------K--A------.------GGKIM---E-I-.N---N----FTKP-L-T-I-------..--S--F...---Q--Y-.NE....I--DI-------NKTL--D-    345
CONSENSUS B         lk-nd-k-n-t..-p-k---t.-------r-v-------.---l--.eevvirse-ft.d-akt-ivq-nes-e-n-t-pnnnt.?rksihi???-pgrafyttgg....ii-di-q-h--israk-nn-    336
B.AU.MBC18          ----E-G---K..---K----.-----------------.------.-----K-D---.N---------------T-V-------..-RS-P-...-------Y-T.E....I--DI---Y---TK-N-TD-    347
B.CN.RL42           I-----K----..--------.-----------------.------.-----KFS---.---RV------E----K-I-------..--S-HL....---K-WY-T-Q....I--DI------L-ST-----    333
B.DE.D31            ---KD-R---K..-Q-K----.-----------------.------.-------D---.----------KE-----------Y-..S------...-AR---Y-K--....I--DI--------G---DS-    340
B.DE.HAN            ----D-K---K..---K----.-----------------.------.K------D---.--T-----H--E-------------..--G-H-...------VY-T-R....IV-DI-L--------R--K-    342
B.ES.89SP061        ----D-K---K..--------.-----------------.------.-------D--S.----V---H--E--------L--I-..-RS-HVGHV-----IY-T-.....I--KI---------------    348
B.GA.OYI            ----D-K----..--------.-------K--------.------.---I---S---.N---I------K-------------..-N--S-...-------H-TKQ....I--DI------L---T-EK-    345
B.GB.CAM1           ----D-K---K..--------.-----------------.------.K------E---.N---------KEP-------L----..--S-A-...----TVYATDR....I--DI------L-ST-----    341
B.GB.WB             ---KD-K----..---K---I.-----------------.------.-------E---.N-I-N------E---------G----..-RS-T-...-------HAT-R....I--DI-R----L---Q----    343
B.JP.ETR            ----------..---K----.-----------------.------.-G-----E---.--V--------ET-K---I----K-..---VTM...-----VYY-T-E....I--DI----------E--K-    347
B.JP.JH32           I---D-K----..--------.-------K--------.------.-------E---.----------KEP-V------SKT-..-R--H-...-------Y-TKQ....IA-DL-------N--R--A-    352
B.KR.WK             -Q--D-K----..---S----.-----------------.------.--I-L--E---.N----------A--------L----..--S---...---ST-YAT-A....I--DI---------E---D-    347
B.NL.3202A21        ----D-K----..---K----.-----------------.------.-------A--S.N---------E--A----------..--G-H-...---K--YAT-Q....I--DI------L--------    343
B.TH.93TH067        ------K----..--------.-----------------.------.-------S---.---RV------E----------..--S-PL...---Q-WY-T-Q....I--DI------L-ST--R--    333
B.TT.QZ4589         ----D-K----..---K----.-----------------.------.-------E---.N-----L----E--V----------..--S-H-...-------Y-.-E....I--DI------L--T---D-    344
B.TW.TWCYS          I----------..-T------.-----------------.---I--.--IL-K-E-I-.N------I---K--K-------IS.K-RSMH-...-T--V-Y-X.Q....T---I------L-K-E----    336
B.UA.UKR1216        ----D-K----..--------.-------K--------.------.-G---------.--T-------KEP----------..G-G-H-...-------YAT-N....I--DI--------IT-----    337
B.US.DH123          ---KD-K----..--------.-----------------.------.-------S---.----I------ET----------..--G-TL...-----V-Y-T-E....IV-DI-K------KV--H--    340
B.US.JRCSF          ----------K..-Q-K----.-----------------.------.-K-----D---.----------E--K------S---..--S-H-...-------Y-T-E....I--DI----------Q----    338
B.US.JRFL           ----D-----K..---K----.-----------------.------.-------D---.N---------KE----------..--S-H-...-------Y-T-E....I--DI----------D-    337
B.US.MNCG           ----D-K-S-K..-S-K----.-----------------.------.-------E---.--------H--E--Q-------Y-K.-----H-...-------Y-TKN....I--TI----------D-    345
B.US.RF             ----D-K----..---K----.-----------------.------.-------E---.--V--------A--Q----------..--S-TK...-----VIYAT-Q....I--DI-K----L---Q----    353
B.US.SF2            ----------K..--------.---------I-------.------.-------D---.N----------E--A----------..--S-Y-...-------H-T-R....I--DI-K--------Q----    343
CONSENSUS C         ---nnkt-n-t..-p-nn---.v-----iK--v-------.n--l--.eeii--se-L-.n-a-t---h-nqs-e-?-t-pn---..-kSi--...---QtfYatgd....I--DI---h---srdk-n?-    328
C.BI.BU910112       ----------..-Q-N----.-------K--I------.------.--II---E-L-.N-------H-KDP---E---------..--S---...---QILYAT-D....I--DI-------NET---K-    335
C.BR.92BR025        ----------..---N----.I-----TK---------.------.--II---K-L-.--V-----H--E----------..--S---...---Q--YAT-E....I--DI---------TA--K-    338
C.BW.96BW0402       ----------S..---N----.-------K---------.S------.--I----E-L-.N---I---H--KT-R-V----G---..-RSM--...---QT-YAT-E....I--DI-$------ERD--D-    341
C.CN.AF268277       ----D------..---H----.-------K---------.------.R-II---E-L-.N-V-----H--Q----V---------..--S---...---QT-YAT-D....I--DI--------KD---E-    338
C.DJ.DJ373A         ---RDTK----..---K----.-------K---------.---T--.-DII---E-L-.N---I------Q----------..-QS---...---QT-YAT-D....I--DI----------Q--KE-    337
C.ET.ETH2220        ---RD---T--..---H----.-------K---------.---I--.G-TI--FE-L-.N---I------E----T----S---..-ES---...---QT-YAT-D....I--DI--------EE---K-    337
C.IN.93IN101        ----------..---N----.-------K---------.------.G-II---E-L-.N-V-----H--Q----V---------..--S---...---QT-YAT-D....I--DI---------D---E-    341
C.SO.SO145A         -----E----..----RY---.-------K---------.------.G-IM---E-L-.N-------H--Q----R---YA---..--SV--...---QT-Y-.ND....I--DI-------GD---R-    340
C.UG.UG268A2        ----------..---N----.-------K--I------.---I--.--II---E-L-.N---I------K------A-------..-ES---...---QT-YAT-D....I--DI---Y-----NE--I-    345
CONSENSUS D         ---?d-k----..-p-kn--t.-------r--------l.-------.-eii---Enl-.N-a-?--vhl-e-v?-n-t--ynnt.?-QrTp-...-lgQaly-tr?.???i?gdIrq-h-n--gaa-nkt    326
D.CD.84ZR085        ----D-R----..---K---S.-----------------.------.--I----E-L-.N---I---H--Q----------YKKE..-Q-TP-...-Q-Q-LY-TRK.....TTRIIG--Y----GV-----    347
D.CD.ELI            ---RD-K----..--------.-----------------.------.---I---E-L-.N---N--AH--E--K-T-A--YQ--..-Q-TP-...-L-QSLY-TRS.....RSIIG---------Q-SK-    340
D.CD.JY1            ---KD-K----..---KK---.-----------------.------.--II---E-L-.N-V-----H--E----------D-KI.T-QSTP-...-L-Q-LY-TR.....IK-DI---Y----A-A--K-    352
D.CD.NDK            ---RD-K----..---S----.-----------------.------.--II---E-L-.N-V--------A-IV------YKY-..-Q-TS-...-LRQSLY--TG.KKKKT-YIG----K----E--KA    338
D.CI.CI13           ---TD-K----..-------S.-------K---------.------.--II---E-L-.N---I------E--T----------..-QGTH-...---Q-LY-TRV....I.-DI-K------GKA--K-    332
D.SN.SE365A2        ---KD-K----..---K----.-------K---------.------.--II---E-L-.N-------F-E--M------Y--K..-Q-TP-...-L-QVLH-TRV....K.-DI--------KER--K-    347
D.TZ.87TZ4622       ----D-K----..---K----.----------------X.------.-D-I---ESL-.N---I------E--T------YIKY..-QKTS-...-Q-Q-LY-TRL.EPKPT-KIEP-Y----G-E--K-    347
D.UG.92UG024-D      ------M----..---N----.-----------------.------.--II---E-I-.N---I---H--E--V------Y--I..-Q-TP-...-L-Q-LY-TRR....I.EDI-R------E-A-KK-    340
D.UG.WHO15-474      ---KD-K----..-T-K----.-----------------.------.--II---E-L-.N-------H--E-IQ------YH--..-Q-TH-...-T-Q-PY-TR.....I--DI-K------E-G--K-    350
F1.BE.VI850         ----D-R----..---K----.-------K---------.------.-GI----Q-IS.N-------H--E--Q----------..--G-HL...---QT-YAT-A....I--DI-K------GTQ----    331
F1.BR.93BR020-1     ----E-N-T--..-S-K----.-------K---------.------.G-I----Q-IS.--------H--E--Q----------..----SL...-----V-Y-T-E....I--DI-K----V-GTQ-R--    342
F1.BR.BZ126         ----D-K----..---K----.-------K---------.------.-DII---Q-IS.--------HF-E--Q--------Y-..--S-YF...-------H-A--....I--DI-K----V-GTQ--K-    341
F1.FI.FIN9363       ---KD-R----..---R----.-----------------.----S-.GGII---Q-LS.--------H--E--Q----------..--S---...---QS-YAT-E....I--DI-K------GEQ--K-    341
F1.FR.MP411         ---RDPR----..---K----.-----------------.------.-DII---Q-IS.--------H--E--Q----------..--S-HL...---Q--YAT-D....I--DIKK-Y-E-NGTQ-SK-    341
F2.CM.CA20          ----D-K----..---R-I--.-----------------.------.-DII---E-IS.--T------F-R----------..--S-H-...-------YAT-D....I--DI-K-Y-V-N-TL--D-    330
F2.CM.HIM277819     ----D-G----..-L-K----.-----------------.------.-NIT---E-IR.K-I-I------R----------..--S---...---QV-YAT-D....I--DI-K-Y-S-NITL--E-    344
F2.CM.MP255         ----D-R----..---EK---.-----------------.-----Q.-DII---K-I-.--T-N----F-R--I-D-R-------..--G---...---QT-FAT-E....I--DI-K-Y---N-TL--E-    340
F2.CM.MP257         ----------..-L-R----.-------K---------.------.-KMI---E-IS.--T------FKNP-K----------..-RS-H-...-------YAT-E....I--DT-K------EKQ-YD-    341
```

```
                                                                                                  |-          V3 neutralizing loop          -|
                    * ^^^         * ^^^    *                ^^^             ^^^            ^^^   ^*^   ^^^                                  *^^^     ^^
B.FR.HXB2           LKCNNKTFNGT..GPCTNVST.VQCTHGIRPVVSTQLLL.NGSLAE.EEVVIRSVNFT.DNAKTIIVQLNTSVEINCTRPNNNT..RKRIRIQ.RGPGRAFVTIGK.....IGNMRQAHCNISRAKWNNT    341
CONSENSUS G         -k-rd-efn-t..-p-kn--t.v------K---------.-------.eeI?irsE-it.d-t-v---ql-etie-n-t--nn-t??-ksihi...---QafYaT-?....i--dI-q---nVs?tk-?em    325
G.BE.DRCBL          ---VD-K----..-T-N----.-------K---------.-------.KDII-S-E-IS.----V---H--R-------------..-RSVA-...----Q--Y-T-E....V--DI-K----V-WT---E-    340
G.FI.HH8793-12-1    ---KE-E----..---Q----.-------K---------.-------.G-II-K-E-I-.--T-V------ET---T-V------..--S-HL...---Q-LYAT-D....I---I-----DV-GRN-S-M    341
G.GA.LBV217         -T-RD-K----..---K----.-------K---------.-------.--IM---E---.N---N----F-K-ID-V--------..--S-H-...---Q-LYAT-A....I--DI------V-ETD-RDM    350
G.NG.92NG083        ---RD-EY---..---K----.-------K---------.-------.-DIR---E---.--T-V------N-I----I-------..--S-P-...---Q--YAT-D....I--DI------V--I--REM    338
G.NG.NG1937         ---RE-N-T-K..-Q-K---S.I------K---------.-------.--I---TE-I-.--T-V------ETI-------S---..--S-T-...---Q--YAT-A....I--DI------V-KT--SEM    349
G.NG.NG1939         ----E-N-T--..-T-RD---.-------K---------.-------.--I----E-I-.N---V------ETI--T-----D-K..--S---...---Q--YAT-D....I--DI------VNGT---EM    344
G.SE.SE6165         ---RD-E----..---K----.-------K---------.-------.GKIKV--E---.--T-V------KT------------MK-I-MG-...---QT-YAT-A....I--DI------VTKR--KEA    347
H.BE.VI991          ----G-K----..--------.------------------.------V---I---K-I-.--T-N------EP-Q-----TG---..--S---...---Q--YAT-D....I--DI-R-Y----GKQ--E-    347
H.BE.VI997          ------------..--------.-------K---------.-------.GQ-I---K-IS.--T-N-----DSPI--T---------..--G-HF...---Q--YAT-D....I---I------V-EE---K-    345
H.CF.90CF056        ------------..-L------.------------------.-------.-QII--TK-IS.--T-N-----K-P-N-T---------..-TS-HL...-------YAT-D....I--DI---------TD--K-    336
J.SE.SE7022         ------------..---K----.-------K---------.---V--.GDII---E-IS.----N------DT---V---------..--G-HM...----QVLYAT-E....I--DI-K-Y-----KD----    339
J.SE.SE7887         ----D-N----..-S-K----.-------K---------.---I--.GDII---E-IS.----N------KT---V-Y-------..--G-HM...----QVLYAT-E....I---I-ET-----ERD-S--    337
K.CD.EQTB11C        -----N-C---..--------.-------K---------.-------.--II---EDI-.K-T-N------EA--------S---..--S-H-...-------YAT-D....I--DI--------GGQ--K-    348
K.CM.MP535          ---KD-R----..---K----.-------K---------.-------.--II---E-I-.--T-N------ET-Q----------..--S-HM...----K--Y-T-D....I--DI--------GE---M-    342
N.CM.YBF106         M---EGN-S-N..-S------.-------K--I----I-.----DT.DDI---H......HGGNLL--W-ET-S------G---..GGQVQ-...--AMT-YN-E-....IV-DV---Y--V-E.E-GSM    326
N.CM.YBF30          M---EGN-S-N..-S------.-------K--I----I-.----NT.DGI---N......-SHSNLL--W-ET-P------G---..GGQVQ-...--AMT-YN-E-....IV-DI---Y--V-KEL-EPM    339
CONSENSUS O         F---nte----..-t-h-itV?-T-----k-T---q-I-.n-TlSk.gkIRiMgKnIs?dsgkN---Tl-stinmT-e-pg?ht..vQei?i...--mawySmgl????nn?nnS-?-y-?YN?td-e?a    314
O.CM.ANT70          F---STE----..-T-R-ITV.-T-------T-----I-.--T-SK.GKIRMMAKDIL.EGG-N---T--STLNMT-E--QI.D..IQEM--...--MAWYSMGIG..GTAGNSS-A-Y-KYNATD-GKI    333
O.CM.CM4974         F---DTN----..---H-ITVX-T-------T-----I-.S-TISK.G-IR-MGK-IIXGTGNN---T--ATINMT---NGTRD..VQQ-GV...--MAWYSIAPN....DLN-S-V-N-TYNETG-IEA    335
O.CM.HIV1CA9EN      F----TN----..-K-H-ITV.-T-----K-T---H-I-.--T-S-.GKIR-MGK-I-.NS--N---T--YT-D-T-E--G-H-..VQE---...--LAWYSMGIE..KNSKNSS-L-Y--YNTTD-KRA    342
O.CM.MVP5180        F---DTD----..-L-H-I-V.-T-----K-T-----I-.--T-SR.-KIR-MGK-I-.ES--N---T---PINMT-I-EGIAE..VQD-YT....-M-WRSMTL-RSNNTSPRS-V-Y-TYNKTV-E-A    340
O.GA.VI686          F----TE----..-T-K-ITV.-T-----K-T-----I-.--T-SK.-KIR-MGK-IS.-SG-N---T--S-ISMT-E--G-H-..VQEMK-...--MAWYSMGLE...ENKT-S-R-Y-RYNATD-EKA    342
O.GQ.193HA          F----SE----..-T-N-ITV.-T-----K-T---H-I-.--T-SK.GKIR-MGK-IS.--TQN---T---TLNMT-R--AI-...VQQVGL....--LAVYSYSLG...VENITS-T-Y-EYNT-A-ER-    327
O.SN.MP1300         F----TG----..------TV.-T-------T-----I-.--TIS-.GKIR-MGK-IS.-TG-N---TI-STINMT-E--G-Q-..VQK-LT....--VAWYSMGL-...NNLT-S-A-S-KYNSSV-EEA    341
CPZ.US.CPZUS        ----E-D-K-K..-E-K----.-------K---T---II.-----T.KN-TV--K--....-DI-L--FSEG-NMT-I--G----.VGNVQL....--MT-YN-P-....IV-DV-E------KLT-EKQ    322
CONSENSUS 01        ----d-N----..---kn--S.-----g-K---------.-----e.eeII--sEnL-.n-------H--K-ve------sknm..rts?ri....--grvfyktgd....ii-dirk-Y-e-NgT--ne?    332
01_AE.CF.90CF11697  ----E-N----..---K---S.-------K---------.-------.-DII---E-L-.N-------H--K---------FK-M..-TSA--...---QV-YKT-S....IT-DI-K-Y-E-NGT---E-    346
01_AE.CF.90CF402    ----D-N----..---K---S.-------K---------.-------.--II---EDL-.--------H--K-I-------FKKV..-ISA--...-----V-H-T-N....IN-DI-K-Y-E-NKT--KE-    349
01_AE.CF.90CF4071   ----D-N----..---KK--S.-------K---------.-------.--II---E-L-.N-------H--K---------FKKM..-TSV--...-----V-YKT-S....IT-DI-K-Y-E-NGT---E-    346
01_AE.CM.CA10       ------N----..---R---S.-------K---------.-------.--II---E-L-.N-------H--K---------STTI..-TSS--...---QT-YK-EG....IS--IGA-Y-E-NGT---E-    352
01_AE.TH.93TH065    ----D-N----..---K---S.-----E-K---------.-------.--II--FE-L-.N-------H--K---------S---..-TS-TM...---QV-YRT-D....I--DI-K-Y-E-NGT---EV    338
01_AE.TH.CM240      ----D-N----..---K---S.-------K---------.-------.--II---EDL-.N-------H--K---------S---..-TS-T-...-----V-YRT-D....I---I-K-Y-E-NGT---KV    343
01_AE.TH.KH03       ----D-N----..---K---S.-------K---------.-------.G-II---E-L-.N-------H--K---------SDRM..-TPM--...-----V-YKT-N....IM-DR-T-Y-E-NGT---RV    343
01_AE.TH.TH022      ----D-N----..---K---S.-------K---------.-----K.--II---E-L-.N-------H--K--K------TI.Y..K-KTTM....-A-VYHRT-D....V--DI-K-Y-Q-NGT---KV    339
CONSENSUS 02        ---ndkk-n--..-p-K----.-------k---------.-------.eev----E-it.N---t------vkp-r---t--nn--...--sVr-...----Qt-yat-d....ii-DI-q----V-rse--k-    330
02_AG.FR.DJ263      ----D-K----..-L-K----.-------K---------.-------.-------E-I-.N---N-----VAP-R---P-------..--SV--...----QT-YAT-D....I--DI------V--SE--R-    348
02_AG.FR.DJ264      ----D-K----..---K----.-------K---------.-------.-------E-I-.N---N-----V-P-R----------..--SV--...----QT-YAT-D....I--DI-K----V--S-----    338
02_AG.NG.IBNG       ---KD-G----..---K----.-------K---------.-------.G------E-I-.N---------ANP-K----------..--GVH-...----Q--YAT-D....I--DI------V-KTE--K-    337
02_AG.NG.NG1921     ----DTE----..---K----.------------------.-------.K------E-I-.N---------VKP-R------GK--..--SVP-...----QT-LAR-G....IM-DI------V---E--E-    351
02_AG.SE.SE7812     ----DEK----..---K----.-------K---------.-------.--I----E----.N---I-----HE--K------G---..--SV--...----QT-YAT-D....I--DI------V-WQQ--K-    335
02_AG.SN.MP1211     ----E-D-S--..-S-K----.-------K---------.-------.GDI----E-IS.N----------KP-W------S---..--SV--...----QT-Y-T-A....V--DI------V--E---S-    334
03_AB.RU.KAL153-2   ----D-K----..--------.-------K---------.-------.------------.--T-------KEP-------------..--G-H-...-------YAT-D....IT-DI--------IT-----    331
03_AB.RU.KAL68-1    ----D-K----..--------..--R----K---------.-------.------------.--T-------KEP-------------..--G-H-...-------YAT-D....IT-DI--------IT-----    330
03_AB.RU.RU98001    ----D-K----..--------.-------K---------.-------.------------.--T-------KEP-------------..--G-H-...-------YAT-D....I--DI---Y-----T-----    329
04_cpx.CY.94CY032-3 ----E-N-T-L..-------S.-R-----K---------.-----T.-------K-I-..--T-N-----AKA-K------G---..--SVH-...----LTWYAT-E....I--DI--------GND--D-    345
04_cpx.GR.97PVCH    ----D-NSP-L..-S-----P.-------K---------.-----T.GG-----K---.--P-N-----DKA-K----GL----GGSE--G-...----HTWYAT-N....IV-DI--------GSD--EA    358
04_cpx.GR.97PVMY    -Q--E-R---S..---K---S.-------K---------.----ST.-G--L--K---.--T-N-----AEA-K----------..--GVH-...----KTWFAT-E....V--DI-K------EKD--T-    346
05_DF.BE.VI1310     ----E-K-S--..---K----.-------K---------.-----K.-GII---Q-IS.----N---H--E--H----------..--S-HL...----Q--YAT-D....I--DI-K----V--EQ--K-    344
05_DF.BE.VI961      ---KE-R----..---K----.-------K---------.-------.-SII---Q-IL.--T-----H--E--Q----------..--S-PL...----Q--Y-T-D....I--DI-K----V-G-Q--K-    348
06_cpx.AU.BFP90     ---RD-D----..---K----.-------K---------.-------.GNIT-KTE-I-.--T-N------QP---R----G---..--S-SF...----Q--IAT-D....I--DI------V---N-TDI    351
06_cpx.ML.95ML127   ---RD-N----..---K----.-------K---------.-------.--II-K-K-L-.--T-I------K----S-S-------..--S-H-...----Q--YAT-E....I---I-K----V--KA--SM    340
06_cpx.ML.95ML84    ---RD-N----..---K----.-------K---------.-------.D-II-K-E-H-.N---I------KT-Q-R----S---..--S-PL...----Q--YAT-D....I--DI------V--TA-KE-    338
06_cpx.NG.NG3670a   ---RD-N----..---K----.-------K---------.-------.GKMI-KYE-L-.--T------F-K----R---------..--S-SF...----Q--YAT-D....I--DIK-----V-KTD--K-    338
06_cpx.SN.97SE1078  ---RD-N----..-A-K----.-------K---------.-------.--II-KTE-L-.--S-N------K-I--K---------..--S-SF...A--Q--YAT-D....I--DI------V--TD---M    342
10_CD.BFL061        I---DRK----..---K-I--.-------K---------.-------.--II---E-L-.----N------E--T---V-------..-RS---...---QT-YAT.D....I---I---YY---GTE--K-    330
10_CD.BFL071        ------------..---D----.---I---K---------.-------.--II---E-I-.SSL-------E--T----------..--S---...----QT-YAT-D....I--DI--------GTE--K-    336
10_CD.BFL110        ----D-K----..---K----.------------------.-------.--II---E-L-.--------H--E--T----------..--S---...----QT-YAT-N....I---I---Y-E-NGTE--K-    338
CONSENSUS 11_cpx    l-cndkef---..gp-k--s-.-------K-v---q--l?-----e.?evk---e-f-.?-a-t----?nst-r------n---..-kg-hi.....-qafYaT-?....I--DIrq-h-n?-?ad-nn-    317
11_cpx.CM.MP818     ---RD-E----..---K----.-------K---------.-------.-K-K---E---.N--------F-NT-R------G---..--S-HL.....H--YAT-A....I--DI-----KV-K-E-L--    332
11_cpx.CM.CA1       ----ETN----..---K--T-.-------K---------.-------.-KIM---E---.--T-N----F-S--R----------..--G-H-...------IYAT-D....I--DI------V-T-D----    336
11_cpx.FR.MP1298    ----DEE----..E--K----.-------K-I------I.-----K.G--K---E---.----------S--M----------..--G-Q-...-----TVYAT-A....I--DI-K------G..-K--    339
11_cpx.FR.MP1307    X-GGE-EX---..---N----.-------K-----K---N------.G--K---D---.----------ET-K----------..--G-HT....-Q-LY-T-A....I--DI---Y----GKA----    336
11_cpx.GR.GR17      ---RD-E----..---K----.-------K---------.-------.G--R---E-L-.N----------ST-R----------..--G-H-...---Q--YAT-D....I--DIK-----V---D----    328
CPZ.CM.CAM3         M---MPN----GT-R-N-I--.----------T---I-.---V--.NKTIA-R......-GYXFLXXFQXX-S------G-XS..-GQ-Q-...---MT-YN-EN....IV-DT-R-Y-Q-N-TV-DER    338
CPZ.CD.CPZANT       -R-EDED-T-V..-M-K---V.-H-----S-M-A-W---.--TYQT.NTS-VMNGRKN.ESVLVRFGKEFENLT-T-I--G-R-..VRNLQ-...---MT-YNVEI....AT-DT-K-F-TVNKTL-EQA    340
CPZ.CM.CAM5         M---DQK----..-----I--.----------T--VI-.------EGPA-A--..ED.LRH-I--I-FKEGIN-I-M--G--S..-GQLQ-...---VS-YN-EN....IV-DP------V-K-Q-MKK    334
CPZ.GA.CPZGAB       ----D-D-S-K..-K------.-H-----K---T----I.-------.GNITV-VE-KS.K-TDVW----VEA-SL--H--G----..-GEVQ-...---MT-YN-EN....VV-DT-S-Y-K-NGTT--R-    338
```

HIV-1/SIVcpz proteins

HIV-1/SIVcpz proteins

```
                                                CD4 CD4          |-                V4 loop                -|       CD4 CD4 binding
                                                | |        *     *^^^   ^^^   ^^^      ^^^                 *       |  |        |       *  ^^^
B.FR.HXB2          LKQIASKLR..EQFGN....NKTIIFKQSS.GGDPEIVTHSFNCGGEFFYCNSTQLFNSTW.FNSTWSTEGSNNTEGS........DTITLPCRIKQIINM.WQKVGKAMYAPPISGQIRCSSNITGLLL    454
CONSENSUS A        lqkva?q-r..kyf??....-k--i-an-s?---l-iT------g-------tsg----t-..???ngt???n??s??n??.....d----p----q----.--Ra-Q-m-----q-v-r-e------l-    430
A.GB.MA246         -REVVKQ--..TY-.............DN--.---L--T--------------TSR------.....ND-TSML-DTKPN.......----------------.--RA-Q-I-----Q-V---E------I-    448
A.KE.K89           -QKVVT---..-Y----------K-AN--.---L--T---------------SG------.....TNMQE--S--SN........----------.-----.--RA-Q-------Q-I-K-V------I-    446
A.KE.Q23           -QEV-E----.TY--.............AN--.---L--T--------------TSG------...YVN--W.NDTDSTQESN.....----------------.--RA-Q-------P-V-K-E--------    444
A.RW.PVPI          -QKV-TQ-G..KH-.........----T-TK--.---L--T--------------T-G-------..KNVNG-..MQESSSTKLN.....-----T---------.--RA-Q-------E-I---D--------    454
A.SE.UGSE8131      --KV-IQ---..KYW......-T----TN--.---L--T--------------TSG----S-...NEND-...KVNYNTESN.....-----Q---------.--RT-Q-T-----P-V-Q-R--------    461
A.UA.ukr970063     -QK-STQ---..RY-N.........-N-.P---L-VT--------------T-D------...DEHG-V.T-.STMAN.......G--------------.--R--Q-------K-S---E--------    451
A.UG.92UG037       -H-VVEQ---..KYWN.....-N----NS--.---L--T------A-------TSG--------....VNG-...T.SST-N.......G--------------.--R--Q-------Q-V-K-E------I-    448
A.UG.U455          IQ-V-EQ-K..KK-N...........-------AS--.---I--T--------------TSG----I-.....NGSM.---.DM-PN.......G----Q---------.--R--Q-------Q-V---E--------    447
A2.CD.97CDKS10     -QKV-EQ---..KH-P.........-----N-TKP-.---L--T--------------T-S--------...KNGAT..IQENSTETN.....GIM---------VD-.--E--Q-------A-V-Y-T-----II-    452
A2.CD.97CDKTB48    --KVVEQ---..-H-P.........------NS--.---L--T--------------T-G--------....ENG-...-KQNYTESN.....-----Q---------.--R--R-------A-V-K-T-----MI-    442
A2.CY.94CY017-41   -QKV-EQ---..-K-P.....K-----TN--.------T-L----A-------T-G---G---...WNNG-W.NGPYTPNNTN.....GS-I-----------.--R--R-------A-I-K-T-----II-    457
CONSENSUS B        -kqiv?-lr??eq-g?....nkt-v-nqss.---p-i-m-s-n-g------nttq-f-s-w..???ngtw???tegsnnt???...dtit-p-r-k--i-m.--e----------s-q-r-s-n----l-    442
B.AU.MBC18         -QKV-I---..------------A--P--.--------------------------G--.....NG-WINGTWKS-YGNDT...TN-------------.--E----------R---K-T--------    459
B.CN.RL42          ----TK---..----------V-N---.-------M-------------TS--------...NDTG-WND..-T-N........S-----------V--.--E----------E--------------    441
B.DE.D31           -R--VK---..-R---------V-N---.----------------------A--------....ND-.....K-SNNTN.....G--------------.--E--------------------------    446
B.DE.HAN           -N--FR---EIR--E.......V-NR--.-------M--------------K-------...NNT--W...-DN-N........---I-----------.--E----------G-L----------I-    451
B.ES.89SP061       ----VT---..---K.......V-N---.-------M-----------------------....NGTAW...SNNTEGNEN....----------------.--E----------R------------I-    458
B.GA.OYI           -E---T----..K--R.......A-DR--.-------M-------------TS--------....ND-......-RANSTE.....V-----------V--.--E------------------K------    451
B.GB.CAM1          ----VT--K..----------N---.-------M-------------T-----T--...LFNG-WND.TEGLNNTE.....RN------------R.--E----------T-T-S----------    453
B.GB.WB            -----T---..----E.......V-N---.-----V-L-I----------T----------....L-NGAR.KE-SNNNE.....T--I----------R.--E------------Y-N--------I-    454
B.JP.ETR           -E---N----.K--E.......V-N---.-------M-N----------D-S------H...LSNG-WWN.G-GP.........EN------------V--.--E----------R--------------    455
B.JP.JH32          ----VG----.K--V.......V-NR--.-------M----------------------...LSN--WND.TEGSNNTGGN...----------------.--E----------E--------------    466
B.KR.WK            ---LVI--G..--------.SNI-V-----.-------M---I---------T---S-----..QRSDG-WNR.TGGLNETK.....EN------------R.--E------------L------------    461
B.NL.3202A21       ----V-----.K---------V-S-PL.-------M----------------------....ND-GNV.TERSNNN......EN----------------.--G----------------------------    452
B.TH.93TH067       -E--TE--K..----------V-N---.-------R-T-----------TS----G-------....NN-IEN.............S--------R--V--.--E----------R--------------    435
B.TT.QZ4589        ----VI---..-----E.....V-N---.-------M----------------------...NANGNGNG.TGGSNDN......------------V--.--E----------R-Y-N----------    456
B.TW.TWCYS         -P--VK-F-..----------V-N---.---L---M-------------A-P-------...NAT--LNA.TN-EN........EN---L----------.--R----------E-L-K--------M-    445
B.UA.UKR1216       ----VI---..---E.......V-N---.-------M-------------T-K--------....ND--WNG.TG-VNNTEG....-IV-------------.--E----------A--------------    449
B.US.DH123         --RVVE----..-K-E.......V-NK--.-------M-------------TKK--------....NG-.....EGSYNIEGN...----------------.--E--------------W----------    448
B.US.JRCSF         ----VE---..---N.......V-TH--.-------M----------------------.....ND-.....EKS-GTEGN...---I-----------.--E----------K--------------    446
B.US.JRFL          ----VI---..---E.......V-NH--.-------M----------------------....NNN-.....EGSNNTEG....N--------------.--E----------R--------------    445
B.US.MNCG          -R--V---K..----K.......V-N---.-------M-------------TSP-------...NGNN-WNN.T-GSN........NN---Q-K--------.--E----------E--------------    454
B.US.RF            ---VVT---..---D.......V-TS--.-------L-------------T----------....N--.....EGSNNTGGN...-------------V--.--E--------------K-I--------    461
B.US.SF2           -E--VK----..----------V-N---.-------M-----R-------T-----N----.RLNH-..E.G-K-N.........---I-----------.--E----------G---S----------    451
CONSENSUS C        -qrVgk--a..eh-p......--t-k-aps-.---l--t------r-------TS?---sty?????nn?t??.?s?sn???.....?n--lp---------.--g--r-m-----?-nit-?--------    426
C.BI.BU910112      -QDVSE--A..KY-P.......----N-A-P-.---L---------R-------TSK-----D........RS-.S--S.........AN--I-----------.--G--R-------K-K-T-K--------    438
C.BR.92BR025       -QEVGK--A..-H-P......-A-K-AKH-.---L--T------R-------TSS-----Y...TPN---NI.TGTEN.......SI--I-----------.--G--R-------E-ILT-R--------    447
C.BW.96BW0402      -NRVSK--A..-H-P.......----E--P--.---L--T------R-------TSR---ESY...-FNES.....HWSNDTN.....A--------------.--G--R-I-----A-N-T-I--------    449
C.CN.AF268277      -QRVGK--A..-H-P.......----E-AS--.---L--T------R-------TSI---G-Y...MPNG-.....EGNS-........SI--I-----------.--E--R-------E-N-T-K--------    444
C.DJ.DJ373A        -Q-VKG--K..-H-.........----K-AP--.---I--T--------------TSA---DSI.....TNAPIN.-SSTD.......AN---Q---------.--E--R-------Q-I-A-K--------    443
C.ET.ETH2220       -QKVKE--Q..KH-P.......----E--P--.---L--T--------------TSN-----K.......L-LF.-SSTN........LN---Q---------.--G--R-------E-I-M-R--------    442
C.IN.93IN101       -QRVGK--A..-H-H.......----K-AS--.---L--T------R-------TSG---G-YMPTYMPNGT...ESNSN........S---I-----------.--E--R-------A-N-T-T--------    451
C.SO.SO145A        -Q-VGK--A..-H-P.......----K-EP--.---L--T------R-------TSG-----Y....NPNNT...GS-SN........NN---Q---------.--E--Q-------K-N-T-T--------    446
C.UG.UG268A2       -QWVRE--K..RH-P.......----N-T-P-.---L--T------R-------TSS----SD....NNNST.................I--------------.--G--R-------K-K-T-R--------    445
CONSENSUS D        lq-vakk-g..nLln.....ktt-i-kpss?------tt-sf--g-e-----TS?---stw....??ns???nstgsn????..?.tnitlp-r----inm?w-?v-k-m-----egl-k-s--------    428
D.CD.84ZR085       -R-V-R--G..nLLN.....QTK----P--.------T--------------TSG----A-......NIS-HSTGLND........TI--I-----------.--E----------E---N----------    454
D.CD.ELI           -Q-V-R--G..TLLN.....KTI-K--P--.------T--------------TSG----------....NISAWN-I--SNNSTN....TN---Q-------K-X--A.-X-I-----ERN-L----------    452
D.CD.JY1           -Q-V-K--G..DLLN.....QT-----PP.A------T--------------TSR----------.....NSSTW--DTLNSE.......G--K----------.--G----------E-L-K-T--------    461
D.CD.NDK           -Q-V-T--G..NLLN.....KT--T--P--.------TS-ML----D-----TSR----------....NQT..-S-GFN......N.G-V---------V-L.--R----------E-L-K----------    444
D.CI.CI13          -Q---I----..NLLN.....KT----NS--.------I--------------TSK----------.....SIE--I--SND.......S--I--------V--.--R----------E-I----------    439
D.SN.SE365A2       -Q-VVR--K..DL-.........------EP--.------A------R-------TSG----------------.K-KSG.........-N-----------L.--R----------E-V-S----------    447
D.TZ.87TZ4622      IQ-VTKR-K..HL-P.......----N-QP--.------T--------------TSE---I-N...........S-GVNES......SI-R-----------.--G----------A---N----------    450
D.UG.92UG024-D     -E-V-K--K..VL-S.....RT--D--P--.------T--------------TSK---HSV....GENNTL-ISNGVN........TN-----K----V--.C-G----------Q-L----------    449
D.UG.WHO15-474     -Q-V-K--G..HLLN.....KT-----P--.------T--------------TSK----------.....VDGTA-G-RPD........NN--I-----------.--E----------E-L-K-T--------    457
F1.BE.VI850        -EYVKAE-K..SH-P.....-TA-K-N---.---L--TM-----R------DTSG---D-G......-NN................G----------V--.--G--R---TS--A-N-T-N--------    431
F1.BR.93BR020-1    -AKVKA--G..SY-P.....-A--K-NS--.---L--TR-N---M-------TDE---D-K.....FNDT.....GFN........G----------V--.--E--R----N--A-N-T-N--------    445
F1.BR.BZ126        -ERVRA--K..SH-P.....-T--K-NS--.---L--TM-----R-------TSG---D-G...............SN.........G----L---R--V--.--E--R----A--E-N-T-N--------    439
F1.FI.FIN9363      -DRVKAE-K..LH-.........-----Q-NS--.---L--TM-----R-------TSL---N-V......PNN................G----------FV--.--E--R----A--A-N-T-N--------    439
F1.FR.MP411        KT-VQE----..AL-.........----K-N---.---L--TM-----R------DTSG---ESE......KYN................G--I---K--------.--G--Q---SA--A-R-N-N-T------    439
F2.CM.CA20         -NKVVEAFQ..RKS......-L-VT-NR-.A---L--T------R-------TDE---NSL.....LNNT......KK........EN---------FV--.--R--Q-------A---Q-N-S------    431
F2.CM.HIM277819    -T-VVEEFK..KLDH-....ITN-T-SP--.------T------R-K---Y-T-D---NME.............KTNN........N--I---K-R-FV--.--R--R-------A---Q-N-K------    445
F2.CM.MP255        --NVSGEFK..KH-.......-FSVA-NS--.---V--T------R-----Y-TSG---E-E....VANNT.......N.......EN-------R-FV--.--RI-R-------E-E-Q-T--------    441
F2.CM.MP257        -IK--TEFK..D-Y.........---VG-QP-.A---L--T------R-------T-I---H-R....VNDILSN-H-REN........------------V--.--R--Q-------A-K-Q-N--------    449
```

```
                                                   CD4 CD4            |-                V4 loop                -|       CD4 CD4 binding
                    ^                   ^^^          | |       *      *^^^   ^^^   ^^^      ^^^                 *         |  |       |      *  ^^^
B.FR.HXB2           LKQIASKLR..EQFGN....NKTIIFKQSS.GGDPEIVTHSFNCGGEFFYCNSTQLFNSTW.FNSTWSTEGSNNTEGS........DTITLPCRIKQIINM.WQKVGKAMYAPPISGQIRCSSNITGLLL    454
CONSENSUS G         lqkVkaq-???kifn?...?nkn-t-ns-??---L--tt---n-r------nTSG---nsi....?????????sisN??......?t-t-p-ki---VRm.--r--Q---ap--agnit-?-n----i-    421
G.BE.DRCBL          -RDVQA--Q..-Y-I.....--S-E-NS--.---L--T-------------TSG---NSI........LK--IS-NN.......-----N-K----VR-.--R--Q-------A-N-T-R------I-    445
G.FI.HH8793-12-1    IEKVKAQ--..KI-.........----T-DS-.A---L--T------R-------TSG---NET..............ISN........G------GX---VRL.--R--Q---S---ARN-T-K--------    439
G.GA.LBV217         -QKVKAQ-Q..GIY.......--S-T-NS--.---L--TA-----R-------TSG----SI........LSN--APSNSS......KN----------VR-.--R--Q-------A-N-T-K------I-    456
G.NG.92NG083        --NVTAQ--..KIYN.......--N-T-NS-.A---L--T------R-------TSG---NNI............SNINN........E------K----VR-.-----Q----L--A-NLV-K------I-    439
G.NG.NG1937         -QKVRKQ-E..KV-NI...SEGN-T-NS-T.---L--T------R------STSG---TSL....LNTSGSD-SSTTNDS......TP-----K----VR-.--R--Q-------A-E-T-R--------    462
G.NG.NG1939         VEKVKAQ-GKL-L-K-...KST--K-GT-T.---L--A--------------TSG---NSA................SN........---E---K----VR-.--R--Q-------E-N-T-R--------    446
G.SE.SE6165         -QNV-AE-G..KI-NK...SSEN-T-NS-.A---L--T----I-R-------TSG----SL.........LRS-SSEN........G------K----VR-.--R--Q-------A-N-E-N-S----I-    453
H.BE.VI991          -HKVIT--G..SY-D..........-----LQPP.A---I--I--------------T-K---------....-N-SYTNDTYNSNSTEDIT..GN---Q-K----V--.--R--Q-------R-N-T-I------I-    462
H.BE.VI997          -Q---TQ-S..KY-V......-R-L---PH-.---L-VT------R-------TSG----S-.........T-D-INMPNDTG.....KN----------V--.--R--Q-------K-S-T-V------I-    453
H.CF.90CF056        -H-VVTQ-G..IHLN......-R--S--PN-.---M-VR------R-------TSG----S-.........-MHT-YTSNDTKGN...EN----------V--.--R--R-------Q-N-M-V------I-    446
J.SE.SE7022         -RRV-K---..-H-..........----D-TSP-.---I--T---------------TST----S-.....DENNIKDTNSTNDN......T---I--K----VR-.--RT-Q-I-----A-N-T-K--------    448
J.SE.SE7887         -RRV-T---..-H-..........----N-TSP-.---I-------------L---TSK----S-....DKNSIEAT-DTSX........A---I--K----VR-.--RT-Q-I-----A-N-T-T--------    445
K.CD.EQTB11C        VN-VKKE-G..KH-..........------QP--.----QVTR-I---R---S--DT-DTVDD-E..............E-ED........T---I---------------.-----Q-I----TA-N-T-R-----MI-    447
K.CM.MP535          -SRVKE--K..-H-K......-G--T--PPNP------L--M---A-------T-K---E-G................EN........G-----------------.-------I-----A-S-N-------MI-    441
N.CM.YBF106         WNKTKK-IK..RLL-......-N-TFKA-DKN---L-VTHLM---X-------TSR---ESE..............NKTN........K--I--------VXL.-TR-X-GI-----R-NLS-X-S----I-    427
N.CM.YBF30          WNRTREEIK..KIL-K...N-I-FRARER.NE--L-VTHLM---R-------TSK---EEL....LNE-G................EP-------R--V-L.-TR---GI-----R-VLN-T------V-    442
CONSENSUS O         -k-taeryL.eLvNnT?.??nvnii-nqss?g--?-vthlH---h------nTs?M-nYtF.???sct?ttc??s?nisne?n??.g-??ip-klk-vVRS.-irGgSgl---ip-nlT-?-n---MIl    420
O.CM.ANT70          ---T-ERYL.ELVNNT...GSINMT-NH--.---L-VTHLH---H-------TAKM--Y-F....SCNGTTCS.VSNVSQG.NN..G-..---KLR-VVRS.-IRGQSGL-----K-NLT-M-----MI-    447
O.CM.CM4974         -N-T-Q-YL.KLVNHT...G-MS---T-GT.---I--N-LH---N------DTSSM--YNF.SCVNSICSLNSTHNAIYNETNKN.A-..I--K---VVRS.-IRGESGL-----K-N-T-Q-----MI-    455
O.CM.HIV1CA9EN      ---T-ERYL.ELINYT...NSVNMT-NH-IA---D-TTYMH---H-------TSKM--Y-F....ECNNTTC.QGNNITND.N...G-DWI---L--VVRS.-MRG-SGL-----R-NLT-R-----MIM    458
O.CM.MVP5180        -Q-T-IRYL.NLVNQT...E-V----SRT-.---A-VSHLH---H-------TSGM--Y-F...INCTKS-C.QEIKGSNETNKN.G-..I--KLR-LVRS.-M-GESRI-----P-NLT-H-----MI-    457
O.GA.VI686          ---MTERYL.ELVNNTS.NKTV-M--N--.ND----VTHLH---H-------TSRM--YSF....SC-GTSC-SSQNNNATHN...--.QI---L--VVRS.-MRG-SGL-----R-NLT-I-H---MI-    460
O.GQ.193HA          -Q-T-ERWL.ELVNNT..KI-VS---N---G---A-TTHLH---H-------TS-M-AYNF.TCNKTTCSVN-SSRN-S.......H-.QI--KL--MVRS.-IQG-SGI-----P-NLT-T-D---MI-    444
O.SN.MP1300         ---T-ERYL.ELMNNT...NTVN-T-.NH-T-----VTHLH---H-------TS-M--Y-F....SCTRTNCI.RQS...NSSIN.G-..IS-----VVRS.-IQG-SGL----RP-YLT-N-S---MI-    454
CPZ.US.CPZUS        R-YTLEIIK..KEANL.....TKVELIPN.A-----V-NMML-----------TIP---M-Y..............-NTDN........T----K---R--V-Q.-MR---GIF----K-VLS-N-----MI-    423
CONSENSUS 01        -k-vt?--k..-hfn?....nk--i-qPps?---l--TM-h---R-e-----T-k---stw....?gNet???meg?n????....gt-t---k---iin-.--gv-q-------sg?iN-v-----il-    434
01_AE.CF.90CF11697  ---VTK---..-H-K..........------QP--.------TM-H---R-------T-K---------....-TNEI...ME-FKGTNS....S----------VV--.--E-----------EAVN-V-----II-    456
01_AE.CF.90CF402    ---VTR---..-HLNG....TM--S-RP--.------TM-H---R-------T-A---------.....INGTMQEVNGTNS.......GN----------V--.--E--R--------EV-N-V-----I--    459
01_AE.CF.90CF4071   -Q--IR--E..-H-N..........----Q--PPYS---L--TM-H---R-------T-R---IST......NGT......TN........G----------V-K-.--E--Q-------DEA-N-V-----I--    448
01_AE.CM.CA10       -N-VTK---..-Y-N..........----K-QP-.P------TM-----R-------T-K---------....RINKT.ETME-VK...........---M---K----V--.--R--Q---------V-N-V--------    460
01_AE.TH.93TH065    ---V-G--K..-H-N..........-------PP-.---L--TM-H---R-------T------------....-GNET...MEGSN........G------K-----R-.--GA-Q---------I-N-V-----I--    444
01_AE.TH.CM240      ---VTE--K..-H-..........------QPP-.---L--TM-H---R-------T-K---N-C....LGNET...MAGCN........-------K--------.--GA-Q---------R-N-V-----I--    448
01_AE.TH.KH03       ---VMG--K..-H-N..........----T-QPP-.---L--TM-H---R-------T-R---N-C....IGNET...KEGCN........G------K-----K-.--G--Q---------I-N-L-----I--    449
01_AE.TH.TH022      ---VTE--K..-H-..........------QPP-.---L--TM-H---R-G-----T-K---------....RGNETIESREGYN........K--I---K--------.--GA-Q-------N-T-N-I-----I--    447
CONSENSUS 02        -qqVatq-r..k?-??....ntt---anps.---l--T--------------Tse---s--??????nnns?n?tsnnkt?n????dt-t-q-r-k--vn-.--k--q-m---ip-e---e-----l--    436
02_AG.FR.DJ263      -Q-V-TQ--..KH-..........------AN--.---I--T--------------TSE------.----DN.....SSNHIESNHTE.GN---Q------V--.-----L---------E---K--------    460
02_AG.FR.DJ264      -Q-V-IQ--..KH-......-T----ANP-.---I--T--------------TSE--------.......DNSL-H--SNHTE.....-N---Q------V--.-----R-------P-E---E--------    446
02_AG.NG.IBNG       -H-VVTQ-K..TY-K......-T----ANPL.---V--T--------------TSK--------....DN-NSTA-H-GSN..............-----Q------V--.-----Q-------Q-I---D--------    446
02_AG.NG.NG1921     ---VSRQ-N..RY-V-.....-TD---AKPT.---L--T--------------TSG------.......NNNNW-T-SNNSSKSN...EI-K-Q---R--V--.--R--Q------TL-V---E--------    464
02_AG.SE.SE7812     -HDV-T---..-Y-N......-T----DEP-.---L--T--------------TSN---R--NH-G--NAP-PF-DTEDKTINGTEDK----Q------VR-.-----Q-------P-E---E--------    456
02_AG.SN.MP1211     -QKVVT--G..KH-N.....SSK---TN--.---L--T--------------TAE------AS-TNGIWASNI-ASNNKDAN...-----K-K--------------.-----Q-I-----E-V---D-----I--    451
03_AB.RU.KAL153-2   ----VI---..K---..........----V-N---.-------M-------------T-K-----------.....NGT....E-LNNTEG....-IV------------.--E----R-----A-------------    438
03_AB.RU.KAL68-1    ----VI---..K---..........----V-N---.-------M-------------T-K--D---------.....NGT....E-LDNTEG....-IV------------.--E----------A-------------    437
03_AB.RU.RU98001    -E--V----..K--R..........----V-N---.-------M-------------T-K-----------.....NNT....E-SNNTKG....-IV------------.--E----------A-------------    436
04_cpx.CY.94CY032-3 --V-SEE-K..RL-P..........----K-APP.V---L--T------K-------T-P-----H....MQNGT..-I-STDSTN.....S----Q--L--FVR-.--E--Q----S--A-S-N---D---II-    455
04_cpx.GR.97PVCH    -QKVVV---..-H-P..........-------N---.---L--T--------------TSG-----Y....MFNST.NRTNTTNGTN.....S-----------VR-.--G--Q-------A-S-N---D---II-    469
04_cpx.GR.97PVMY    -QK-VDE--..KH-P.......--N-T-AP-.A---V--T----RL---------TSD---R-Y....MVNKNET-S-NTTDE......KI-R--------V-R.--E--Q-I----LQ-SLTATQV---II-    457
05_DF.BE.VI1310     -I-V-KE-Q..SH-P..........----K-NS--.---L--TM-----R------DTSK---A-V.....FNDTVF-A-MFNNDSD....KN-I---K----VR-.--G--Q----A--A-N-A-N-T---I--    456
05_DF.BE.VI961      -E-VKEE--..AHIKDI..G----V-NS-.A---L--TS-I---R-------TSG---V-V......PNN...............E----------------.--G--Q----A--A-N-T-N-----I--    450
06_cpx.AU.BFP90     -GEVKV--E..-V-N......-TH-T--S-.A---L--T--------------TSN---TSD....LFN-S..RGNDTN........T------K----VR-.--R--Q-------A-N-T-V-----II-    458
06_cpx.ML.95ML127   -QNVTA--K..-L-N.......--N-T-NS-.A---L-VT--------------TS----NNI..............-DSNET......TNF----K----VR-.--R-AQ-------A-N-I-T--------    442
06_cpx.ML.95ML84    -QNVTE--K..QLL.......-TN-T-NP-.A---L--T------R-------TS-----SI.............P-SNET......-I-----K----VR-.--R--Q-I-----A-N-T-I--------    439
06_cpx.NG.NG3670a   -QEV-T--K..-L-NK......N-T-NS-.A---L--T--------------TSE--TN-S....-KNGT.......N........E-P----K----VR-.--R-RQ----L--A-N-T-T--------    439
06_cpx.SN.97SE1078  --NVTT--I..-V-......K-N-T-NS-.A---L--T--------------TS-----SN....LNNNS......SDNN......G--------N---R-.--R--Q-------A-N-T-T--------    446
10_CD.BFL061        -Q-V-K--G..DL-NR.....P-----P--.------T--T-----------TSE---------....-SNNT...S-GVNG.......S--------R------.--G----I-----A-L-N---------V-    437
10_CD.BFL071        -Q-V-K---..DLLN.....KTK----P--.-----VT--T-----------T-I---NSE...W-SNNT-NTSSGVND.......S--II-------V--.--------I-----A-L-S----------    447
10_CD.BFL110        -Q-V-N--G..KLLN.....KT--H-RP--.------TE-T-----------TSK----SI.............EISNTSIGNT..SI-K---K---------.--G----I-----A-L-N----------    444
CONSENSUS 11_cpx    -qqVatq-r..ekf?.....nkt-i-???-?---l-iTt-----g-------nTSg--nst-...??wn?t?n?tmelNdt????..d?-t---r----vrm.--R--Q-m-----q-?ir-?-n---l--    417
11_cpx.CM.MP818     -Q-V-TQ--..GK-..........------DNP-P---I--TS-----R-------TSG---N--.....LFNSTWNSSQ-LNGTE.....PN----------VR-.--R--Q-------Q-E---D--------    444
11_cpx.CM.CA1       -QKV-KQ--..-IY..........------TKH-.---V--T--------------TSR--QG--...DKKNDTWN.YPGSNDN......------------VH-.--R--Q-------Q-K-E-K--------    446
11_cpx.FR.MP1298    -E-V-MQ--..K--N.....KTN---NST-.---I--T-------------DTSG----S-....S-NET.N-TMTLNGTIA....-N----------VR-.--R--Q-------P-K-S-N-S------    451
11_cpx.FR.MP1307    ---V-TQ-W..RK-..........----V-TN--.---L--T--------------TSG---------...I-NNS....M-PNDTKLTD..SN-V---K----VRL.--R--Q-------Q-I---E-----I--    446
11_cpx.GR.GR17      -Q-V-EQ-H..NN-..........----V-NEH-.---L-VT--------------TSG---------....YANDN...TSTQNDMQSN...----------------.--R--Q-V-----Q-EL--D--------    438
CPZ.CM.CAM3         -NETGQA--..-L-T-...LTQVNFTVSP..-----VTNMM-----------T-T---Y--..............KNNNITKGDN...-TFF----R--V-S.-MR---GI-----R-V-S-T--V--II-    444
CPZ.CD.CPZANT       RNKTEHV-A..-HWKKVDNKTNAKTIWTF.QD----VKV-W---Q------DI-PW--A-Y......TGNLI....TN........GALIAH------V-H.-GI-S-GI-LA-RR-NVS-T-S---IM-    448
CPZ.CM.CAM5         -NET-LAIA..KYDKS...IA-VTFTPNP..------TNMM-----------T-P---KFN..YTCDNQNNCTRNNTNDTNSDN..S-L-IS---R--V-S.-MR---GI-----R-N-T-I--------    452
CPZ.GA.CPZGAB       VEEVKKA-A..TSSNR...TAAN-TLNRA-.-----VTH-M-----------TS-I-TDNI...............TN........GI-I-----R--VSS.-MR--RGI-----R-N-T-N--------    438
```

HIV-1/SIVcpz proteins

```
                  CD4   |-     V5      -|                                        gp120 \/ gp41
                    |         ^^^                                                |-  fusion peptide  -|
B.FR.HXB2         TRDGGNS......NNESEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRRVV.....QREKRAV.G.IGAL.FLG.FLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQLT    569
CONSENSUS A       -----vn...???nssn-T---g--d-----r---------ki--l-v-p-k-k-r--.....e-----v.g.lg-v.-l-.---a-------a-i-------q---------s------e----l-k--    545
A.GB.MA246        -----GN.....SSVS--T----------------------------------------.....E------.-.L--V.-I-.-------------I----------------S-------------K--    564
A.KE.K89          -----DN......-S---T----------------------------------------.....E------.-.---V.---.-------------I----------------S-------------K--    561
A.KE.Q23          -----KD......--VN-T---------------------E---------R--------.....E------.-.---V.---.------------T-I-----------------------------K--    559
A.RW.PVPI         -----VG...DNST-GT-T------N------------------------T---K-----.....E-----I.-.---V.---.-------------I----------------S-------------K--    572
A.SE.UGSE8131     -----VT.....N-TNN-T----------------------L----------Q---.....K------.-.LA-V.-F-.-------------I---------------------------M-R--    577
A.UA.ukr970063    -----SG.....T-SSN-T---I--N------------------I-L-R-R-------.....E-----I.-.L--A.---.------------------------------S------K------K--    567
A.UG.92UG037      -----VN......SSD--T----------------------------------R---.....E------.T.L--V.-I-.---T---------I-------K---------S-------------K--    563
A.UG.U455         -----TN......-TKN-T------------K------------------R-------.....E------.-.L--I.---.-------------I----------------S-------------K--    562
A2.CD.97CDKS10    -----S-......-TN-------------------------------SR--------.....E------.-.---V.---.-------------I----------------S---K---------K--    567
A2.CD.97CDKTB48   -----KN.......SIN-T---------------------------I---E-R-------.....-------.-.L--V.---.-------------I----------T------S---K------QM-R--    556
A2.CY.94CY017-41  ------N.......GTN-T----------------------L--------R--------.....E------.-.L--V.---.-------------L----------------S---Q--------K--    571
CONSENSUS B       -----nn.?????tn?t-i--pg--d-r-n--se--k---vki-p--v---k-k-------.....------v.g?i--m.f--.----------aasmt----a-q--s------n------ea--hl-q--    556
B.AU.MBC18        -------...NTTD-ST-T-----------------R----QV----I---R-------.....-------.-.---M.---.------------V----------------------------M----    577
B.CN.RL42         ------N.....ESKPT-T---------------------------------R-------.....-------.-T---M.---.------------I----------------R----------------    558
B.DE.D31          -----KN.....KD--T-T------N--------------------------------.....-------.-LL--V.---.-----------R--A--------------------------------    563
B.DE.HAN          ------D.....NSSTT--------N------N---------------------------.....-------.-ML--M.---.-----------R-L--------------------------------    568
B.ES.89SP061      -----IN.....QT-TT----------K----------------------------------.....-------.-I---M.L--.-----------------------------------------H--    575
B.GA.OYI          -----KN.....TT-GI-----A-----------------------------R-------.....-------.-ML--M.---.-----------R----------------------------------    568
B.GB.CAM1         -----RG......E--T-T-----------------------------------------.....-------.-A----.---.-----------VAL-----T----------------------------    569
B.GB.WB           -----TT....NNT-AT--------N-K--------R-----------------------.....-------.-TL--M.---.------------L-----------------------------------    572
B.JP.ETR          ------T.QNNNT-SSI-------------------------------------------.....-------.-.---V.---.------------AV---------P-------------D---------    575
B.JP.JH32         -----DN......Q--T-T------N-----------------L--------------.....-------.-.---V.---.------------.---------L----------------G--------    580
B.KR.WK           ------E...NNGT-GT-T---E--N-K-----K-------R-----I---R-R-------.....-------.T.F---.---.-----------CT---------L-------------------------    579
B.NL.3202A21      -----KD....ENKTGT----------K-----------------------R-------.....-------.-.----.---.------------I-----------------------------------    569
B.TH.93TH067      ------N.....GS-TT-T------N----------R-----------------------.....-------.-T---M.---.------------I----------------R----------------    552
B.TT.QZ4589       -----IG......T-GT-----------------------E-----------------.....-------.-T---M.---.------------L-------L-------------------------    572
B.TW.TWCYS        -----TN......DS-V-V--------K-I--N---------V----L---R-R-------.....-------.-.----.---.------------L-----------------------------M----    560
B.UA.UKR1216      ------Q......S-VT-----------------------------------R-------.....-------.-.---V.---.------------I-----------------------------------    564
B.US.DH123        -----KN.......SST-----------------------RV----I----------.....-------.-.---V.---.------------I-----------------------------M----    562
B.US.JRCSF        -----KN......ES-I-----------------------------------------.....-------.-.----.---.-----------R-----------------------------M----    561
B.US.JRFL         -----IN......E-GT-----------------------------------------.....-------.-.---V.---.----------------------L--------------------RM----    560
B.US.MNCG         -----KD....TDT-DT-----------------------T-----------------.....------A..----.---.------------V-------L---------------------M----    570
B.US.RF           -----ED.....TT-TT----L---N---------------R---------R------.....-------.-T---M.---.-----------G-I-------H-------------------------    578
B.US.SF2          -----TN.....VT-DT-V--------------------I------I----------.....--------.-IV--M.---.-----------V-L-----------------------------------    568
CONSENSUS C       tr--gtt???????n?t-i-r-g----r-------------EiK---?---kak-------.....E-----v.g.ig-v.--g.-------------I---v-a-----------s---r---aq--m----    538
C.BI.BU910112     -----TT......-DST-A---A----K-------------E-K----------------.....E------.-.---V.--R.-------------I----------------S-------T---M----    553
C.BR.92BR025      -----TG......MHDT-----E------------------E-K---I----------.....E------.-.---V.---.------------I-----V----------S----------M----    562
C.BW.96BW0402     -----TT.....R-------G--------------------E-K---I---E-R-------.....E------.-.---V.---.------------I----------------S----------M----    565
C.CN.AF268277     V----KE......T-DT-T----------------------E-K---I---A-------.....E------.-.---V.---.------------I-----S----------S----------------    559
C.DJ.DJ373A       -----TN.......STD-T----------------------E-K---L----------.....E-----A...LA-V.---.------------I----------------S---K---VH--M----    556
C.ET.ETH2220      ----AKE.....PHSTK-----E------------------E-K--------P------.....E-----A...L---.---.------------I----------------S---K-------M----    557
C.IN.93IN101      VH---IK..ENDTE-KT----------------------E-K-------A-------.....E------.-.---V.---.------------I---A------------S----------------    570
C.SO.SO145A       L-----DPKRTEEQ-NT-R----------------------E-K---I---G-------.....E------.-.---V.---.------------I----------------S----------M----    567
C.UG.UG268A2      -----ET....SET-ST-T----------------------EVK----------------.....E------.-.---V.---.------------I----------------------------M----    562
CONSENSUS D       t----an....??nssneTf------mr----s-------v?i--l-v---r-k-------.....e-----I.-.L-am.---.-----------a-?t--v-a-qll----q--n-----------l----    541
D.CD.84ZR085      -----AN......-TQNDT---------------------------------------.....E-----I.-.L--M.---.----------------------V----------------------    569
D.CD.ELI          -----IN......-STN-T----------------------Q---------R------.....E-----I.-.L--M.---.-----------R-V---------M----------------------    567
D.CD.JY1          -----VN......-STN-T---------K----N--------R-----I---R-------.....E-----I.-.L--V.---.-----------V-VA--G----------------------M----    576
D.CD.NDK          -----AN......-SSH-TI-------------------------I-------R------.....E-----I.-.L--V.---.------------V---------M----H-----------------    559
D.CI.CI13         L----VN......-HSN-T-------I--------------RL--------R-------.....E-----I.-.L--M.---.-----------V----------L----------------------    554
D.SN.SE365A2      L----IH....NTS--T-T----------------------RL--------T------.....K-----I.-.L-VM.---.------------A-----------------------------    564
D.TZ.87TZ4622     -----AN.......SSD-T----------------------RL--I-L----------.....E-----I.-.L--V.---.------------L-----T-K---------D----------------    564
D.UG.92UG024-D    ------T.......SQN-T---------------------I------L--R-------.....A-----I.-.L--V.---.------------L---------M----------------------    563
D.UG.WHO15-474    I----AN.....NGSNN-T---------------------------I---T------.....E-----I.-.L--M.---.------------L-----------------------------------    573
F1.BE.VI850       ------E.......SNI-T---E--N-K-------------E-------------Q--.....------A..L---.---.---DSREH-----I-----------------------------------    545
F1.BR.93BR020-1   -----LN.......STN-T------N-K-------------E-------------Q--.....K--R---.-.L---.---.------------I----------------S----------------    559
F1.BR.BZ126       -----QN......S-QT-T------N-K-------------E---------R---Q--.....K------V-.----.---.--S----------I-----------E----------------------    555
F1.FI.FIN9363     -----Q-.....N-SD--T--------K-------------E---------RP--P--.....R--R---.A.---V.---.--S----------L----------------------Q-------M----    555
F1.FR.MP411       -----Q-....NDT-RT-T---E--N-K----N--------E-----------R------.....---R---.-.----.--R.-------NI----I-----------------------------    556
F2.CM.CA20        ----NKN........G--TL---R----------------------------R---Q--.....K--R---.-.---V.---.------------I----------------S---K------------    544
F2.CM.HIM277819   -----EN......T-G--TL--------------------R-------------Q--.....--G----.-.---V.LF-.------------I------------------K------------    560
F2.CM.MP255       ----SKD......IDGK--L--I------------------R---V----------------.....--A----.-.M--V.LF-.------------AI---A------------S---K------------    556
F2.CM.MP257       -I---EG.......----TL---------------------------------Q--.....-------.-.M--M.---.------------I-------N---------S---K------------    563
```

```
                    CD4   |-     V5      -|                                    gp120 \/ gp41
                     |          ^^^                                          |-  fusion peptide  -|
B.FR.HXB2           TRDGGNS......NNESEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRRVV.....QREKRAV.G.IGAL.FLG.FLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQLT    569
CONSENSUS G         -----nn..?????st?-t---g--d-rd---s------?vk-Kp--v---r-R----.....e-----v.g.l--v.fl-.---a---------it--v-V-----------s-------a--------    533
G.BE.DRCBL          -----DN......-ST----------KN----------T---KS--I---R-R----.....E------.-.V--I.---.---T---------I-----V-----------S----------------    560
G.FI.HH8793-12-1    ------A....NNAS-T-T---A--N-K----N----------K---------R----.....G------.-.V--V.---.-------------I-----V-----------S----------------    556
G.GA.LBV217         ----------.....T-GN-T---A-----------------I--K---------R----.....G-----I...T--V.---.-------------LA----V-----------------------------    570
G.NG.92NG083        ------N.....NDSTE-T--------------------T---KS------R-R----.....E------.-.L--V.---.-------------I---A-V-----------S----------------    555
G.NG.NG1937         -----TN.......KTN-T---A------------------R-K---I-----R----.....E------.-.---V.LF-.-------------I-----V-----------S-------V--------    576
G.NG.NG1939         ------I........TE-T--------------------I---K-------R-R----.....E------.-.L--V.L--.---T---------I-----V-----------S----------------    559
G.SE.SE6165         ------N..NNTNTS-----------------------T---KS------R-R----.....E------.-.L--V.---.-------------I-----V-----------G----------------    572
H.BE.VI991          -F-RN-T........NNVT--------------------------------E-R----.....E------.-.M--F.---.-------------I----------------S------Q----M----    575
H.BE.VI997          -Y-EDKG......--DNVT--------------------------------E-R----.....E------.-.M--F.---.-------------I----------------S------Q----M----    568
H.CF.90CF056        -I-E--A......SA-NYT---------------------------I----TR----.....E------.-.M--S.---.-------------I----------------S------Q-R--M----    561
J.SE.SE7022         ------R...NGSE-GT-T---T--N-K-------------EL-----------------.....E------.-.---V.---.---T---------I-----V-----------S---K--X------K--    566
J.SE.SE7887         ------R..GNGSE-GT-T---T--N-K-------------E------------------.....E------.-.---V.---.---T---------I-----V-----------S---K---------K--    564
K.CD.EQTB11C        ------D....NNTRTE-T---------------------Q-----I---R-R----.....-------.-.----.---.-------------I---------------------------QM----    564
K.CM.MP535          ------N.......THN-T---------------------Q-----I---R-R----.....-------.-.L--V.-F-.-------------I----------------S----------------    555
N.CM.YBF106         EHS-E.........-GNKTVY-S--N-VNL--Q-------S---I----G-----T-.....S-----AF-.L---.---.-------------I-------T--------------------------S    540
N.CM.YBF30          EYS--PD.......TKET-VY-S--N-VNL--Q-------S---I----G-----T-.....S-----AF-.L---.---.-------------I-------T----------I-------------S    557
CONSENSUS O         qldnpwN....??sn???tfr-i----k-I--te-fn----rvK-FS----kia-PvIgtgt?h------.-.l-Ml?---.V-S---------ATt-a-qtHtlmK------Dn-----Qa--q--R-S    536
O.CM.ANT70          QM-NTWN.....SS-NNVT---I----K-I--T--FN----RVK-FS----RIA-P-ISTRT.H------.-.L-M-.----.V-S---------AT--A--THT--K------D------Q--Q--R-S    567
O.CM.CM4974         ELHNPYN.....KS-PNTT-G-T------I--A--CN----QVK-FS-----VS-PIISLNNPH------.-.L-M-.----.V-S---------AT--A-RTHS-IK------DS-----QT-----R-S    576
O.CM.HIV1CA9EN      QL-RPWN......SSRDNTL--A------I--T--F------VK-FS-----IA-PTIGTGTPH------.-.L-M-.----.V-S---------AT--A--THT-MK------D------Q---E--R-S    578
O.CM.MVP5180        QL-QPWN......STGENTL--V----K-I--TK--N----Q-K-FS-----MS-PIINIHTPH------.-.L-M-.----.V-S---------ATA---RTHSV-K------D------Q------R-S    577
O.GA.VI686          QM-MPWN.......KSNTT---V----K-I--T--F-----RVK-FS-----IA-P-IGTGT.H------.-.L-M-.----.V-S---------ATA-----HT-IK------D------Q---Q--R-S    578
O.GQ.193HA          QL-NQWN.....TSTNEAK---I------I--K--FN----RVK-FS----HIS-P-IGTRT.H------.-.--M-.----.V-S---------ATA----THT-MK------D------Q---Q--R-S    564
O.SN.MP1300         QL-KTWN....RT--SEST---I----K-I--T--F-------K-FS-----IA-P-IGTGT.R------.-.L-M.L----.V-S---------AT--A--THT-MK------D------Q---Q--R-S    575
CPZ.US.CPZUS        DISISAV....NNDSRNITVM-T----TAL-KN--H-----S---I----G----HT------.K-----AF-.L---.---.-------------VV-----------------------------S    541
CONSENSUS 01        t----?n...???nttn-T---g--NiK----s-------vq-----i--tr-k----.....E------.-.---m.IF-.-------------i---------l-------S----a-----------    548
01_AE.CF.90CF11697  ------A........TN-T------N-K-------------Q-----I---R-R----.....E------.-.---M.IF-.-------------L----------------S----------------    569
01_AE.CF.90CF402    -----IN...QNQT-KN-T------NIK-------------Q-----I-----R----.....E------.-.---M.IF-.-------------I----------------S----------------    577
01_AE.CF.90CF4071   V----KI.....E-ETI-T------N-K-------------Q---------G----------.....E------.-.---M.IF-.-------------I----------------S----T-----------    564
01_AE.CM.CA10       -----GN......STDT-T------N-K------------IQ-----I-----R----.....E------.-.----.IF-.-------------L---------M-------S----------------    575
01_AE.TH.93TH065    -----SA.....N-TNN-T---E--NIK-------------Q-----I---R----------.....E------.-.---M.IF-.-------------I----------------S----------------    560
01_AE.TH.CM240      -----VN......-TDN-T------NIK-------------Q-----I---R----------.....E------.-.---M.IF-.-------------I----------------S----------------    563
01_AE.TH.KH03       -----PN......-MTN-T------NIK----N--------R-----I---R----------.....E------.-.---M.IF-.-------------I----------------S----------------    564
01_AE.TH.TH022      -----DN.....N-TIN-T------NIK-------------Q-----I--S----------.....E------.-.---M.IF-.-------------I----------------S----------------    563
CONSENSUS 02        -----?n......nstn-T-------------------------p--v--tha-----.....E-----v?-.L--v.-l-.-----------a-I----------------s---?---------k--    549
02_AG.FR.DJ263      -----SN......-STN-T--------------------------------RP----------.....E------.-.L--V.---.-------------I----------------S-------------K--    575
02_AG.FR.DJ264      -----SN......-STN-T----------------------------I---H----------.....E------.-.L--V.-F-.-------------I----------------S-------------K--    561
02_AG.NG.IBNG       ------N......STN-T--------------------------------R----------.....E------.-.L--V.---.-----------R-I----------------S---K---------K--    560
02_AG.NG.NG1921     -----KH......-STN-T-------------------------A------Y----------.....E------.-.L--V.---.-------------I----------------S-------------K--    579
02_AG.SE.SE7812     ------D......--NT-T------------------------------SH----------.....E-----LV-.L--F.-F-.-------------I----------------S---K---------R--    572
02_AG.SN.MP1211     -----DN......T-GD-T------------------------------SH----------.....E------.-.L--V.---.-------------I---------------------K---------K--    566
03_AB.RU.KAL153-2   ------Q......S-VT----T--------------------------R----------.....-------.-.---V.---.-------------I-----------------------------    553
03_AB.RU.KAL68-1    -----YQ......S-VT---------------------------------R----------.....-------.-.---V.---.-------------I-----------------------------    552
03_AB.RU.RU98001    ------Q......--VT---------------------------------------------.....-------.-.---V.---.-------------I-----------------------------    551
04_cpx.CY.94CY032-3 -----TN........NT----------------------------I----N--R----------.....----$--.-.---M.---.-----------------------------S-------------R--    567
04_cpx.GR.97PVCH    -----I-..NNNET-DN-TL-------------V-----------V-I---G-R----------.....-K-----.-.L---.---.-------------L----------------S-------------K--    588
04_cpx.GR.97PVMY    ------R.....SDTGN-T---------------------Q---V-----R-R----------.....-------.-.---V.-P-L-------------I----------F------S----------Q--R--    574
05_DF.BE.VI1310     A-----G...NDSS-DT-----------------------E-Q---I---R---Q--------.....K-----M.-.---M.---.--A----------IA--------------------Q------------    574
05_DF.BE.VI961      -----VN.....IT-DT-T---------------------E----------R---Q--------.....-------.-.V---.LI-.-----------------------------Q------------    566
06_cpx.AU.BFP90     ----N-E........-V--T-----------------------K---I---W-R----------.....G------.-.L--M.---.---T---------I-----V-----------S----------------    572
06_cpx.ML.95ML127   -----RN......DSN--T-------------N--D-------K---I---R-R----------.....GK-----V-.L--V.---.---T---------I-----V-----------S----------------    558
06_cpx.ML.95ML84    ----NTN.......TT---I------------N----------K---I---E-R----------.....G------.-.---F.---.---T---------I-----V-----------S----------------    553
06_cpx.NG.NG3670a   ----N-D.......SHT-T-----------------------K---I---W-R----------.....G------.-.L--V.---.---T---------I-----V--------H--S----------------    553
06_cpx.SN.97SE1078  ----H-D.........T--------N-K---------------K---I-----R----------.....G------V-.L--V.---.---T---------I-----V--------H--S----------------    559
10_CD.BFL061        -----AN.....N-SQN-T--------------------------I-L---T----------.....E-----I.-.L--V.---.-------------L--------V----------------------    553
10_CD.BFL071        ---RDGG...ANNTSQN-T--------------X-------------L----------.....E-----I.-.L--M.---.-------------L---------------------K------------    565
10_CD.BFL110        -----AN....NS-QSG-------------------------R-----L----------.....E-----I.-.L--V.---.-------------L-----------------------------M----    561
CONSENSUS 11_cpx    -----nn....???ttn-T---T--d-rd---s---k----E-k-------r----------.....e-----v?-.i--v.L?-.---a---------I-----a-----------sn--k--ea--?l-K--    529
11_cpx.CM.MP818     -----LN.......STN-T---T-----------------E-K-------R----------.....E------.-.---V.LF-.-------------I-----V-----------S---K---------K--    558
11_cpx.CM.CA1       ------K.......TQN-T---T-----------------E-R----------------.....-------.-.---V.LF-.-------------I----------------S---K------Q--K--    560
11_cpx.FR.MP1298    -----PN......-TT--T---T--N----------R----E-K-------R----------.....E------V-.---V.L--.-------------I-----T-----------S-------------K--    567
11_cpx.FR.MP1307    ------M.....NDTD--T---T--N------N--------E-K-------R----------.....E-----AV-.L---.LF-.-------------I---------------------K------QM-K--    563
11_cpx.GR.GR17      -----EG....NDTIGK-T---T-----N-----------E-K----------------.....E------.-.---V.L--.-------------I---------------------K---V-----K--    555
CPZ.CM.CAM3         ETGH-IN.....NSITNITLY-T--N-V-L--L--H-----S---I----S----HT------.S-----AF-.L---.---.-------------VV-----------------------------S    561
CPZ.CD.CPZANT       EGQIY-E.........TVKVS-.AARVA-Q--A--SR-Q--E-X--S----TX--PEIKQH..S-Q--GI.-.--LF.----.L-S----------IA--A-T-N-XH------A---Q---T-------S    562
CPZ.CM.CAM5         EAG-KWE.......-NTITVY-S--E---L--L--NR----S---I---------HT------.A-----AF-.L---.---.-------------VV------H-----------------------S    567
CPZ.GA.CPZGAB       -S-TPVT.....N-SGNLT---T--N-K-I----------R----S------R-HT-ARQK.D-Q---AF-.L---.---.------------AV--------------------K----------S    559
```

HIV-1/SIVcpz proteins

HIV-1/SIVcpz proteins

```
                                   immunodominant region
                                   |-        *     * -|   ^^^   ^        ^^       ^^^          ^^^                                  ^^^
B.FR.HXB2          VWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWS.N........KSLEQIW..NHTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIKLFIMI    688
CONSENSUS A        ----------Vl-v----r---l---------l--t-n----Ss--.---------sqse--..?NM--lQ-dk-isn--di-Yq-i-e--nq--k--qd--a-----n--n--d-sn-----ki-im-    663
A.GB.MA246         ----------V--L----R---I--------------N----S---.---------QSE--..ENM--LQ--K--S---HI-YT-L----I-------D--A-----N-----D---------I----    683
A.KE.K89           ----------V-------R-------------I----N----S---.---------QSE--..ENM--LQ--K--S---DI-YN-L------------D--A-----N-----D-S-------I----    680
A.KE.Q23           ----------V-------R------------------N----S---.---------DE--..NM--LQ--K------Q--YR-------------K---------N--S--D-S-------I--I-    678
A.RW.PVPI          ----------V-----------------------P-T----S---.---------T-SE--..DNM--LQ--K--S---D--YQ---D----------D--A----------D-S------RI----    691
A.SE.UGSE8131      ----------VM--------------------I---------ST--.---------YT---..DNM--LQ--K--S---DI-YQ------Y------K---------N-----D-S-------I----    696
A.UA.ukr970063     ----------V-----------------------N----SN--.---------QSE--..DNM---Q----VIY--DI-YD---K---P------D--A--------S--D-S-------I--I-    686
A.UG.92UG037       ----------V-------R---------------P-N----S---.---------DE--..ENM--LQ--K--S---IK-YE------I---R--KD--------------D-SK------I----    682
A.UG.U455          ----------V-------Q------------------T----S---.---------Q-D--..NM--LQ-EK--SS--GI-YQ-------------LD--A-----N-$-----S------R--VI-    680
A2.CD.97CDKS10     ----------V--L----Q----------------T----S---.---------TY-E--..NM--LQ-----D---NI-YN-L------------D--A---------S-------RI----    686
A2.CD.97CDKTB48    ----------V--L----Q----------------A-D-R--S---.---------TQ-----.KNM--LQ--K--ST--DI-YM-L------------D--A-----N-----D--R------I----    675
A2.CY.94CY017-41   ----------V-----------------------A-T----T---.---------QDE--..DNM--LQ--K--S---NI-YR-L------------D--A-----D--S----SH-----RI----    690
CONSENSUS B        ----------v-------k--q-l-i-----kl--t-a----a---.n........slde--.nNm---e-er-id--tsl-Yt-i-e--nq-----qe--e-dk-as--n-fd-tn-----ki--m-    675
B.AU.MBC18         ----------V-------R------L-----------T----T---.---------T--K--..DNM---Q-ET------NI-YT-L-----------KD-----Q--N-----T-S-------I----    696
B.CN.RL42          ----------V-------------------------------.---------HE--..NM-----E---D---RE-YT-------------L---------------D--K------I----    677
B.DE.D31           ----------V-------R-----------------------.---------MDM--..NM-----E---D------YT------------------N--EN--S--D-S-------I----    682
B.DE.HAN           ----------V-------R------------------T--------.---------T-D---..NM-----E---D------YT---Q----------------------S-YD-S-------I----    687
B.ES.89SP061       ----------V-------R------F----------------T---.---------S---..DNM---Q-E---D------YN--------------------N--------S--------I----    694
B.GA.OYI           ----------V-----------------------T--------.---------NE--..DNM---Q-E---D---H--YT----------------------G--S--S--------RI--I-    687
B.GB.CAM1          ----------V-------------------------------.---------DK--..NM-----E---D---N--YT-------------KD-----T--------D---------I----    688
B.GB.WB            ----------V-------Q--R---------R-----------.---------NE--..DNM-----E---E---G--YN-L-T------------------PN-------------------    691
B.JP.ETR           ----------V------------M--------F---------T---.---------FNE--..DNM-----E-------N--YN--------------D--A----D------S--K------I----    694
B.JP.JH32          ----------------------------------------------.---------E--..DNM-----E---D------YT------------------G----------T--------RI----    699
B.KR.WK            ----------V-------R------------------N----T---.---------ENE--..DNM-------------N--YD-L-K---------------------$------K------I----    697
B.NL.3202A21       ----------V-------------------------------.---------DK--..NM--------VS------YT-----------------------------D--K------I----    688
B.TH.93TH067       ----------V-------------------------------.---------NE-$..NI-----E---D---RE-YT-------L-----L-----N---------D--K------I----    670
B.TT.QZ4589        ----------V-------R------------------T----T---.S........N-SD--..DNM-------------KF-YT-L-------------------------S--D--Q------I----    691
B.TW.TWCYS         ----------V-------R---------------N-N----I---.---------TYDE--..NM---Q-E---D---DH-YN-L-K----------D-------------D--------RI----    679
B.UA.UKR1216       ----------V-----------------------------T---.---------DK--..DNM-----E------ID--YN-----------------A-----------D-SK------I----    683
B.US.DH123         ----------V-------Q------------------T----T---.---------DT--..GNM---Q-EK------G--YN-----------------A-------------S-------I----    681
B.US.JRCSF         ----------V------------M------------------T---.---------DS--..NM-----EK--E---NT-YT------I---------------------G--K------I----    680
B.US.JRFL          ----------V-------G-----------------------.---------DR--..NM-----E---D----E-YT---------------------------D--K------I----    679
B.US.MNCG          ----------V--------------F-----------T--------.---------DD--..NM---Q-E---D------Y--L-K--T---------------------D---------I----    689
B.US.RF            ----------V-------R------------------T--------.---------NM--..NM---Q-E---D---GI-YN-L---------------------N-----D--Q-----RI----    697
B.US.SF2           ----------V-------R-----------------------.---------D--..DNM---Q-E---D---NT-YT-L---------------------------S---------I----    687
CONSENSUS C        -------qt-V--I--y-k------------------a-p--s---?........ksqee--..nnm---q--r-is--tnt-Yr-L-d--n--ek--kD--A--k-qn--s--d--n-----kI--m-    657
C.BI.BU910112      --------T-V--I--------------------------D---N---------T-GD--..-S----Q-----S---NT-YR-L-D------Q--KD--A----QN--S--D---------I----    673
C.BR.92BR025       --------T-V--I----R----------------------S---.---------R-Q-D--..NM---Q-----S---NT-YR-L-D----------D--A----QN--T--G---------I--K-    681
C.BW.96BW0402      --------T-V--I----------------------N----N---.---------DY--..GNM---------S--SNT-YR-L-D---------KD--A--S-NN-----G--K------I----    684
C.CN.AF268277      --------T-V--I--------------------------H--S---.---------R-Q-E--..NM---Q-----S---NT-YR-L-D------R--KD--A--N-KN--S--D--------RI----    678
C.DJ.DJ373A        --------T-V--I--------------------------S---.---------Q-E--..DNM---Q---------ET-YR-L-V--T---Q--KD--A----QN--S--D--K------I----    675
C.ET.ETH2220       --------T-V--I--H-R--------------------S---.---------Q-E--..DNM---Q-----S---DI-YN-L-V-----D---KD--A----EN----------------I----    676
C.IN.93IN101       --------T-V--I--------------------------S---.---------TQSE--..NM---Q----VS---NI-Y--L-----------KD--A--S-KN--S--D---------I----    689
C.SO.SO145A        -------R--V--I----E----------------------S---.---------RTQ-E--..ENL---Q--K--S---DT-YR-L-V---------KD--A--S-G------D--K------I--I-    686
C.UG.UG268A2       --------T-V--I----Q----------------------S---.---------GD--..DNM---Q-----S---GT-YR-L-D----------KD--A----QN--S--D--K------I----    681
CONSENSUS D        ----------?l---r--k------i-----rH--t-t----s---.---------k-lde--..?NM--me-Er--d---Gl-Ys---e--t------ke--e-----------s--q-----ki--mi    658
D.CD.84ZR085       ------------------------------H----T----S---.---------R-V-Y--..GNM---Q-E---D---G--YN------I------K---------------S--Q------I----    688
D.CD.ELI           ------------------------------H----N----S---.---------R--NE--..QNM-----E---D---G--Y-------T------K---------------S--Q------I----    686
D.CD.JY1           ----------V----S---------------H----T----S---.---------E--..-NM--I--E---D---GV-Y----N--I-------D--Q-----------S--K------I----    695
D.CD.NDK           ----------V-------R------------RH----N----S---.---------R--DE--..QNM-----E---D---G--Y-------I------K---------------S--K----------    678
D.CI.CI13          ----------V-------------------RH----T----S---.---------DE--..ENM-----E---D---G--Y-------T------K---A-----------D--Q------I---V    673
D.SN.SE365A2       ------------------R-------------RH----N----S---.---------TE--..QNM-----E---D---G--Y-------T----------A----------D---------I---V    683
D.TZ.87TZ4622      ----------XP---S--R------------RH----T--------.---------GE--..GNM-----E---D---G--Y-------T-------D--Q-----------S--Q------I----    683
D.UG.92UG024-D     ----------V----S--------------RH--P-Q----S---.---------DT--..GNM-----E---S---G--YD------I------KD--------------D--------R---I-    682
D.UG.WHO15-474     ------------------Q------S-----RH----T----S---.---------DD--..KNM-----EK--D---GV-YR------T----------Q-------------Q------I----    692
F1.BE.VI850        ----------V-----------------------N----S---.---------Q-E--..NM-----EK--S--SNI-YK-----------------A-----------D-S-------I----    664
F1.BR.93BR020-1    ----------V--------------L----------N----S---.---------E--..GNM-----EK-VS--SKE-YR---D-------------A-------------D--Q------I----    678
F1.BR.BZ126        ----------V-------Q------L----------N----S---.---------Q-E--..ENM-----EK--S--SYE-YR---Q-------------A-----------DMS-------I----    674
F1.FI.FIN9363      ----------V--------------L----------N----S---.---------QDE--..NM---Q-EK--S--SKT-YM---K--S---R-----------D---S--D---------I----    674
F1.FR.MP411        ----------V-----------------------N----T---.---------HDE--..NM---Q-EK-----SNT-YR-----------------A--------S--D-S-------I----    675
F2.CM.CA20         ----------L---------------------I----T----T---.---------TQ-E--..KNM--IQ-EK--S---DT-YR---GA---------N--A----DI--S--S--------RI---V    663
F2.CM.HIM277819    ----------L----------F--L----------N----S---.---------QDE--..NM---Q-EK--S---GT-YR---VA-----Q------A----DN-----T---------I----    679
F2.CM.MP255        ------------------------------------N-R--S---.---------YDD--..DNM---Q-EK--D---KT-Y----DA-----R------A----D---S--S--------I----    675
F2.CM.MP257        ---------------------------------P-T----L---.---------QDE--..GNM-----EK--G---DT-YR---SA---------D--A----DN-----S--R-----EI----    682
```

```
                                   immunodominant region
                                    |-        *     * -|   ^^^   ^         ^^       ^^^          ^^^                                   ^^^
B.FR.HXB2            VWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWS.N........KSLEQIW..NHTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIKLFIMI    688
CONSENSUS G          -------qa-v--v--y-k------------------N----t---.-........ksyne--..dnM--ie-e--is---qq-Ys-i-e--n-------D--A--k-as--?--?-tn-----kI-i-i    650
G.BE.DRCBL           -------R--V--L-----------------------N----T---.---------..YNE--..ENM--I--E---D---YH-Y----Q--I-------D--A--Q-----S--S-S-------RI-V--    679
G.FI.HH8793-12-1     ----------V--L----R------------------N--------.---------..TYND--..DNM--IQ-----S---QQ-Y---------------D--A--N-----T--D--K------I----    675
G.GA.LBV217          ----------V-------Q------------------N----T---.---------..FNE--..DNM--I--E-------HQ-Y--L----I-------D--A----E---S--S--Q------I----    689
G.NG.92NG083         --------S-V--I-----------------------N----T---.---------..YNE--..DNM--L--E---H---QH-Y----------------D--A-----------D-S-------RI----    674
G.NG.NG1937          ----------V-----F--------------------N----T---.---------.R-YDE--..DNM--I--E--VS---QE-YK--------------D--A-----G-----S--K------I---V    695
G.NG.NG1939          ----------V-----------------------N----I---.---------..Y-D--..DKM--I-----VS---QQ-Y---------------D--A--------S--D---------I----    678
G.SE.SE6165          ----------L-----------------------N----V---.---------..YNE--..DNM--I--E-------YQ-Y--L------------D--A--Q--------G--R------I----    691
H.BE.VI991           ----------V-----------------------N----S---.---------...DE--..DNM------KQ-----DE-YR-L-V----------D--A-----N-----S--------RI----    694
H.BE.VI997           ---V------V-----------------------N----ST--.---------...AE--..DNM-------Q-D---EV-YR-L-L--T---Q---D--A----D------S--------I--I-    687
H.CF.90CF056         ----------V-------R----------------N----S---.---------..QSE--..DNM------KQ-S---EE-YR-L-V--T-------D--A--------T--D-SH------I----    680
J.SE.SE7022          ----------V-----------------------N--------.---------..Y-D--..ENM--IQ-E-------GI-Y-----A-----T--KD--A----TN-------S-------I----    685
J.SE.SE7887          ----------V-----------------------N--------.---------..Y-D--..ENM--IQ-E-------GI-Y-----A-----N--KD--A----TN-------S-------I----    683
K.CD.EQTB11C         -------R--V-------R----------------N----S---.---------..QSE--..ENM---Q-EK--S-H--T-YR------I-------D--A-----------D-S-------I----    683
K.CM.MP535           -------R-----------------------------N----S---.---------..W-E--..NM-----EK--G--SDT-YK------T-------D--A-----------D--K------I----    674
N.CM.YBF106          I------R-KV--I----R---I-SL------T--Y-T----DX--S........T-YDT--..XNL--QQ---KVR--SGV-FD---QA-E--NT--KA-----Q--------D--K------IA--V    660
N.CM.YBF30           I--------KV--I----R---I-SL------T--Y-T----ET--N........T-YDT--..-NL--QQ--EKVR--SGV-FG---QA-E--NT--KS-----Q-D---S--G--K------IA---    677
CONSENSUS O          v---R--R--Ll-L-TLiQN--l-nL--cK-rli-Y-S-k--?t-???......????nds--..dnl--Qe--qqi?-iSst-ydE-QkA-v---q-ekkl----E---iw--lD--K------iA-I-    649
O.CM.ANT70           X---R--R--L--L-TL-QN----SL---K---V-Y-S-K--RT-I.G..........N-S--..DTL--Q----Q-S-IS-T-YEE-QKA-V---Q--KK-----E---I---LD--K------IA-I-    684
O.CM.CM4974          ----R--R--LQ-L-TLIQN--R-NL---K-R---Y-S-K--TT-G.----------.YNDS--..DNY--QQ--QQV--VS-I-YDE-QKA-D---R-VKAV----E-------LD--K------IA-I-    694
O.CM.HIV1CA9EN       ----R--R--L--L-TLIQN----NL--YK-R---Y-S-K--DT-YK-.........TTDD--..DKL--Q---QQ---VS-I-YEE-QKA-V---E--KK-----E---IG--PD--K------MA-I-    697
O.CM.MVP5180         ----R--R--LQ-L-TLIQN--R-NL---K-----Y-S-K--T---.G.......RYNDDS--..DNL--QQ--QH---VS-I-YDE-QAA-D-----VKA-----E--------D--K------IA-I-    697
O.GA.VI686           ----R--R--L--L-TLIQN----NL---K-R---Y-S-Q--KT-TNK......SNVD--N--..GNL--Q---QQ-D-IS-T-YDE-QKA-V---Q---K-----E---I---LD--K------MA-I-    700
O.GQ.193HA           ----R--R--L--L-TLIQN----NL---K--IV-Y-S-E--KT-I.G..........NDS--..-DL--Q---QQ-S-ISTT-FEE-QKA-V---Q--KK-----E---I---LD--K------IA-I-    681
O.SN.MP1300          ----R--R--L--L-TLIQN----NL---K-R-V-Y-S-K--RT-TN-.......NTD-DT--..GNL--Q---QQ-S-ISAT-YDE-QKA-V---H--KK-----E---I---LD--K------IA-I-    696
CPZ.US.CPZUS         ----------V-----------I--L------T--Y-T----DT--.---------.NL-YDA--..GNL--Q----KVR--SGT-F----QA-E--NT--KS-----Q-S-------D---------I-LIV    661
CONSENSUS 01         ----------V----------kF--L------Ii---a----s?--.---------.k-yee--..nnM--ie-er--?n--nq-YE?LTe--n--?r--kD--e--k-as-----?-tn-----kI--mi    662
01_AE.CF.90CF11697   ----------V----------KF--L------I----N----T---.---------..YAE--..DNM--I---K------NQ-YE-LTK----------D--A-----N-----S---------I---V    688
01_AE.CF.90CF402     ----------V----------KF--L------I----S----S---.---------..Y-A--..-NM--I-------E--NQ-YE-LT---D---R--KD--A-----------D--R------I--I-    696
01_AE.CF.90CF4071    ----------V-----------F--L------I----S----S---.---------..YNE--..-NM--I----------NQ-YE-LTK--D---R--KD-------T------D--R-----RI----    683
01_AE.CM.CA10        ----------V-----------F--L------I----S----S---.---------..Y-E--..ENM--I--E-------KQ-YE-LT-------R--KD--A--S--------S-------RI----    694
01_AE.TH.93TH065     ----------V----------KF--L------I---------ST--.---------.R-Y-E--..-NM--T--E---S---NQ-YEILT---D--DR--KD--------N-----S-S-------I----    679
01_AE.TH.CM240       ----------V----------KF--L------I---------ST--.---------.R-F-E--..-NM--I--E---S---NQ-YEILT------DR--KD--------------D---------I----    682
01_AE.TH.KH03        ----------V----------KF--L------I---------ST--.---------.R-Y-E--..DNM--I--E---S---NQ-YEILT------DR--KD--------------S---------I----    683
01_AE.TH.TH022       ----------V----------KF--L------IV--------ST--.---------..Y-E--..DKM--TQ-E---S----T-YEILT------DR--KD--K-----------D-SR-----RI----    682
CONSENSUS 02         ----------V--L-r--r-------------l----t----s---?.--------.Tynd--..dNM--lQ-dK-iS---di-Yn-i--s--q-----qD--A--k-as--n--d-tn--w--?I-imi    667
02_AG.FR.DJ263       ----------V--L-S--R------------------N----S---.---------.TYND--..DNM--LQ--K-VS---DI-YN--------------D--A--Q--N-----G--K------I----    694
02_AG.FR.DJ264       ----------V--L----R------------------T--------.---------.TYND--..DNM--LQ--K--S---DI-YN-------------KD--A----E---------K------I----    680
02_AG.NG.IBNG        ----------V--L----R------------------T----S---.---------.TFND--..DNM--IQ-EK--S---DI-YN-------R------D--A-----------D--------RI----    679
02_AG.NG.NG1921      ----------V--L-G--R-------------I---------S---.---------.TY-Y--..DNM--QQ--K-VS---NT-YR-L------------D--A--S-----S--D-S---R--QI----    698
02_AG.SE.SE7812      ----------V--L-A--------------------T----S---.---------.TYDH--..GNM--LQ--K--S---HI-YD--------------D--A--------D--S-SS-----RI--I-    691
02_AG.SN.MP1211      ----------V--L----------------------T----S---N---------.TFKD--..DNM--LQ--K--S---EK-YT----A---------D--A-----------D--------LI-L-V    686
03_AB.RU.KAL153-2    ----------V-----------------------------T---.---------.P-DE-$..-NM-----E-------G--YN---------------I-A-----------D-SK------I----    671
03_AB.RU.KAL68-1     ----------V--A--------------------------T---.---------.P-DE--..-NM-----E-------G--YN---------------I-A--------S--D-SK------I----    671
03_AB.RU.RU98001     ----------V-----------------------------T---.---------...DK--..-NM-----E-------G--YN-----------------A-----------D-SK------I----    670
04_cpx.CY.94CY032-3  ----------V--L-S--------------------N----S---.---------..YND--..DNM--LQ--K------QI-YG-L------------D--A-----------S--K------I----    686
04_cpx.GR.97PVCH     ----------V--L-S-------------F-------N----S--N.---------..YND--..GNM--LQ--K------QI-YE-L----I-------D--AF----N-------S-------I----    707
04_cpx.GR.97PVMY     ---V------V--L-S--R------------R-----N----S---.---------..Y-K--..DNM--LQ--K------EI-YD-LG-A---------D--A-----N--S--D-SH------I----    693
05_DF.BE.VI1310      ----------V--------------L--S---H--S-S----S---.---------..EGE--..DNM-----EK--S--S-T-YR---Q--I----------S--Q--------D--K------I----    693
05_DF.BE.VI961       ----------V-------R------L----------S----S---.---------..Q-E--..ENM---Q-EK--GQ-SNT-YR---Q---------KD--A-----------S-S------RI----    685
06_cpx.AU.BFP90      ---------------------------------P-N-L--T---.---------.TYDE--..GNM--I----------QQ-Y----L--T-------D--A-----N--S--D---------I----    691
06_cpx.ML.95ML127    ---------------------------------P-N--------.---------.TYNE--..DNM--I----------QQ-Y---------------D--A--------S--D-S-------RI-VI-    677
06_cpx.ML.95ML84     ----------V----------------------P-N--------.---------.RTFNE--..-NM--I----------QQ-Y-------S-------D--A-----------D-S-------I----    672
06_cpx.NG.NG3670a    ----------V--L----R------L--Y------P-N--------.---------.TYNE--..DNM--IQ-----Y---QQ-Y-------S-------D--A----T---S--D-S-------I----    672
06_cpx.SN.97SE1078   ----------V--L--------------------P-N----I---.---------.TYSE--..DNM--I----------QQ-YN--------------A--------S--D-S-------I----    678
10_CD.BFL061         ----------V----S---------------RH----N----S---.---------...GE--..DNM-----E---D---G--YN------I----------Q--------S--S---------I----    672
10_CD.BFL071         ---------------S--------------------N----D---.---------...DA--..NM---D-E---D---G--Y---------------K---------------S-------RI----    683
10_CD.BFL110         ----------V----S----------------H---------S---.---------....E--..-NM-----EK--D---G--Y-------T----------Q-----------S--K------I----    680
CONSENSUS 11_cpx     ----------V--v-r--k-q----------------n----v---?.--------.syde--??dNM--ie-e-e-d---qt-YT-l----n-------d--?--------n--d-sn-----?I---i    646
11_cpx.CM.MP818      ----------V-----------------------N----I---.---------..YDE--..DNM--I--E---D---QT-YT-L----T-------D--S--------S--D---------I----    677
11_cpx.CM.CA1        ----------V-----------------------T----N---H---------..VDD--..QNM--L--E---S---QT-YT-L---------------S--------S--E-S------RI----    680
11_cpx.FR.MP1298     ----------V----S--R------------------N----I---K---------..YNE--..-NM--IQ---------Q--YT-L----T----------S----------D-SS-----I---V    687
11_cpx.FR.MP1307     ----------V--L----R-R----------------N----V---.---------..YDE--..DNM--IQ-----H---QT-YT-L----S-------D--A--------D--S-S------RI----    682
11_cpx.GR.GR17       ----------V-----------------------N----V---.---------.TYNE--ENENM--IK-E---D---QT-YT--------------D--A-----------D-S------RI----    676
CPZ.CM.CAM3          ----------V-------R---I--L------A--Y-T----NT--A........T-FDE--..-NL--QD--KRVK--SGV-F----QA-E--NT--KS-----Q-S------D--R----------    681
CPZ.CD.CPZANT        ---V------M----K--R-----SL---AD-VT-H-T----N--VNFTQTCAKNSSDIQC--..ENM--Q----LVQ-S-GQ-YNILQIAHE---R-KK--Y-----S------D--Q------I----    690
CPZ.CM.CAM5          ----------V-----------I-SL------A--Y-T-----T--A........T-YDE--..-NL--QD--KKVK--SGV-F----QA-E--NT--KD-----Q-S---S-----Q------I-LIV    687
CPZ.GA.CPZGAB        I--V------L-------Q---I--L------AV-Y-T----N--P.G.......SN-TDD--..GNL--QQ--KLVS---GK-FG-L--A-S------RD-----Q--------D--K------I-L-A    679
```

HIV-1/SIVcpz proteins

HIV-1/SIVcpz proteins

```
                                                                                <- tat cds end    *                                   rev cds end <-
B.FR.HXB2          VGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTHLP.TPRGPDRPEGIEEEGGERDRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELLGRR.......GWEALKYWWNLLQYWSQELKNS   810
CONSENSUS A        ----I----v-a--siIn-v--------f-tht-.n-rgld--gR-e--g--Qgrd------s-F-ala-d---s--l---h----Fil-aart---l-hsslkglrlgweGl--L?---l--gr--ki-   791
A.GB.MA246         ----I------T----I---------------T-.N-G-L---RR-------Q---------G-F-T-V-----------------FT--AA-T-----HSSLKGLRL---G---LG---L--GR--RI-   812
A.KE.K89           ----I------T----I---------------T-.A---L----R-------Q---------T-F---A-----------------FI--A-ET-----HSSLRGLRL---G---L----L--G----I-   809
A.KE.Q23           ----I----------VI---------------T-.N---L----R----D--QG-G------S-F---A-----------------FI--AA-T-----HSSLKGLRL---GI--L----S--GR---I-   807
A.RW.PVPI          ----I----I------I-------------I-T-.N-G-----GR-------QG--------S-F---A-E---N-----------FI--AA-T-----HSSLKGLRL---G---L----L--GR---I-   820
A.SE.UGSE8131      ----I---------TVIK------------I-T-.S--D----GR-------QG--------S-F---A--------R--------FIS-A--------Q...........G---LG---L--IR---I-   814
A.UA.ukr970063     ----I-----------I--A--------L--LT-.H-E-----GR-K-----QG--------S-F---A------------R----FIS-TA-T------SSLKGLRL---G---LG---G--G----S-   815
A.UG.92UG037       ----I----------VI---------------T-.N---L---GR-------Q--G------S-F---A-----N-----------FI--AA-T---P-HSSLKGLRL---G---LG---L--GR---I-   811
A.UG.U455          ----I------T----I--------------LA-.I-E-LG--GR-------QGK-------S-F--IA-----N-----------FA---A-A------SSLKGLRL---G---L----L--GR---I-   809
A2.CD.97CDKS10     ----I-----M-II-V--------------IPT-.N-E-L--HGR---G---Q--T------S-F-G-A-----------------CI---A-T-----HSSLKGLRL---G---LG---L--GR-----   815
A2.CD.97CDKTB48    ----I-----I-II-V-K------------IPT-.N-E-L---GR-------QG--------S-F---A-----------------CI--AA-----V-HSSLKGLRL---G--HL----V--G----T-   804
A2.CY.94CY017-41   ----I-------IITV----------V---IPT-.S-E-----R-T--G---QG----------FF--A----------------CI--AA-T-----HCSLKGLRL---G--NL----L--GR-----   819
CONSENSUS B        v---v--riv-avl-i--r-------l-f-trl-.ap---dr-eg-ee---erdrdr-grlvd-flali-d--rs-cl-s-hrl---l-ivtri--l--rr..........-a-k-wwn-----s---kn-   797
B.AU.MBC18         ----I------T-F--------------------.------------------G-GS-T---H-F---F--------------------------------G.......................R---QK-   818
B.CN.RL42          ---------------------------L--RF-.AQ---G----------------E---T-F-S---E--------------------A--------..........---V-R---------I------   799
B.DE.D31           ------------------S------------R--.A---------------D------N---K-F-------S-----L----------AA---------..........---V------------------   804
B.DE.HAN           ----I--------------------------L--.AT---RQ--E----------G--V---S-F---F------------R---------AK---T----..........---V------------------   809
B.ES.89SP061       -----------T------------------R--.AR--------------------GQ--D-F--I--V-------------------------------.......................I------   816
B.GA.OYI           ------------------------------R--.-Q--------------------G---D-F----------------------I---A---------..........---V------------------   809
B.GB.CAM1          I---I------TI--L---------------RF-.V---------------G----T-G---T-F--------------------------A---------.......------------------R--   810
B.GB.WB            ----I------TI-----------------I-P-.AQ------G---------------P-AT-F-----------------------------F----.......------------------------   813
B.JP.ETR           -----------T------------------R--.A$--------------------GP--D-L-----V----------------------T------K..........---V---L--------------   815
B.JP.JH32          -----------T------------------R--.A---------------D------GQ--D-L--I--V-------------------------------.......------------------------   821
B.KR.WK            ---------I-F--------------------F-.S-------G---G----G-SGG-SH--D-F-T---V-----------L----------S-----L-..........---I---L--------------   819
B.NL.3202A21       ---------I--------------------R--.A----------------------G-----F-----------------------------------..........---V------------------   810
B.TH.93TH067       ---------I----------------V-L--RF-.AQ---G----------------K---D--WT---E--Q----------------A---------.........---R--$------I------   791
B.TT.QZ4589        ----I--------------------------L--.--------G--------------V---S-FS-IV-------------------AA----I----.................GS----------Y-   813
B.TW.TWCYS         ----I--------------------------.-----------D-----------E-S-D-F--I--V-------------------------------.......-------L------I------   801
B.UA.UKR1216       ---------I--------------------R--.-Q-------------------T-------F--------------L--H-------AA---------.......------------------I------   805
B.US.DH123         ----I------S-------------------RF-.AS--------------D------SP--D-F--I--V---T-F-------------------------.........---L---L--------------   803
B.US.JRCSF         ----I------S-------------------L--.AT---------------------GQ----F-----V-----F------------T-----------.........---I------------------   802
B.US.JRFL          --------L--T-------------------L--.A---------------------G-----F-----V-----------------T-----------.........---V------------------   801
B.US.MNCG          ----------------------------L--RP-.V--------------------T-G---H-F--I--V-----F-----H.------AA---------.........---V----------------S-   810
B.US.RF            -------K--------------------------.A-----------G----------GGA---F-T------WT--S-------------V---------.......------------------------   819
B.US.SF2           ------------------------------R--.V-------D--------------V---D-F-----E----------R--------AA-T--I--H-...................S-----I------   809
CONSENSUS C        ---lI----I-av--------q---------Lt-.n-r----lg?-------Qdrd--i--vs-F--la-----s-cl---hr---liliaa-av----rsslrglqrg--a-kylgs-vq--gl---k-   785
C.BI.BU910112      ----I----I--I-------------------LT-.S------LGR-------Q---------S-F---A-----------------I--A--V------SSLRGLQR---I---LGS-V---GL---K-   802
C.BR.92BR025       ----I----I---------------------LT-.N------LG--------Q---------S-F---A-----------------I--AA-A------SSLRGIQR---I---LGG-V----L---K-   810
C.BW.96BW0402      ----I----I--A-------------------LT-.N------LG--------Q---------S-F---V-----------------FI--AA-AA-----SSLRGLQK---T--FLGS-V---GL---K-   813
C.CN.AF268277      ----I----I---------------------LT-.N-G----LGR-------Q-KN--------F---A-----N--R----L--------A---------.........------R------K--V------   800
C.DJ.DJ373A        ----I----I---------------------LI-.N-------R--------Q-KG--T---S-F---A-------F----------I---A-V-----Q-..........T---LGS-V---GL---K-   797
C.ET.ETH2220       ---VI----I---------------------LI-.H------LG--------QG---------F--IF----------------I--AA-T------SSLKGLQR---T---LGS-V---GL---K-   805
C.IN.93IN101       ----I----I---------------------LT-.N------LGR-------Q-K---------F---A-----N----------FISVAA-V------S.......S------LGS-V---GL---K-   811
C.SO.SO145A        ----I----I-V--------------------LT-.N------LGR-------Q-K---------F---A----------------FI-VAA-A------SSLRELQR-------LG--G---GL---K-   815
C.UG.UG268A2       ----I----I----------K---------LT-.--------GE-------Q-----V--IS-F---A------------R--------AA-A------SSLRGLQR-------LGS-V---G----K-   810
CONSENSUS D        vg--I----v-a---l-----------s---Ll-.a-------egi--e---qg-dr-i--v--fs--i--d--n---fs--r--dli--Aa-i-e---r-.........ealk-L-n--q--iqe-k--   780
D.CD.84ZR085       ----I--------------------------L--.A----------------Q--GT-------FS--------N----------E-I--AA---------.........--------L--------R------   810
D.CD.ELI           I---I----------L---------------L--.A---------T-------G----V--L--FS----------------------I--AV---------........DI---L-----------R--   808
D.CD.JY1           ----I------T---L---------------L--.A----------------QG----------FS--F-----N------------I--A---------.........----I--L-S-----T------   817
D.CD.NDK           ----I----------V---------------L--.V--------E--------G----------LF--F-----N-----------SI--AA---------.........--------L-----------R--   800
D.CI.CI13          ----I----------L---------------L--.A----------------QG----------FS--------N---------------AA---------.........---V---L-------I------   795
D.SN.SE365A2       -E--I----I-----L---------------L--.A----------------QG----------FS-----G--N------------I--AA---------.........---V---L-S-----I-K----   805
D.TZ.87TZ4622      ----I----------L---------------L--.A----------------Q--G--------LS--------N------------I--A---------.........---TI--L-------I------   805
D.UG.92UG024-D     ----I----------L---------------L--.--------GET------QG-G-----L--LS--------N----R--H----I--AA---A-----.........----I--L-------I------   804
D.UG.WHO15-474     ----I----------V------------P---LF-.A--------E---G---QG-G--------FS--------N---L---H----I--AA-V-----H-..........IR-L-S--K--I------   814
F1.BE.VI850        ----I----------------K------L--LI-.S------------G---QGK---V---T-F---A-----N------RH---FI--AA---DRGL--.................LG--TR---------   786
F1.BR.93BR020-1    ----I------T---------K----------I-.S--E---------G---QGK---V---T-F---A-----N------RH---FI--AA---DRGLK-.................LG--T---G------   800
F1.BR.BZ126        ----I----------------K------L---I-.S--E---------G---QGNE--V-----F---V-----N------RH----I--AA---NRGL--................LLG--TL--G------   796
F1.FI.FIN9363      ----I----------------K------L--LI-.A-TE---------G---QGK---V-----F---V-----N------RH---FI--AA---DRGL--.................LG-II----------   796
F1.FR.MP411        ---MI----------------K------L--LI-.S------------G---Q--N--V-----F-S-V-----N------RH---FI--AA-T-DRGLT-..........---L---A---------   797
F2.CM.CA20         ----I----------V-K----------L--LI-.S--X--X-GX-------QG--------S-F---A---I--------RH---FI--AA-T-DRGLKG.......-------L---T---G------   785
F2.CM.HIM277819    ----I--------------------------L--L--.----------------Q--G--------LS--------N---------------AA-S-D-----.......-------L-------------R   801
F2.CM.MP255        ----I----------V------------L--LI-.N-------G---------P-----M---S-F-P-T--------S---RH-------AA-T-DRGVKG.......-------L---T-H-GR-----   797
F2.CM.MP257        I-S-I------T----I-----------L--LI-.NS---E--G--------Q-K-------S-F---A---F----V----C--NFI--AA-T-DKGLK-..........---V---L---A---G------   804
```

```
                                                                                <- tat cds end    *                                   rev cds end <-
B.FR.HXB2           VGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTHLP.TPRGPDRPEGIEEEGGERDRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELLGRR.......GWEALKYWWNLLQYWSQELKNS    810
CONSENSUS G         ----Ig-r-v-----------q--------tLtH.HQ-E---pe?i--g---Qdrd--i---S-F-ala---l---------r---fil-aa-Tve---rsSLkGLRL---g---L----l--gr-----    778
G.BE.DRCBL          ----I--------------------------L-H.HQ-E----A----G---Q---------S-F---A-----------------FI--AA-T------NSLKGLRL-------L----L--AR-----    808
G.FI.HH8793-12-1    ----IS-K-I-----------K---------LTH.HQ-E-----R-------Q-K-------S-F---A-----------------FI--AA-T-----HNSLKGLRL---G---L----L--GR-----    804
G.GA.LBV217         ----I--K-----------------------LAH.HQ-E-----R---G---Q-----V---S-F-SIA-------------Q---FI-----T------NSLKGLRL---G---L----L--G------    818
G.NG.92NG083        ----I--------------------------LTH.HQ-E---LGKT--G---Q-----T---S-F---A-----------------V--AA-T------SSLKGLRL---G---L----L--GR-----    803
G.NG.NG1937         ----I-------------------------ILPH.HQ-E-----R---G---Q-K-------S-F---A-----------------FIS-AA-T-T----SSLQGLRL---G---L----L--G------    824
G.NG.NG1939         ----I--------------------------LTH.HQ-E---------G---QGK-------S-F-S-A---V-------------FI--A--TLQ----SSLQGLRL---G---L----V---------    807
G.SE.SE6165         ----I--------------------------LTH.HQ-E---------G---QG-G--V---S-F-P-------------------SI---A-T------SSLKGLRL---G---L----L--GR-----    820
H.BE.VI991          ---II-------------L--LI-.NQ------RE-------Q-----------F-P-V-E---N------R------S--A-T-------.............LLG---L--G------    816
H.BE.VI997          --A-I----I-------G-------------LI-.N----------------Q--G--V-----F-PIV------------RL---S----I-T-------........-R-----L-------G------    809
H.CF.90CF056        ----I----I--------------------LV-.N---------T--G---Q-----V-----F-PVV-------S----RL--------V-T-------........-R-----L-------G------    802
J.SE.SE7022         I---I----I----A---------------LI-.N-TEA---G----G---QG-T--------F---A-----------------FV--AA-T-GT--L-..........I---LV--VW--G------    807
J.SE.SE7887         I---I----I----A---------------LI-.N-TEA---G----G---QG-T--------F---A-----N-----------FV--AA-T-GT--L-..........I---LV--VW--G------    805
K.CD.EQTB11C        ----I------T---V---------------LT-.S------------G---Q-K---V---S-F---A-----N------RH----V--A---LD.......RGLKGS------L---IL--G--I---    805
K.CM.MP535          I---I----A-----V---------------LI-.-S--A------------Q-KN--V---S-F---A-----N------RQ--N-I------L-.......RGLRG-------L---V----------    796
N.CM.YBF106         -A-II-I--IS-IIT-IA----------L--LI-TAA-------ET--GV-GQ--G--V---S-F-----E---N-LI-L----A-S---IR-TL-I--QSLSRGLQLLN-LRIRL-GIIA--GK---D-    790
N.CM.YBF30          -A-I--I--ISI-IT-IA----------L--LI-.-A-------ET-GGV--Q--G--V---S-FS--V-E---N-LI-L----T-S---LR-TL----QSLSRGLQLLN-LRTHL-GI-A--GK--RD-    806
CONSENSUS O         --A-i-v-vvMIvlNl-rNI-q--Q---l-iPih?hq?EA?t-GrTGeg--eeg-pr?ipspq-F-plLYT---TIilWt--L-sN-aSgIqtvish-gLGLwILGQkiIsaCRiCaAVt-y-l---qn-    775
O.CM.ANT70          --A---V-VIMI--N--KNI----Q---L-IPNH.HQEEAGT-GRTGGG---EG-P-W-PSPQ-F-P-LYT---TII-WT--L-SN-ASGIQKVISY-RLGLWILGQKIINVCRICAAVT---L---Q--    813
O.CM.CM4974         --A-I-I-V-MILINL-KNI----Q---L-IPI-.HHQEAET-GRTG----DGG-QXWTPWPP-F-Q-LYT---TII-WI--L-SN-ASGIQ-VISY-SLGLWILGQKTI--CRLCGAVT---L----T-    823
O.CM.HIV1CA9EN      --A-I-V-V-MI-IN--RNI----Q-----IPIH.HQSEAGT-GSTG-G---EG-P-L-PSPQ-F-P-LYT---TII-WT--L--N-ASVIQTVISH--LGLSILGQRIIN-CRVCIAVI---L---Q--    826
O.CM.MVP5180        --A-I-I-VIMII-NL-KNI----Q---L-IPV-.HRQEAET-GRTG-----G--PKWTA-PP-F-QQLYT---TII-WT--L-SN-ISGIR-LIDY--LGLWILGQKTI--CRLCGAVM---L------    826
O.GA.VI686          --A---I-V-MII-NL-RNI----Q---L--PIR.HQSEAET-GRTG-G---E--P-L-PSPQ-F-P-LYT---TII-WI--L-SN-ASGIQKAISH--LGLWILGQKIIS-CRICAAVT-F-L---Q--    829
O.GQ.193HA          --A-I-V-VIMI--NL-RNI----Q---L-IPIQ.HQAEAGT-GRTG-G---EG-P-L-P-QQ-F-T-LYT---TITVW---L-SN-ASGIQTVTRH--LGLWILGQKIIS-CRVCAAVT---I---Q--    810
O.SN.MP1300         --A-I-V---MI--NL-RNI-H--Q------PTHHQQPEAQA-G-TG-G------L---PSPQ-F-P-LYT---TII-W---L-SN-ASGIQTVISH--LGLWILGQKIIS-CRICIAVI---L---Q--    826
CPZ.US.CPZUS        -AS---I---GVIF-L-AK---------L--LF-.-T-E------T--GA-KT-NV--T---S-F---V-E---N-LI-L----E-----LR-T-QI--QNINKGLQLLN-LRARC-GVIA--AR---V-    790
CONSENSUS 01        ----I----v------------------f--ltH.hQre-----r---gg--qgr---v---S-F-a-a----------s------fi--AA-tv----hSSLkGLRRg--G---Lgn--s--g---k?-    790
01_AE.CF.90CF11697  ----I--------------------------LTH.HQ-E-----R---G---EGK-------S-F-S-A-----------------FI--AA-T-----HSSLKGLRR---G---LG---S--G----T-    817
01_AE.CF.90CF402    ----I--------------------------LTH.QQ-E-----R----S--QG--------S-F---A-----------------FI--AA-T-----HSSLQGLRRR--G---LG---S--V---RI-    825
01_AE.CF.90CF4071   ----I--------------------------LTH.HQ-E-----R---G---Q-K---V---S-F---V-----------------FI--AA-T-----HSSLEGLRR---G---LGS--S--G----T-    812
01_AE.CM.CA10       ----I--------------------------LPH.HQ-E-----R---G---Q-----V---S-F---A-----------------FI--AA-A-----HSSLKGLRR---G---LG---S--G----T-    823
01_AE.TH.93TH065    ----I----I------------------L--PTH.HQ-E-----R---G---QG----V---S-F-T-A-----------------FI--AA-T------SSLKGLRR---G---LG---L--G----I-    808
01_AE.TH.CM240      ----I----I---------------------PSH.HQKE---------G---QG----V---S-F---A-----------------T--AA-T-----HSSLKGLRR---G---LG---L--G----I-    811
01_AE.TH.KH03       ----I----I---------------------PTH.HQ-D-----R---G---QG----V---S-F---V-----------------FI--AA-TL----HSSLKGLRR---G---LG---A-------T-    812
01_AE.TH.TH022      ----I--------------------------LSH.HQ-D-----R---G---QG----V---S-F---A----------L------FI--AA-T-----HSSLKGLRR---G---L----V--G----I-    811
CONSENSUS 02        ----I-----f--ltiIn-------------Lth.HQ-epd---ri--G--eQ-r-r-v---S-F---A-------c--s--r---fvl-aa-?--l--hsslkglrl--e--kyL?--ls--g------    794
02_AG.FR.DJ263      ----I---------T-I--------------LTH.HQ-------R---G---Q-----G---S-F---A-----------------FV--AV-A-----HSSLKGLRL-----Q-LG---L--G------    823
02_AG.FR.DJ264      ----I-----L---TVIK-------------LTH.HQ-E-----R---G---Q-K---V---S-F---A-------------H---FV--AV-A-----HSSLKGLRL-------LG---T--G------    809
02_AG.NG.IBNG       ----I---------T-I--------------LTH.HQ-E-----R---G---Q-K---V---S-F---A-----------------I--AA-T-----HNCLKGLRL--G----L---IS--V------    808
02_AG.NG.NG1921     ----I--------FT-I--------------LPL.HQ-E-----RT--G---Q---K-----S-F---A-------F----------I--AA-A--.-.......LRL-------LG---L--G------    820
02_AG.SE.SE7812     ----I---------A-I--------------LTH.HQ-E-G---R---G--GQ-----V---S-F---A-----------------FVS--A-T-----.......HR-------L----S--G------    813
02_AG.SN.MP1211     ----I---------A-I--------------LTH.HQ--A--------G---Q-----V---S-F---A----------L------FV--AA-T-----HSSLKGLRL------HL----S--G------    815
03_AB.RU.KAL153-2   ---------I---------G----------R--.AQ------------D------T------RF------------F-I--H-------AA-------.......RR---------------I----S-    793
03_AB.RU.KAL68-1    ---------I--------------------R--.-Q-------------------T------RF-------------I--H-------AA-------.......RR---------------I----S-    793
03_AB.RU.RU98001    ---------I--------------------R--.-Q-------------------T-------F--------------I--H-------AA-T-------.......RR---------------I----S-    792
04_cpx.CY.94CY032-3 ----I----I------------------L--LI-T-Q--L---G-T------Q--S--------F-P-------N------RH--N-----A-T-------.....IR-------L--F-L--G------    809
04_cpx.GR.97PVCH    ----I----I--------------------LI-.-Q-E-G----T------Q--S--------F-P-------N-------H--N----AA-T-------.....IR-------L----L--G---R--    829
04_cpx.GR.97PVMY    ----I------------K----------S--LI-T-Q--------T-GG---Q--IE-------F-P-V-----N------RQ--N-----AKT-------.....IR--GT---L----L--G--IRS-    816
05_DF.BE.VI1310     ----I------T---------K---------PF-.A--------ET--G---Q--G--------FS--------N-------H----T---V-------.....RR-------L-S-P----R------    815
05_DF.BE.VI961      ----I------T---V-----K---------L--.G---------T------QG----V--LT-LST-------N------------I--AA-------.....RR-------L-S------------    807
06_cpx.AU.BFP90     ----I--------I--------------L--LI-.N-T-A---GE---G---QG-T--------F---A-----------------FG--AA-T--I--.......RR---I---LG--IC--G---Q--    813
06_cpx.ML.95ML127   ----L--------F--------------L--LI-.N-T-A---GE-------QG-T--------F---A-----------------FV--AA-T-GT--.......HR---I---LGS-VW--G------    799
06_cpx.ML.95ML84    ----I----------------------L--LI-.N-AEV---G----G---QG-N--------F---A-----------------FV--AA-T--T--.......RR---I---L---VC--G------    794
06_cpx.NG.NG3670a   ----I------------R----------L--LI-.N-T-A---GE---G---QG-T--------F---A-E---N-------Q----V--AA-T--T--.......HR---I---LG--VC--G------    794
06_cpx.SN.97SE1078  ----I--K-------L-K----------L--LI-.NSA-V---GE--GG---Q--T--------F---A-E------R----L---FV---L-T--T--.......HR---I---LG--VC--G------    800
10_CD.BFL061        ----I-----C----L------------LH-L--.A-----------------G-G------D-F-T-------N------------I--A--------.......RM----I--L-------I------    794
10_CD.BFL071        ----I---V------L---------------L--.A----------------PG----------FS-IV-----N------------I--A--------.......RR-------L-------I------    805
10_CD.BFL110        ----I----------VA--------------LF-.A----------------QG-G--------LS--------N------------I--A--T--------.......RR----I--L----K--I------    802
CONSENSUS 11_cpx    ----ig---i-a---i---C------l---tLt-.nq?Ea---g-ie-Gg--Q?-tr-----s-F--lA-d---n---fs--r---fIl-?A---eT--------.rRg--i--yLgn-aq--G----n-    765
11_cpx.CM.MP818     ----I----I-----V---C-----------LN-.-QQEA---G----G---QG-T------S-F--IA-----N------------FI---A----T--------.HR---I---LG--T---G------    799
11_cpx.CM.CA1       ----IS---I---------C-----------LI-.SQQE----R----G---Q--T--------F--IA-E---N------------FI--AA---GT--------.RRV--I---LG--A---G------    802
11_cpx.FR.MP1298    ---------I---------C-----------LT-.NQKET---G----GS--Q--S--------F---A---------CL------FIF-AA----T--------.HR---I--LLG---L--G------    809
11_cpx.FR.MP1307    ----I----I-T-------C-----------LT-.NHKEA---R----GD--QG-TN-----S-F---A-----N------------FI---A----T--------.RR---I---L-S-T---G------    804
11_cpx.GR.GR17      ----I----I---------C----------ALT-.SQQE----G-TK-G---QG-TS-----S-F---A-----N----L--Q---FI---A----T--------.RR---S---LG--A---G----S-    798
CPZ.CM.CAM3         -A----I---G-II-F-AK---------L--LI-.-T-E----G-T--DV--PGKG------S-F-----E---N-L--C----------LG-TL-N--QSLNKGLQQLRNFSR-L-GVIT--GR--QT-    810
CPZ.CD.CPZANT       --AI-----LLVLV-CLRK-----H-----IPTQ.NQQD-EQ--E-R----RK--I-WRA-QH-FF--L-V--T-IIQWI-QIC-TC--NLWAVLQH-......CRITFRLCNHLEN--STL-.TIIRTE    812
CPZ.CM.CAM5         -A--I-F-LIGI-M-VIAK---------L--LI-.-T-E---RG-T--D---PG-G--T---S-F---V-E---N-LI-F----I---S-LW-T-QI--QSINKGLQQLKDFST-LKGV-T--GR-IQI-    816
CPZ.GA.CPZGAB       ---II----IMT-F-V-R----------L--LI-.VQ-EQG-LGE-D-G---Q--S--V---E-C-P-------N-GIW--QS-TS-ACN-W-QLKT--HLILHSLRLLR-R-CLLGGII---GK---I-    808
```

HIV-1/SIVcpz
proteins

```
                     ^^^                 *
B.FR.HXB2            AVSLLNATAIAVAEGTDRVIEVVQGACRAIRHIPRRIRQGLERILL$   856
CONSENSUS A          -in-lDti-iav-GWt-----ig-rig--iln--r-----f-rA-l-   837
A.GB.MA246           -IN--DTI--II-GW------IG-RF----LN--------F--A---   858
A.KE.K89             -SN-VDTI--V--GW------IA--IG---L---------F--A---   855
A.KE.Q23             -IN-VDTI-----GW------IA-RIG---L---V--------A---   853
A.RW.PVPI            -IN-FDTI--V--GW-------G-RIG--FLN--------A--A-I-   866
A.SE.UGSE8131        -I--FDTI-----GW------IG-RIG---L---------F--A---   860
A.UA.ukr970063       -IN-IDTI-----GW------IG-RF-----N--X-----A-KA-Q-   861
A.UG.92UG037         -IN--DTI-----GW------T--RLG---LN--------F--A---   857
A.UG.U455            -IT--D-V-V---GWI-----IG-TIG---LN-----------A---   855
A2.CD.97CDKS10       -I----S-------W------IG-R-----LN--------F--A---   861
A2.CD.97CDKTB48      -IR--DTI-V----W------IG-R------N-----------A---   850
A2.CY.94CY017-41     -I--FDTI-V----W------IG-R-F---LN-----------A---   865
CONSENSUS B          avsllnat-i--a-----vievvqRay-aiLh-prri---le-a-l-   843
B.AU.MBC18           ----F----------------AL-RTY---L------------A---   864
B.CN.RL42            -IG--------------------R-Y---L---T--------A---   845
B.DE.D31             ------------------------R-W---L---V--------A---   850
B.DE.HAN             ----F-TI----------------R-----L-----V------A---   855
B.ES.89SP061         ----------------------L-R-----L------------A-Q-   862
B.GA.OYI             VI------------------I--R-Y--FLN-----------A---   855
B.GB.CAM1            ----FDTI----------------R-----L------------L---   856
B.GB.WB              ----F-T--V-------------HRIW-G-L------------A---   859
B.JP.ETR             ---------------------IL-RTY---L---VK-------A---   861
B.JP.JH32            ----F--I-----------LKIL-R-F---L---T--------A---   867
B.KR.WK              ---------V--------I--IL-R-Y---LN-----------A---   865
B.NL.3202A21         ------------------------R----VL---V--------A---   856
B.TH.93TH067         ------T------------L----R-Y---L---T-----FK-A---   837
B.TT.QZ4589          -------I----------I-KA--RI----L---------F--A---   859
B.TW.TWCYS           ----F--I----------------RVF---L---T--------A---   847
B.UA.UKR1216         ---------V-------------RR-F--FLR--------F--A---   851
B.US.DH123           ------------G-----I--IL-R-G---LN--T--------A---   849
B.US.JRCSF           ------------------I-----RVY---L---T--------A---   848
B.US.JRFL            ------------------I--AL-RTY---L---T--------A---   847
B.US.MNCG            ----------------------L-R-G---L---T--------A---   856
B.US.RF              ------T-----------I---A-RIL--FL------------A---   865
B.US.SF2             ---W--------T---------A-R-Y---L--H---------L---   855
CONSENSUS C          -is-ldtt--a-a-----i--?i-ric-aicn--r-----f-AA-q-   830
C.BI.BU910112        -IN---I-----------I--II-RI----Y---------F-AA-Q-   848
C.BR.92BR025         -I--FDTI----------I---I--IW---CN--------F-AA-Q-   856
C.BW.96BW0402        -I---DT-----------I--IA-RI----CN--T-----F-AA-Q-   859
C.CN.AF268277        -------------------------Y---L---------F-AA-Q-   846
C.DJ.DJ373A          -----DT-----------I--II-RIW--FCN--------F-AA-Q-   843
C.ET.ETH2220         -IN---T---V-G-----F--LI-RIW--FCN----------AA-Q-   851
C.IN.93IN101         -I--FDSI--V-------I--L---F-----N--T-----F-AA-Q-   857
C.SO.SO145A          -I---DTI--T-------I--LA-RI--GV-N--T-----F-AA---   861
C.UG.UG268A2         -I---DTI----S-----I---G--IG---L---------F-AA-Q-   856
CONSENSUS D          ai--fdat-ia---g---iie?vqRacr-vlni-t-i--g---a---   825
D.CD.84ZR085         -I--VD------------I-DI-RR--K-VL---T--------A---   856
D.CD.ELI             -S--FD-I-------------II-R----VLN-----------S---   854
D.CD.JY1             FI----------------I--LIRR-F--VL-----V------A---   863
D.CD.NDK             -S---DTI------R---------R-----LNV----------L---   846
D.CI.CI13            -I--FD---V--------IL-A--RVW--VL---T--------A---   841
D.SN.SE365A2         -I---DTI----------I-D---R-----L---T--------A---   851
D.TZ.87TZ4622        -I--F-T------------DI--RTV--VLN--T----S---A---   851
D.UG.92UG024-D       -I--F-T---V-------I--L--RIG---LN--T--------L---   850
D.UG.WHO15-474       -I--F-----T-------I---G-R------N-----------A---   860
F1.BE.VI850          -I--F-T---V-------I---L-R-G--VLN--------A--A---   832
F1.BR.93BR020-1      -I------------W------AL-R-G---LN-----------A---   846
F1.BR.BZ126          -I----T---V-T-----I-AAL-RIG---LN--T-----F--A---   842
F1.FI.FIN9363        -I--F-T---V----------AL-R-V--VLN-------RV--A-I-   842
F1.FR.MP411          -I----T---V-----------L-R-G--VLNV----------S---   843
F2.CM.CA20           VI----T---V-------IV--L-R-G--VL---------F--A---   831
F2.CM.HIM277819      -I---D------------I--II-RTF---LN-----------A---   847
F2.CM.MP255          -I--FDTI----------I---L-R-G--VL---------A--F---   843
F2.CM.MP257          -I---DR-----------I--IL-R-G--VLN-----------A---   850
```

```
                         ^^^                  *
B.FR.HXB2               AVSLLNATAIAVAEGTDRVIEVVQGACRAIRHIPRRIRQGLERILL$   856
CONSENSUS G             -In-lDTi---v-Nw------va-R-c--iLNi-r-i---l--A-l-   824
G.BE.DRCBL              -IN--DTI-----NW-------A-R-G--VLN-----------A---   854
G.FI.HH8793-12-1        -IN--DTI-----NW------I--R-F--FLN--T--------A---   850
G.GA.LBV217             -IN--DTV---T-NW-------A-R-----LN-----------A---   864
G.NG.92NG083            -IN--DTI---T-N--------A-R-Y---LNV-T--------A---   849
G.NG.NG1937             -IN-FDTI-----NW-------A-R-----LN--------A--A-Q-   870
G.NG.NG1939             -IN-FDTI-----NW-------A-R-R---LNV---V------A---   853
G.SE.SE6165             -I---DTV-----NW-------A-R-----LN--T--------A---   866
H.BE.VI991              -I----T-----------I--L--R-W---L---------F--A---   862
H.BE.VI997              -IN---T---V-------I--I--R-W--VL----------------   855
H.CF.90CF056            -ID---T----------GI-VI--R-W---L---------F--S---   848
J.SE.SE7022             -I----T-----------I--IA-R-F---L------------A---   853
J.SE.SE7887             -I----T-----------I--IA-R-F---L------------A---   851
K.CD.EQTB11C            -IN---T-----------I--I-YR-F--LL---------F--L---   851
K.CM.MP535              -I----T------G----I--IG-R-F--LL------------A---   842
N.CM.YBF106             -I----T---V-------F--LA-RIG-G-L------------A---   836
N.CM.YBF30              -I----T---V-------I--LA-RIG-G-L------------A-I-   852
CONSENSUS O             -ts-lDt?-Va--NW--gI-lGi-RIGrGiln--r-----l--?-l-   819
O.CM.ANT70              -T---DTL-V---NW--GI-AGI-RIGTG--N-----------S---   859
O.CM.CM4974             -IT-FDSI-VS--NW--GI-LGI-RIGQG-LN--------A--F---   869
O.CM.HIV1CA9EN          -T---DTF-V---NW--SI-SGI-RIG-G-LN-----------G---   872
O.CM.MVP5180            -TN--DTI-VS--NW--GI-LGL-RIGQGFL---------A----V-   872
O.GA.VI686              -T--IDTL-VT--NW--GI-LGI-RIG-G--N--T--------S---   875
O.GQ.193HA              -T---DTI-V---NW--GI-SGI-RIG-G-WN----------------   856
O.SN.MP1300             -T---DTL-V---NW--GI-LGL-RIG-G-LN-----------A---   872
CPZ.US.CPZUS            -T---DT-----------I--LTRRLFLG-I------------S---   836
CONSENSUS 01            aIs-ld-t-iav-GW---v--va-?-w--il---------l--a---   835
01_AE.CF.90CF11697      -IT-FD-I-V---GW---------R-W--LI------------A---   863
01_AE.CF.90CF402        -IT--D----T--GW------I--R-W---L------------A---   871
01_AE.CF.90CF4071       -IT--D----T--GW---A--IA-R-----L------------A---   858
01_AE.CM.CA10           -I---D----T--GW------GA-R-Y--FI---------F--A---   869
01_AE.TH.93TH065        -I---D-------GW-------A---W---L------------A---   854
01_AE.TH.CM240          -I---D-----A-GW-------A---W---L------------T---   857
01_AE.TH.KH03           -I---D-------GW-------A---W--FL------------A---   858
01_AE.TH.TH022          VI--------V--GW-------A---W---L------------A---   857
CONSENSUS 02            -In--dTi-iav-NW---v--ig-RvG---rN--r-----l--a-l-   840
02_AG.FR.DJ263          -IN--DTI-L---NW------IG-RVG---LN-----------A---   869
02_AG.FR.DJ264          -IN--DTI-----NW------IG-RVG----N--V--------A---   855
02_AG.NG.IBNG           -IN---TI--V--NW---A--IG-RVG----N--------F--A---   854
02_AG.NG.NG1921         -IN--DTI--FT-NW------IG-R-G----N-----------S---   866
02_AG.SE.SE7812         -I---DTI--V--NW------L--R-G---LN--------F--A---   859
02_AG.SN.MP1211         -IN--DT--V---NW------I--RTG---CN-----------A-Q-   861
03_AB.RU.KAL153-2       -IN-IGTI-----GW------IG-RF---M-N--------A-KA-Q-   839
03_AB.RU.KAL68-1        -IN-IGTI-----GW------IG-RF-----N--------A-KA-Q-   839
03_AB.RU.RU98001        -IN-I-TI-----GW------IG-RF-----N--------A-KA-Q-   838
04_cpx.CY.94CY032-3     -IN-F-T-----------I--A--R-----CN-----------A---   855
04_cpx.GR.97PVCH        -IN--DT-----------IL-A--R------N--------F-KA---   875
04_cpx.GR.97PVMY        -I---DT--V--------I--A--RI----LN-----------A---   862
05_DF.BE.VI1310         -I----T--VV----------AL-R-G---LN-----------A---   861
05_DF.BE.VI961          -I----T--VV-------IL-AL-R-G--VLN-----------A---   853
06_cpx.AU.BFP90         -I--FD-A-----NW----A----RIF--FLNV-------F--A---   859
06_cpx.ML.95ML127       -IN--DT------N-------I--R-F--VLN--T-----F--A---   845
06_cpx.ML.95ML84        -I--IDT------NW---------R-F--VLN--------A--A-I-   840
06_cpx.NG.NG3670a       -IN-VDT------NW-------G--IL--FLN-----------A---   840
06_cpx.SN.97SE1078      -I---DT------NW------I--RVF--FLNV-------F--A---   846
10_CD.BFL061            -I---DTI--E--GW------I--R-V---LN--T--------A---   840
10_CD.BFL071            -I---DT--V--------A--I--R-V-TVIN--------A--AF--   851
10_CD.BFL110            -I----T---V-------A--I--R-V--VLN--T--------A---   848
CONSENSUS 11_cpx        -is-l-a-----------I--i??rvL-ailh--------?--A---   808
11_cpx.CM.MP818         -IN---T-----------I--I--RVL-G-L------------A---   845
11_cpx.CM.CA1           -----------------I---AHRVL--FL---------F--A---   848
11_cpx.FR.MP1298        -I--F-------------I--IAHR-L--FL------------A---   855
11_cpx.FR.MP1307        -I--F-------------I--I--RVL---L---------F--A---   850
11_cpx.GR.GR17          -I----------------I---AHR-L---LN--------F--A---   844
CPZ.CM.CAM3             -I---D------------IL--A-IIG-G-L------------S---   856
CPZ.CD.CPZANT           IIKNIDRL--W-G-K--SILLAL-TIV-I--EV---------IA-N-   858
CPZ.CM.CAM5             -T---DT------N----I--F--IVG-G-L------------S---   862
CPZ.GA.CPZGAB           -I---D------------I--AF-VTL-I--N-----------A---   854
```

HIV-1/SIVcpz
proteins

HIV-1/SIVcpz proteins

```
                                                                                                  heart of CD4 binding site         PKC
                                                                                                                  acidic
                                   N-terminal alpha-helix                                       ||                          |-  -|
                myristoylation     MHC-1 downregulation + protein kinase recruitment            ||   cluster  SH3 binding       polypurine tract  PAK binding
                  |-  -|            |--|                                                         ||   ||| |   |  |  |  |                |-     -|   ||
B.FR.HXB2         MGGKWSKSSVIGWPTVRERMRRAE...............PAADRVGAASRDLEKHGAITSSNTAATNAACAWLEAQ.EEE.EVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQ.DILD    111
CONSENSUS A       --g----ssivg-pev-ermrrtp?????.....??????p--kG---V-Q--dkh--v--s-i??nhpSca-l---?eee?e-----r---------y-g-?--sh---e----d--iy-rk--.E---    108
A.FR.HIV232956    ---------IV---DI------TP................---TG---V-Q--D-Y--V----I..NHPS-------.-D-.------R---------F-G-L------------D---Y-R----.E---    109
A.KE.Q23          ---------IV---EI-------P................---PG---V-Q--D----V--K-I..NHPSY------.-D-.D-----R-----------G---------K----D--VY-RK--.E---    109
A.SE.SE6594       ---------IV---EI-------P..............SA--PG---V-Q--A----V----I..N-PS-V-----.---.------R-------------------------D---Y-RK--.E---    110
A.SE.SE8538       -------R-KE--SE---KI-QTP................---KG---V-Q--DR--------I..NHPS-T-----.-D-.------K-----------G-F------------D---Y-RK--.E---    109
A.SE.SE8891       ---------IV---K-----A-TP................---KG---V-Q--D----V----I..NHPS---V---.---.------R-----------G--------------D---Y-R----.E---    109
A.SE.UGSE8131     --N----.....-E----I-Q-RAPA........HTPAPT--TG---V-Q------------I..NHPS-T-----AQ-DE------R-----------G-L----------------Y-RK--.E---    115
A.UG.92UG037      --N-----CIV---E----I-QTPTAARE.....RTRQAPT--KG---V-Q--D----V----V..NHPS-V-----.---.------R-------------F--GF--------D---Y-KK--.E---    120
A.UG.U455         -------K-RVE--E--K---ETP................A--KG---V-Q--D-Y--V-----SS---S-------.--G.D-----R-------------F---F--------D------K--.E---    111
A2.CD.97CDKTB48   -------RTIV---EI------TPPAAE......GVRPTP---EG---V-Q--AR---V-------N-PD-------.---.------R------A----G--------------D---Y-----.----    121
A2.CY.94CY017-41  -------R-IP---AI------TPPTAQ......RTEAVS---PG---V-Q--AT---V---------PD---V---.---S------R---------F-G-F---F-------D---Y--K--.----    122
CONSENSUS B       --gk-s-?s?g-wptv-e-mrrae??????????????pa-dG---v-r--ekhgai-ss-taat--dcawl-aq?eee?ev----r-------mtykaal-l----kek---e-liy-qkrq?d---    109
B.AU.MBC200       -------C..A------------APGER.......RRRAE----G---------R---L-------N--D-------.---G------R-----------------------------Y--E--.----    119
B.CN.RL42         -------H-MF---S-----K---................----G---V-----------------------------.---.------R-----------G-L------R---------Y-----.----    111
B.DE.D31          ----------V---AI----K---................---EG---V---------------P------------.---.------K-----------------------------V---K--.----    111
B.DE.HAN          -------C..S---------KQ--...............PE----G--------K-----------TN----------.---.-E----R-----------G-L----------------Y-PK--.E---    111
B.ES.89SP061      ---------K------------------------------G---V----------------TN--D-------.---.------R-----------G-L----------------W--K--.----    111
B.FR.NE100        -------N-MS--T-I-----Q--................----G---V--------------------D-------.---.------R---------RR--I------R--------V---K--.E---    111
B.FR.SWB884       ---------MS---AI---K---................----G---V--------------------D-------.---.------R---------RR--I------R--------V---K--.E---    111
B.GA.OYI          -------C-MK----I----K---LQ...........PPE---EG-----------------------D-------.-D-.------R-----------G-L----------------Y--K--.----    116
B.GB.CAM1         -------R-LG--SA-----Q---P............RAE---EG---V--------------------D-------.---.------R-------------L-I------------I-Y-----.----    115
B.IN.HIVP35A      --------KMG---A----I--TDPAER.......IRQPE---AG---V-----R-------------D-------.-D-.D-----R-------------G-L------RK------I-Y--K--.----    120
B.IT.B-IT-R5      --NX-X---XS-X-A--------------------------....G---V-----R-----X----TN--D------H.-D-.------R-------X-CR--I---------------VW--K--.----    111
B.JP.nef<7>-a     -------R-GE---K--------R.............PAE----G---V-----R------N----N--D-------.---.------------------G-----------------V---Q--.----    114
B.KR.CSR9412d     -------R...----A-----E---................----G----------Y----------N--DL------.---.------R-----------------------G--V---K--.----    108
B.KR.WK           -------R--P--N-I-K-----.................---EG---------QR----T----SN------Q---.---.------R-----------S-L----------------VY--K--.----    111
B.NL.3202A21      ----------V---AI----K---................----G---V--------------------D-------.-D-.------K-----------G-L----------------Y-----.----    111
B.SE.AF047085     -------C--V---------K-----------------P--G---V----G-----------TN--D-------.-D-.N-----R-------------F--------Q-------------.----    111
B.TH.28-19        --A----F..G---A--------QPETRERRRPTQQPATE---VG---V----------------E---D--R--S-.---.------K-----------G-L-------------------K--.E---    125
B.TH.AF082839     -------R-GG--NA-------T.................P--G---V-Q--G----------P-N--D-------.---.------R-----------------------------HK--.----    110
B.TW.TWCYS        -------R-IP--SNI---I-Q--P...............A---G---V--------------------D-------.---.------R--------D--G-I------RKE------VY--KKEX----    113
B.US.JRCSF        -------H--P--S----------PATD......RVRQTE---VG---V--------------------D------Y.-D-.------R-------------I----------------Y--K--.----    121
B.US.JRFL         -------R--P--S----------PAAD......RVRRTE---VG---V--------------------D-------.-D-.------R-----------G---------------------K--.----    121
B.US.LM1          --S-----IRA--S----------................----G---V--------------A-HN--D-------.---.------R-----------G-L--------E-------Y-----.----    111
B.US.MN           -------R.-T-------------P.............AEL---G-------------L----------D-------.---.------K-------------L------------D---Y--K--.----    113
B.US.NC7          ---------IV----I----K--G................----G-----------R------------D-------.---.------R-------------GI---------------VW-----.----    111
B.US.RF           --------KMG---A-----QK--................----G------------T--------N----T-----ED-DE------R---------F----------------D--VF--K--.----    113
B.US.SF2          -------R-MG--SAI--------P............RAE----G---V--------------------D-------.---.------R-------------L-I--------------W------.E---    115
CONSENSUS C       --gkw--csivgw-ai-e-mrrte???........?????p--eG--aa-q-ldky--l-s-n-dtt-adcawleaq.eeeee-g---r---------y-gafd-sf--kek---egliY-kK--.e---    112
C.BR.92BR025      --N----C-TV-R-AI-------.................---EG---P--Q-SD-Y--L---S-P-N--D-------.Q--E------R-----------V----F--E---------Y-KK--.----    111
C.BW.96BW01B21    ---------IV---A----I--T-................---EG-----Q--D-Y--L-T---DPN-------R--.---G------R-----------G-F--GF------------Y-KK--.E---    112
C.BW.96BW0402     ---------IV---A-------T-................---EG--T--Q--D----L-T-----N----------.---E------R----------------F----------I-Y-KK--.----    112
C.BW.96BW0502     -------C..S---A-------TRPAV........EGRTES--EG-----Q--D-Y--L-----STN--D-------.---GD-----R-----------G-F--GF-----------VY-KK--.E---    118
C.BW.96BW1104     -------R-K-E---I-D----T-PAA.........EGVGA--EG-----Q--D----L-------N--D-------.---.------R-------------FG--F------------Y-RK--.E---    118
C.ET.ETH2220      ---TM--C-PV---AI---I---A................---EG--------D-Y--L-----P-N-PD-------.---E------R-------------F---L------------Y-KK--.E---    112
C.FR.HIV232980    --N----C-PV---A-------TVPE..........AERAG--VG-----Q--D----L-----P-N--D-------.---E------R----------------F---------D--TY--K--.----    118
C.FR.HIV232996    --S----C-PV----------QTN.............PA---EG-----Q--D-Y--L-T---P-N--D-------.---GG-----R-----------G-F---F------------Y--K--.E---    114
C.IN.93IN101      -------C-IV---AI--------................---EG-----Q--D----L-----DT---D----R--.---G----------------S-F---F------------Y-KK--.E---    112
C.IN.93IN904      -------C-IV---DI------TQ................---EG-----Q--D-Y--L-----DI--PD----K--.---E------R---------F-E-----F------------Y-KK--.E---    112
C.IN.93IN999      -------C-IV---A-------T-................---EG-----Q--D-Y--L-----DT---D----R--.---E------R---------F-G-F---F------------Y-KK--.E---    112
C.IN.94IN11246    -------C-IV---EI------TQ................----G-----Q--D-Y--L-----DT---D----RT-.---E------R---------F-G-F---F------------Y-KK--.E---    112
C.IN.95IN21068    -------C-IV---DI------T-................---EG-----Q--DRY--L-----DT--TE........----.G-----R---------F-G-L---F------------Y-KK--.E---    104
C.IN.HIVY15117    -------C-IV---DI------T-................---EG-----Q--D----L-----DT---D.........-DG.--A---R-----------G-F---F----E----A--Y-KK--.E---    104
C.IN.HIVY17884    ---------IV---AI------T-P............AAE---EG-----Q--DR---L-----DT---D----R--.--DE------R---------F-G-F---F------------Y-KK--.E---    116
C.IN.HIVY17891    -------C-IV---AI------T-................---EG-----Q--D----L-----DT---D----RT-.---E------R---------F-G-F---F------------Y-KK--.E---    112
C.IN.HIVY17892    ---------IV---AI----K--Q................---EG-----Q--D-Y--L-----DT---D----R--.---EG-----R---------F-G-F---F------------Y-KK--.E---    112
D.CD.84ZR085      ---------IV---AI---I-KTDPRE........RRRPE----G---V-----R----------T-------V---.--DE------R---------------------RK-------VY--K--.----    120
D.CD.ELI          ---------IV---AI---I--TN................----G---V----------------S---D-------.--SD------R-----------E-L----------------W-KK--.E---    112
D.CD.NDK          ---------LV---AI---I-KTD................----G---V----------------S--DT-------.--SE------R-----------E------------------W-KK--.E---    112
D.UG.94UG1141     ---------IV---A-------T-................---EG--------------------Q--D-------.---.------R-----------E----------------VW-PK--.E---    111
F1.BE.VI850       ---------IV---A-G----QTP................T--EG---V----DRR--------RT--PDL------.---.------R----V-----------------------Y-KK-G.-T--    111
F1.BR.93BR020-1   ---------IV---AI-----TPP..............TP---EG---V-Q---RR--------R-N-PDL------.--D.------R-----------G------------------Y-K----.E---    114
F1.FI.FIN9363     ---------IV---AI------PP..............PA--EG---V-Q---RR--------G---PDL------.-D-.------R-----------G-F---Q-XX--------XY-KK--.E---    112
F1.FR.MP411       ---------IV---A-------TP................---EG---V-Q--DRR--V--------PDL------.-D-.------R---------F--------------------Y-KK--.E---    111
F2.CM.HIV232985   ---------IV---N-------TP................V--EG---KV-Q--D----------R----DL------.-D-.------R-----------------------------Y-KK--.E---    111
F2.CM.MP257       ---------IV---AI---I--TP................V--EG---V-Q--D-R----N---G----DL------.-D-.------R-------------L----------------Y-RK--.E---    111
```

```
                                                                                heart of CD4 binding site         PKC
                                                                                ||   acidic           |-  -|
                                N-terminal alpha-helix                          ||   cluster  SH3 binding       polypurine tract  PAK binding
                myristoylation  MHC-1 downregulation + protein kinase recruitment ||   ||| |   |  |  |  |                |-     -|    ||
                 |-  -|         |--|                                             ||   ||| |   |  |  |  |                |-     -|    ||
B.FR.HXB2       MGGKWSKSSVIGWPTVRERMRRAE...............PAADRVGAASRDLEKHGAITSSNTAATNAACAWLEAQ.EEE.EVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQ.DILD    111
CONSENSUS G     --gk---ssiv-w-av---i-qtP??...........???pa-EG---v-Q--?r------s--atn-pdc------.e-dSe-----R--v-l----y-a-f---F--------D--iy-Kk---.d-l-    111
G.BE.DRCBL      --N----RK-A---E----L-QHP...............A--EG---V-Q--DR------R---G--PD-------.--DS------R-------------F---F--------D--VY-KK--.E---    112
G.FI.HH8793-1-1 ---R---R.....-A----I-QTPIR...........QPP---EG---V-Q--AR----------TN-PD-------.--DSD-----R----V--------F---F--------D---Y-KK--.----    114
G.ML.HIV232990  ---------IV---Q----I--TP...............A--EG-----Q--D----------P-N-P--------.Q-DS------R-------------L---F--------D-----KQ--.----    112
G.NG.92NG083    ---------IV---QI---I-QTP...............V--EG---V-Q--AR----------TN-PD-------.--DSD-----R-------------F---F--------D---Y-K---.----    112
G.NG.IKCSW22    ---------IV-R-A----I--TP................---EG---F-Q--GR----------G--PDS------.Q--S------R--L--------G-F---F--------D--VY-KQ--.--P-    112
G.NG.MACSW39    ---------IV---A----I-QTP...............-V-EG-----Q--AR----------T-----------.Q-DSD-----R-------------G-F---F--------D--VY-KK--.E---    112
G.SE.SE6165     ---------IV---E----I-NTP...............T--EG---V-Q--DR-----------N-PD-------.--DS------R---------F-G-F---F--------D---Y-KK--.E---    112
H.BE.VI991      -------GCIS---A----I-QT-................---EG---V-Q--DRR--V-IN-I-SN--DS------.---E------R-------------G-F----------------D---Y-KK--.E---    112
H.BE.VI997      ---------IV---A----I---Q................----G---V----DRR--V-IN------PDV------.--AE------R-------------L------------D---Y-KK--.E---    112
H.CD.HIV232994  --N----....-C-AI---I---QPAAE......RIRQAE---EG---V-Q--SRR--V-IN-M-SN-PDSP-----.---E------R-----------E-F-F-F-------------Y-KK--.E---    118
H.CD.HIV232995  ----L--F..S---AI------TRTAAE.........GEEK--EG---V----DRR--V-VN------PDS------.---E------R----V------G-L-F----------------Y-KK--.E---    117
H.CF.90CF056    --------RMG--S-I-------................-V-EG---V----DRR--V-IN---S--RDA------.-DGE------R-----------G-F------------D---Y-KQ--.----    112
J.SE.SE7022     --N-----....-Q--D-----A...............PA----G---V-Q--A--------------DD-------.T--.------R--X--------G-F---F--------D---Y-KK--.E---    109
J.SE.SE7887     --N-----....-Q--------P...............A----G---V-Q--A---------------D-------.T--.------K--I--------G-----F--------D---Y-KK--.E---    108
K.CD.EQTB11C    ---------IV--S-------KTP...............----G---V-Q--D----V------FN-PD-------.-D-.D-----R---------F-G-F--GF--------D---Y-K---.E---    111
N.CM.YBF106     --KI-----LV---EI---I--QT..............PE--V.G---V-Q--ANR----T---KDN-QTV------.--X.------R-------------F---F----------VW-RK--.E---    112
N.CM.YBF30      --KI-----LV---EI------QT...........QEPAVE--V.GA----Q--ANR----IR--RDN-ESI------.---E------R-------I---Q-F---F---D-------VW-RK--.----    117
O.CM.ANT70      --NALR-GKFE--AA-------TRTF..........PESE-C-PG--QI--E-AAR-G-P--H-PQN---L-F--SH.Q--.------A-----------G-F---F------------Y-HK-A.E---    117
O.CM.MVP5180    --NA----KFA--SE--D----SSSD...........PQQ-C-PG---V--E-ATR-G-S--H-PQN---L-F-DSH.KD-.D-----R---------F---F--------D---Y-HK-A.E---    116
O.SN.MP1299     --NVLG-DIFK--SA------GTSP............DPE-C-PG--QI--E-AAR-G-P--Y-PQN---L-F--SH.QD-.------R-----------G-F---F--------D---Y-HK-A.E---    115
O.SN.MP1300     --NVLG-DKFK--SA------KTSP............EPE-C-PG--QV--E-AAR-G-SN-H-PQN---L-F--SH.QD-.D-----R-----------G-F---F--------D---Y-PE-A.E---    115
K.CM.MP535      ---------IV---AI-------RPAAD.......RVGTQ----G---V-Q--AR---V-----SHN-PD-------.---.------R-------------F--GF--------D---Y-KK--.E---    120
CONSENSUS 01    --?--s-ssivg--Qv--ri?qTP???........?????p-aEG--av-Q--D----v----m..nnadcv--ea-.-e-.e-----r----l------G-F--sF--k-----D--iy-Kk--.E---    107
01_AE.CF.90CF11697 -------NRIV---Q----I--TP...............A--EG-----Q--D---------M..N--D-------.---.------R-----------G-F---F--------D---Y-KK--.E---    109
01_AE.CF.90CF402 --------CIV---Q----I-QTPVAE........ERQTPA--EG---V-Q--D----V----I..N--DNV-----.---.------R-----------G-F---F--------D-----K---.E---    117
01_AE.CF.90CF4071 --N-----....-QI---I-QTP...............V-TEG---V-Q--D----V----M..NS--SV-----.-D-.------R----T------G-F--GF--------D---Y-KK--.E---    105
01_AE.FR.HIV232982 --S------IV---Q---KIEQTP................--TEG---V-Q--D----V----M..N--DSV--R--.-D-.G-----R-----------G-F---F--------D---Y-K---.E---    109
01_AE.TH.93TH057 ---------IV---Q---KIKQTP................---EG--TV-Q--D----V----M..N-DD-V--G--.---.------R-----------G-F---F--------D---Y-KK--.E---    109
01_AE.TH.98-4   --S--P--....-Q----VKQTP................---EG---V-Q--D---------M..N--D-V--SS-.---.G-----R-----------G-F---F--R-----D--VY-KK--.E---    105
01_AE.TH.CM240  --S------IV---Q---KIKQTP................--TEG---V-Q--D---------I..D--D-V--R--.-D-.------M-----------G-F---F--------D---Y-KK--.E---    109
01_AE.TH.TH022  --N-----....-R----IKQTP................---EG---V-Q--D----V--T-M..D--DSV--R--.--DE------E-----------G-F---F--E-----D--VY-KK--.E---    106
CONSENSUS 02    ----w---siV---?vre-irqTp????.....???????p--tG---a-Q--Drh-----s--aa--ad-------?-ee?e-----r---------y-a-v--sh----k---?--iY-KK--.E---    109
02_AG.CM.97CMMP807 ---------LV---Q--D---QTP...............---EG-----Q--DR--------------D-------.--G.------K---------F---F------------D---Y-KK--.E---    111
02_AG.FR.DJ263  ---------IV---K----I-QTP................---TG-----Q--DR----------D---D-------.---.------R-------------F--GF--------D--VY-KK--.E---    111
02_AG.FR.DJ264  ---------LV---K----IIQTP...............---TG-----Q--DR----------E---D-------.-D-.------K-----------G-F--GF--------D---Y-KK--.E---    111
02_AG.GH.G829   ---------IV---Q----I-QTP...............T--KG-----Q--D-Y---------F---D-------.---.------R-----------G-L----------------Y-RK--.E---    111
02_AG.NG.IBNG   ---------IV---K-MK---QTP...............T--TG-----Q--DR----------Q--PD-------.-D-.N-----R-----------G--------------D---Y-KK--.E---    111
02_AG.SE.SE7812 ---------IV---QI-D-I-QTP...............---RG-----Q--DR--------------------.---.------R-----------------------Y-KK--.E---    111
02_AG.SN.MP1211 ---------LV---Q----I--TQ.............PTPS--IG---V-Q--DR------R---H---D-------.---.------R----------------------Q-------Y-KK--.E---    114
02_AG.SN.98SEMP1213 ----C----.IV---K----I-QTPVRQS.....PPAARGAG--TG-----Q--DR---------S----D-------E---E------R-----------------------Y-KK--.E---    123
03_AB.RU.KAL153-2 ---------IV---Q----I---P...............A---RG--PV-Q--D-Y--V-------N--D-------.---.------R-----------G-F-------X----D---Y-KK--.E---    112
03_AB.RU.RU98001 ---------IV---QI---I---PA................---RG--PV-Q--D-Y--V-------N--D-------.K--.------R-----------G-F------------D---Y-KK--.E---    112
04_cpx.CY.94CY032-3 ---------IV---EI-------RAEPER....MRRAQAE---AG---V-Q--D------IN------PDKT------.---E------R---------F-G-L------------D---Y-KK--.E---    124
04_cpx.GR.97PVCH ---------LV---AI-------RAEP........AAQAE---AG---V----D--------------PDK------.--DE------R---------F---L------------D---Y-KQ--.----    120
04_cpx.GR.97PVMY --N-----....-A-------RAEP.........ARAE---VG---V-Q--D-Y----T-------PDK------.---E------R-----------------------Y-KK--.E---    115
05_DF.BE.VI1310 ----------V---AI--K---TP...............---EG---VFQ--DNR--V-IR--V---PDM--V---.---E------K-----------------------Y-KK--.----    112
05_DF.BE.VI961  -------NRIV---AI------TPP............AAGA--EG--SV-Q---RR-----N-------DL--V---.---.------R---------F-G-L------------D---W-RK--.K---    115
06_cpx.AU.BFP90 --N-----....-SQ-------TPP............TER--EG---V-Q--D----------T-------V---.T--.------R----V--------F---F--------D---Y-KK--.E---    110
06_cpx.ML.95ML127 --S------IV---Q----I-QTP............PTEG--KG---V-Q--D-------N--PT---------E-.T-D.------R---------F-G-F--GF--------D---Y-KK--.E---    114
06_cpx.ML.95ML84 ---------LV---Q----I-QTPP............TEG--EG---V-Q--DT---------VT-----------.T-D.------R---------F-G-F---F--------D---Y-KK--.----    114
06_cpx.SN.97SE1078 ---------IV---K------QTPPAA........ERQTP---EG-----Q--A---------------D-------S-DNE------R-------------F--GF--------D---Y-KK--.E---    121
10_CD.BFL061    ---------IV---Q-------TS...............---EG-----Q--A-F--L-----PGN--D-------.-A-E------R---------F-G-F--GF--------D---Y--K-K.----    112
10_CD.BFL071    ---------IV---AI---I--TDP...........RRTE---EG----------Y---------Q--PD-------.A-GD------R-------------F---F----------------K---.----    117
10_CD.BFL110    ---------FC---A----T--TD...............---EG--------------------IQ--PD----A--.--DG------R-----------G-L----------------Y-K---.----    112
11_cpx.CM.MP818 ---------IV---EI---L--TPPA.............A---G---V-K----F---------H--DT-------.-D-.------R-------------F--GF--------D---Y--K-K.E---    113
11_cpx.FR.MP1298 --------GIV---EI---L--TRPV..............---EG---V-K-------V--T-Q..N---ST-----.--AE------R-------------F--GF--------D---Y-KK--.E---    112
11_cpx.FR.MP1307 ---------IV---EI---I--TRPT.............A--EG-----K----F--L-T----Q--E--------.---.------K-----------G-F---F--R-----D---Y-KK--.E---    113
11_cpx.GR.GR17  ---N-----IV---EI---L--TPPT.............A--EG-----K---R---------SQN----------.-D-.G-----R-----------G-F---F--------D---Y-KK--.E---    113
CPZ.CD.CPZANT   --SA---IKWV-ARQAIRKIHETN................P-DI-PCGNE-ASR--L---TIGTEKDVITYS-DH..T-EGT----R----M----E-L-----X--------DX--XXLK-A.A---    109
CPZ.CM.CAM3     --NR-----LV---D----LSELN...............LL-PG--PV-Q---A---V-TRH-PQN-QTL----EMQ-HNE------R---------F---F-------K--------Y-IXXX.XXX-    113
CPZ.CM.CAM5     --N------LV--AA--Q-I-QTP...............---SG--PV-Q---AR----TR-HTKY.QTL----EMENH-E------R-----------Q-F---------------VY-TK--.E---    112
CPZ.GA.CPZGAB   --T------LV---E--R-I-E-P...............T--EG--EV-K---R------R--PE--QTL----EM.DN-.------R----T--------F---------------VY-R---.E---    111
CPZ.US.CPZUS    --N------IV---E--N-L-QTQ.............TTA--EG--PV-Q--AE-----TR--PQN-QTL---DEMTNH-S------R-----------Q-F--GF-----------VY-R---.E---    115
```

HIV-1/SIVcpz
proteins

HIV-1/SIVcpz proteins

```
                                                   COP recruitment
                                                              ||        di-leucine based AB recruitment
                                beta turn      SH3 binding ||          ||         V-ATPase and RAF - 1 binding
                                ......                  |  | ||          ||         ||
B.FR.HXB2          LWIYHTQGYF.PD QNYTPGPG.VRYPLTFGWCYKLVPVEPDKIEEA.NKGENTSLLHPVSLHGMDDPE.REVLEWRFDSRLAFHHVARELHPEYFK......NC  205
CONSENSUS A        --Vyht-gy-.--Wq-------.i-y-------f-----dpdEV-ka.te---Ns----icQ--m--E-.r-t-m-K--sr--lk-r-r-l---fy-......dc  203
A.FR.HIV232956     --V-N-----.--W--------.T---------F-----D-VEV-K-.-E---NC----M-Q-----E-.--T-I-K------LR-I-Q-M---FY-......D-  204
A.KE.Q23           --V-------.--W--------.T-F-------F-----D--EV-K-.TE---N-----ICQ-----E-.----K-K------LK-R-------WY-......D-  204
A.SE.SE6594        --VH-I--F-.--W--------.I-F-------F-----D--EV-KD.TE---N-----ICQ-----E-.K---K-K------LK-L-C-K---FY-......D-  205
A.SE.SE8538        --V-N--X--.--W--------.----------F-------EEV-K-.-E---N-----ICQ-----E-.--T-M-K---K--LK-R-H--------.......--  204
A.SE.SE8891        --V-N-----.--W--------.T-F-------F-----D--EV-K-.TE---N-----ICQ-----E-.--T-M-K------LT-R-------FY-......D-  204
A.SE.UGSE8131      --V-------.--WH-------.I---------F-----D--EV-K-.TE---N-----MCQ-----E-.--T-M-K--PH---K-R-F------Y-......-.  209
A.UG.92UG037       --V-------.--W--------.I---------F-----DE-EV---.TG---N-----ICQ-----E-.K-T-R-K---S--RV-K-------FY-......D-  215
A.UG.U455          --V-----F-.--W--------.I---------------D-AEV---.TG---N-----ICQ--V--E-.K---M-K---T--LK-R-Y-----FY-......D.  205
A2.CD.97CDKTB48    --V-N-----.--W--------.A-F-------F-----D-SEV---.TE---N-----ICQ--AE---.----K-K------LR-L---Q---FY-......D-  216
A2.CY.94CY017-41   M-V-------.--W--------.I---------F-------SEV---.TQ---N-----ICQ--V----.----R-E--RS--RR-R--------Y-......D-  217
CONSENSUS B        --vyh---yf.--Wqn--p---.i-y-ltf---F--v-vepekveea.negenns----msl-gmd-p-.k-v-v-kfd-r-afhhm-rel---yy-......dc  204
B.AU.MBC200        --V-D-----.--W--------.T-W-------F-------E-----.-A---NC----M-Q-------.--G----------------------Y-......-.  213
B.CN.RL42          --V-------.--W--------.T-F-------F-----D-EQV---.-E---NC----M-Q--A----.----M-K------I--M---M----H-......D-  206
B.DE.D31           --V-------.--W--------.T-F-------F-----K-EQV---.-E---N-----M---------.----V---------K-M--------Y-......-.  205
B.DE.HAN           --V-------.--W--------.----------F--------E.......E--S-----A----TE-T-.----K-K---H-----K--------Y-......D-  200
B.ES.89SP061       --VH------.--W--------.I---------F-------E-V---.-E---NR----------E---.K---V-------VLR-M--------Y-......D-  206
B.FR.NE100         --V-------.--W--------.I-W---L---F-------EQV---.-E---N-----M-Q---E---.K---V-K---S-----K--------Y-......D-  206
B.FR.SWB884        --V-------.--W--------.I-W---P---F-------EQV---.-E---N-----M-----E---.K---V-K---S-----K--------Y-......D-  206
B.GA.OYI           --V-------.--W--------.I----C----F----MD--QV---.-E---N-----I---------.K---V-K-------R-M---V----Y-......D-  211
B.GB.CAM1          ----------.--W--------.I---------F-------EQV---.--R--A-----M-Q-------.K---M-K---------M---K---FY-......D-  210
B.IN.HIVP35A       --V-------.--W--------.I---------F-----D-E-V-KY.-E---NC----M-Q-------.K---Q-K-----------------FY-......D-  215
B.IT.B-IT-R5       ----------.--W--------.T---------F--E----E-V---.-E---N-----------E---.K---V----------SC-CVF----Y-......D-  206
B.JP.nef<7>-a      --V-------.--W-K------.I-F-------F-------E-----.-E---NR----MA--------.----V-K---------M---------.-------  209
B.KR.CSR9412d      --V------L.--W-G------.I---I-----F------KE-V---.-E---N-----M----IE---.K-------E----------D-----Y-.-------  203
B.KR.WK            --V-----F-.--W--------.T-F-------F-------E-V---.TV-K-NC----MN--------.G---V-----------M---K----Y-......D-  206
B.NL.3202A21       ----------.--W--------.I---------F------QE-----.-E---N-----M-Q-------.----------------M-------Y-......D-  206
B.SE.AF047085      --V-------.--W--------.T-F-------F-------E-----.-E---NC----M-Q---E---.----M-K---------I--------Y-.-------  206
B.TH.28-19         --V-------.--W--------.----------F-------ELV---.-E---N-----M-Q-------.----M-K---------M--------Y-......D-  220
B.TH.AF082839      --V-------.--WH-------.I---------F-------E-----.-E---NC----I----I----.----Q-K------------------Y-.-------  205
B.TW.TWCYS         --------F-.--W--------.----------F-------EQV-K-.-E---XC----M-Q-------.K---V-----T-----R--------Y-......X-  208
B.US.JRCSF         ----------.--W----A---.--F-------F-----D-E-V---.-E---NC----M-Q-------.K---V-K---K--L----------Y-......D-  216
B.US.JRFL          --V-------.--W--------.I-F-------F-------E-V---.-E---NC----M-Q--IE---.K---------K-------------Y-......D-  216
B.US.LM1           --V-------.--W--------.T---------F----I--EEV-K-.-E---NC----M-----E---.K---V-K-------N-M--------Y-.-------  206
B.US.MN            --V-------.--W--------.I---------F-------E-----.-----NC----M-Q-X-----.----V-KS--H---Q-Y--------Y-.-------  208
B.US.NC7           ----------.--W--------.I---------F---------V---.-E--DNI----MC----E-A-.K---V-----K-------------Y-......D-  206
B.US.RF            --V-------.--W--------.T---------F---------V---.TE---N-----IC--------.K---V-K-------------K----Y-......D-  208
B.US.SF2           ----------.--W--------.I---------F-------E-V---.-E---N-----M-----E-A-.K---V-----K-----M--------Y-......D-  210
CONSENSUS C        --V-h---y-.--Wq-------.vry-l----cfk----D-reV-ea.ne--dNc----vcq--me-eh.r--lk-kfdsq--?r-marelh--?yk.....?d-  205
C.BR.92BR025       --V-N-----.--W--------.--F-------F-----D-REV---.-T---N-----M-----E-SH.----Q-K---L--RR-M--------Y-......D-  206
C.BW.96BW01B21     --V-N---V-.--W--------.----------F-----D-REV---.SE--DNC----M-Q---E-EH.----K-K---Q--RR-M--------Y-......D-  207
C.BW.96BW0402      --V-N-----.--W--------.----------------D-KEV---.-E---NR----M-----E-AD.G---R-K---H--HR-M--------Y-......D-  207
C.BW.96BW0502      --V-----F-.--W--------.-K-------F-----D-GEV---.----DNC----M-Q---E-EH.G---K-K---Q--RR-M----Y---Y-.......  213
C.BW.96BW1104      --V-----F-.--WH-------.----------F-----D-GEV---.-E--XNC----IAQ---E-EH.K---K-K---Q--RR-L-------FY-......D-  213
C.ET.ETH2220       --V-N---F-.--W--------.----------F-----D-SEV--I.-E---NC----A-----E-ED.----K-K---H--RR-M--------Y-......D-  207
C.FR.HIV232980     --V-----F-.--W--------.----------------D-REV--N.-E---NC----M----IE-AH.K---Q-K---L--HR-M--------Y-......D-  213
C.FR.HIV232996     --V-N-----.--W--------.--H-I----LF-----D-KEV--N.TE---NC----M-----E-EH.G---M-K---S--HR----------Y-......D-  209
C.IN.93IN101       --V-----F-.--W--------.----------F-----D-REV---.-E--DNC-----CQ---E-EH.----K-K---Q--RR-M-------FY-......D-  207
C.IN.93IN904       --V-------.--W--------.----------F-----D-REV---.-E--DNC-----CQ-----EH.----Q-K---L--HR-R-------FY-......D-  207
C.IN.93IN999       --V-----F-.--W--------.--F-------F-----D-REV---.-E---NC-----CQ---E-EH.----Q-K---H--HR-M--------Y-......D-  207
C.IN.94IN11246     --V-------.--W--------.T-F-------F-----D-REV---.-E--DNC-----CQ---E-DH.----M-K..XQ--HR-I-------FY-......D-  205
C.IN.95IN21068     --V-------.--W--------.--F-------F-----D-REV---.----DNC-----CQ---E-DH.N---V-K---Q--HK-R-------FYN.....KD-  200
C.IN.HIVY15117     --V-------.--W--------.--F-------F-----D-KAV---.-E--DNC-----CQ---E-EH.---FM-K---Q--RR-MG-G----FY-......D-  199
C.IN.HIVY17884     --V-------.--W--------.T-F-------F-----D-REV---.-E--DN------CQ---E-EH.----K-K---Q--RR-W-------FY-......D-  211
C.IN.HIVY17891     --V-------.--W--------.----------F-----D-REV---.-E---NC-----CQ---E-EH.----Q-K---Q--HR-R-------FY-......D-  207
C.IN.HIVY17892     --V-------.--W--------.T-F-------F-----D-REV---.-E--DNC-----CQ---E-EH.----R-K---Q--LR-T-------FY-......D-  207
D.CD.84ZR085       --V-----F-.--W--------.I---------FE----D-EVV-K-.TE--DNC----ICQ---E-E-.K---V---N-----E-K-K-KY----------  215
D.CD.ELI           --V-N---I-.--W--------.I----------E----D-QEV--D.TE--TN-----ICQ---E---.-Q--K---N-----E-K---M---FY-------.  206
D.CD.NDK           --V-N---I-.--W--------.I---------FQ----D-QEV---.TER-DNC----MCQQ--E---.-Q--M---N----LE-K-------FY-......D-  207
D.UG.94UG1141      --V-----F-.--W--------.I---------FE---M--KEV--N.TE--DNC----INQ---E---.----V---N-----E-K-KMK----Y-......D-  206
F1.BE.VI850        --V-------.--W--------.I-----L---F-----D-EEV-K-.-E---NC----M-Q---E-ED.----R-K---S--LR-I---R---FYQ......D.  205
F1.BR.93BR020-1    --V-------.--W--------.I-----M---F-----D-EEV-K-.-E---NC----M-Q---E-ED.K---K-E------LR-I---R----YQ......D.  208
F1.FI.FIN9363      --L-------X--W--------X----------F-------EEV-K-.-E---NC----M-Q---E-ED.----K-K------LK-I---R---FYR......D.  208
F1.FR.MP411        --V-----F-.--W--------.I-F-------F-----D--EV-K-.-E---NC----M-Q-----ED.----R-E-------R-I---K---FYQ......-  205
F2.CM.HIV232985    --V-------.--W--------.I-----L---F-----D-EE--K-.-E---NC----I-----E-ED.----R-K---S--LR-----R---FYQ......D  205
F2.CM.MP257        --V-------.--W--------.P-F-------F-----D-EEV-K-.-E---NC----M-----E-DD.K---K-Q------LR-I---R----Y-......D  205
```

```
                                                   COP recruitment
                                                              ||        di-leucine based AB recruitment
                                   beta turn      SH3 binding ||          ||        V-ATPase and RAF - 1 binding
                                     ......                  |  |  ||      ||        ||
B.FR.HXB2           LWIYHTQGYF.PD QNYTPGPG.VRYPLTFGWCYKLVPVEPDKIEEA.NKGENTSLLHPVSLHGMDDPE.REVLEWRFDSRLAFHHVARELHPEYFK......NC   205
CONSENSUS G         --V-n---f-.-dW--------.t-f-------F--v-md-aEv---.nkg--n-----iCQ-gmE-eD?r---v--fd-S--Rr-l-r---p-yYk......d-   206
G.BE.DRCBL          --V-----F-.--W--------.T-V-------F----M--SEV---.-----N-----ICQ---E-ED.G---V-----S--RR-L--------Y-......D-   207
G.FI.HH8793-1-1     --V-N---F-.--W--------.T-F-------F--E--D-AEV---.ITE--I-----ICQ-E-E-EDX----R--X--S--RR-L-Q-----FY-........--   210
G.ML.HIV232990      --V-N---F-.--W--------.T-F-------F----MD-AEV---.-N---N-----ICQ---E-TD.----V-----S--RR-I--------Y-......D-   207
G.NG.92NG083        --V-N---F-.--W--------.T-L-------F----MD-AE----.-----I-----ICQ---E-ED.----V---N-S--RR-L--------Y-......D-   207
G.NG.IKCSW22        --V-N------.-EW--------.I-F-------F----LD-TE----.-----N------CQ--LE-ED.----V-----S--RK-I-----L--YT......D-   207
G.NG.MACSW39        --V-N---F-.--W--------.I-F-------F----MD-AE----.-----N-----ICQ---E-ED.----V-----S--RR-L--------Y-......D-   207
G.SE.SE6165         --V-N------.--W--------.T-F-------F----MD-AEV---.-----N-----ICQ---E-ED.----V-----S--RR-I--------Y-......D-   207
H.BE.VI991          --V-N------.--WH-------.E---------F-----D-QDV-K-.-E---N-----MCQ--IE---.----M-K------LR-R-K-----FY-......D-   207
H.BE.VI997          --V-N------.--W--------.EG--------F--I--D-QEV-R-.-E---NC--Y-ICQ---E-E-.G---M-K-------T-T---K---FY-......D-   207
H.CD.HIV232994      --V-N---F-.--W--------.----------F-----D-QEV-Q-.-E---N-----ICQ---E-ED.----K-T------RR-M-------FY-......D-   213
H.CD.HIV232995      F-V-N------.--W-K------.----------F-----D-LEV-R-.-E---N-----ICQ---E-ED.----M-K---H---R-R-------FY-......D-   212
H.CF.90CF056        --V-N------.--W--------.E-F-------F-----N-QEV-Q-.-E---N-----M-----E-DG.----M-K------LT-L--VK----.-......D-   206
J.SE.SE7022         --VHN------.--W--------.TX--------F-----D-SEV---.-E---NC----ACQ--IE-E-.----K-K---S--RR-I-------FY-......D-   204
J.SE.SE7887         --VHN------.--W--------.I---------------D-SEV---.-E---NC----ICQ--IE-E-.----Q-K---S--RR-I-------FY-......D-   203
K.CD.EQTB11C        --V-----F-.--W--------.I---------------D-REV---.TE---NC-----NQ---E-EH.----K-K---S--RK-----M----Y-......D   205
N.CM.YBF106         --V-----F-.--W--------.----XC----F----LSEEAV---.-E-D-NA----ICQ--V--DH.KQ--V-----S--RR---K----DFY-........--   207
N.CM.YBF30          --M-----IL.--WH-------.I---V-----F----LSAEEV---.-E-D-NA----ICQ--A--DH.K---V-----S--RR---------FY-........--   212
O.CM.ANT70          --V-N---F-.--W--------.T-F------LF-----SEEEA-RLG-TC-RAN----ACA--FE-TH.K-I-M-K--RS-GNT---MIT---L-Q......KD   213
O.CM.MVP5180        --------F-.--W-C------.P-F------LF-----SAEEA-RLG-TN-DA-----ACN--AE-AH.G-I-K-Q--RS-GLT-I-LQK---L-P.....XSN   213
O.SN.MP1299         --V-----F-.--W-G------.P-F------LF-----SEAEA--LG--C-RA-----ACN--FE-NH.GQI-K-Q--RS-GST---MVTN--L-N......KD   211
O.SN.MP1300         --V-----F-.--W--------.T-F------LF-----SEAEA--LG--CDRAK-----CN--FE--H.K-M-K-Q--RS-GST---LIT---L-L......KD   211
K.CM.MP535          --V-----F-.--W--------.I---------------D-AEV--T.TE--DNC----INQ---E-EH.--I-M-K---S--RR--------D-Y-......D   214
CONSENSUS 01        --Vyn---f-.--Wq-------.i-y--C----f-----DprEV--D.--g-nNC----m-Q--ie-e-.----M-K---a--Rr-i---lh-e-Y-......Dc   202
01_AE.CF.90CF11697  --V--------.--W--------.I---C----------D-KEV--D.--D-SNC----M-Q--V--E-.----M-K---S--RR-I----R---Y-......D-   204
01_AE.CF.90CF402    --VHN------.--W--------.-----C----F-----D-REV--D.-----NC----M-Q--I--D-.----M-K---S--RR-I--------Y-......D.   211
01_AE.CF.90CF4071   --V-N---F-.--W--------.T----C----------D-REV--D.-----NC----M-Q--IE-E-.----M-K---S--RR-I--------Y-......D-   200
01_AE.FR.HIV232982  --V-N---F-.--W--------.I----C----F-----DTGEV--D.-----NC----M-Q--LE-E-.----M-K---A--RK-I------D-Y-......D-   204
01_AE.TH.93TH057    --V-N---F-.--WH-------.I----C----F-----D-REV--D.-----NC----L-Q--IE-E-.----M-K---A--RK-T--------Y-......D-   204
01_AE.TH.98-4       --V-N---F-.--WH-------.I----C----F-----D-GEV--D.-----NC----M-Q--VE-E-.----M-K---A--RR-I----R---Y-......D-   200
01_AE.TH.CM240      --V-N---F-.--W--------.I-F--C----F-----DQREV--D.-----NC----M-Q--IE-E-.----M-K---A--RK-----Q----Y-......D-   204
01_AE.TH.TH022      --V-N---F-.--W--------.--L--C----F-----D-REV--D.-----N-----M-Q--IE-E-.----I-K---S--RK-L--------Y-......D-   201
CONSENSUS 02        --V-h---F-.--Wq-------.t-?-------fk-v-mD-aev-ea.nkg--Ns----ICQ--mE-eD.r---v-----r--ft-?-r-l---fY-......Dc   202
02_AG.CM.97CMMP807  --V-----F-.--W--------.T--------------MD-QEV---.--E--NC----ICQ---E-ED.----V-----H--LK-R---I---FY-......D-   206
02_AG.FR.DJ263      --V-----F-.--W--------.T---------FN-E-ID-AE----.--E--N-----ICQ---E-ED.----V-----S--RT-R--------Y-......D-   206
02_AG.FR.DJ264      --V-----F-.--W--------.T---------F--E-MD-AEV---.-----N-----ICQ---E-ED.----V-----S--RR-I---R---FY-......D-   206
02_AG.GH.G829       --V-----F-.--WH-------.T-F-------F----MD-KAV---.TE---N-----ICQ--IE-ED.----V-----S---T-R---M---FY-......D-   206
02_AG.NG.IBNG       --V-N---F-.--W--------.T-F-------F----MD-AEV---.-E---N-----ICQ---E-DD.----I---------R-T--------Y-......D-   206
02_AG.SE.SE7812     --V-----F-.--W--------.T-F-------F----MD-AD--KD.TE---N-----ICQ---E-ED.----V---------T-K---M---FY-......D-   206
02_AG.SN.MP1211     --V-----F-.--W--------.I---------------MD-AEV---.-Q---N-----ICQ---E-ED.K---V---------R-T--------Y-......D-   209
02_AG.SN.98SEMP1213 --V-----F-.--W--------.T-F-------F----MD-AD----.--E--N-----ICQ---E-ED.----V---------T-T-K-----FY-......D.   217
03_AB.RU.KAL153-2   --V--------.--W-.......-.I-F-------------D-AEV---.TE---N-----ICQ-----E-.K---M-K------LT-R-------FY-......D-   202
03_AB.RU.RU98001    --V--------.--W--------.I-F-------------D--EV---.TE---N-----ICQ-----E-.K---M-K------LT-R-------FY-......D-   207
04_cpx.CY.94CY032-3 --V-----F-.--WD-------.E-F--C----F-----D-QEV---.TE----C----I-Q---E-E-.----K-K------YK---------FY-......D-   219
04_cpx.GR.97PVCH    --V-N------.--W--------.E-F--C----F-----D-QEV---.-E---NC----I-Q---E-E-.----K-K-------K-I-------FY-......D-   215
04_cpx.GR.97PVMY    --V-N------.--W--------.T-F--C----F-----D-QEV--I.TA--DNC----I-Q---E-E-.----K-K---L--YR-M-------FY-......D-   210
05_DF.BE.VI1310     --V-N----S.--W--------.I-----L---F-----N-EEV-K-.-E--DNC----M-----E-DD.----Q-K---S--LR-I---R---FYQ......D.   206
05_DF.BE.VI961      --V-N------.--W--------.T---------F-----D-EEV-K-.-E---KC----MHQ-----ED.G---R-K---S--LK-I---RR--FYQ......D.   209
06_cpx.AU.BFP90     --V-N---F-.--W--------.T-F-------------D-EEV--L.T----NC----ICQ--AE-E-.----K-K---S--RR-I---K---FY-......D-   205
06_cpx.ML.95ML127   --V--------.--W--------.I---------------D-REV--E.T----N-----MCQ--AE-E-.G---M-K---S--RR-I------DFY-......D-   209
06_cpx.ML.95ML84    --V-----F-.--W--------.I---------------D-REV--D.T----NC----M-Q---E-E-.----M-K---S--RR-T---M---FY-......D-   209
06_cpx.SN.97SE1078  --V-----F-.--W--------.I---------F-----D-KEV--E.T----X-----MCQ--V-----.X---M-K---S--RR-I-X-X---FX-......D-   216
10_CD.BFL061        --V-N---F-.--W--------.T---------F-----D-REV---.-----NC----M-----E-KH.G---V-K---T--HK-I--------Y-......D-   207
10_CD.BFL071        --V-N---F-.--W--------.I---------------D-KEV---.-E---N---D-M-----E-AH.G---I-Q--IK--HT-M--------Y-........--   212
10_CD.BFL110        --V-N---F-.--W--------.I---------------D-REV---.-E---N-----M-----E--D.----K-K---Q---K-M--------Y-......D-   207
11_cpx.CM.MP818     --V-----F-.--W-E------.-----C----------D-REV---.-E---NC----M-Q-----E-.----K-V---S--RK-I------DFY-......D.   207
11_cpx.FR.MP1298    --V-----F-.--W--------.T----C----F-----D-REV---.-E---N-----ICQ-----E-.G---M-K---S--RK-I------DFY-......D-   207
11_cpx.FR.MP1307    --VH-------.--W--------.--F--C----F-----D-KEV---.-E---NC----I-Q-----E-.G---M-K---S--RR-I-------WY-......D.   207
11_cpx.GR.GR17      --VHN---F-.--W--------.I----C----------D-REV---.-E---NC----M-Q--IE-ED.----R-K---S--RR-I------DFY-......D.   207
CPZ.CD.CPZANT       M-MFN---I-.--W----EE--.------CR---F------SPPDD-R......NI----ACT--DG--H.K-I-R-E--AS-MRR-I---R-----R......D.   198
CPZ.CM.CAM3         HGY-----F-.--W--------.------Y-G-F----LT-EEV-Q-.-E-D-NI----ICQ---E-AH.G---V--------LR-I---Q----YQ.......-.   207
CPZ.CM.CAM5         --V-----I-.--W--------.----------F----LT-EEV-K-.-E-DTNI----ICQ---E-EH.G---V--------LK-L---Q----YGDHNPAPAQ   213
CPZ.GA.CPZGAB       --V-----F-.--W----T---.T-F--C----F----LTEEQV-Q-.-E-D-NC----ICQ---E-ED.K---V--------LR-I---Q----Y-......D   205
CPZ.US.CPZUS        --V-----I-.--W--------.------Y---F----LTEEEV-Q-.----TNI----MCQ---E-EH.G---I-Q--TE--RR-R-K-------R.......N   209
```

HIV-1/SIVcpz proteins