V

HIV-1/SIVcpz Proteins

Contents

V-1   Introduction

The HIV-1/SIVcpz protein alignments are based on the complete genome nucleotide alignment, but in some cases a few sequences were removed because they were too short to be informative (especially in Nef), had many stop codons, or a problematic segment of amino acids. As with the other alignments in this compendium, they are intended to display the genetic variation of the world-wide HIV epidemic in a compact form. They are annotated in more detail than the complete genome nucleotide alignment.

HIV-1 Proteins

V-2   Annotated features

Features of HIV-1 annotated in the alignment that follows.

Feature	Protein	Location	Page
Gag start, p17 start	Gag	1	318
membrane binding	Gag	1-30	318
phosphorylation site	Gag	112	318
p17 end	Gag	132	318
p24 start	Gag	133	318
CyPA binding	Gag	205-241	320
major homology region	Gag	285-304	320
p24 end	Gag	364	322
p2 start	Gag	365	322
p2 end	Gag	377	322
p7 start	Gag	378	322
Zn motif	Gag	392-404	322
Zn motif	Gag	413-425	322
p7 end	Gag	432	322
p1 start	Gag	433	322
p1 end	Gag	448	322
p6 start	Gag	449	322
Vpr binding	Gag	455-460	324
Vpr binding	Gag	489-494	324
p6 end, Gag end	Gag	501	324
Pol start	Pol	1	326
Gag-Pol TF start	Pol	1	326
Gag-Pol TF end	Pol	56	326
protease start	Pol	57	326
protease end	Pol	155	328
p66, p51 RT start	Pol	156	328
M41L	Pol	196	328
D67N	Pol	222	328
K70R	Pol	225	328
D110 catalytic site	Pol	265	328
polymerase motif	Pol	337-342	330
M184V	Pol	339	330
T215Y	Pol	370	330
K219Q	Pol	374	330
p51 RT end	Pol	595	332
p15 RNase H start	Pol	596	332
p66 RT, p15 Rnase H end	Pol	715	334
p31 Integrase start	Pol	716	334
p31 Integrase end	Pol	1004	338
Pol end	Pol	1004	338
Vif start	Vif	1	340
Vif end	Vif	193	342
Vpr start	Vpr	1	344
oligomerization	Vpr	1-41	344
amphipathic α-helix	Vpr	17-33	344
H(S/N)RIG motifs	Vpr	71-83	344
frameshift in HXB2	Vpr	72	344
Vpr end in HXB2	Vpr	79	344

HIV-1 Proteins

Feature	Protein	Location	Page
Vpr end	Vpr	98	344
Tat start	Tat	1	346
C-rich region	Tat	22-37	346
nuclear localization	Tat	49-57	346
exon 1 end	Tat	72	346
exon 2 start	Tat	73	346
Tat end	Tat	102	346
Rev start	Rev	1	348
exon 1 end	Rev	25	348
exon 2 start	Rev	26	348
NLS	Rev	34-49	348
Leu-rich effector domain	Rev	75-83	348
Rev end	Rev	117	348
Vpu start	Vpu	1	350
transmembrane domain	Vpu	1-27	350
cytoplasmic domain	Vpu	28-82	350
α-helix	Vpu	43-51	350
phos	Vpu	53	350
phos	Vpu	57	350
α-helix	Vpu	58-70	350
Vpu end	Vpu	83	350
Env start	Env	1	352
signal peptide end	Env	30	352
gp120 start	Env	31	352
glycosylation NVT	Env	88-90	352
CD4 binding	Env	124	352
V1	Env	131-156	352
glycosylation NDT	Env	136-138	352
glycosylation NSS	Env	141-143	352
glycosylation NCS	Env	156-158	354
V2	Env	158-196	354
glycosylation NIS	Env	160-162	354
glycosylation NDT	Env	186-188	354
glycosylation NTS	Env	197-199	354
CD4 binding	Env	196	354
glycosylation NKT	Env	230-232	354
glycosylation NGT	Env	234-236	354
glycosylation NVS	Env	241-243	354
glycosylation NGS	Env	262-264	354
glycosylation NFT	Env	276-278	354
CD4 binding	Env	279	354
glycosylation NTS	Env	289-291	356
glycosylation NCT	Env	295-297	356
V3	Env	296-331	356
glycosylation NNT	Env	301-303	356
V3 tip	Env	312-315	356
glycosylation NNT	Env	339-341	356
glycosylation NKT	Env	356-358	356
CD4 binding	Env	365	356
V4	Env	385-418	356
glycosylation NST	Env	386-388	356
glycosylation NST	Env	392-394	356
glycosylation NST	Env	397-399	356

HIV-1 Proteins

Feature	Protein	Location	Page
glycosylation NNT	Env	406-408	358
CD4 binding	Env	425	358
glycosylation NIT	Env	448-450	358
CD4 binding	Env	455	358
V5	Env	460-471	358
glycosylation NES	Env	463-465	358
CD4 binding	Env	469	358
fusion peptide	Env	512-527	358
gp120 end	Env	511	358
gp41 start	Env	512	358
immunodominant region	Env	588-607	360
glycosylation NAS	Env	611-613	360
glycosylation NKS	Env	616-618	360
glycosylation NHT	Env	624-626	360
glycosylation NYT	Env	637-639	360
transmembrane domain	Env	685-704	360
gp41 cytoplasmic tail start	Env	705	360
glycosylation NGS	Env	750-752	362
glycosylation NAT	Env	816-818	362
cytoplasmic tail end	Env	857	362
gp41 end	Env	857	362
Env end	Env	857	362
Nef start	Nef	1	364
myristoylation	Nef	2-7	364
acidic cluster	Nef	62-65	364
poly-P helix	Nef	69-78	364
phosphorylation	Nef	77-81	364
HXB2 premature Nef end	Nef	124	364
normal Nef end	Nef	207	366

V-3 Sequences

Sequences included in the HIV-1 protein alignments.

Name	Accession	Proteins	Author	Reference
B.FR.83.HXB2	K03455	All	Wong-Staal, F.	*Nature* **313**(6000):277-284 (1985)
A1.CD.02.LA01AlPr	KU168256	All	Berg, M.G.	*J Clin Microbiol* **54**(4); 868-82 (2016)
A1.CM.08.886_24	KP718928	All	Luk, K.-C.	*PLoS One* **10**(11); e0141723 (2015)
A1.ES.15.100_117	KY496622	All	Cuevas, M.T.	Unpublished
A1.KE.11.DEMA111KE002	KF716474	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
A1.PK.15.PK034	KX232624	All	Chen, Y.	*PLoS ONE* **11**(12):E0167839 (2016)
A1.RW.11.DEMA111RW002	KF716472	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19

HIV-1 Proteins

Name	Accession	Proteins	Author	Reference
A1.SE.13.078SE	MF373181	All	Neogi, U.	*Sci Rep* **7**(1):6371 (2017)
A1.UG.11.DEMA110UG009	KF716486	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
A2.CM.01.01CM_1445MV	GU201516	All	Carr, J.K.	*Retrovirology* **2010** Apr 28;7:39 doi: 101186/1742-4690-7-39
A3.SN.01.DDI579	AY521629	All	Meloni, S.T.	*J Virol* **78**(22):12438-12445 (2004)
A4.CD.97.97CD_KCC2	AM000053	All	Vidal, N.	*ARHR* **22**(2):182-187 (2006)
A6.BY.13.PV85	KT983615	All	Sasinovich, S.	Unpublished
A6.IT.02.60000	EU861977	All	Riva, C.	*ARHR* **24**(10); 1319-25 (2008)
A6.RU.11.11RU6950	JX500694	All	Baryshev, P.B.	*ARHR* **30**(6); 592-7 (2014)
A6.UA.12.DEMA112UA024	KU749403	All	Hora, B.	Unpublished
B.AR.14.DEMB14AR003	KY658685	All	Hora, B.	Unpublished
B.BR.10.10BR_MG035	KJ849804	All	Pessoa, R.	*Transfusion* **55**(5); 980-90 (2015)
B.CU.14.14CU005	KR914676	All	Blanco, M.	Unpublished
B.DE.13.366396	KT124767	All	Tully, D.C.	*PLoS Pathog* **12**(5); e1005619 (2016)
B.DE.13.947915	KT124808	All	Tully, D.C.	*PLoS Pathog* **12**(5); e1005619 (2016)
B.DO.11.DEMB11DR001	KY658702	All	Hora, B.	Unpublished
B.ES.15.100600	KY989956	All	Cuevas, M.T.	Unpublished
B.FR.11.DEMB11FR001	KF716496	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.JP.12.DEMB12JP001	KF716498	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
B.PH.15.DEMB15PH003	KY658690	All	Hora, B.	Unpublished
B.RU.11.11RU21n	JX500708	All	Baryshev, P.B.	Unpublished
B.SE.15.102SE	MF373204	All	Neogi, U.	*Sci Rep* **7**(1):6371 (2017)
B.TH.10.DEMB10TH002	KP109514	All	Hora, B.	Unpublished
B.US.16.2609	KX505536	All	Bruner, K.M.	*Nat Med* **22**(9); 1043-9 (2016)
C.BR.11.DEMC11BR035	KU749393	All	Hora, B.	Unpublished
C.CN.10.YNFL19	KC870038	All	Wei, H.	Unpublished
C.DE.10.622166	KT124786	All	Tully, D.C.	*PLoS Pathog* **12**(5); e1005619 (2016)
C.ES.15.100_118	KY496624	All	Cuevas, M.T.	Unpublished
C.ET.08.ET104	KU319528	All	Amogne, W.	*ARHR* **32**(5); 471-4 (2016)
C.IN.15.NIRT008	KX069226	All	Aralaguppe, S.G.	*J Virol Methods* **2016** Oct;236:98-104 doi: 101016/jjviromet201607010 Epub 2016 Jul 19
C.MW.09.703010256_CH256.w96	KC156214	All	Parrish, N.F.	*PNAS USA* **110**(17); 6626-33 (2013)
C.NG.10.10NG020523	KX389612	All	Heipertz, R.A. Jr.	*Medicine(Baltimore)* **95** (32):E4346 (2016)
C.NP.11.11NP016	KU341724	All	Bhusal, N.	*Curr HIV Res* **14**(6):517-524 (2016)
C.PK.14.DEMC14PK009	KU749412	All	Hora, B.	Unpublished
C.SE.15.100ET	MF373202	All	Neogi, U.	*Sci Rep* **7**(1):6371 (2017)

HIV-1 Proteins

Name	Accession	Proteins	Author	Reference
C.TZ.08.707010457_CH457.w8	KC156220	All	Parrish, N.F.	*PNAS USA* **110**(17); 6626-33 (2013)
C.US.14.M0118v1d14_5M04_C5	KT982201	All	Rainwater-Lovett, K.	*PLoS ONE* **12**(2):E0170548 (2017)
C.ZA.13.DEMC13ZA152	KU749417	All	Hora, B.	Unpublished
C.ZM.11.DEMC11ZM006	KF716467	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.BR.10.10BR_RJ095	KJ787684	All	Pessoa, R.	*Genome Announc* **2**(3):e00586-14 (2014)
D.CD.03.LA17MuBo	KU168271	All	Berg, M.G.	*J Clin Microbiol* **54**(4); 868-82 (2016)
D.CM.10.DEMD10CM009	JX140670	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.CY.06.CY163	FJ388945	All	Kousiappa, I.	*ARHR* **25**(8); 727-40 (2009)
D.KE.11.DEMD11KE003	KF716476	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.SE.12.077UG	MF373180	All	Neogi, U.	*Sci Rep* **7**(1):6371 (2017)
D.TZ.04.CO6405V4	KX907406	All	Billings, E.	*ARHR* **33**(4); 373-381 (2017)
D.UG.10.DEMD10UG004	KF716479	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.UG.11.DEMD11UG003	KF716480	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
D.YE.02.02YE516	AY795907	All	Saad, M.D.	*ARHR* **21**(7):644-648 (2005)
F1.AO.06.AO_06_ANG32	FJ900266	All	Guimaraes, M.L.	*Retrovirology* **6**, 39 (2009)
F1.AR.02.ARE933	DQ189088	All	Aulicino, P.C.	*ARHR* **21**(2):158-164 (2005)
F1.BR.11.DEMF111BR037	KU749396	All	Hora, B.	Unpublished
F1.CY.08.CY222	JF683771	All	Kousiappa, I.	*ARHR* **27**(11); 1183-99 (2011)
F1.ES.11.VA0053_nfl	KJ883138	All	Delgado, E.	*PLoS ONE* **10**(11):E0143325 (2015)
F1.FR.04.LA22LeRe	KU168276	All	Berg, M.G.	*J Clin Microbiol* **54**(4); 868-82 (2016)
F1.RO.03.LA20DuCl	KU168274	All	Berg, M.G.	*J Clin Microbiol* **54**(4); 868-82 (2016)
F1.RU.08.D88_845	GQ290462	All	Fernandez-Garcia, A.	*ARHR* **25**(11):1187-1191 (2009)
F2.CM.10.DEMF210CM007	JX140673	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
F2.CM.11.DEURF11CM026	KU749422	All	Hora, B.	Unpublished
G.CD.03.LA23LiEd	KU168277	All	Berg, M.G.	*J Clin Microbiol* **54**(4); 868-82 (2016)

Name	Accession	Proteins	Author	Reference
G.CM.08.789_10	KP718925	All	Luk, K.-C.	*PLoS One* **10**(11); e0141723 (2015)
G.CM.10.DEMG10CM008	JX140676	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
G.CN.08.GX_2084_08	JN106043	All	Liu, W.	*Zhonghua Liu Xing Bing Xue Za Zhi* **34**(1); 53-6 (2013)
G.CN.16.224GX	KY275364	All	Zhang, M.	Unpublished
G.ES.14.EUR_0033	KU685592	All	Cuevas, T.	Unpublished
G.GW.08.LA57LmNe	KU168300	All	Berg, M.G.	*J Clin Microbiol* **54**(4); 868-82 (2016)
G.KE.09.DEMG09KE001	KF716477	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
G.NG.12.12NG060409	KX389648	All	Heipertz, R.A. Jr.	*Medicine(Baltimore)* **95** (32):E4346 (2016)
H.CD.04.LA19KoSa	KU168273	All	Berg, M.G.	*J Clin Microbiol* **54**(4); 868-82 (2016)
H.CF.02.LA25LeMi	KU168279	All	Berg, M.G.	*J Clin Microbiol* **54**(4); 868-82 (2016)
J.CD.03.LA26DiAn	KU168280	All	Berg, M.G.	*J Clin Microbiol* **54**(4); 868-82 (2016)
J.CM.04.04CMU11421	GU237072	All	Yamaguchi, J.	*ARHR* **26**(6); 693-7 (2010)
K.CD.97.97ZR_EQTB11	AJ249235	All	Triques, K.	*ARHR* **16**(2);139-151 (2000)
K.CM.96.96CM_MP535	AJ249239	All	Triques, K.	*ARHR* **16**(2);139-151 (2000)
01_AE.CM.11.1156_26	KP718930	All	Luk, K.-C.	*PLoS One* **10**(11); e0141723 (2015)
01_AE.CN.12.DE00112CN011	KP109508	All	Hora, B.	Unpublished
01_AE.IR.10.10IR.THR48F	AB703616	All	Jahanbakhsh, F.	*ARHR* **29**(1); 198-203 (2013)
01_AE.JP.11.DE00111JP003	KF859741	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
01_AE.MM.14.fKSDU26	KU820849	All	Chen, X.	*Virulence* **2016** Aug 30:1-7 doi: 101080/2150559420161226722
01_AE.PH.15.DE00115PH012	KY658688	All	Hora, B.	Unpublished
01_AE.SE.12.070SE	MF373173	All	Neogi, U.	*Sci Rep* **7**(1):6371 (2017)
01_AE.TH.10.DE00110TH001	KP109513	All	Hora, B.	Unpublished
01_AE.TH.11.40436v09_04	KY580697	All	Kijak, G.H.	*PLoS Pathog* **13**(7); e1006510 (2017)
01_AE.TH.90.CM240	U54771	All	Carr, J.K.	*J Virol* **70**(9):5935-5943 (1996)
02_AG.CM.10.DE00210CM013	KF859739	All	Sanchez, A.M.	*J Immunol Methods* **2014** Jul;409:117-30 doi: 101016/jjim201401004 Epub 2014 Jan 19
02_AG.DE.09.701114	KT124792	All	Tully, D.C.	*PLoS Pathog* **12**(5); e1005619 (2016)
02_AG.KR.12.12MHI11_10746	KF561437	All	Cho, Y.-K.	Unpublished
02_AG.KR.12.12MHR9	KF561435	All	Cho, Y.-K.	Unpublished
02_AG.LR.x.POC44951	AB485636	All	Baesi, K.	*PLoS One* **9**(9); e105098 (2014)

HIV-1 Proteins

Name	Accession	Proteins	Author	Reference
02_AG.NG.12.12NG060304	KX389647	All	Heipertz, R.A. Jr.	*Medicine(Baltimore)* **95** (32):E4346 (2016)
02_AG.NG.x.IBNG	L39106	All	Howard, T.M.	*ARHR* **10**(12):1755-1757 (1994)
02_AG.PK.15.PK032	KX232622	All	Chen, Y.	*PLoS ONE* **11**(12):E0167839 (2016)
02_AG.SE.14.098GN	MF373200	All	Neogi, U.	*Sci Rep* **7**(1):6371 (2017)
02_AG.SN.13.9580	KT223760	All	Aman, S.	Unpublished
03_AB.RU.97.KAL153_2	AF193276	All	Liitsola, K.	*ARHR* **16**(11):1047-1053 (2000)
04_cpx.CY.94.94CY032_3	AF049337	All	Gao, F.	*J Virol* **72**(12):10234-10241 (1998)
05_DF.BE.x.VI1310	AF193253	All	Laukkanen, T.	*Virology* **269**(1):95-104 (2000)
06_cpx.AU.96.BFP90	AF064699	All	Oelrichs, R.B.	*ARHR* **14**(16):1495-1500 (1998)
07_BC.CN.98.98CN009	AF286230	All	Rodenburg, C.M.	*ARHR* **17**(2):161-168 (2001)
08_BC.CN.97.97CNGX_6F	AY008715	All	Piyasirisilp, S.	*J Virol* **74**(23):11286-11295 (2000)
09_cpx.GH.96.96GH2911	AY093605	All	McCutchan, F.E.	*ARHR* **20**(8):819-826 (2004)
10_CD.TZ.96.96TZ_BF061	AF289548	All	Koulinska, I.N.	*ARHR* **17**(5):423-431 (2001)
11_cpx.CM.95.95CM_1816	AF492624	All	Wilbe, K.	*ARHR* **18**(12):849-56 (2002)
12_BF.AR.99.ARMA159	AF385936	All	Carr, J.K.	*AIDS* **15**(15):F41-F47 (2001)
13_cpx.CM.96.96CM_1849	AF460972	All	Wilbe, K.	*ARHR* **18**(12):849-56 (2002)
14_BG.ES.05.X1870	FJ670522	All	Cuevas, M.T.	*ARHR* **26**(9); 1019-25 (2010)
15_01B.TH.99.99TH_MU2079	AF516184	All	Viputtijul, K.	*ARHR* **18**(16):1235-1237 (2002)
16_A2D.KR.97.97KR004	AF286239	All	Gao, F.	*ARHR* **17**(8):675-688 (2001)
17_BF.AR.99.ARMA038	AY037281	All	Carr, J.K.	*AIDS* **15**(15):F41-F47 (2001)
18_cpx.CU.99.CU76	AY586540	All	Thomson, M.M.	*AIDS* **19**(11):1155-63 (2005)
19_cpx.CU.99.CU7	AY894994	All	Casado, G.	*JAIDS* **40**(5):532-537 (2005)
20_BG.CU.99.Cu103	AY586545	All	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
21_A2D.KE.99.KER2003	AF457051	All	Dowling, W.E.	*AIDS* **16**(13):1809-1820 (2002)
22_01A1.CM.01.01CM_0001BBY	AY371159	All	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
23_BG.CU.03.CB118	AY900571	All	Sierra, M.	*JAIDS* **45**(2):151-160 (2007)
24_BG.ES.08.X2456_2	FJ670526	All	Cuevas, M.T.	*ARHR* **26**(9); 1019-25 (2010)
25_cpx.CM.02.1918LE	AY371169	All	Kijak, G.H.	*ARHR* **20**(5):521-530 (2004)
26_A5U.CD.02.02CD_MBTB047	FM877782	All	Vidal, N.	*ARHR* **25**(8):823-832 (2009)
27_cpx.FR.04.04CD_FR_KZS	AM851091	All	Vidal, N.	*ARHR* **24**(2):315-321 (2008)
28_BF.BR.99.BREPM12609	DQ085873	All	Sa Filho, D.J.	*ARHR* **22**(1):1-13 (2006)
29_BF.BR.01.BREPM16704	DQ085876	All	Sa Filho, D.J.	*ARHR* **22**(1):1-13 (2006)
31_BC.BR.04.04BR142	AY727527	All	Sanabani, S.	*ARHR* **22**(2):171-176 (2006)
32_06A6.EE.01.EE0369	AY535660	All	Adojaan, M.	*JAIDS* **39**(5):598-605 (2005)
33_01B.ID.07.JKT189_C	AB547463	All	SahBandar, I.N.	*ARHR* **27**(1); 97-102 (2011)
34_01B.TH.99.OUR1969P	EF165539	All	Tovanabutra, S.	*ARHR* **23**(6):829-833 (2007)
35_AD.AF.07.169H	GQ477446	All	Sanders-Buell, E.	*ARHR* **26**(5):605-608 (2010)
36_cpx.CM.00.00CMNYU830	EF087994	All	Powell, R.L.	*ARHR* **23**(8):1008-1019 (2007)
37_cpx.CM.00.00CMNYU926	EF116594	All	Powell, R.L.	*ARHR* **23**(7):923-933 (2007)
38_BF1.UY.03.UY03_3389	FJ213783	All	Ruchansky, D.	*ARHR* **25**(3); 351-6 (2009)
39_BF.BR.04.04BRRJ179	EU735535	All	Guimaraes, M.L.	*AIDS* **22**(3):433-435 (2008)
40_BF.BR.05.05BRRJ055	EU735537	All	Guimaraes, M.L.	*AIDS* **22**(3):433-435 (2008)
41_CD.TZ.05.CO6577V5	KX907411	All	Billings, E.	*ARHR* **33**(4); 373-381 (2017)
42_BF.LU.03.luBF_01_03	EU170151	All	Struck, D.	*ARHR* **31**(5); 554-8 (2015)
43_02G.SA.03.J11223	EU697904	All	Badreddine, S.	*ARHR* **23**(5):667-674 (2007)
44_BF.CL.00.CH80	FJ358521	All	Delgado, E.	*ARHR* **26**(7); 821-6 (2010)
45_cpx.FR.04.04FR_AUK	EU448295	All	Frange, P.	*Retrovirology* **2008** Aug 1;5:69 doi: 101186/1742-4690-5-69
46_BF.BR.07.07BR_FPS625	HM026456	All	Sanabani, S.S.	*Virol J* **2010** Apr 16;7:74 doi: 101186/1743-422X-7-74

HIV-1 Proteins

Name	Accession	Proteins	Author	Reference
47_BF.ES.08.P1942	GQ372987	All	Fernandez-Garcia, A.	*ARHR* **26**(7); 827-32 (2010)
48_01B.MY.07.07MYKT021	GQ175883	All	Li, Y.	*JAIDS* **54**(2):129-136 (2010)
49_cpx.GM.03.N26677	HQ385479	All	de Silva, T.I.	*Retrovirology* **7**(1):82 (2010)
50_A1D.GB.10.12792	JN417240	All	Foster, G.M.	*PLoS One* **9**(1); e83337 (2014)
51_01B.SG.11.11SG_HM021	JN029801	All	Ng, O.T.	*ARHR* **28**(5); 527-30 (2012)
52_01B.MY.03.03MYKL018_1	DQ366664	All	Tee, K.K.	*JAIDS* **43**(5):523-529 (2006)
53_01B.MY.11.11FIR164	JX390610	All	Chow, W.Z.	*J Virol* **86**(20):11398-11399 (2012)
54_01B.MY.09.09MYSB023	JX390976	All	Ng, K.T.	*J Virol* **86**(20):11405-11406 (2012)
55_01B.CN.10.HNCS102056	JX574661	All	Han, X.	*Genome Announc* **1**(1):E00050-12 (2013)
56_cpx.FR.10.URF5_patient_A	JN882655	All	Leoz, M.	*AIDS* **25**(11):1371-1377 (2011)
57_BC.CN.09.09YNLX19sg	KC899008	All	Han, X.	*PLoS ONE* **8**(5):E65337 (2013)
58_01B.MY.09.09MYPR37	KC522031	All	Chow, W.Z.	*PLoS ONE* **9**(1):E85250 (2014)
59_01B.CN.09.09LNA423	JX960635	All	An, M.	*J Virol* **86**(22); 12402-6 (2012)
60_BC.IT.11.BAV499	KC899079	All	Simonetti, F.R.	*Infect Genet Evol* **2014** Apr;23:176-81 doi: 101016/jmeegid201402007 Epub 2014 Mar 3
61_BC.CN.10.JL100010	KC990124	All	Li, X.	*Genome Announc* **2013** Jun 27;1(3) pii: e00326-13 doi: 101128/genomeA00326-13
62_BC.CN.10.YNFL13	KC870034	All	Wei, H.	*ARHR* **30**(4):380-383 (2014)
63_02A.RU.10.10RU6637	JN230353	All	Baryshev, P.B.	*Arch Virol* **157**(12); 2335-41 (2012)
64_BC.CN.09.YNFL31	KC870042	All	Hsi, J.	*ARHR* **30**(4); 389-93 (2014)
65_cpx.CN.10.YNFL01	KC870027	All	Feng, Y.	*ARHR* **30**(6); 598-602 (2014)
67_01B.CN.11.ANHUI_HF115	KC183779	All	Wu, J.	*PLoS ONE* **8**(1):E54322 (2013)
68_01B.CN.11.ANHUI_WH73	KC183782	All	Wu, J.	*PLoS ONE* **8**(1):E54322 (2013)
69_01B.JP.05.05JPMYC113SP420	LC027100	All	Kusagawa, S.	*Genome Announc* **2015** May 28;3(3) pii: e00196-15 doi: 101128/genomeA00196-15
70_BF1.BR.10.10BR_PE004	KJ849758	All	Pessoa, R.	*PLoS One* **9**(11); e112674 (2014)
71_BF1.BR.10.10BR_PE008	KJ849759	All	Pessoa, R.	*PLoS One* **9**(11); e112674 (2014)
72_BF1.BR.10.10BR_MG002	KJ671534	All	Pessoa, R.	*Genome Announc* **2**(3):e00386-14 (2014)
73_BG.DE.01.9196_01	AY882421	All	Harris, B.	*ARHR* **21**(7):654-660 (2005)
74_01B.MY.10.10MYPR268	KR019771	All	Cheong, H.T.	*PLoS ONE* **10**(7):E0133883 (2015)
77_cpx.MY.14.14MYNBB090	KX673820	All	Tee, K.K.	*Genome Announc* **2017** Jun 29;5(26) pii: e00459-17 doi: 101128/genomeA00459-17
78_cpx.CN.13.YNTC19	KU161143	All	Song, Y.	*ARHR* **2016** Apr 8
79_0107.CN.15.SX15DT013	KY216146	All	Li, Y.	*ARHR* **2017** May 30 doi: 101089/aid20170066
82_cpx.MM.13.mSSDU12	KU820825	All	Chen, X.	*Virulence* **2016** Aug 30:1-7 doi: 101080/2150559420161226722
83_cpx.MM.13.mSSDU94	KU820847	All	Chen, X.	*Virulence* **2016** Aug 30:1-7 doi: 101080/2150559420161226722
85_BC.CN.14.14CN_SCYB2	KU992929	All	Su, L.	*ARHR* **2016** Jun 13
86_BC.CN.13.15YNHS18	KX582249	All	Li, Y.	*ARHR* **2016** Oct 18
87_cpx.CN.12.DH32	KF250408	All	Wei, H.	*Sci Rep* **2015** Jul 2;5:11323 doi: 101038/srep11323
88_BC.CN.05.05YNRL25sg	KC898979	All	Han, X.	*PLoS ONE* **8**(5):E65337 (2013)
90_BF1.BR.11.BRGO6043	KY628221	All	Reis, M.N.d.G.	*PLoS ONE* **12**(6):E0178578 (2017)

HIV-1 Proteins

Name	Accession	Proteins	Author	Reference
O.CM.96.LA51YBF35	KU168294	All	Berg, M.G.	*J Clin Microbiol* **54**(4); 868-82 (2016)
O.CM.96.LA52YBF39	KU168295	All	Berg, M.G.	*J Clin Microbiol* **54**(4); 868-82 (2016)
O.CM.99.99CMU4122	AY169815	All	Yamaguchi, J.	*ARHR* **19**(11):979-988 (2003)
O.CM.x.pCMO2_3	AY618998	All	Tebit, D.M.	*Virology* **326**(2):329-339 (2004)
O.ES.01.Read25_HIV_GroupO	KX228804	All	Ragupathy, V.	Unpublished
O.FR.06.RBF206	KY112585	All	Mack, K.	*J Virol* **2017** Feb 28;91(6) pii: e02177-16 doi: 101128/JVI02177-16 Print 2017 Mar 15
O.GA.10.10Gab1190	JX245014	All	Liegeois, F.	*ARHR* **29**(7); 1085-90 (2013)
O.GA.11.11Gab6352	JX245015	All	Liegeois, F.	*ARHR* **29**(7); 1085-90 (2013)
O.SN.99.99SE_MP1300	AJ302647	All	Toure-Kane, C.	*ARHR* **17**(12):1211-1216 (2001)
O.US.10.LTNP	JN571034	All	Buckheit, R.W.3.	*ARHR* **30**(6); 511-513 (2014)
N.CM.15.S4858	KY498771	All	Rodgers, M.A.	*Virology* **504**, 141-151 (2017)
N.FR.11.N1_FR_2011	JN572926	All	Delaugerre, C.	*Lancet* **378**(9806); 1894 (2011)
P.CM.06.U14788	HQ179987	All	Vallari, A.	*J Virol* **85**(3); 1403-7 (2011)
P.FR.09.RBF168	GU111555	All	Plantier, J.-C.	*Nat Med* **15**(8); 871-2 (2009)
CPZ.TZ.06.TAN5	JN091691	All	Rudicell, R.S.	*J Virol* **85**(19); 9918-28 (2011)
CPZ.US.85.US_Marilyn	AF103818	All	Gao, F.	*Nature* **397**(6718):436-441 (1999)
GOR.CM.12.SIVgor_BQID2	KP004991	All	D'arc, M.	*PNAS USA* **112**(11); E1343-52 (2015)
GOR.CM.13.SIVgor_BPID15	KP004990	All	D'arc, M.	*PNAS USA* **112**(11); E1343-52 (2015)

HIV-1 Proteins

membrane binding

phosphorylation site

Gag start, p17 start

p17 end p24 start

```
B.FR.83.HXB2                    MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSL.QTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEEEQN...KSKK......KA...........QQA...AADTGHS....NQ..VSQNYPIVQNIQG 140
```

Gag

membrane binding phosphorylation site

Gag start, p17 start p17 end_p24 start

```
B.FR.83.HXB2                    MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSL.QTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEEEQN...KSKK......KA...........QQA...AADTGHS....NQ..VSQNYPIVQNIQG 140
02_AG.SN.13.9580                ----------K--A--------------R----L----G-----L---A--Q-LME--ST--K----------K-F--I---W---------R-------V--V----QKTQP--T...............AA-S-...............A-- 141
03_AB.RU.97.KAL153_2            ----------K--A---------E--RI--L-------L-S------A--Q-LME----ST--K--------K-F-I---W-----R--------I-S-------Q------T.................TG--S-...SK-...........A-- 140
04_cpx.CY.94.94CY032_3          ----------K--A--------------R----L----------K-IS---AI----T--K-----------IT-W----DVQ------------I-S------Q------X...............AA-G-..SN.............A-- 140
05_DF.BE.x.VI310                ----------A--V-----------------------------L-----------K-------------------E--V---------E---------Q------------T..............E-GA-N-...S....A........A-- 141
06_cpx.AU.96.BF90               ----------K--E----------------M--L-----------A--Q--IE---ST--K------K----------V---------------KVT--Q-...........H......A-P-..S...A........A-- 139
07_BC.CN.98.98CN009             ------I-R--K--K---------H-M--L-------------L----------MK---A---T-----------TE-DVR---------------I-K----IQQ......T..................KEAD...GK.............L-- 137
08_BC.CN.97.97CNGX_6F           .#----I-R--K--K---------H-M--L-------------L-----K--IK--A---T------------EE--VR---------------I-K----IQQ......T..................KKAE...EK.............P-- 135
09_cpx.GH.96.96GH2911           ----------K--A-----------------D-S--TG--Q--L----------K--A-----------------K--VR---------I-G----Q------T.................T-N--NN...SK.............A-- 140
10_CD.TZ.96.96TZ_BF061          ----------K--E--------------Q----------D--I------A----I---AI--------IK----------K-VA----------T..........T---N-...SK.............A-- 140
11_cpx.CM.95.95CM_1816          ----------K--A-------R-Q-R--L-------L-S----Q--MV-I--A-N--T-------------Q--VR----------------Q------T.................T--N--...SK.............A-- 140
12_BF.AR.99.ARMA159             ------------------------R-------------------K-I------------------------F---V--K--VR----------E-L--QQ......T..................K-..............L-- 136
13_cpx.CM.96.96CM_1849          ------------K----------R-RM----L----T--Q--M------A---T--I--F----------------V------------EEV-KV--S-QQ...RT...............KNE-N................A-- 140
14_BG.ES.05.X1870               ----------K--S--------------R---L----L-D---A--Q--M------A----T-I--F----------V------EEV-KA-K...-Q---T..............K-E-NN...S....A........A-- 139
15_01B.TH.99.99TH_MU2079        ----------K--A--------------I-S--SA---Q---A----A---T--K-----------W--QN----V----------NQQ......T..................-G--S-...ST....G........A-- 140
16_A2D.KR.97.97KR004            ----------K--A---------R--R----I-------------K-I--E--I-K----IK----F-A-V------VR----------L------H-Q--T..............P-............S-....G........A-- 138
17_BF.AR.99.ARMA038             .--------------------R--R---------------AD--K-I------------I-V-Y--KVDV------E-L----RQ...TQQAAADKT...............-...............A-- 142
18_cpx.CU.99.CU76               ----------K--K----------------R--L----------L------A--Q--IE---S-I-K----IK----F------W---GV----EV-KL-K...-QQ......T............--AA-S-...S.............A-- 140
19_cpx.CU.99.CU7                ------L----K--A-------S-Q-RM--L------L------A--Q-IE---ST--K-----F------W-----GV----EV-KL-K...-Q------T..............E--T---RN....SH...........HKL-- 140
20_BG.CU.99.CU103               ----------K--A-------------R----L----L---V--Q-IN---A-K--T--K----F--------A-DV------------------N...............-T-GE-N-...S..............A-- 140
21_A2D.KE.99.KER2003            ----------K--A-------Q-R--L-----L----------K-I------I-----K------------A-DV------------N.................P....N--N-...S.............A-- 138
22_01A1.CM.01.01CM_0001BBY      ------------K--A--------RM-----L---D--LD---Q--IE-----A-K---I--IK----------V------------------N.................T-GE-N-...S.............SN-- 138
23_BG.CU.03.CB118               ----------K--A-------------RM--L----A--Q--MS---A--K--T-IK--F-A----------EV-KL-K...-QQ......T............-----GE-N-...S.............A-- 140
24_BG.ES.08.X2456_2             ----------K--A--------------R----L----AG--Q--N--A---K--T-------F------------EV-KL-K...-QQ......T...............-T-GE-S-...S.............A-- 140
25_cpx.CM.02.1918LE             ......--I--R--A--------R-R-R--L-S--NG--Q------SA--K------------I----------------I---N-Q...-K............-E--T---SNS..S..............A-- 139
26_A5U.CD.02.02CD_MBTB047       .......--I--R--A----------K-SI---------A--Q--E--AV-K------K-T--S----------NV---------V-------N.................-DN--S...SK.............A-- 137
27_cpx.FR.04.04CD_FR_KZS        ----------K--A--------------RM--L-----AAD--LD---A-----K-F-L------------DV------------EEV-KV-K..NRQQ......T............-E--E-S-...S..............H-A-- 137
28_BF.BR.99.BREPM12609          ---------------------------------------I----------------------V------------K....................G--N-...S..............L-- 140
29_BF.BR.01.BREPM16704          --------------------------N-------------------------------------V-------------E----------T..........-QQAAAGTGNSR-----G--N-...S..............L-- 151
31_BC.BR.04.04BR142             ------I-R-EK--T-------Q-MM----LD-------G-K-IQ----A--K-TK--V-----------EK--VR-------K---QQ......T...............E.-AAAG..KGN................L-- 139
32_06A6.EE.01.EE0369            ----------K--E--------------R-------------G--VE-V--A-R----F-------------KVA------K---QQ......T................TA--N-.................L-- 137
33_01B.ID.07.JKT189_C           ----------K--A--------------R----L----------Q--ID--ST--K---Q-H--IV--W--------M--T---RQQ......T................-G--SN...SK.............L-- 139
34_01B.TH.99.QUR1969P           ----------K--A--------------N-QM--L----L--Q--IE--ST--K---F--I--W---DV------------L--V--Q------T.................-G--SN...SK.............L-- 138
35_AD.AF.07.169H                .----I-R--K--T-------------R----L----L-S----T---ME----A-K------------Q-A-----------T................-G--SG...SK.............A-- 138
36_cpx.CM.00.00CMNYU830         ----------K--A------------RI--L------D----SA--Q--IE----T--K----------DV------------KL-S-..N...Q------T.............-G--S-...ST....R........A-- 138
37_cpx.CM.00.00CMNYU926         ------T--K--A-------Q-RM--L------L----------A--Q--IE--ST--K----F------------DVR---------------Q------T..............-G--S-...CK-I...........A-- 136
38_BF1.UY.03.UY03_3389          --------------------------R---K-MAH------------------------VDV-------E-L---K..-Q------T..............K-.................L-- 137
39_BF.BR.04.04BRRJ179           --------------------------R-------------IA---------------I-V-Y--KV-V--------R.................NN.................A-- 140
40_BF.BR.05.05BRRJ055           ----------K--A------------Q-----------L----------I---AI---FK------------EK--VR------------E-L----QA..................A--A-DN...SK.............A-- 141
41_CD.TZ.05.CO6577V5            ----------K--A-----------------------------------------------FK------------EK--VR------------E-L.................QA..............AA-S-..S..............A-- 139
42_BF.LU.03.luBF_01_03          #.-------------------------------M--A---A---------K-DVR.....................N................-L-- 139
43_02G.SA.03.J11223             ----------K--A--------------RM--L------I----SA--Q--ME-X-ST--R----F--I---W------V--M---Q------T...............A--S-.................L-- 137
44_BF.CL.00.CH80                ------I--K--A--------------RM-LI------D------P--K-I------------VV--F--V-V-----E-L----QQ......QQKT..............A.A--S-.................A-- 142
45_cpx.FR.04.04FR_AUK           ----------K--M------------R-----L----D--L--Q--IE---N--K----IK------V---DV------------------E-L----QQ......NT...............A--S-...SS................-- 140
46_BF.BR.07.07BR_FPS625         ----------K--A--------------RM-LDS------P--K-IA--S--A-----------------------DV------------E-L----QQ......T..................G--N-.................L-- 136
47_BF.ES.08.P1942               ---------------------T----------------------A------------F-A-----K-DV.............V...................NN................L-- 139
48_01B.MY.07.07MYKT021          ------I--K--A--------------R-----L----------A--Q-LIE--ST--K-----F--IV--L------------M-K...-QQ......T.................E-G--S-...SK.............A-- 140
49_cpx.GM.03.N26677             ------I--K--A--------R-Q-R--L------L----------Q--IE---ST--K------F------------DV------------------GT.......................ANI..SS...I........-T-- 140
50_A1D.GB.10.12792              ......--I------A--------R-----L----L-D---A--L--MD-I--A-R--T-----------------------E-L--R-Q.......Q.............P....----AASC..S..............A-- 140
51_01B.SG.11.11SG_HM021         .#.......X--K--A-----------R-----------A--Q--IE------A-K---V--W--------MDV.....................N--N-...S.............A-- 138
52_01B.MY.03.03MYKL018_1        -------X--K--A--------------QM--L------A--Q-IE--ST--K------V--V--W-------M--A---Y...-T.................N--NX...SK.............A-- 139
53_01B.MY.11.11FIR164           ----------K--A--------------XI-------------L--A--Q--IE-I--X--K----H-L---V--W-----VX-----M--X-K-.GQQ......T...............A-- 140
54_01B.MY.09.09MYSB023          ------I---Q-----------------R-----------Y------S----K----A------GF---F------V------A-K-.GQQ......K...................A-- 140
55_01B.CN.10.HNCS102056         ----------K--A--------------R-----------A--Q-LIE--ST--K--------V-------------DV------------------V-K-.R-QQ......T...............GA--S-...SK.............A-- 140
56_cpx.FR.10.URF5_patient_A     ----------K--A--------------R----D--LD----SA--Q-LIE--------K-----F--I--L--N-------------A--S-...S.................L-- 140
57_BC.CN.09.09YNLX19sg          ------I-R--K--A--------R-RM--L-------L----A-K-LMK------A--T-GFK--F-----------AE-DVR------R--K-...IQQ......T..................QEAD...GK.............A-- 137
58_01B.MY.09.09MYPR37           ------T--K--A--------------M---------L----------A--Q--IE--ST--K----F------------RL--A---Q..G.Q.....-T..................GK.............L-- 139
59_01B.CN.09.09LNA423           ------I--K--A--------------RI--L------D----K--Q-IE--A--K--T-----F------------M--A-K-....IQQ......T..................GA--S-...K-I...........A-- 136
60_BC.IT.11.BAV499              ------I-R-EK--T--R-K-K-----H-M----L----LD-------G-K--MN------A----K--V-H-A--------EK--VR-----K--R...QQ......TT...............R.-EAAD..KGK.............-- 136
61_BC.CN.10.JL100010            ------I-R--K--A---------H-M--L--------------L--------------------SE--DVR.....N.QQ......R................-KEAD...GK.............L-- 137
62_BC.CN.10.YNFL13              ------I--K--A---------H-M--L-------------D-K-LMK------A--T-----F------------KG-DVQ.................E-AD...GK....G........L-- 133
63_02A.RU.10.10RU6637           ------I-R--K--A---------H-M--L----L----A--Q-LME---ST--K-----F--I---W---------DVR......................KA-D..GK.............A-- 137
64_BC.CN.09.YNFL31              ------I-R--K--A---------H-M--L----------A--MK-----A---T------------EE-DVR---------------I-K-...IQQ......T...............-E...K-AD..GK.............A-- 137
65_cpx.CN.10.YNFL01             ------T--K--A--R-------------RM--L------A--Q-IE----TV-K------F-A-VV----------DV--------------------QR...T.................-S--SG...GK.............L-- 140
67_01B.CN.11.ANHUI_HF115        ----------K--A--------------Q--M--L----------A--Q-IE--ST--K----F--IV-W----------A---------------V-----GA--S-...SK.............A-- 140
68_01B.CN.11.ANHUI_WH73         ----------K--A--------------RM--L-------A--Q--IE----ST--K-----F--I-V-W..................-V-K..R-QQ......T...............GA--S-...SK.............A-- 140
69_01B.JP.05.05JPMYC113SP420    ----------K--A--------------R-----------A--Q-LIE-I--AI-K----------F--I---W--------------M-K...-QQ......T...............GK.............L-- 136
70_BF1.BR.10.10BR_PE004         I-----I--------------------Q-R--V---------A----------R------------DV--------E...................G--N-...S-SSQSG.........-M-- 143
71_BF1.BR.10.10BR_PE008         --------------------------R----------------A--------------------M-DV------------E...................GADN-...S..............H-- 140
72_BF1.BR.10.10BR_MG002         --------------------R----N-------------------I-----------------------F--I-V------DV------------E..................E----KG..SP.............L-- 140
73_BG.DE.01.9196_01             ----------K--A------------RM--L------L-E---T--Q--M------K-T-I--F------------K--V------------EEV-KA-K..XQE......K...............-M-E-NN...S....A........A-- 139
74_01B.MY.10.10MYPR268          ----------K--A------------RM--L--------L-E----T--Q--M----IE--ST--K------IK-------------W--------L--AR....-QQ......T...............-G--S-...SK.............A-- 140
77_cpx.MY.14.14MYNBB090         ------I-R--K--A--Q-K--------H-M--L----L-D---------K--IE---A---V----TD-VK---F------------TG-DVR.................VQQ..............GQ...K-QA.AD...GK.............L-- 139
78_cpx.CN.13.YNTC19             ----------K--A--------------RM--L------A--Q-LIE---ST--K----F--I---W---------A-------E---A-K..R-QQ......T...............-G--S-...SK.............A-- 123
79_0107.CN.15.SX15DT013         .......--I-----K--A------------R-----L-------------A--Q-IE--------K-H----V-W--------------DV.................PV...V-G--SN...K...H-------A-- 140
82_cpx.MM.13.mSSDU12            ----------R-EK--K--------------H-M--------------A-----------T------------EG-NVR.................-P..............KEAA...MG-- 130
83_cpx.MM.13.mSSDU94            ----------R--A--------------RM--L------L--------T-----G-K--F------K-F-LIV--W----A---------------L--A-....-QQ......T...............-NN.................A-- 133
85_BC.CN.14.14CN_SCYB2          ----------R-EK--A-------------H-M----L------DK--L--------K--IK--H-A--K-T---------TG-DVR...............-M-....IQQ......NT...............KEAD...R.............A-- 137
86_BC.CN.13.15YNHS18            ---------------C-------------H-M--L----------A-----------------------EG-DVR.................-V......AEQ......T...............-GK....H-----L-- 120
87_cpx.CN.12.DH32               ------I-R--K--K---------H-M--L----------------K--IK--H-A--K--T-A--------AG-DVQ.......................MEQ.................K-- 134
88_BC.CN.05.05YNRL25sg          ------I-R--K--A--------------H-M--L------AD--K--IR--H-A---T-----------AG--VR.........................-I-Q.AQN..S...............KEAD...GK.............A-- 140
90_BF1.BR.11.BRG06043           ---------------------R-Q---------------------------------------N--K--------N.....................QA............N------F-------L-- 142
O.CM.96.LA51YBF35               ----------T-SK--A--Q-----K--S--T-R--L---------D--C--E------A--NEKL-Q--E-A--K-----S-Q--W---V-W----N-FKVD--QQ-IQ-LKQVME....NR-S...........AST....KEETS....R...TG----VIT-A-- 138
O.CM.96.LA52YBF39               ---------T-K-SK--A--Q--S-----R--L---------K--CD-E------A--NERL-Q--ESA--K-----S-Q--W---V-W----N-FKVE--QQ-IE-LK-VME....RR-P...........ADT....KEDTN-....R...TC----V-S-A-- 138
O.CM.99.99CMU4122               ----S-----K-SK--A--Q---R-----R--L---------D--C-E------A------NEKL-Q--E-A---DS-Q--W--IVV-W----N-YR-E--QQ-IK-LR-VMG....SX-S...........ASD....KEDTSA....R..TG----IP-A-- 138
O.CM.x.pCMO2_3                  ----S-----K-SK--A--Q--K--C--R-----------D--C-E------A--NEKL-Q--E-A--K-----DS-H--W--ITV-W----N-YKVG--QQ-IQ-LK-VMG....NR-S...........AG----KEDTSA....R...PG----VIA-A-- 138
O.ES.01.Read25_HIV_Group0       ---S------T-SK--A--Q--S-----R--L----------D--C-E------A--NEKL-Q--E-A--K-----S-Q--W---IV-W----N-YK-G--QQ-IQ-LK--VMG....SR-S...........AGT....TKEDTSA....K...TG----P-A-- 138
O.FR.06.RBF206                  ---R-SK--A--Q--S-------LI--------NY-CD-E--DNA--NEKL-K--E-A------N-K--W--I-V-W----NKYQ-G--QQ-IQ-L-AVMK....SRQS...........AST....KEETS-....G...TG----V-A-A-- 138
O.GA.10.10Gab1190              ---S------T-SK--A--Q--S-----R--------------C-E------A------NE-L-Q--E-A--K-----KG-I--W---V-W----N-KGYKVR--QQ-VQ-LK-VMG....SK-S...........AG----KEDTSA....R...TG----P-G-- 138
O.GA.11.11Gab6352              ---S------R-SK--S--Q--S-----R--L---------D--C-E------A--NEEL-Q--E-A------DS-Q--W--I-V-W----NGFR-N--QQ-IQ-LK--VMG....NK-S...........AG----KEDTSE....R...TG----IP-A-- 137
O.SN.99.99SE_MP1300            ---S------SK--T-Q--Q--K--C--R--L----------C-E------A--NEEL-Q--E-A--K-----F-KVE--QQ-IQ-LK--VMG....SR-S...........AG----KEDTSA....R...TG----V-A-A-- 138
O.US.10.LTNP                   --------T-SK--A--Q--S-----R--L---------C-S-E------A--NEKL-Q--E-A--K-----S-Q--W--I-V-W----N-FKVE--QQ-IQ-LK--VMG....SR-P...........AG----KEDASA....R...TG----P-A-- 138
N.CM.15.S4858                  ----------T--K--Q--STY----------QM--L---------D--C---MDSA--A-LMN--E-A--K-----G--------L-V--------R--SVQN-Q---.-VHG-KE....QHN-P....-N................SE-...G-AAAVD...SS...I-R--L--TA-- 141
N.FR.11.N1_FR_2011             ----------T--K--Q--A-Y---------RM--L---------C---MD-A--XV-L-N--E-A--K--------G--------L-V-------S--S-Q--QA-----.-VHG-KE....QH-QPEP..-N................PE-...G-AAAPD...N...I-R--L--A-- 143
P.CM.06.U14788                 ----------T--R--A------P---------V----------C-----M--A--EKL-E--E-A-R--DG-Q--W--LVV-W----K-----S-...QQ-IT-WK--MQ....R-P..................TSEG...TG...T--------A-- 133
P.FR.09.RBF168                 ----------T--R--A--------------------C-----M--A--EKL-E--E-A-R--DG-Q--W---R-VD-S--QQ-IT-WK--MQ....RR-................ALEG-T...TG-------A-- 135
CPZ.TZ.06.TAN5                 ------------R-DK--T--S---KSR-R---LI--L---GS--Q--M----M-NV----WK-IL-----V.DI--P--IV--F--ICV........AGERVQ--E--VKIVKMKKLCIQ..DSSTATS.SGGQSNAGGKERETVTPSGK-ENAGRA....TEPPSG-RL--VITDA-- 164
CPZ.US.85.US_Marilyn          ------------T--R--A----------------MM--L---------C----M-AD--L-L-K--E-A--K-----G-Q--W--L-V-W----N-S-VTVE--QQ--V-LK-VVQLQKQQEE....E............-Q...QEASG-....I..G-S---VI--A-- 142
GOR.CM.12.SIVgor_BQID2        ----------SK--E----------------A---------TEKL-Q--E-A--K-----S-Q--W--IVV-W----N-FKVG--QQ-IQ-LK-VKG....HKNT...S..........T-QRS..Q...TG--------P-A-- 138
GOR.CM.13.SIVgor_BPID15       ----------T--R--A--R--K------T-M--V------C-E-M-A--EKL-----E-A--K-----G-Q--W--LVV-W----K--DVA--QQ-IT-WK--MQ....RR-................EKES...PG...T--------A-- 133
```

Gag

CyPA binding major homology region

```
B.FR.83.HXB2              QMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGW.MTN.NPP..I.PVGEI.YKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTL 304
A1.CD.02.LA01AlPr         --T--L----------A----------T------------M--I-------D--------Q----P--L-------------------------.G.---D-----------------------V-----------------F-I- 310
A1.CM.08.886_24           --T--SL------------G------------------M--I-------D----------L-AQ---LP--L-------------P-----.G.---D-----------------------V-----------------F-V- 304
A1.ES.15.100_117          ------NL----------I------------------M--I-------D----------L-IQ---P--L-------------P-----LQ--.A.---D-----------------------V-----------------F-A- 304
A1.KE.11.DEMA111KE002     --I--SL-----------AI----------I------M--I-------D--L-IQ---P--L-------------P-----.G.---D-----------------------V-----------------F--- 304
A1.PK.15.PK034            -FIY--V----------AI----------I------M--I-------D--L--IQ---V-----------------S-----.G.---A.---D-----------------------V-----------------F-V- 304
A1.RW.11.DEMA111RW002     ----SL------------AI----------I------M--I-------D----L--P--VP----I-D-------S---R--.S.---D--R-----------------------V-----------------F--- 305
A1.SE.13.078SE            ------TL-----------I------------------M--I-------D----------P----A-.G.--A.---D-----------------------V-----------------F--- 304
A1.UG.11.DEMA110UG009     ------V-----------I---------G--M--I-------D----L--L-------I--------P-.S.--V.---D--R-----------------------V-----------------F-C- 305
A2.CM.01.01CM_1445MV      --------------------------------------M--I-------D-------P------------.S.---------------------------V-----------------F--- 300
A3.SN.01.DDI579           ----SL------------I------------T----M--I-------D--------M---I--------.S.--.V.---R-----------------------V-----------------F-C- 302
A4.CD.97.97CD_KCC2        ------L-----------I------------C--------M--I-------D------------------P-.S.---D-----------------------V-----------------F--- 304
A6.BY.13.PV85             ------SM-----------I------------------M--I-------D--------D---T--TQ---LP---.S.---------------------------V-----------------F--- 304
A6.IT.02.60000            ------SL-----------I------------------M--I-------D----L--Q---LP--L-----A.-.S.---------------------------V-----------------F--- 304
A6.RU.11.11RU6950         ---I--SM-----------I----------T------M--I-------D----L--AQ---FP-----.S.--.V.---D-----------------------V-----------------F--- 300
A6.UA.12.DEMA112UA024     ------SM-----------I----------T------M--II-------D----T-AQ---FPA----.S.--.V.---------------------------V-----------------F--- 304
B.AR.14.DEMB14AR003       ------L-----------I----------T----------L--Q--V-----L----------N----A.-.H.-----D-------M-------------------V----------- 304
B.BR.10.10BR_MG035        ------PL---------------------------------------------L----------------------------.V.---------------------------V----------- 305
B.CU.14.14CU005           ------L-----------I----------T----------KAGGG-TPQ-PSTY-LPKCF--R-L-S---D-MS-A--D-GDCAL.TSD.-GL...T-SP-.HGGE--QRW-SR------------- 304
B.DE.13.366396            ------L-----------I------------------D----L----------------------T-.-.S.---------M-----------------VG---------------- 304
B.DE.13.947915            --------------------------------------------D--L-----------------------------------------K-------------- 307
B.DO.11.DEMB11DR001       ------PLT-------------------------------L--------------------------A.-.----------------------------------------------- 309
B.ES.15.100600            ------L-----------I----------T----------L--Q--V-----I----------A.-.H.-.V.----------M--------------V----------- 304
B.FR.11.DEMB11FR001       -------------------------------------------L----------------------A.-.-----------M-----------------V----------- 304
B.JP.12.DEMB12JP001       --------------------------------------------L------NP--------------A.-.----------------------------V----------- 304
B.PH.15.DEMB15PH003       ------------------I-----------------------L--------------------------A.-.-----------------------------V----------- 304
B.RU.11.11RU21n           --------------------------------------------L--------------------------------------------------V----------- 309
B.SE.15.102SE             ------P----------------------------------L-----------------------------------------------------V----------- 304
B.TH.10.DEMB10TH002       ------P---------------------------------L--A--V----L-D--------------A.-.H.----------------------------V----------- 304
B.US.16.2609              ---------------------------------------------------------S-------A.-.S.--A.----------------------------------- 309
C.BR.11.DEMC11BR035       --T--P-----------I-------------------D----------------------A.-.G.-----------------------S----------F--- 302
C.CN.10.YNFL19            ------P-----------I--------------I-DV----------------------A.-.G.-.V.----------------------------F--- 301
C.DE.10.622166            ------P-A-------------I----------T----------D----------------A.T.A.----------V-------------------F-C- 301
C.ES.15.100_118           ------P-A-------------I----------T----------D----L--V--------A.-.S.-.V.---D------------V--------F-C- 299
C.ET.08.ET104             ------PL----------I---G-N-----------D----L--Q--V----------N----A.-.H.----------------------------F--- 302
C.IN.15.NIRT008           ------SL----------I--R------------I------D----L--Q--V----------N----A.-.-.V.-----------------------I--F--- 301
C.MW.09.703010256_CH256.w96 ------PL----------I---G-N-----------T----D----L--Q--V----I----------N----A.-.-.V.---D------------------------V----------- 301
C.NG.10.10NG020523        ------L-----------I----------I--T----D----L--A--V--D--------N----A.-.G.-.V.--------------------------F--- 301
C.NP.11.11NP016           ------P-----------I----------T----------D----------------T---.-.G.----------V-----------------F--- 297
C.PK.14.DEMC14PK009       ------------------I----------T----------D----L--V------A---A-.I.G.-.V.-----------------X----------F--- 302
C.SE.15.100ET             --I---------------I----------T----------D----L--V----------N----.S.-----------------------X----------F--- 302
C.TZ.08.707010457_CH457.w8 --D-----L--------I----------T----------D-----P--V--L-D----------N--S-A.-.-.V.---------------------------V----------- 300
C.US.14.M0118v1d14_5M04_C5 ------L-----------I----------T----------D------NP--------A.-.S.----------V-----------------F--- 305
C.ZA.13.DEMC13ZA152       ------L-----------I----------T----------D----L--Q--V---V------A.-.S.---D------M--------------F-C- 302
C.ZM.11.DEMC11ZM006       ------L-----------I----------T----------D----L--Q--V---V------A.-.S.---D------M--------------F-C- 302
D.BR.10.10BR_RJ095        ------------------I----------------------D--------------A.-.S.-A.------------------------- 304
D.CD.03.LA17MuBo          ------L---------------N------------------T----V--------------A.-.S.-A.------------------------- 304
D.CM.10.DEMD10CM009       ------L-----------I--D--------A----------T----L--Q--V--------A.-.S.---D------------------------ 303
D.CY.06.CY163             ----------I-S-------I------------H--------L--V--------------D-.-.S.-.V.------------------------ 304
D.KE.11.DEMD11KE003       ------------------I----N--I---------A----------L--Q----HP-------------A.-.S.---D------------------------ 304
D.SE.12.077UG             ------------------I----------------------L----Q----HP----------------A.-.S.------------------------ 304
D.TZ.04.CO6405V4          ------------------I----N--------------A-----------L--Q----HP--HL----------A.-.-------------------------- 305
D.UG.10.DEMD10UG004       ------------------I--------------------D--------------A---S----.S.------------------------ 304
D.UG.11.DEMD11UG003       ------------------I----N------------------D----------------S----.S.------------------------ 303
D.YE.02.02YE516           ------------------I----N------------G---L-AQ--------I-------V--E--R--.S.------------------------ 302
F1.AO.06.AO_06_ANG32      ------------------I---------------------------PQ----I-------------Q--.S.-.V.--------L-----------F-C- 301
F1.AR.02.ARE933           ---Y-SL-----------I----------------------D--L-A--P-------------P---Q--.S.-.V.---------------------K-----------F-A- 300
F1.BR.11.DEMF111BR037     ------SL-----------I--------A--D--------D----------VP----------Q--.S.-.V.---D----------------K-----------F--- 301
F1.CY.08.CY222            ------S-----------I----------T--------D----D--L------------------Q--.S.-.V.---D-----------------------F--- 300
F1.ES.11.VA0053_nfl       ------L-----------I----------------------D--L-----------P--I-----Q--.S.-.V.--------------------K-----------F--- 300
F1.FR.04.LA22LeRe         ------P-----------I----------------------D--D--L-------P--I----Q--.G.-.V.----------------------F-V- 300
F1.RO.03.LA20DuCl         ------S-----------I----------------------D--L--PV----P-----------Q--.S.-.V.---------------------K-----------F--- 302
F1.RU.08.D88_845          ------S-----------I------------------D--------P----I------------Q--.S.-.V.--------------------K-----------F--- 300
F2.CM.10.DEMF210CM007     ----------------V-----------------------D----------Q----I----------A.-.S.-.V.---------------------K-----------F--- 300
F2.CM.11.DEURF11CM026     ------X-----------I----------------------D----L-AQ--------------A.-.S.-.V.-----------------------K-----------F--- 297
G.CD.03.LA23LiEd          ------------------S--I-------D--I-PQ-----P--I-------------R--.S.-.V.---------------------------V----------F-A- 304
G.CM.08.789_10            -PI--L-----------------------S-----------D--M--PQ----I-----------R--.-----------------------------V----------F--- 304
G.CM.10.DEMG10CM008       -P--L-------------I----------------------D--L--PQ----FP--I-D----N----A.-.-----------------------V----------F--- 304
G.CN.08.GX_2084_08        ------SL----------I----------T----------D--L--PQ----FP--I----------S----T-#--.S.-------RN#----------V----------F--- 301
G.CN.16.224GX             ------SL----------I----------T----------D--L--PQ----FP--I----------S----T-#--.S.-------RN#----------V----------F--- 305
G.ES.14.EUR_0033          ------P-------------------------------------D--L--Q--------P--I------------R--.------------------------V----------F--- 303
G.GW.08.LA57LmNe          ------P-----------------------T----------I----PQ----FP--I-D----N----A.-.-----------------------V----------F--- 300
G.KE.09.DEMG09KE001       ----------A-----------------T----------D--M--QQ-----I-------------T--.S.------------------------V----------F-V- 304
G.NG.12.12NG060409        ------------------I----------T----------D--L--PQ----I--------N----T-.I.G.--------------M------L----V-----------K----F--- 303
H.CD.04.LA19KoSa          ------------------I--R--------------D--T---D--L--------Q----------N----A.-.-A.---------------------------V----------K----F--- 307
H.CF.02.LA25LeMi          -P----------------------------------A----D----L----------------N----A.-.S.----------------------V----------K----F-C- 301
J.CD.03.LA26DiAn          -P--L-------------I----------A---------D----------L------V--------N----A.-.-A.---------------------------V----------K----F-A- 304
J.CM.04.04CMU11421        -P--L-------X------------------------D--------X-----P--A--------A.-.G.---------------------------V----------F--- 304
K.CD.97.97ZR_EQTB11       ------L-----------I------------------D--M--Q----P--I----------T--.S.----------------------V----------FRV- 300
K.CM.96.96CM_MP535        ------L-----------I------------------D--------L----------------------V--------A.-.----------------------V----------F--- 300
01_AE.CM.11.1156_26       --I---------------G-N-----------------M--I-------D------------NP--I---------A.-.-------------------V-----------------F--- 304
01_AE.CN.12.DE00112CN011  ------PV-----------G-N-----------------M--I-------D----------Q--P--V--------A.-.-------------------V-----------------F--- 304
01_AE.IR.10.10IR.THR48F   ------SL----------G-N------D-----M--I-------Q----P--V----------A.-.-------------------I----- 304
01_AE.JP.11.DE00111JP003  --T--PV-----------G-N-----------------M--I-------D----------Q--P----XXS-I-X---X--.S.-------------------------- 304
01_AE.MM.14.fKSDU26       ------------------G-N-----------------M--I-------D----------P--A.-.------------------------V----------------- 297
01_AE.PH.15.DE00115PH012  ------PV-----------G-N-----------------M--I-------IQ---P--L----.G.-.S.--D------------------V-------- 304
01_AE.SE.12.070SE         ------PV-----------G-N-----------------M--I-------D----------Q--P--FPA----.S.---D--------------------V----------F--- 301
01_AE.TH.10.DE00110TH001  ------P-A---------G-N-----T--------A-----V------A--------T----P--.G.-A.---------------------------- 301
01_AE.TH.11.40436v09_04   ------PV-----------G-N-----------------M--I-------D----------P--------X--------A.-.-------------------------- 304
01_AE.TH.90.CM240         --A--PL-----------G-N-----------------M--I-------D----------P------.S.------------------------------- 304
02_AG.CM.10.DE00210CM013  ------SM-----------------------------------M--I-------D------------P--.S.---------------------------V----------F--- 306
02_AG.DE.09.701114        ------V-----------I----------T----------D----------P--------N----.S.-.V.---D-----------------V----------------------- 306
02_AG.KR.12.12MHI11_10746 --T--PM---------AI-K--------------------M--I-----R--P----------.-------------------V----------K----I--F--- 301
02_AG.KR.12.12MHR9        --T--PM-----------I----------T----------D----------P--------A.-.S.-.A.------------------------------FR-- 301
02_AG.LR.x.POC44951       --T--S------------AI------------------M--I-------D----------Q----P--.S.-.V.-------------------------V----------F--- 301
02_AG.NG.12.12NG060304    ------SM-----------AI----------T----------M--I-------D----------P--D------.S.---------------------------V----------F--- 301
02_AG.NG.x.IBNG           --T--SM-----------I---G-------------------M--I-------D----------P------.S.-.V.--------------------------V----------F--- 301
02_AG.PK.15.PK032         ------S-----------I----------T----------M--I-------D--I-AQ---FP--L----.S.-.V.--------------------VG---K----F--- 301
02_AG.SE.14.098GN         ------VL--K-F-V-MQ-L.KN----------AD--------V--I-------D------------I-----P--.S.-.V.-------------------V-----------------F--- 299
```

CyPA binding

major homology region

```
B.FR.83.HXB2          QMVHQAISPRTLNAWVKVVEEKAFSPEVIPMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGW.MTN.NPP..I.PVGEI.YKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTL   304
02_AG.SN.13.9580                                                                                                                                                                         305
03_AB.RU.97.KAL153_2                                                                                                                                                                     305
04_cpx.CY.94.94CY032_3                                                                                                                                                                   305
05_DF.BE.x.VI1310                                                                                                                                                                        305
06_cpx.AU.96.BFP90                                                                                                                                                                       305
07_BC.CN.98.98CN009                                                                                                                                                                      301
08_BC.CN.97.97CNGX_6F                                                                                                                                                                    299
09_cpx.GH.96.96GH2911                                                                                                                                                                    304
10_CD.TZ.96.96TZ_BF061                                                                                                                                                                   304
11_cpx.CM.95.95CM_1816                                                                                                                                                                   304
12_BF.AR.99.ARMA159                                                                                                                                                                      300
13_cpx.CM.96.96CM_1849                                                                                                                                                                   304
14_BG.ES.05.X1870                                                                                                                                                                        303
15_01B.TH.99.99TH_MU2079                                                                                                                                                                 304
16_A2D.KR.97.97KR004                                                                                                                                                                     302
17_BF.AR.99.ARMA038                                                                                                                                                                      306
18_cpx.CU.99.CU76                                                                                                                                                                        304
19_cpx.CU.99.CU7                                                                                                                                                                         304
20_BG.CU.99.Cu103                                                                                                                                                                        304
21_A2D.KE.99.KER2003                                                                                                                                                                     302
22_01A1.CM.01.01CM_0001BBY                                                                                                                                                               302
23_BG.CU.03.CB118                                                                                                                                                                        302
24_BG.ES.08.X2456_2                                                                                                                                                                      304
25_cpx.CM.02.1918LE                                                                                                                                                                      303
26_A5U.CD.02.02CD_MBTB047                                                                                                                                                                301
27_cpx.FR.04.04CD_FR_KZS                                                                                                                                                                 301
28_BF.BR.99.BREPM12609                                                                                                                                                                   304
29_BF.BR.01.BREPM16704                                                                                                                                                                   315
31_BC.BR.04.04BR142                                                                                                                                                                      303
32_06A6.EE.01.EE0369                                                                                                                                                                     301
33_01B.ID.07.JKT189_C                                                                                                                                                                    303
34_01B.TH.99.OUR1969P                                                                                                                                                                    302
35_AD.AF.07.169H                                                                                                                                                                         304
36_cpx.CM.00.00CMNYU830                                                                                                                                                                  303
37_cpx.CM.00.00CMNYU926                                                                                                                                                                  303
38_BF1.UY.03.UY03_3389                                                                                                                                                                   300
39_BF.BR.04.04BRRJ179                                                                                                                                                                    303
40_BF.BR.05.05BRRJ055                                                                                                                                                                    305
41_CD.TZ.05.CO6577V5                                                                                                                                                                     303
42_BF.LU.03.luBF_01_03                                                                                                                                                                   303
43_02G.SA.03.J11223                                                                                                                                                                      301
44_BF.CL.00.CH80                                                                                                                                                                         306
45_cpx.FR.04.04FR_AUK                                                                                                                                                                    304
46_BF.BR.07.07BR_FPS625                                                                                                                                                                  300
47_BF.ES.08.P1942                                                                                                                                                                        303
48_01B.MY.07.07MYKT021                                                                                                                                                                   304
49_cpx.GM.03.N26677                                                                                                                                                                      304
50_A1D.GB.10.12792                                                                                                                                                                       303
51_01B.SG.11.11SG_HM021                                                                                                                                                                  302
52_01B.MY.03.03MYKL018_1                                                                                                                                                                 304
53_01B.MY.11.11FIR164                                                                                                                                                                    304
54_01B.MY.09.09MYSB023                                                                                                                                                                   304
55_01B.CN.10.HNCS102056                                                                                                                                                                  304
56_cpx.FR.10.URF5_patient_A                                                                                                                                                              304
57_BC.CN.09.09YNLX19sg                                                                                                                                                                   301
58_01B.MY.09.09MYPR37                                                                                                                                                                    303
59_01B.CN.09.09LNA423                                                                                                                                                                    300
60_BC.IT.11.BAV499                                                                                                                                                                       300
61_BC.CN.10.JL100010                                                                                                                                                                     301
62_BC.CN.10.YNFL13                                                                                                                                                                       297
63_02A.RU.10.10RU6637                                                                                                                                                                    301
64_BC.CN.09.YNFL31                                                                                                                                                                       301
65_cpx.CN.10.YNFL01                                                                                                                                                                      304
67_01B.CN.11.ANHUI_HF115                                                                                                                                                                 304
68_01B.CN.11.ANHUI_WH73                                                                                                                                                                  304
69_01B.JP.05.05JPMYC113SP420                                                                                                                                                             304
70_BF1.BR.10.10BR_PE004                                                                                                                                                                  307
71_BF1.BR.10.10BR_PE008                                                                                                                                                                  304
72_BF1.BR.10.10BR_MG002                                                                                                                                                                  304
73_BG.DE.01.9196_01                                                                                                                                                                      303
74_01B.MY.10.10MYPR268                                                                                                                                                                   304
77_cpx.MY.14.14MYNBB090                                                                                                                                                                  303
78_cpx.CN.13.YNTC19                                                                                                                                                                      287
79_0107.CN.15.SX15DT013                                                                                                                                                                  294
82_cpx.MM.13.mSSDU12                                                                                                                                                                     294
83_cpx.MM.13.mSSDU94                                                                                                                                                                     297
85_BC.CN.14.14CN_SCYB2                                                                                                                                                                   301
86_BC.CN.13.15YNHS18                                                                                                                                                                     284
87_cpx.CN.12.DH32                                                                                                                                                                        298
88_BC.CN.05.05YNRL25sg                                                                                                                                                                   304
90_BF1.BR.11.BRG06043                                                                                                                                                                    306
O.CM.96.LA51YBF35                                                                                                                                                                        303
O.CM.96.LA52YBF39                                                                                                                                                                        303
O.CM.99.99CMU4122                                                                                                                                                                        303
O.CM.x.pCMO2_3                                                                                                                                                                           303
O.ES.01.Read25_HIV_Group0                                                                                                                                                                303
O.FR.06.RBF206                                                                                                                                                                           303
O.GA.10.10Gab1190                                                                                                                                                                        303
O.GA.11.11Gab6352                                                                                                                                                                        302
O.SN.99.99SE_MP1300                                                                                                                                                                      303
O.US.10.LTNP                                                                                                                                                                             303
N.CM.15.S4858                                                                                                                                                                            305
N.FR.11.N1_FR_2011                                                                                                                                                                       307
P.CM.06.U14788                                                                                                                                                                           297
P.FR.09.RBF168                                                                                                                                                                           301
CPZ.TZ.06.TAN5                                                                                                                                                                           331
CPZ.US.85.US_Marilyn                                                                                                                                                                     305
GOR.CM.12.SIVgor_BQID2                                                                                                                                                                   303
GOR.CM.13.SIVgor_BPID15                                                                                                                                                                  297
```

HIV-1 Proteins

```
                                                                    p24 end_p2 start        p2 end_p7 start         Zn motif          Zn motif        p7 end_p1 start      p1 end_p6 start
B.FR.83.HXB2              RAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQV.......TN.SATI.MMQR....GNFRN..QRKIVCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCT...ERQ..ANFLGKIWPSYK....GRPGNFLQ 450
A1.CD.02.LA01AlPr         -----T-A------------------------S--R---TG-S----------------S---------A....QG..TN-----------KG..-KR.I----------------L-----------------------------R---N-......P-  454
A1.CM.08.886_24           -----T---H------------------------S--R---TG-S---------------------------QQ..TS-----------P-G...-I------------------L------------------------------------P-  448
A1.ES.15.100_117          -----T---G------------S----S--R----TG------------------------------------QQ..SN--V---------KG.I---------L-S--------------------------------------------P-  448
A1.KE.11.DEMA111KE002     -----CT-----D-----------S--R-RG---G-------------S---VNV-----------------QQ.HTM------------KG.I-------------------------------------------------H--------P-  448
A1.PK.15.PK034            -----T---G----D--I-------S--R--GG---------------S----------A--------------QQ..TN-----------KG..-KR.I---------------------------------------R---N-......P-  449
A1.RW.11.DEMA111RW002     -----T-D------I--------S--R---GG----------------------H----------------QH..TN-----------G..-K-.I--------------K-------------------------------L--N-......P-  449
A1.SE.13.078SE            -----T-D------I--------S--R---SG----------------S----A-------------------QH..TN-----------GG.-K-.I----Y----------K-------------------------R---N-......P-  449
A1.UG.11.DEMA110UG009     -----T-D------------------R--AG---------------A.IS--V-E-RRM-------M-..T----LKG....-KR.I----------------------------------------------------------PN-......P-  450
A2.CM.01.01CM_1445MV      -----T-D------------------S--R---TG---------------------------------QH..TN-----------G..-KR.I--------------------------------------------------N-......P-  446
A3.SN.01.DDI579           -----T-D-------------D-----R--AG----------------------------------A.IS--V-E-RRM-------M-..T----LKG....-KR.I----------------------------------------N-......P-  446
A4.CD.97.97CD_KCC2        -----T-D-------------D-----R--AG----------------------------------A.IS--V-E-RRM-------M-..T----LKG....-KR.I----------------------------------------N-......P-  448
A6.BY.13.PV85             -----T-D------------------R--G---------------------S----K.S--G..PKR.I----------------L-------------.T----------------R---N-......P-  449
A6.IT.02.60000            -----T-D------------------R--G---------------------QK..TN----K.S--G..PKR.I----------------L-----------------------R---N-......P-  448
A6.RU.11.11RU6950         -----T-D------D-------------R--SG------------------Q..-N----K.S--G..PKR.I-----------R------------------------R---S-......P-  444
A6.UA.12.DEMA112UA024     -----T-D-------------D-----S------G---------------M-..Q..TN-M--K.S--G..PKR.I-----------L------------------------R---S-......P-  449
B.AR.14.DEMB14AR003       -----T-------------------------G----------------------S---K------------PI-----------I---------------------------L--H-......  450
B.BR.10.BR_MG035          -----------D---------D-----R---G------------------R..S-V-----------PI-----M-K---R-----E---K------------------L--H-......  450
B.CU.14.14CU005           -----------D--X-------------------G-N--------------S---A-------H-V---K------I-K---R------------------------------H-......  450
B.DE.13.366396            --------------------------------------G-----------------S----------------K------I-------------------------------------H-......  450
B.DE.13.947915            -----T------------------------S---------------------L----TAV--K-R-G----TI---L----------R----------------------Q-......P-  453
B.DO.11.DEMB11DR001       -----------D----------S-----------------------------L-------K-----K--I-K-----R-----------------------------------H-......  455
B.ES.15.100600            -----------D----------S------------------------------PTAV--K------K----I-K----R------------------------------------H-......  450
B.FR.11.DEMB11FR001       -------------------------------------------------------------K----KRSI------I-K----------------------------------H-......  450
B.JP.12.DEMB12JP001       -----------------------------------------------------PT------------KRTI----I-----------------------------------H-......  450
B.PH.15.DEMB15PH003       -----------D----------I--S------------S---------A---K--S----------T-----D-I--------R-----------------------H-......  450
B.RU.11.11RU21n           ---------------------------------------------------------YK------T-------D-I-K--------------------------H-......  455
B.SE.15.102SE             -----------D-------------------------------S---I----TNSAA-NV----S--K-----D-I-K---D-----------------------------H-......  452
B.TH.10.DEMB10TH002       -----------D----------------------------------------PT--L----K-.-KR------D-I-K--------RF---H-......  450
B.US.16.2609              -----T--------------------------------------------AT.PP--S--------------D-I---------------------------H-......  455
C.BR.11.DEMC11BR035       -----T-D-----D-----------R---G-SI---------------A-----N-.TS----K.S--KG.P-RTI----V-L-----D-------------------447
C.CN.10.YNFL19            -----T-D-----D-----------R---G-SI---------------A----NS--N-L----S--KG.PKR--------I-K------------------445
C.DE.10.622166            -----T-D---------D--RG--TG---------------------A----N.--N-LI---S--KG.PKRT-----I-K---------------------445
C.ES.15.100_118           -----T-----------RG--G-----------------------------KS.TN----K.S--KG.PKRTI----R--L--K--------Q-------444
C.ET.08.ET104             -----T-D-----D-----------RG--G-S---------------A----H-.-TN----K.S--KG.P--PI-------I-K----------------447
C.IN.15.NIRT008           -----T-D-----D-----------RG--G-S--------------A----N-.-TN--L---S--KG.PKRP--------I-K----------------445
C.MW.09.703010256_CH256.w96 -----T-D-------------S-----S--I------A----N-.TN-L----S--KG.T-R--------I-K------------------445
C.NG.10.10NG020523        -----T-D-----D--I----I--G-TG-S-----------------A----NH..-N-----N-KG.AKR-----------I-K-----R--------446
C.NP.11.11NP016           -----T-D-----D------------R--G-----------------A----NS----A----S--KG.SKRTI-------I-K----------------441
C.PK.14.DEMC14PK009       -----T-D-----D------------R--G----------------T---------A----S--KG.PKRTI----L--------I-K-----N------445
C.SE.15.100ET             -----T-D-----D------------R--G----------------A----N-.-TN-----K.S--KG.P-RTI-------I--------------------447
C.TZ.08.707010457_CH457.w8 -----T--D----D--I--------G----------------------A----N-.TN----S--KG.TKR-I--------I---------------447
C.US.14.M0118v1d14_5M04_C5 -----T-D-----D----------TG----------------------A----NT.-TN--L----S--KG.PKR-I-----I-K----R---------450
C.ZA.13.DEMC13ZA152       -----ST-D----------------R--QG-S----------------A----G-.HPS---L---S--KG.P-R-I-----I-K------------------448
C.ZM.11.DEMC11ZM006       -----CT-D----------------S---------------------A----NH..-N----S--KG.AKR---------V--I-K----Q---------L--H-......  447
D.BR.10.10BR_RJ095        -----------D-----------------Q---------------------TGSTT-A---------KG.---S-------I-K----E-----452
D.CD.03.LA17MuBo          -------------G--D-----------------------------S------SA--A-------KG..-S---------I-K----R-----------451
D.CM.10.DEMD10CM009       -----------D-----------------G-----------------S------I--A----N-.--V-----SKRT------I-K----R-----------448
D.CY.06.CY163             -----------D-----------------G-----------------KA----AT.NTA------KS.PKR------I-K---------------S-H-450
D.KE.11.DEMD11KE003       ------------N----M--G-----------------------S----A-----SVN-A----N-.-KG.PKR------V-K---R--------Q-----451
D.SE.12.077UG             -----------D-----------------G---------------S------A----QA--AV---K----KS.P-R-I-------I-K----R-------451
D.TZ.04.CO6405V4          -----------D-----------------Q---------------RA-----GAA--V--K.P---I---------I-K-----R-------452
D.UG.10.DEMD10UG004       ------------------------------------------------S-----------VN-A-------KYTG.P--SI-------L-K------RF---451
D.UG.11.DEMD11UG003       -----G--D---------------------M-----------------S-----AN-A------KG.PKR-I------I-K----R--R----------N--R--S-449
D.YE.02.02YE516           -----------D-----------------G-S----------------S------IA-----SNN-A-------K-.P--------I-K----------449
F1.AO.06.AO_06_ANG32      -----T-D---G--D----------SGGSI-------------------Q-.-N-.A-TAM------S--KG.-TI----------R-------N-----446
F1.AR.02.ARE933           -----T-D---G--D-----------G-S-----------------A-----S.T-V--K----S--KG.-KR----V--I-K-----R--------N---445
F1.BR.11.DEMF111BR037     -----T-D---G--D-----------G-S-----------------A-----G.-AV.LL-K--S--KG.----I-K--------R------N-----446
F1.CY.08.CY222            -----T---G--D---------------G----------------I--A-----TA--K---S--KG.-T--------I-K-----------N----444
F1.ES.11.VA0053_nfl       -----T---G--D-----------------G------------------TK--L-N--S--KG.-T-I-K-----E--------------N---445
F1.FR.04.LA22LeRe         ------D-G--D------------TG-----------------S-----KA------T---L-N--S--R----I-K-------------------N--444
F1.RO.03.LA20DuCl         ------T---G--D---------------G----------------A----S.TAV----K---S--KG.-K-----V-K----R--------------N--444
F1.RU.08.D88_845          -----CT--G--D----------R------------S-----------I--A-----S.TA---K.S--KG.-K----R----I-K----R------R---P-451
F2.CM.10.DEMF210CM007     -----T-D---G--D-----------G-S----------------S----I--A-----S.-N-.-K---S--KG.-R-----I-K------------L-PS-444
F2.CM.11.DEURF11CM026     -----T-D---G--D-----------G-S--------------------I--KA----Q.-SA---K.S--KG.--I-K-----R----QG------442
G.CD.03.LA23LiEd          -----T---G--------A-------R----------------------N-A...GA..TAM------S--KG.PKRNI---------S--------I-449
G.CM.08.789_10            -----T-D------D--------------RG--SG-----------------SGAG-A----K...S--KG.PKRNI-------R-----------L--H-451
G.CM.10.DEMG10CM008       -----T-D------------------------TGTA-----------------SG.A-A----K...S--KG.P-RXI---------R-----------------450
G.CN.08.GX_2084_08        -----T-D---------------#--RG--TG--------------S----------NA----K...S--KG.P-RNI---------R-----------------451
G.CN.16.224GX             -----T-D-------------#----RG--TG-----------------S-----NA----SGAA-A----K...S--KG.P-RNI---------R-----------------451
G.ES.14.EUR_0033          -----T-D------D--------------R--G--------------S-----KA------SG.--V---K...S--KG.P-RMI------L-K-------448
G.GW.08.LA57LmNe          -----T-D------D--------------R--G--------------S----A......SG..---K...S--KG.P-RMI----------E----S-445
G.KE.09.DEMG09KE001       ------D-S--D--------------QG--------------------S----A....SL..PS---K.S--KG.P-RTI----R------H-449
G.NG.12.12NG060409        -----T-D------D----------AG-------------S-----------SSTA-A----K...S--KG.-KRNIR-------I-------R---------450
H.CD.04.LA19KoSa          ---------G--D-------------R--QG-SV------------A----AN-A-......-K...S--KG.P-RTI------R----------N------454
H.CF.02.LA25LeMi          -----T-D------D--------------R--QG--G--------------S----A------TAN-A-......-KG.-I-K-----R----------N------448
J.CD.03.LA26DiAn          -----T-D-----------SG--------------------------A...TNV...---KG.-H-R----I-K-------------V-------449
J.CM.04.04CMU11421        -----T----D--I-----R-----------M------------------.-.TN-X----KG.-KRMI----P-L---------Q-----------N------450
K.CD.97.97ZR_EQTB11       -----T----------------R-----SG--------------------SAV...---KG.--R-I---L-----------S-------F-LN...E------445
K.CM.96.96CM_MP535        -----T------D-----------G-S----------------S--------PVV...-K....S--KG.H------I---L----------------H-445
01_AE.CM.11.1156_26       -----T----------S---TG---------------------M......Q-.-N----G.-GR.I-----------------------S----P-448
01_AE.CN.12.DE00112CN011  -----T----------S---TGV-----------------------S.......Q-.-N----KD..-KR.I----R--L------------P-----V---RL--N-.....P-448
01_AE.IR.10.10IR.THR48F   -----T----------S---TG-S-----------------HT......QQ.-N-L----KG.-KR.I------------------------L--N-.....P-448
01_AE.JP.11.DE00111JP003  --X---T----------S---TG-S--------------H-.....QQ.TN----KG.P-R.I-----------------R---L--N-.....H-448
01_AE.MM.14.fKSDU26       -----T---------S---TG-S--------------T......QQ.-VN----KG.-KR.I-----------------------RL--N-.....P-441
01_AE.PH.15.DE00115PH012  -----T----------S---TG----------------S......QH.-NV----KG.PKR.I--------------------------N-.....P-448
01_AE.SE.12.070SE         -----T----------S---TG--------------------QH..TN----KG.-KR.I---------------------------N-.....P-446
01_AE.TH.10.DE00110TH001  -----T----------S---TG--------------H-......QH.-NV----KG.-KR.I-----------------------------N-.....P-448
01_AE.TH.11.40436v09_04   -----T----------S---TG-----------------S......QH.-NV----KG.-KR.I--------R----------R---L--N-.....P-448
01_AE.TH.90.CM240         -----T----------S---TG--------------HA......QH.----KG.-KR.I--------------Q-----------N-.....P-448
02_AG.CM.10.DE00210CM013  -----T----------S--R--G--------------S......QQ..TN----G.-R-T.I--------------K----------------H-.....P-450
02_AG.DE.09.701114        -----T----------S--R--SG-------------S......QQ..TN----G.-T.I---------K------------------N-.....P-450
02_AG.KR.12.12MHI11_10746 -----T----------S--R--SG-----------GNFR---G.-T.I---------K---------R---------------N-.....P-445
02_AG.KR.12.12MHR9        -----T----------S---A-R-G--------------S......QQ..PN----G.-T.I---------K---------P---------------N-.....P-445
02_AG.LR.x.POC44951       -----T----------S--R--SG-------------S......QQ..TN----G.-R-T.I--------------K---------------S-.....P-445
02_AG.NG.12.12NG060304    -----T---A------S--R--SG------------Q......QQ..SN----G.--T.I---------K-R--R-----------------S-.....P-446
02_AG.NG.x.IBNG           -----T----------S--R--SG-------------S......QQ..SN----G.-T.I---------K----------------------S-.....P-445
02_AG.PK.15.PK032         -----T----------S--R--SG---------------G......QQ..SN----G.-T.I---------K---------R----------------S-.....P-445
02_AG.SE.14.098GN         -----T----------S--R--SG--------------------QQ..SS----G.-T.I---------K---------R------------S-.....P-443
```

Gag

HIV-1/SIVcpz Proteins

	p24 end_p2 start	p2 end_p7 start	Zn motif	Zn motif	p7 end_p1 start	p1 end_p6 start

```
                                                     p24 end_p2 start           p2 end_p7 start              Zn motif              Zn motif        p7 end_p1 start     p1 end_p6 start
B.FR.83.HXB2                    RAEQASQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQV.......TN.SATI.MMQR....GNFRN..QRKIVKCFNCGKEGHTARNCRAPRKKGCWKCGKEGHQMKDCT...ERQ..ANFLGKIWPSYK....GRPGNFLQ 450
02_AG.SN.13.9580               ------T-D---------------------------R----G--------------------A----QQ..SNV---------G...--T.I--------------L----N---------------------R---FS-.-------P- 449
03_AB.RU.97.KAL153_2           ------T-D-----------------------------SG--------------------------Q...-N---...-K...S--G..PKR.I--------------L------------------------R---S---.-------P- 448
04_cpx.CY.94.94CY032_3         ------T----------S--------TG------------------------------S------A--S-AA-A----K--SK-KG.--RTI--------------L----------------------------RM---S-.-------P- 451
05_DF.BE.x.VI1310              ------T----------------------QQ--------------------------------A-----SA--A--...-K...--G..P--.I--------------L------------I-K----G----R---------------V--H- 452
06_cpx.AU.96.BFP90             ------T-----D--------------R----G-------------------------S------A---SV..GA---K...S--KG.PKRSI--------------L------------------------------------------H- 448
07_BC.CN.98.98CN009            ------T-D--------------------R-G-SI----------------------S--------A----NS...-L--K...S--KG.SKR----------------------------------------------------------H- 445
08_BC.CN.97.97CNGX_6F          ------T-D--------------------R--SG-S----------------------S------T----N...--L--K...S--KG.SKR--------------I-K---------------------------F---H- 443
09_cpx.GH.96.96GH2911          ------T-D----*--------------------------------------------------A--Q-..TN--L--...-----XL----------------------------------------------N---------------P- 447
10_CD.TZ.96.96TZ_BF061         ----------------------------------------------S--------A----G.GN---------KG..PK-SI------------RR-----R----L--P..ER---------N------------451
11_cpx.CM.95.95CM_1816         ------T----Q-----D------------S-R--TG----------------QQ..TN--...-K...S--KG..KR.I--------------L------------------------------------------------N---------P- 448
```

```
                                                                                     Vpr binding
                                                                                     p6 end, Gag end
                                   Vpr binding
B.FR.83.HXB2                   SRPE.............PTAPPE...ESFRSG.....VETTT.PPQ....KQ....EPID...KE......LYPLTSLRSLFGNDPSSQ*     500
A1.CD.02.LA01AlPr              ------------------A...-VGM......E-I.--SP.....-.....-QKE.............PP--I--K-------L---     502
A1.CM.08.886_24                ------------------M...-ICGM.....E-I.G.S-L....---.....QK----R-Q.....VP--V--K-------L---     499
A1.ES.15.100_117               N-L---------------A...-LGM--G-M.-S-L.........-QK....-R....AP--V--K-------L---             499
A1.KE.11.DEMA11KE002           -----------------A...-ICGM.....E-I.--L....R-....-QK....R-.....HP-SI--K-------*---          497
A1.PK.15.PK034                 --L---------------A.-I-GMR....E-I.A.-P........-QNS...RDQ....NP-SV--K------LL---            500
A1.RW.11.DEMA111RW002          -------------S---V.-LLGM......E-I.A.S-P....R-....-QR---R-Q....GP--V--K-------L---          500
A1.SE.13.078SE                 -----------A.-DLLGM...-E-I.A.S.P....-....-QK--RNW....EP-SV--K------S-L*-W-                 499
A1.UG.11.DEMA110UG009          --T---------------A.-I-GMR....E-I.I.S-P........-QSN...RDP....SP-SV--K------LL---           501
A2.CM.01.01CM_1445MV           --T---------------A.-NVGM......E-I.A.SSS.....-LRN...R-Q.....HT-TI-------L---               497
A3.SN.01.DDI579                -----------------A.-DLLGM......E-I.-.SSP.....-QK....R-.....PP--V-------L---                496
A4.CD.97.97CD_KCC2             KGL---------------LGM-.....-G-I.-SSL........-IK....RDQ.....QS--I--K-------*---             498

A6.BY.13.PV85                  --------------S---A.-NWGTTG...-E-M.--SL.....-QK....R-Q.....RP-SI--K-------L---             501
A6.IT.02.60000                 -----------------A.-DLWM......E-I.N.-SL.....-QK....R-Q.....HP-AI--K-------L---             499
A6.RU.11.11RU6950              --------------S---A.-N-GM......E-I.--SL.....-QK....R-Q.....HP-SI--K-------L---             495
A6.UA.12.DEMA112UA024          --------------S---A.-N-GM......E-I.I.-SL.....-QK....R-Q.....RP-SV--K-------L---            500

B.AR.14.DEMB14AR003            -----------------------.....-E--I.--S-.......-R--QD-..........-V--K-------L---             501
B.BR.10.10BR_MG035             ----PTAPPA...........-A.....-E----S-.......-Q....-Q-..........-A-----------                507
B.CU.14.14CU005                N-L----------------GP-M.....E-I.--.......-....-GT-....-D.....-A--K-------LL--             501
B.DE.13.366396                 --------------------F.....E---S-.......-....-SK-....-L.....YP----K-------L---              502
B.DE.13.947915                 --------------------F.....E------S-.....-K...DQG-...QG........-A--K----Q*---               503
B.DO.11.DEMB11DR001            -----------APPE---F.....E------S-........-QT-....-E....K----------S--LL---                 509
B.ES.15.100600                 N-----------------M.....E---S-........-....-V-.......---K----------L---                   499
B.FR.11.DEMB11FR001            ------------------F.....E---.-S-........-QTE.......K-.......---L---                       499
B.JP.12.DEMB12JP001            -----------APPE---F.....EK-E-.-S-........-Q....-Q--A--K--------*---                       503
B.PH.15.DEMB15PH003            --L-------------GF.....E-I.--.......-....-M-.........P-FA............-L---                 498
B.RU.11.11RU21n                ------------------F.....E---S-.......-QK-.RDRDKE.......-A----------                       510
B.SE.15.102SE                  N-T---------------A.....-SF-.G-----S-.......-MN.........-N--K--------*---                  502
B.TH.10.DEMB10TH002            ----SRPE.............-A.....---S-.......-....-T-........V---A-----------                   505
B.US.16.2609                   ------------------F.....E-I.--S-.......-....-K-.......M--A--K--------E-                    506

C.BR.11.DEMC11BR035            ------------------A.....-F.....E-I.--SP......-QEN.--D.........-S--L----                    494
C.CN.10.YNFL19                 N-T---PTAPPAE.....-A.....-F.....EET-.--S-.....-S..............-K--------L---               496
C.DE.10.622166                 N---PTAPPAE.......-A.-I-GF...G-.-.HAP....E....GSK-...R-.........-A-K----S--LL--             499
C.ES.15.100_118                ----PTAPPAESFRPA.-A.....-F.....E---.-S-......-SP-..R-....-QR-.........-K----S--L---        501
C.ET.08.ET104                  -----------------A.....-F.....E---.-SL.......-KE-....-QK-.........-K----S--L---            498
C.IN.15.NIRT008                N-T---------------A.-D---F.....E---.LAP....-K....-SK-....-Q-..........-K----S--L---        492
C.MW.09.703010256_CH256.w96    N---PTAPPLE.......-A.....-S..........-AL.......-QK....R-.........-SI--K--------L---        494
C.NG.10.10NG020523             --T---------------A.-KF......E---.--TP........-K-...R-.........-A----K----S--W---          493
C.NP.11.11NP016                ------------------A.-KF......E---.-AP.........-QK--R-.........-A----K----S--L---           488
C.PK.14.DEMC14PK009            ------------------A.-KF......E---.-AP.........-QK--R-.........----K----S--L---             493
C.SE.15.100ET                  ----PTAPTV........-A.-KF......E---.-AP.........-QK--R-.........----K----S--L---            499
C.TZ.08.707010457_CH457.w8     N-T---------------A.-EF......E---.-H-P.......-QKN--R-..........-T----K----S--L---          492
C.US.14.M0118v1d14_5M04_C5     N-T-PTAPPLQSRPE...-A.....-F.....E--.A.FAP.....-K-....R-.........-L----S--L---              508
C.ZA.13.DEMC13ZA152            ----ATAPLQ........-A.....-F.....E-P.-.-AP......-LR-....RD.....P.-I--K----S--L---            502
C.ZM.11.DEMC11ZM006            ----PTAPLQ........-A.....-F.....E---.-AP.......-K-....R-.........-AI----------L---          499

D.BR.10.10BR_RJ095             ------------------A.-GF......E-I.A.-S-........DQK-.............--------S------             501
D.CD.03.LA17MuBo              ------------------A.-GF......E-I.A.-S-......-QG-.KD-.........----A--K-------L---           503
D.CM.10.DEMD10CM009            ----PTAPLE........-A.-GF......E-I.-LS-.......-QR-...-.KD-A--K---------L---                 504
D.CY.06.CY163                  N-----------------A.-GF......E-I.--S-........-QR-..-.........----------L---                500
D.KE.11.DEMD11KE003            G-----------------A.-LGF......E-I.--S-......-QKG.......QC------K---------W---              501
D.SE.12.077UG                  ------------------A.-GF......EKI.N.-.......-QT-....-Q..........-A--K-------L---            501
D.TZ.04.CO6405V4               ------------------A.-GF......E-I.--S-.....-N...-QK-.........--AA--K----S--L---              502
D.UG.10.DEMD10UG004            ------------------A.-C-GF....E-.-N.-S-....-Q...-QK-.........----------------L---           501
D.UG.11.DEMD11UG003            ------------------A.-GF......E-I.--.......-QK-.....EKD....E---------K----S--L---           501
D.YE.02.02YE516                R-----------------A.-GF......E-I.--S-........-QK-.KD-D..............------L---              501

F1.AO.06.AO_06_ANG32           ------------------A.---GFR....E-M.P.--P.......-QKE...EGQGQG.LYP--A--K------Q*---            500
F1.AR.02.ARE933                ------------------A.---GF......E---.-SP.......-QK-....EGL....YP--A--K-------*L--            495
F1.BR.11.DEMF111BR037          --T---------------A.--C-GFR...E---.-AP.......-QKA...EGQ....YP--A--K------*---              496
F1.CY.08.CY222                 --------------S---A.---GFK....E-I.A.-SP....E....-QK-..-GL....YP--A--K-------*L*-            493
F1.ES.11.VA0053_nfl            N-----------------A.---GFR....E-I.--SL.......-QK-..-.EGP.......YP--A-----------             495
F1.FR.04.LA22LeRe              -----------------A.---GFR....E-I.--SP.......-QK-..-.EGQ.......YP--A-----------             495
F1.RO.03.LA20DuCl              NQNRPE.........--S---A.---GF......E-I.--SP.......-QK-..-.EGQ....YP--A--K--------*---        499
F1.RU.08.D88_845               --L---------------S---A.-GF......E-I.--SP....R-....-QK-..-.DQ....CP----K------Q*---         501
F2.CM.10.DEMF210CM007          N-----------------A.-GF......E-I.--SP........-QK-..--.........-P.---A---A----S---L---       494
F2.CM.11.DEURF11CM026          N-----------------A.-GF......E-I.............-QKK...ENL....YP-----K--------L---            487

G.CD.03.LA23LiEd               N-----------------A.-LGF......E-I.A.-AP.......---KE-.............----A--K--------*---        498
G.CM.08.789_10                 N-----------------A.-LGF......E-I.A.-T--....EP....-KE-..-N.....P-----K----S--*-V-           499
G.CM.10.DEMG10CM008            N-----------------A.-GF......E-I.A.-N--.......---KE-....-.....-A----K----S--*-M-           490
G.CN.08.GX_2084_08            N-L---------------A.-LGF......EGI.--.-SP.......---RE-..-.Q-PREK.G--P--A--K-------*---       501
G.CN.16.224GX                 N-L---------------A.-LGF......EGI.A.-SP.......---.G--KE-..............-P--A--K----S---       499
G.ES.14.EUR_0033              N-----------------A.-GF......E-I.APSSP.....RP...--KE-..........T----A----S-----           498
G.GW.08.LA57LmNe              N-L---------------A.-GF......E-I.V.SSP.......---KE-.............----A--K----S--L---         493
G.KE.09.DEMG09KE001            N-----------------A.-GL......E-I.--SP.......-DQKG...R-........-----K----S--L---            497
G.NG.12.12NG060409             ------------------A.-GF......K-I.A.-LP.......---KE-.............TP-------K--------*L--       496

H.CD.04.LA19KoSa               --T---------------A.-GF......E-M.N.-SL........---K-..R-.......----K-------S--L---           503
H.CF.02.LA25LeMi               --T---------------A.-GF......E-M.--SP.......---SR-..-.........S-------------S--L---         497

J.CD.03.LA26DiAn               ------------------A.-GF......E-I.A.SSP.........---KE-............-------K------              497
J.CM.04.04CMU11421             ------------------A.-GF......E-I.--.XXP.......---AKE...............---K------X--X---        497

K.CD.97.97ZR_EQTB11            N-----------------A.-GF......EKI.--SL....R-...-MK-....Q-Q....GP-----K----S--L---           496
K.CM.96.96CM_MP535             ------------------A.-GF......E-I.--SP....R-...-TK-....-.....SP-----K---------L---          496

01_AE.CM.11.1156_26            ------------------A.-LGM......E-I.--SIP......---.KQR-...RS....PN-SV--K------LLL--           498
01_AE.CN.12.DE00112CN011       --T---------------A.-NWGM......E-A.--SSL.......---QK--....H.....HP-SV--K-------L---         497
01_AE.IR.10.10IR.THR48F        ---------------S---A.-NWGM......E-I.A.SST....EH..........PP-SI--K--------L---              492
01_AE.JP.11.DE00111JP003       --L---------------A.-NWGM......E-I.--SSL.......---QK--....DH.....PP--V--K-------L---        498
01_AE.MM.14.fKSDU26            ------------------A.-NWGR......E-M.--SL........---QK--....DQ.....PP--I--K-------L---        492
01_AE.PH.15.DE00115PH012       ------------------A.-NWGME.....E-C.A.S-L.......---QK--....DQ.....HP-SV--K-------L---        498
01_AE.SE.12.070SE              ----PTAPLE........-A.-NWGME.....EGK........................HP-SV--K-------L---             490
01_AE.TH.10.DE00110TH001       ------------------A.-NWGM-.....E-.-.SLL....R-....-QK-....-S.....PP--V--K-------L---         499
01_AE.TH.11.40436v09_04        ------------------A.-NWGM......E-I.........-......-QK-....EGH.....PP--V--K-------L---        495
01_AE.TH.90.CM240              ------------------A.-NWGM-EEITGE-I.--.SLP.........-QK-....-H.....PP--V--K-------L---        504
02_AG.CM.10.DE00210CM013       -----------------LA.-LGM......E-R.P.SSP.........---K-....-GL....YP--A--K-------*L--         500
02_AG.DE.09.701114             --T---------------A.-I-GV......E-I.--SSP......---.-....-R-....-GL....YP--A--K-------L---    500
02_AG.KR.12.12MHI11_10746      --Q---------------A.-NLGM......E-I.--SSP....Q-....-K-....-R-....-GL....YP--A--K-------L---  499
02_AG.KR.12.12MHR9             ------------------A.-LGM......E-I.--S-P......-EQRE.-R-....-GL....YP-----K-------L---        499
02_AG.LR.x.POC44951            ------------------A.-GM......E-I.--S-P........---K-....-GL....YP-----K-------L---           495
02_AG.NG.12.12NG060304         ------------------A.-LGM......E-I.--SSL.....Q.....---K-....-GL....YP-----K-------L---       496
02_AG.NG.x.IBNG                ------------------A.-GM......E-I.P.-SP........Q.....---K-....-GL....YP-----K--------*L--     495
02_AG.PK.15.PK032              ------------------A.--GM......E-I.--S-P........---QR-....-GL....YP-----K-------L---         495
02_AG.SE.14.098GN              --Q---------------A.-TWGMR.....E-I.--.L-P.........---R-....-L....YP-----K--------*---       493
```

Gag

Vpr binding   Vpr binding
                    p6 end, Gag end

```
B.FR.83.HXB2                  SRPE.............PTAPPE...ESFRSG.....VETTT.PPQ....KQ....EPID...KE......LYPLTSLRSLFGNDPSSQ*   500
02_AG.SN.13.9580              -KT-PFPQSRTE.......----A....----.-LGM----E-V.-.-S-----------RG.....-L....YP-----------Q*---   507
03_AB.RU.97.KAL153_2          -----------------S----A....-N-GM----E-I.-.-SL.........-QK....R-Q....HP-SI--K----D-L----   499
04_cpx.CY.94.94CY032_3        N-----------------------A....-CLERK----E-I.A.SSP.........-QK....EGL....YP-A--K------S-L---   501
05_DF.BE.x.VI1310             -------------------S----A....----F.....-E-I.A.-SP.........-QK....EGL....YP-A--K------S----   503
06_cpx.AU.96.BFP90            N-----------------------I....-GF-----E-I.A.-SP.........-SKE....-EK..G----A-K--S------S----   500
07_BC.CN.98.98CN009           --------------------------F.....-EET.-.-S.-----------------T-.-----------------K-------   496
08_BC.CN.97.97CNGX_6F         --------------------------F.....-E-.-.-AP.........---K-.-R.----------------------S-L---   490
09_cpx.GH.96.96GH2911         -------------------S----A....-GMR----E-.-.-SSX.....XKD....G-X....PP--X--K-------L----   497
10_CD.TZ.96.96TZ_BF061        -----------------------A....--GF----E-I.-.-S-----------QK.-..--------PH--A--K------L----   501
11_cpx.CM.95.95CM_1816        -----------------------A....--GF----E-V.A.-S-----------KE.....------------K------S------   497
12_BF.AR.99.ARMA159           N----GPTAPPESRPE.......----A....--GF----E-I.A.-SSP....-R....-QK....EGQ....YP-A--K------*-*-   494
13_cpx.CM.96.96CM_1849        ---GPTAPPESRPE.......----A....--GF----E-I.A.-S-----------KE.....--------T---AA-K--S--*-*-   508
14_BG.ES.05.X1870             N-L----------------------A....--GF----E-I.A.-SP.........-KE.....---------M---A--K--S------   497
15_01B.TH.99.99TH_MU2079      --------------------S----A....-NWGM----E-.-.-SLL.........-QK....R-H....TP--V--K-------LL--   498
16_A2D.KR.97.97KR004          --T--------------------A....-D-GM----E-.-.--L.........-LKN....R-Q....HT-AI--K-------LL--   499
17_BF.AR.99.ARMA038           N-----------------------A....----L.....-ETI.-.-QP.........---K-.-EGL....YP-A--K------*-K-   501
18_cpx.CU.99.CU76             K-----------------------A....-LGF----E-I.A.-SP.........-SKE....-EK......-FS--K--S------   497
19_cpx.CU.99.CU7              --------------------------A....--GF----E-I.A.-SP.........---S-.------------I--K--S--*---   501
20_BG.CU.99.CU103             --------------------------A....--GF----E-I.A.-SP.........-GQKE...-D.......-I--K--S--*---   497
21_A2D.KE.99.KER2003          --S----------------------A....-M-GF----E-I.-.-S.........-QK....-.-.......-SV--K-----LL--   499
22_01A1.CM.01.01CM_0001BBY    G----------------------A....--VGM----E-I.A.-S.........-QRE....-.-.....PP--V--K-----L----   500
23_BG.CU.03.CB118             N-L----------------------A....--GF----E-I.A.-SP.........-G-KE...QD.......-I--K---S--*---   498
24_BG.ES.08.X2456_2           N-L----------------------A....-LGF----E-I.A.-SP.........-GQPE...-D.......-I--K---S--*---   497
25_cpx.CM.02.1918LE           N-----------------------A....-LGF----E-I.-.-SP.........-QK-.-.......---MA--K--S--*---   499
26_A5U.CD.02.02CD_MBTB047     --LV---------------------A....-NLGME----EG-.-.-SP.........-KEGE...-Q.....SP--I--K------   496
27_cpx.FR.04.04CD_FR_KZS      N-----------------------A....-LGIM----E-M.-.-SYP.........-LK-.-.......-P--V--K------L---   497
28_BF.BR.99.BREPM12609        N-----------------------A....--GF----E-.-.-SP.........-QKE....EGL....YP-A--K------*---   499
29_BF.BR.01.BREPM16704        N-------------------S----A....-LGF----E-R.-.-SP.........-QK-.-EGK....YP-A-----------   510
31_BC.BR.04.04BR142           ----PTAPPASRLE.......----A....----.-FE----E-R-.-S-----------QR-.-R.-----------K---S-L---   506
32_06A6.EE.01.EE0369          N-----------------------A....----FE.....-E--AA.-SS.........---KK.....---------SA-------   495
33_01B.ID.07.JKT189_C         --------------------------A....----F.....-EQ.-.-T-S-----------Q-.-.-----------VA--K------   497
34_01B.TH.99.OUR1969P         --------------------------A....----F.....-EG.-.-T-S-----------.-.-......---SA--K------LL--   498
35_AD.AF.07.169H              G-----------------------A....--GF----E-I.K.-S.........-QK....-G....---------K---------   497
36_cpx.CM.00.00CMNYU830       -------------------------A....-N-GI----E-M.S.SS.........-SR-.-GL....YP--------K------*L--   498
37_cpx.CM.00.00CMNYU926       --------------------------A....--EVK----E-V.A.SSL.........---.--G--L....YP-A--K------S--*P-   493
38_BF1.UY.03.UY03_3389        -------------------S----A....--GF----E-I.-.-PKQEQ-----QK-.-EGM....YP-A--K------*---   499
39_BF.BR.04.04BRRJ179         --T-PTA.............----A....-LMF----E-.-A.-S.........-R-.-.-.---------------------   502
40_BF.BR.05.05BRRJ055         N-L----------------------A....--GF----E-I.-.-S-----------QKN....EGL....YP-----K------L---   498
41_CD.TZ.05.CO6577V5          N-S-PTAADRPE.......----A....--GF----E-I.-.-SS.........-QKN....-.-.-----------------L---   508
42_BF.LU.03.luBF_01_03        N---PRPE............----A....--GF----E-I.A.-SL.........-GQK....EGL....CP-------------   503
43_02G.SA.03.J11223           -------------------------A....--GF----E-I.A.-S-----------RE.....---------SP-A--K------W-K-   498
44_BF.CL.00.CH80              N-L----------------------A....--GF----E-I.-.-SP.........-QK-.-EGK....YP-FA--K------*-K-   501
45_cpx.FR.04.04FR_AUK         -------------------S----A....--GM----E-I.-.-SCP.........-QK-.-R-Q....VYP-K------------L---   499
46_BF.BR.07.07BR_FPS625       N-----------------------A....--EIR----E-I.-.-S.........-QK-.-EGL....YP-A--K-------W---   496
47_BF.ES.08.P1942             N-----------------------A....-LGF----E-.-.-SP.........-QKE....GL....YP-A-----------   498
48_01B.MY.07.07MYKT021        --------------------------A....----.----E-I.A.-SP.........-QKE....-.-......-A------S-----   499
49_cpx.GM.03.N26677           N-----------------------A....--GF----E-I.-.-SP.........-QKE....-.-......SA------S--*LP-   496
50_A1D.GB.10.12792            --------------------------A....--GF----E-V.I.-S-----------.-.-......-SA--------K--S-----   498
51_01B.SG.11.11SG_HM021       ------------------------F.....-E-.-.-SP.........---K-.-.-.....Q--A--K--S-----LL--   499
52_01B.MY.03.03MYKL018_1      --X--------------------A....-NWGM----E-X.-.-SSL.........-QX-.-DP....PP-SV--K----S--LL--   500
53_01B.MY.11.11FIR164         N-------------------------F.....-EG.-.-T-S-----------.-.-.----------------K-------   499
54_01B.MY.09.09MYSB023        N--------------------------L.----E-.-.-S-----------.-.-.-------------VSA-K------*---   501
55_01B.CN.10.HNCS102056       -------------------------A....-DWGM----E-I.-.-SSP.........-QK-.--GH....RP-VV--K--S--L---   499
56_cpx.FR.10.URF5_patient_A   -------------------------A....--VGM----E-A.-.-YSP....-R....-SR-.-KD--L....YP-----K--S--L---   500
57_BC.CN.09.09YNLX19sg        --------------------------F.....-E-.-.-S-----------.-LK-.-R.-----------------K------   492
58_01B.MY.09.09MYPR37         ------------------X----------F.....-E-.-.-T-X.........---.-.-G----------------K------   498
59_01B.CN.09.09LNA423         -------------------------A....-NWGM----E-S.-.-LL.........-QK-.--Q....PH-SI--K-------L---   492
60_BC.IT.11.BAV499            -------------------------A....----.----E-.-.-P.........-QK-.-R.----------------K---S-L---   491
61_BC.CN.10.JL100010          -------------------------A....----F.....-E--A.SAP.........---MK-.-R.-------------------S-L---   492
62_BC.CN.10.YNFL13            -------------------------A....----.----E-.-.-AP.........---SK-.-R.---------------------S-L---   488
63_02A.RU.10.10RU6637         -------------------------A....--VGM----EET.-.-S-----------Q.-----G-L....YP-A--K-------LC--   495
64_BC.CN.09.YNFL31            -------------------S----A....----F.----E-.-.-S.-----------GT-.---------A--K---------L---   496
65_cpx.CN.10.YNFL01           --L----------------------A....----F.....-E-.-.-AP.-----R.--SK-.-RG.----------------K---S-L---   495
67_01B.CN.11.ANHUI_HF115      -------------------------A....-DWGM----G-I.-.-FSP.........-QK-.-.-.....H-SV--K--S--L---   497
68_01B.CN.11.ANHUI_WH73       -------------------------A....-DWGM----G-I.-.-SSP.........-QR-.-R.-----Q-V--K-------L---   497
69_01B.JP.05.05JPMYC113SP420  -------------------S----L.-E....-ET--.-LT-----------D-----------------------------K--   495
70_BF1.BR.10.10BR_PE004       N-----------------------A....----.----E-M.-.--PK.....-E....-QK-.-EGR....YP-----K-------W---   504
71_BF1.BR.10.10BR_PE008       N-T----------------------A....--GF----E-M.-.-S-----------QKS....EGL....YP-A--K------W-K-   499
72_BF1.BR.10.10BR_MG002       N-T----------------------A....--GF----E-K.-.-SP.........-KE....EGL....YP-A--K-------WL--   501
73_BG.DE.01.9196_01           N-----------------------A....--GF----E-I.A.-SP.........-KE....S-----A--K------S--*---   497
74_01B.MY.10.10MYPR268        N-L----------------------F.....-E-.-.-A-T.........---N.-.-.----------SA--K------*---   499
77_cpx.MY.14.14MYNBB090       --L----------------------A....----.----E-.-.-TP.........---.--.-.----------SI--K--S-L---   484
78_cpx.CN.13.YNTC19           ----PTAPPAELQSRPE.....----V....----F.....-E-.-.-AP.........---K-.-R.-----------K---S-L---   492
79_0107.CN.15.SX15DT013       --T----------------------A....-DWGM----EG-.-.-S-L.........-QK-.--L....PP-SV--K-------L---   499
82_cpx.MM.13.mSSDU12          -------------------------A....----.----E-.-.-AP.........---K-.-R.----------------S-L---   485
83_cpx.MM.13.mSSDU04          --------------------------F.....-E-.-.-AP.........---K-.-RD....A---------S-L---   488
85_BC.CN.14.14CN_SCYB2        N-----------------------A....-G--F----E-.-.-AP.........---K-.-R.------I---------S-L---   492
86_BC.CN.13.15YNHS18          -------------------------A....----.----E-.-.-AP.........-QK-.-.-.----------------S-L---   475
87_cpx.CN.12.DH32             --------------------------A....----.----E-.-.-AP.........--XK-.-R.----------------LL--   489
88_BC.CN.05.05YNRL25sg        -------------------------A....-N-F.....-E-.-.-AP.........---K-.-R.----------------S-L---   495
90_BF1.BR.11.BRG06043         -------------------------A....----.----E-.-.-S.-SL........-EK.-.---------F---K-----*---   505
O.CM.96.LA51YBF35             KQVS.................-S---M....-EVVKR......-Q-K....QGQ...EN.QTQGGDQ...D------FA--K----T-Q.---   504
O.CM.96.LA52YBF39             KQVP.................-S---M....-EVMK----Q-N.-.-.......-E....KGEP...N------FA--K----T-Q.---   499
O.CM.99.99CMU4122             RQVS.................-S---M....-EE-K----P-S.-.-.......-E....KGDP...N------FA--K----T-Q.---   499
O.CM.x.pCMO2_3                RQVP.................-S---M....-EEMK----Q-N.-.-.......-E....KGDP...N------FA--K----T-Q.---   497
O.ES.01.Read25_HIV_GroupO     KQVA.................-S---M....-EEIK----Q-N.-.-.......-E....KGGP...N------FA--K----T-QLQ-R   502
O.FR.06.RBF206                KQVS.................-S---M....-EMEK----Q-K.E.-K.......-N....GGDQ...N------FA--K----T-Q.---   501
O.GA.10.10Gab1190             KQVS.................-S---M....-EV-KE----Q-N.-.-R.......-E....KGDQ...T------FA--K----T-Q.---   498
O.GA.11.11Gab6352             KQVH.................-S---M....-EEVQK----Q-N.-.-.......-E....KGDQ...N------FA--K----T-Q.---   498
O.SN.99.99SE_MP1300           RQVS.................-S---M....-TEEMK----Q-N.-.-.......-E....KGDQ...N------FA--K----T-Q.---   498
O.US.10.LTNP                  KQVA.................-S---M....-EEVNR----Q-N.-.-.......-E....KGDQ...N------FA--K----T-Q.---   499
N.CM.15.S4858                 TTTRKE..............------L....-NYG-Q----E-R.S.KQG....-E....MQE---Q-KTENSLYP---------S-----   511
N.FR.11.N1_FR_2011            TTTRKE..............------L....-NYGFQ----E-R.S.TQG....-E....MQE---Q-RTENSLYP---------S-----   513
P.CM.06.U14788               KQVQ.................------M....-E.....-E-M.-.-QNKQ...XE....-KE--.-EKE.......-A--K----T-Q*Y--   495
P.FR.09.RBF168               KQVQ.................------M....-E.....-D-M.-.-RDKQ...RE....XEEX--.-E.-----------T-Q*Y--   497
CPZ.TZ.06.TAN5                NT--KGRN..........QEQ-.---....------VQ.....-TPVVP.TAP....-PL....-MPMK..DG........F--K-I--S-Q**--   529
CPZ.US.85.US_Marilyn          N-K.................------I....-D-GYQ----E-.-V.TQE.........---GKE.......PFQ----K----S----E-   506
GOR.CM.12.SIVgor_BQID2        KQAQ.................-S---M....-EVMK----K-N.-.-.......-D....GGDQ...G------FA--K----T-Q.---   499
GOR.CM.13.SIVgor_BPID15       KQVQ.................------M....-E.....-E-V.-.-QSRQ...-E....G-KE.......----K----T-Q*YL--   493
```

HIV-1/SIVcpz Proteins

HIV-1 Proteins

Pol

Pol start
Gag-Pol TF start
Gag-Pol TF end protease start

```
                                   FFREDLAFLQ....GKAREFSSEQTR.............ANSPTR...RELQVW.....GRDNN.SPS....EA....GADR...QG......TVSFNFPQVTLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMSLPGRWKPKMIGGIGGFIKVRQYDQILIEICGHKA 127
B.FR.83.HXB2
A1.CD.02.LA01AlPr                  ----N---Q-...-E-GK-------------------S.----CWDG-........-PF----.----E-........-AS-----I----------V-------------DID---K----------------------Q----K-- 124
A1.CM.08.886_24                    ----N---Q-...-E-GK----------------N-GD-WDG--........-R.L----.--E-.........-T-.-GPTLS--I----------V--E-----------DIE---K----------K-E------K-- 127
A1.ES.15.100_117                   ----N---Q-...-E----------G-----------C--S.--T-WDG--....-D.L----.--E-......T-.GPALS--I----A-R-E--R-----DIN---K--R------K----A------K-- 127
A1.KE.11.DEMA111KE002              ----N---Q-...-RE--K-P-------------------S.--D-WDE--.....-.P----.--T-.....-EG-.......-PPT----I------IV--------------DIN---K--R------K-E------K-- 126
A1.PK.15.PK034                     ----N---Q-...-R--K--P----------------S.--D-WDE---G-.S.----.----EQ........-P-.EST-S--I----------------------DIN---K----------K-------K-- 127
A1.RW.11.DEMA111RW002              ----N---Q-...-RE--K-P-------------GT-GDG-.--.S.L----.--T-.....-EG-.......-GPT-S--I---------------------DIN---K----------K-------Q----K-- 127
A1.SE.13.078SE                     ----A---Q-...-E------A------------T--T-S.--S-GDG-.--.S.L----.---E-.....-EL-.GTT-S--I------MV-------------DIN---K----------KE--------K-- 127
A1.UG.11.DEMA110UG009              ----N---Q-...-RE--K--K-NG-------------S.--D-WDE--.....-.L----.--T-.....-EG-.......-P-.EPT-S--I--------------------DIN---K----------K-E------K-- 127
A2.CM.01.01CM_1445MV               ----N-SQ-...-RE--KL-P--N-----------AS.---RWNG-........S.FLV----.--TEK----.-T-...-H-Y---I---------------------DIN---K-R----------V-----EKR- 127
A3.SN.01.DDI579                    ----N---Q-...-E--K----------------------S.--SPGDG--.R-----.LLP----.--EG-.......-A--S--I-------I--VR-----------DIN---K----------------------K-- 126
A4.CD.97.97CD_KCC2                 ----N---Q-...-RE-----L-P-R-------------PGDG--........-.L----.----DKG--.-P-.AI-LDL--I---------V--VRVE-----------DIN---K----------------------K-- 127
A6.BY.13.PV85                      ----N---Q-...-RE-----------------I-.-S.---GDDG-.....-D.PLP-.--T-.....-TE-.-T-.....-ET------I---------V-V--E------------DIN---K----------------------K-- 128
A6.IT.02.60000                     ----N---Q-...-E-------------------S.---GPLDG-.....-K.PL--.--T-.....-E-.--T-.AS-C---I-------------------------DIN---K----------------------K-- 127
A6.RU.11.11RU6950                  ----N---Q-...-RE--K----------I-.-S.--GK-WDG-.....-.PLP-.--T-.....-EG-.--T-.AS--S--I----------AVR----------DIN---K----------------------K-- 127
A6.UA.12.DEMA112UA024              ----N---Q-...-RE------A---------I-.-S.----WDG-.....-.PLP-.--T-.....-TEG-.--T-.AS--S--I-------------------------DIN---K----------------------K-- 127
B.AR.14.DEMB14AR003                ----T---P-...-----------------ANSPTS.----T---S.-------N---LP----.----EG.SG-..........-LSL--L-------R----------D-N------------------VP----T 127
B.BR.10.10BR_MG035                 -------P-...-E-----P-----ANSPTS.----T---S.----------L----.--------P-.........-S---I-----------------N--K--------V-------T 133
B.CU.14.14CU005                    -------P-...--G----A------------GP-DG-.-------------.----DN-..........-S---S--I-----------------N--R----------VPV---- 127
B.DE.13.366396                     ----N-P-...-E---LP--------------------G.------.----L---.--VE-.---T-.VS--D--I-------S-------------I----------K---VP----- 128
B.DE.13.947915                     ----N-PE...-E---L----------------------G.------.----L---E-.-PGG--.-P-............-SL--I--------V--------------DI---K----------K---VC------V 127
B.DO.11.DEMB11DR001                ----N-P-...-R----P----------SPTR-------N.-G-----G--.PL----.--T---G-.-E---L--I------------------N--K----------E----V--- 130
B.ES.15.100600                     ----N----...-E-GL--P-------------.N.-G-----G--.PL----.---LD-............-K----L--I------------------N--K------------VH-------T 126
B.FR.11.DEMB11FR001                ----N----...------------------------.N.-G-----G--.PL----.---LD-............--S-----------------N--K------------VH---- 126
B.JP.12.DEMB12JP001                ----N---Q-...--LP--A-------SPTR---------K-RD-.-L-----.---R------.--.----S--I---------------D--N--K----------K---VH---- 130
B.PH.15.DEMB15PH003                ----N---P-...--GK------------T--AS.--R----.--G.-----------G-..........-PALC--I----I----E----------D--N--K----------K---VSV---- 124
B.RU.11.11RU21n                    ----N-P-...-E--------------------G-F----.--G---L-----------EG-.QG--QG-.A--LS--I----------I------R-------V-D-Y-----------------VP----Q-- 131
B.SE.15.102SE                      ----N-P-...----P---D---------AS.----L----.--G---F-----.---E----.----I--E--I-------I---------I-----I--N--K----------E-P----- 127
B.TH.10.DEMB10TH002                ----I---P-...-E-------EQTR--------------.----G---PL-----.--T---.........-S---I-------------------I--N--K----------------S----- 131
B.US.16.2609                       ----N-P-...-E-----------------------.----G---PL-----.----EG-.........-N--LSL--I-----R-R--T----------D-E------------------T----TV 127
C.BR.11.DEMC11BR035                ----N-P-...-E--K-PP--------------T--S.-------RG-.--L-------GK-.........-IPL--I-------------------I------K----------------------------IK--T----------------K-- 123
C.CN.10.YNFL19                     ----N---Q-...--G----N--------------S-S.----GRDN.--I---V-----.-----------IPL--I----------V-V--I----------I------K-------------------------P----K-- 127
C.DE.10.622166                     ----N-P-...---P-K-A-ANSPTCR--.--H---S.-----N.RG-.--PL-----.---VE-.........-TL-C--I---------------------------------IN--K----------------T------K-- 130
C.ES.15.100_118                    ----N-PE...-E--K-PP----ANSPTSREFQTS.---SSS.--G---I-.RG-.--PL-----.----E-.........--L--I-------RV----I----------------K--N---------K----------------K-- 133
C.ET.08.ET104                      ----N-P-...-E-----P-----------------S.----RG.--PL-----.---E-.--E-T-.LPGTL-C--I----------------------IN--K----------------------------------K-- 127
C.IN.15.NIRT008                    ----N-P-...-E-----P--N---------SS.--Q---.RG-.--PR-----.---E-.-VK--.-P-.........-AL----I-------IPVRV-------------I---K----------------------P----KR- 123
C.MW.09.703010256_CH256.w96        ----N-P-...-E---P--A-ANSPTTR.----T---S.-------.--H.PR-----.---E-.........-TL---I---------RV------------L-------K----------------------S------K-- 125
C.NG.10.10NG020523                 ----N-P-...---P--NG---------------S.----RG.--PH-----.---EG-.........-TL----I------M-S--------I--N--K----------------S----S-KR- 125
C.NP.11.11NP016                    ----N-P-...---PP-----------------S.-------F.RG-.--PS-.--K-----.---E-.........--S--V--I-------T----------LN--K-------------------K-- 123
C.PK.14.DEMC14PK009                ----NM-P-...-E---LP----------------.-D---S.-------RG-.--PS-----.---EG-.........-TL---I-I-S-RV--T----------ID---------------V--Y--K-- 123
C.SE.15.100ET                      ----A---Q-...--LP---ANSPDS.---.-D---S.----RG.--S-----.---TE-.........-TL---I-------VE---R-------------------IN--K------------H-----EK-- 128
C.TZ.08.707010457_CH457.w8         ----N-P-...-E-GK-Q--N----------S.------R.--.RG-.--S-----.---EK-.........-NL---I----S--V---I---------------DIN--K-------------------K-- 127
C.US.14.M0118v1d14_5M04_C5         ----N-P-...-E--LP--D-TNSSTTSEQTR.------S.------.-RG-.S.LR----.----EG-.........-TL---I----S-EV--LI--------N--K----------EE--MK---KQ- 134
C.ZA.13.DEMC13ZA152                ----N---Q-...-RE-G--------GNSPPS.---T-C--S.-------.-R-A-.-PR-.--T-.....-EG-.........-STL---I----S--V---I----------DIN--K----------EE--------K-- 130
C.ZM.11.DEMC11ZM006                ----T---P-...-E----LPP--------ANSSPS.---------.-RG-.--PR-----.---EG-.........--Y-----I----A--V--I---------IN--K----------E-P----K-- 128
D.BR.10.10BR_RJ095                 ----N---Q-...--G---------------------S.---R-R-.-------G---T-.....-PE-.---------DL--I--E-------------------DID---K----------------PL-----V 125
D.CD.03.LA17MuBo                   ----N-P-...--G-L---------------------S.---R-R-.--------S.PL----.--GG_QR-.......--LS--I----------S--V--I---------IN--K------------------------- 128
D.CM.10.DEMD10CM009                ----N---Q-...--GKL------ANSPTR---------S.---R-R-.--------G.PL----.--TEG-.-K-.......-QG-LS--I----F--V----------DID---K----------P---------- 132
D.CY.06.CY163                      -------V-P-...-A----------------------S.---R-R-.--------G-.PL----.--EG-.--R-.............-LS--I-------------------------IN--K----------P----------- 126
D.KE.11.DEMD11KE003                ----N-P-...--G-LP-G-A---------TH---S.-----PW.----G--L-----.--T---T-.-P-.............-S---I-------------------IN--K----------------------G------Y-- 126
D.SE.12.077UG                      ----N-P-...----LP--------------------S.----R-.------E.K.P----.--T-.---T-.-P-.........--S--I-------------------N--K----------------------------------K-- 126
D.TZ.04.CO6405V4                   -------P-...-E--------------------T---S.---R-.-------------P-.----TEG-.........--CS--I---------------------N--K----------------K---------- 127
D.UG.10.DEMD10UG004                ----I-PE...-----LP----------------S.---V-R-.--------G-.K.PL----.--AT-.....-EG-.........-A--LSL--I--------------------------------------TVD----- 126
D.UG.11.DEMD11UG003                ----Y---P-...--LP--KA----------------S.---------.---------T---.--T-.--T-.-RER-.--RA--L--L--I-------V----------N---Q------------H--SVD---- 128
D.YE.02.02YE516                    ----N---Q-...-E-GKL----------C-S.---R-R-.-----PL----.--EG.QR-.....--S--S--I------------------------N---Q------------E-V-V------- 128
F1.AO.06.AO_06_ANG32               ----N-P-...-E-GK--------------------S.--R-Q-.RG-.--G-A-----.---EG-.R-TGTG.TVS-LS--I-------V--V-VE------------DIN---K----------------T-D---Y-- 131
F1.AR.02.ARE933                    ----N-P-...-E-K-P----------------AS.---R-Q-.RG-.--PL-----.---E-R-T-.-VP-LS--I-----A-N----------------DID------------------N-------- 127
F1.BR.11.DEMF111BR037              ----N-P-...-E--P--D-------------AS.---V-R-Q-.RG-.--PS-----.---ES-R-T-.-VP-LSL--I-----------RV-D-----------DID--------------------N--D---- 127
F1.CY.08.CY222                     ----N---Q-...-E-KL-----------I---AS.---R-Q-.RG-.-S.PL----G-.----EG-.---A-.-VS-LSL--I-----V-VV--V----------------DIN---K-R----------K-------T-D---- 127
F1.ES.11.VA0053_nfl                ----N-P-...-E-K------------T---AS.---R-R-.----G.PL----.---EG-.R-T-.-VS-LSL--I-----V---V--E--------------DIN-T-K----------K--EN--------Y-- 128
F1.FR.04.LA22LeRe                  ----N---Q-...-E-K----------------S.---R-Q-.RG-.--PL-----.---EG-.R-T-.-VP--S--I-----------V---------------------DIN---K----------------H---------- 128
F1.RO.03.LA20DuCl                  ----N-P-...-E-K----PEQTR--------------S.---W-Q-.RGN-.--PL-----.---E-.---S-.-VS-LSL--I------------------------------DIN---K----------------T-D------ 129
F1.RU.08.D88_845                   -------T--SQ-...-RE-----X----------I--AG.G--R-R-.--PL----.----E-.--S-.-VP-LD--I------V--V--V--E-------------DIN---K----------------K--------- 127
F2.CM.10.DEMF210CM007              ----T--SQ-...-RE-------------------AS.---R-R-.RG-.--PL-----.----E-.-G--LS-C--I-----------V--V---------DID------------------------Q-- 127
F2.CM.11.DEURF11CM026              ----N-PG-...-----P--------------------S.---R-.RG-.-------R-EK-.REP-.-VS-LD--I-----------V--V----------DID------------------------------------ 121
G.CD.03.LA23LiEd                   ----A---Q-...-E----L-------------S.---PR-R-.---G-.S.PCP----.----EG-.-K-.....-AI-LS--I----------------------------DIN---K----------------------------K-- 126
G.CM.08.789_10                     ----N---Q-...-E--KLY---------------G.----PW-R-.---G-.S.PHP----.--G-.----E-K-.....-IS-L--I----I------SVRV-E--I----------DIN---K----------E-P----K-- 124
G.CM.10.DEMG10CM008                ----N---Q-...-E-G--P--A--------------W-R-.---G.S.PLP----.--T-.....-EG-.--K-.....-NVRV--I----I------SVRV-E--I----------QID------------------K-- 126
G.CN.08.GX_2084_08                 ----N---Q-...-RE---LP---A-----S.---------R-R.RG-.-S.PLP----.--T-.---EG-.T-AEGE.GVT--SL--I---------VR--I--------QID---------R--------K-E------VP----E-K-- 131
G.CN.16.224GX                      ----N---Q-...-E-------------------S.-------R-.RG-.-S.PLP----.--EG-.....-EG-.......-VT--SL--I---------VR--I----------QID---------R----------K---P----E-K-- 125
G.ES.14.EUR_0033                   ----N---Q-...-E------P--------------.---R-R-.RG-.-SPLLP----.-R-EG.-K-.....-DIPLSL--I---------------------DIN---K----------------------------K-- 127
G.GW.08.LA57LmNe                   ----N---Q-...-E----------------------S.---R-R-.RG-.-S.LLP----.---E-.-K-.....-IPLSL--I----------------------DIN---K----------------------------KR- 127
G.KE.09.DEMG09KE001                ----N---Q-...-R------P--A------------C-.---RAR-.RG-.-S.PLP----.-RPEG-.-E--K-.....-SPL--I------------------------DIN---K----------------------K-- 125
G.NG.12.12NG060409                 ----N-P-...-E------L--------------------S.---R-.KG-.-S.--P----.--PEG-.-K-.....-NT--L-LX-I--------------AVRV------------KIE------------------V-L--E-K-- 123
H.CD.04.LA19KoSa                   ----N---Q-...-E-----L---D--------------S.---R-R.RG-.-E.PL----.--TEG-.-.........-I-L--I---------V---------------DIN---K----------E--VSL---K-- 125
H.CF.02.LA25LeMi                   ----N---Q-...-RE-----P--D-----------S.---R-R.RG-.-D.PL----.--IEG-.-.........-I-LD--I----------V-----------DIN---K----------E-VAM---K-- 125
J.CD.03.LA26DiAn                   ----N---Q-...-RE------L-P-------------S.---R-R-.---G-.S.LLP----.--T-.....-.........-I-S--I----I--V-------------DID---K-----------V------NEVPV--E-K-V 125
J.CM.04.04CMU11421                 ----X---Q-...-XE-----L-X--X--------.----R-R-.-----.XXP----.--T-.-RSEG-.......-.........-I--I-----------------I--ID--X----------------SN-P---K-- 123
K.CD.97.97ZR_EQTB11                -------V--SQ-...-RE--K------------------S.----W-R-.-----E.---PL----.--T-.----NE-.-S-T-.--GS----------I----V--V--V------------IN---K----------------------VCM---Q-- 127
K.CM.96.96CM_MP535                 ----N-P-...-E------------------------S.---R-R-.-----G.-.PL----.--DQ-.----T-.--EP-------I----I----V--V------------IN---K----------------------------K-- 127
01_AE.CM.11.1156_26                ----N---Q-...-E--K-----------------------S.---GDG-.-----.LHP----.-E-E-.---E-.----SQP--S--I---------VR-E--------------DIN-------------------T------K-- 126
01_AE.CN.12.DE00112CN011           -------TV--Q-...----QLP--------------------S.---GDG-.--G-.-LLT----.----P-.....-SS---S--I----------V--E--M---------DID---K----------------------K-- 125
01_AE.IR.10.10IR_THR48F            ----N---Q-...-----R--------G------I--AS.---K-GDG-.--.S.LLD-.--G-.............-SS----I----------V--E--R--------DIN---K----------------------K-- 119
01_AE.JP.11.DE00111JP003           -------KA--Q-...-----KL-----------S.---K-GDG-.--.LLT----.--T-.----E-.---A-.---SS-LS--I----I------------V--E----------K-I-K------------------------- 128
01_AE.MM.14.fK5DU26                -------KT--Q-...-----P-------------S.---K-GDG-.--.D.LLT----.--T-.....-E-.---A-.---AP-L---I-------I-----------R----------DIN---K--R--------K---------- 127
01_AE.PH.15.DE00115PH012           -------TV--Q-...--GKL----------------S.---G-GDG-.--.S.L-T-.--T-.----E-.---P-.....--AP-L---I-----------------------DIN---K--R----------K-----------KR- 127
01_AE.SE.12.070SE                  ----N---Q-...-R--KL-----ANSPTR------T----S.---K-GDG-.-----E.---I-----.---E-.............-SS----I----------V--E---------DID---K--R----------K---------KR- 120
01_AE.TH.10.DE00110TH001           ----N---Q-...--GK----------------------S.---GDG-.--------.L----.--T-.....-E-.---SP--S--I--------------------------F-DIN---K--R----------------------- 127
01_AE.TH.11.40436v09_04            ----N---Q-...--G--P----------------S.---K-GDE-.---V-.---L-T---.-E-.---R-T-.---PS-LS--I---------------------------------IN---K----------------M--K-- 123
01_AE.TH.90.CM240                  ----N---Q-...--G--P--------------S.---K-GDGGRDNG-.-----.LLT----.--T-.....-E-.---SS--S--I-------------------------------IN---K-------------------------- 132
02_AG.CM.10.DE00210CM013           ----N---Q-...-E------A---------------S.----PWDG-.--.E.T.LL----.--EG-.---I-.............-IS--S--I---------VR---I----------DIN---------------M----K-- 127
02_AG.DE.09.701114                 ----N---Q-...-E-----NG---------------.T--S-S.-D-WGG-.--N.-L----.---G-.....-T-.-IS-LS--I---------VRV-----------VRV-----------DIN-------------------M--KR- 127
02_AG.KR.12.12MHI11_10746          ----N---Q-...-E------A----------------S.---GDG-.--.LL----.--A---TEG-.--V-.-IS--S--I----------VR-E--I----------DI------------------K-R------K-- 127
02_AG.KR.12.12MHR9                 ----N---Q-...-E------G---------------.T--S-.---RDG-.-------.L----.--GAEG-.-EG-.---IS-L---I---------VR---I---------IN---K--R--------------------- 131
02_AG.LR.x.POC44951                ----N---Q-...-RE--KL-----------------G.-----.----S.-.--PQP----.--T-.....-TEG-.---IS--S--I---------VR-E--I----------DI------------------------P----K-- 127
02_AG.NG.12.12NG060304             ----N---Q-...-E----------G-----------.T--S-.----FRDG-.--.LL----.--A---.---TEG-.--T-.-IS-L--I----------VR-E--I---------IN--------------------H--K-- 127
02_AG.NG.x.IBNG                    ----N---Q-...-E--KL-------------------S.----WDG-.--------.T.-L----.--T-.....-TEG-.--A-.-IS--S--I---------VR-E--I----------DIN---K----------------------K-- 127
02_AG.PK.15.PK032                  ----N---Q-...-E--KL-----------------.T--S.----WDG-.--.LL----.--T-.....-TEG-.--T-.-IS-L---I---------VR-E--I----------IN---K----------------L---K-- 127
02_AG.SE.14.098GN                  ----N---Q-...-E--K-----------G.----T---S.---D-GDE-.-GN.-P----.--K-.....-TEG-.--A-.-IS-L--I---------VR-E--I----------IN---K--------Y------------KR- 127
```

Pol

HIV-1/SIVcpz Proteins

Pol start
Gag-Pol TF start
Gag-Pol TF end protease start

```
B.FR.83.HXB2                    FFREDLAFLQ....GKAREFSSEQTR.............ANSPTR...RELQVW.....GRDNN.SPS....EA....GADR...QG......TVSFNFPQVTLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMSLPGRWKPKMIGGIGGFIKVRQYDQILIEICGHKA 127
02_AG.SN.13.9580                ----N---Q-....-E--------N-TISSEQDR.....T---S.....-GDG.....-G.-.PL-.-...-TEG....-.T.....IS-----I-----VI-V-V-E-PI---------------ID-----K-----------------A----K-- 135
03_AB.RU.97.KAL153_2            -----N---Q-....RE--K---------N.......I---S.....-K-WDG.....PLP...-T.....TEG....-.T.....AS-----I-----R----VR-----IN--K--------------------------A----K-- 127
04_cpx.CY.94.94CY032_3          -----NV--Q-....RE--K----A.........--A-.-GM-REE...-G.-.LL-.-...-EG....R.T.....VP-LS--I-----------L--IR-------IN--K---------------------P----K-- 126
05_DF.BE.x.VI1310               ----S--P-....-E--LPP--G.........-L--AS.....-G-.S.LL-.-...-EG....R.T.....VP-LS--I---------V--V---IN--K--------------------------------Q-- 127
06_cpx.AU.96.BFP90              -----N---Q-....-E---------A........---H.-RFR...-RG-.S.PLP.-T...-VEG..E-GK.GAI-LSL--I---------VR----IN--K--------------------------V--Q-- 129
07_BC.CN.98.98CN009             ----N-LP-....................-....G------GRDN.-.I.......-DN.-...........-I---------D-N--K--------------------E-P------ 127
08_BC.CN.97.97CNGX_6F           ----I---P-....-E----PP------.....-S........RG-.-.PS....-TE-.-.......TL---I-----S-V--I---------VN--K--------------------E-P--K-- 123
09_cpx.GH.96.96GH2911           -----N---Q-....RE--L-X--G.........-IG-XS.....-WDE.....----.X.LL--.XX...XEG....R-T.....AP--X--I-----I-V-V---------------------------------Q-- 126
10_CD.TZ.96.96TZ_BF061          -----N---Q-....R---LP----.......-S-S......-G..-.TL-.-...-T-.-........AA-LS--I-----------------IN--K----M-----------------Y-- 126
11_cpx.CM.95.95CM_1816          -----N---Q-....-E--K-P--A.......---S.....-R-R.RGG.S.PLP.-.-T...-EG....-......AI--S--I-----I--V--A--R-------------IN--K----------------EE----E-K-- 125
12_BF.AR.99.ARMA159             -----N---Q-....-E--K-P--A.......---S......-W-R.RG-.-.LL-.-...-EG....R.T.....VP-LS--I------I--V--V--------------IN--K--------------------NV----K-- 127
13_cpx.CM.96.96CM_1849          -----N---Q-....RE---------TNSPTREQTR.......-S........-RG-.S.PLP.-.......-TEG...........DI-LSC--I--I-IV--E--I---------------K-IN--------------------NMS--E-K-- 134
14_BG.ES.05.X1870               -----N---Q-....-E------P--A.......---R.......-W-R.RG-.S.PLP.-.......-R-EG..K-...D-PLSL--I-------------------------DIN--K-------------------------K-- 126
15_01B.TH.99.99TH_MU2079        ----T-.............................................LLT.........-.....YS--S--I-----------------P----P--DINF--K-N-----------------------P----K-- 127
16_A2D.KR.97.97KR004            -----N---P-....RE--------N.......T----S.....-G-WNG.....-G-.-.PLA.-.......-EK....-.T.....-H-C--I---------V--E--R-------------DIN--K--------------------T----EKR-- 127
17_BF.AR.99.ARMA038             -----N---Q-....-E--K-P--A.........-AI........-R--.-G-.-.P-A.-...-EG....R.T.....VP-LS--I---------E-V---------DIN--K--------------------K---D------K-- 125
18_cpx.CU.99.CU76               -----N---Q-....-E-----P-KA.......---S.-G--R.RG-.S.PLP.-.......-RVEG..GE.....I--L--I--------------------------DIN--K--------------------------KR-- 125
19_cpx.CU.99.CU7                -----N--P-....-E-KLP-----.......---R.......-R--.-G-.-.PL-.-T...-SE-.-..........LSL--I---------------------DIN--K--------------------------Q-- 123
20_BG.CU.99.CU103               ----T---P-....................---R.......-RG-.S.LP-.-.......-R-EG..E....-L-I--I-----------VR----I--------DIN--K--------------------P----K-- 123
21_A2D.KE.99.KER2003            -----N-----....-E-G-L---V.......---S.......-DVR-R.......PL-.-...T-.......-.....-LS------I---------------------------N--K-----------------PL----K-- 126
22_01A1.CM.01.01CM_0001BBY      -----N---Q-....-E---L--G-.......---S.-RGDG.....-LLP.-.....-E-.-T.....AS--S--I----------------------------DID------K----------------------P----K-- 127
23_BG.CU.03.CB118               ----T---P-....-E------A.........---C........-RIR.RG-.S.PLP.-.......-R-EG..A-...-L-L--I-----------VR----I--------DIN--K--------------------P----K-- 127
24_BG.ES.08.X2456_2             -----N---Q-....-E--K-P--A.......---S.......-R-R.RG-.S.PLP.-.......-R-AG..ER.....-L-L--I-----------V-V----I--------DIN--K--------------------------K-- 127
25_cpx.CM.02.1918LE             -----Q-....RE--L-P--A.......-D--S.....-PR-.-...-EG....-......VI-YG--I-----------V-V----II--------DIN--K--------------------E-VL--S-K-- 126
26_A5U.CD.02.02CD_MBTB047       ----I--SQ-....-E-------VS.........-S-S.....-GNG.....RG-.PLP.-T...-ERG....-.......VS-S--I--------------------------DIN--S-K--------------------VT-D--K-- 125
27_cpx.FR.04.04CD_FR_KZS        ----N-VSQ-....RE-----P--A.......--PRDH.....-G--.D.LL-.-...-T...-EG....-......T--S--I-----------V--E--------DIN-Q-K---------------------EEVTVD--K-- 125
28_BF.BR.99.BREPM12609          -----N---P-....-E--K-P--A.........-AS......---R.......-PL-.-...-E-.-...G-T.....VP-LS--I---------V--------DIN--K--------------------K---D--D----- 127
29_BF.BR.01.BREPM16704          -----N---P-....-E--K-P--A-ANSPASEQTR.......-I--AS.......-R-.-...R-E.-...-R-E.....VP-LS--I---------V--------DIN--K--------------------K---D-------- 127
31_BC.BR.04.04BR142             -----N---Q-....-E----PP---ANSPASEQTR....-D--AS........-R........-E-.-..........TL-----I---------------------------IK--N---------------------------K-- 133
32_06A6.EE.01.EE0369            -----N---Q-....-E-------A.........--S.....-G-SS.PLL........-EE....K-....AI--S--I-----I--VR--E--I---------------DIN--K--------------------H------K-- 127
33_01B.ID.07.JKT189_C           -----N---P-....LP-KP------.......-S........-T..-.N-I.......-E-.-...........AI--CS--I---------------I-------DIN--K--------------------------K-- 127
34_01B.TH.99.QUR1969P           ----I---Q-....-----G--P----.....----........G........---.N-I......-T...-X.......-I--S--I------I--------------DIN--K--------------------------K-- 127
35_AD.AF.07.169H                -----N---Q-....RE--K----G.........-----.......-R--.-.......K.AL......-T.....-K-....R-......V--S---I-------V--L--QR---------DIN--K-----------------A------K-- 127
36_cpx.CM.00.00CMNYU830         -----N---Q-....RE--K-----G........----T--S.....-WDR.....-V.LF-.-...-IEG....-.T.....IS--L--I-----------------------DIN--K--------------------P------K-- 125
37_cpx.CM.00.00CMNYU926         -----N---Q-....RE--K----G.......---RGE.....RGS.S.LL-.-...-....G--T.....IP--S--I-----IT-V-------I--------DIN--K--------------------------P------K-- 125
38_BF1.UY.03.UY03_3389          -----N---P-....-E-GK-P--A.........--R.-KFW-R.....RG-.-.P-EAGA-.-...-EG....R-N.....VP-LS--I--------------------DIN--K--------------------K-N-C--D--V 131
39_BF.BR.04.04BRRJ179           -----N---Q-....-E--------N-ANS.......P---S.....-PN-.-.-G-.-.LP-.-........-E-.-.........-L-I--I------V--V--R--------DIN--K-T-----------------IV---EN-P-------- 130
40_BF.BR.05.05BRRJ055           ------S--....-E--------A.......---S........-W-R.RG-.-.P-.-...-EK....R-T.....VS-L--I-----------V-V-----------DIN--K--------------------R-K--ED-P-------- 127
41_CD.TZ.05.CO6577V5            -------N---P-....-E--------V-ANSSGQAR.......-R..-.......-IL-.-...KT.......-EK....R.T.....A------I-----V-VAV-----------ID--K--------------------SV----R-- 134
42_BF.LU.03.luBF_01_03          -----N---Q-....-E--K-PP--A-AKAR.......----........-W-R.RG-.-.PL-.-...-E-.-...R-S.....VP-L--I---------V--E--I------------X----------------------N-S-D---- 131
43_02G.SA.03.J11223             -----N---Q-....-E--LP--A.........----C.......-R-R.RG-.S.-LP.-...-E-.-...E-....IT-ISL--I-----KV---V-I-----------QID--------------------K---P--E-K-- 126
44_BF.CL.00.CH80                -----N---Q-....-E--------A.......----........-W-R.RG-.-.PF-.-...-EG....R-K.....VP-LC--I-----K-V--I---------DIN-----T-----------------K--ND-T---- 127
45_cpx.FR.04.04FR_AUK           -----N---Q-....RE----L--G.........---I--SS.....-WDG.......LL-.-...-E-.-...-A-.....GLS---I---------V-V---------DIN--K----------------------K-- 125
46_BF.BR.07.07BR_FPS625         -----N---P-....-E--K-P--A.........-AS.....-RDQ.......RG-.-.PL-.-...-E-.-...R-T.....LP-LS--I------V--------DIN--K--------------------K---D-------- 127
47_BF.ES.08.P1942               -----N---Q-....-E--------A.........-AS......---R.RG-.-.PL-.-...-TE-.-...-T.....VS-LS--I---------V-V---M----------DIN--K-------------------------I----K-- 126
48_01B.MY.07.07MYKT021          ----KA---P-....-E----LY------.....-S........-N-I.......-.............I--SL--I-----------------------N------R--------------------ND--E-K-- 125
49_cpx.GM.03.N26677             -----N---Q-....-E-----P--A.......---S.-G--R-R.....R--.-.S.PLP.-...-EG-.-..........-I---------------I----------N------K-----------------------K-- 126
50_A1D.GB.10.12792              -----N---P-....R-----PP-----.....-S.........-S.TFP.-.-T...-E-.-..........A---S--I----------VRV----------N-Q-K--------------------P------K-- 127
51_01B.SG.11.11SG_HM021         -----N---Q-....-E--------A.........-S........G--.-.-I.-........-.....-LSL--I---------------------------DID--------------------------VSVD-- 127
52_01B.MY.03.03MYKL018_1        -----N---Q-....R--G------A.......---S.-K-GDG.....-.LLT.........-XG....-S-.....SS--S--I-----------V-V----R--------DID--K--------------------K-------K-- 127
53_01B.MY.11.11FIR164           ----A--SP-....-E-X-LP-XP-....----.......-G--P.-...-N-I.......-.....AI----I-----VA---------------IN--K--------------------E----L---K-- 127
54_01B.MY.09.09MYSB023          ----N-LP-....................-G---G.........-.I.......-.......-I-CVC----I---------E------------------IN--K-------------------------K-- 127
55_01B.CN.10.HNCS102056         ----AV--Q-....-----GK---------.....---S.........-G--.-.-I.-...-K-.....-EG....-.....SP-CS--I-----I---------------------DIN--K---------------------K-- 127
56_cpx.FR.10.URF5_patient_A     ----AV--Q-....-E--------G.......T----S.....-RWDG.......-S-.-.LL-.-...-VKG.QG--T.....IP-L--I---------VR-----I-------------------------SV--KR-- 129
57_BC.CN.09.09YNLX19sg          ----AV--P-....-E--------A.......----........-RG-.-.PSA.........RTE-.-..........TL----I-------S-RV-----I-------------------------P---KR-- 123
58_01B.MY.09.09MYPR37           ----I--SQ-....RE---LP-XP-.......---X----.....-G........-X-.-.N-X.-.-........-.......-IP-D--I---------------------DID--------K---------------------K-- 123
59_01B.CN.09.09LNA423           -----N---....R--K--PP--....---S.....-K-GDG.....-GE.....LL........-TE-.-...-T.....AS-L--I---------------------DIN--K----R---------------------C------K-- 124
60_BC.IT.11.BAV499              -------P-....-E--K-PP---G.......T---T-S........RG-.-.-A-.-...-E-.-..........-I---------T----------V--V---------IN--------------------S--------A---K-- 122
61_BC.CN.10.JL100010            -----N---P-....-E----PP-----.....-AS.........RG-.S.LSP.........-IE-.-........-I---------T------------------IN--K--------------------P----K-- 123
62_BC.CN.10.YNFL13              -----N---Q-....-E--K-PK-A.......---S.........R--.-.PS-.-...-IE-.-..........-I---------T----------P-IV---M--------IN--K--------------------P----K-- 123
63_02A.RU.10.10RU6637           ----T---Q-....RE--K----.......----S.-RGDG.....-.-A.......-EG....-.A.....IS--S--I-----------VR-E--I------------DID--------------------M--Y-K-- 127
64_BC.CN.09.YNFL31              -----N---LP-....................-I---S.-G--R.GRDN.-.-I.......-.....-......IPLS----I-----S-V--I-------------IN--K---------------------A----K-- 127
65_cpx.CN.10.YNFL01             -----N---Q-....-E----PP-----.....-D---S.........-VE-.-.......-VE-.-..........-I---------S-RV-----I-------------DIN--K--------------------------K-- 123
67_01B.CN.11.ANHUI_HF115        -----N---Q-....---X--X---A.......---S.-G-GDG.....-R--.-.LLT.......-E-.-..........-T-LS--I---------V-V------------DIN--K--------------------------K-- 125
68_01B.CN.11.ANHUI_WH73         -----N---Q-....---X-------.......---S.-G-GDG.....-R--.-.LLT.......-R-E.-..........TALS--I---------V-V--------------DIN--K--------------------------K-- 125
69_01B.JP.05.05JPMYC113SP420    -----N---P-....................-I-----G.-.P-.G.-AH-.-...-SN-.-........-.......-S--I------E--------------DIN--K------------------------P------K-- 125
70_BF1.BR.10.10BR_PE004         ---K-N---P-....-E-KK-Y---A.......----S.........-R.....-R---.D.P--E------G....-.A.RRT.....VP-L--I------V--I---------XDIN--XKX--X-X--XX----------N-----X-- 129
71_BF1.BR.10.10BR_PE008         -----NV--Q-....-E--K-P----A.......----S.........-R-R.......-D.PL-.-...-EK....R-T.....VP-LS--I---------------------DID--K--------------------K--ENM----- 128
72_BF1.BR.10.10BR_MG002         -----N---P-....R-E--LP---NG.......-GSAS.......-R-R.RGE-.-.PL-.-...-EG....R-T.....VP-LS--I----M---RV------------DIN--K--------------------------Q---- 128
73_BG.DE.01.9196_01             -----N---Q-....-E-----P--A.......---R.......-R-R.RG-.S.PLP.-.......-E-.-...R-K.....VIPL-L--I-----------VR------I-----DIN--K----------------------G-K-- 126
74_01B.MY.10.10MYPR268          ----A---P-....-E--LP-KP-....----.......-G------G.........-S-N.-.-...K-.-...-K.......-I--S--I---------------------DIN--K--------------------------K-- 123
77_cpx.MY.14.14MYNBB090         ----KT---P-....-----LP---------.....-S.........-RG-.-.PN-.-.......-T-.-.........TL-----I---------S-RV----I-------------DVN--K---------------------KR-- 113
78_cpx.CN.13.YNTC19             ----KN---P-....--------ANSPTSRASEQAR---AS.......-RG-.-.PS-.-.......-TE-.-.........TL-----I---------S-V---I---------------IN--K--------------------------K-- 136
79_0107.CN.15.SX15DT013         ----TV--Q-....-E--------N.......T----S.-G-GDG.....-.-L-.T.......-E-.-...-T.....SS--S--I---------V--V--E--I---------------DIN--K------------------------K---- 123
82_cpx.MM.13.mSSDU12            ----KX--XP-....-E----PP------.....---S-S.........-RG-.-.PS-.-.......-TK-.-.......T-----I---------S-XV----I---------------IN--K---------------------PV----K-- 123
83_cpx.MM.13.mSSDU94            -----N---P-....-E----PP------.....-S----F---.......-RG-.-.PS-.-.......-CL-.-..........-I---------A--V---I---------------IN--K---------------------E-T----K-- 123
85_BC.CN.14.14CN_SCYB2          -----N---P-....-E----PP------.....---S.........-RG-.-.PS-.-.......-TE-.-.........TL-----I---------S-V---I---------------IN--K---------------------E-T----K-- 123
86_BC.CN.13.15YNHS18            -----N---P-....-E----PP------.....---S.........-RG-.-.PS-.-.......-TEG.-.........AL-L--I---------------V---M----------IN--K-R-----------------P----K-- 123
87_cpx.CN.12.DH32               -----Q---P-....-E----P------.....-H--S.........-RG-.-.PS-.-.......-TE-.-.........TL-----I---------S-RV----I---------------VN--K---------------------VP-------K-- 123
88_BC.CN.05.05YNRL25sg          -----N---P-....-E----P------.....-H-SS.........-RG-.-.PS-.-.......-EG-.-.........TL-----I---------S-RV----I---------------IN--K---------------------P----KR-- 123
90_BF1.BR.11.BRG06043           -----N---Q-....-E-G-L------.........-RG-.-.K-.-.-L-.........-G-.-.....-A-PLV-----T--------------------V---------------------HV----K--T 123
O.CM.96.LA51YBF35               --PI--SGG....HE--QL-A-TSV.......PI---N.-GGGSEE.....-AGE.-.AGA-.GK.SDSRGRP....GR.....-LPVCL--IP--D--V-IA-V--H-C-V----------NNIQ-E--T-----------K--NNVT--E-K-I 128
O.CM.96.LA52YBF39               -----I--SGG....RE--QLCA-TST.......PI---N.-GGGKEG.....-T-E.-.......GT.EGGT..KR.....-LPICL--IP--D--V---H-C-V----------NNIQ-E-K-T-----------KE-SDVIV-VE-R-V 123
O.CM.99.99CMU4122               -----I--SGG....HE--QLCT-TSV.......PI-SSD.-GGGKEG.....-T-E.-.......GT.EGGP..ER.....-L-VCL--IP--D--V--ARV--H-C-V----------NNIQ-E-K-T-----------KE-YXVKV--REV 123
O.CM.x.pCMO2_3                  ----I--SGG....HE--QLCT-TSA.......PI--SN.-GGGNEG.....-T-E.-.......GT.EGGS..AR.....-L-VCL--IP--D--V--ARV--H-C-V----------NNIQ-E-K-T-----------KE-NDVTV--REV 123
O.ES.01.Read25_HIV_Group0       --QI--SGG....HE--QLLA-TSG.......PI---N.-GGGNKG.....-TGE.-.......GT.ERGP..DR.....-AI-VCL--IP--D--I-A-V--H-C-V----------TNIQ-E-K-T-----------KE-NNVTV-VE-REV 122
O.FR.06.RBF206                  -----N--SGG....HE--QLCA-TSV.......PI---N.-GGDGEG.....-A-E.G.-EA.....-S.-.GGP..ER.....-AI-VCL--IP--D--VIIA-V--H-C-V----------NNIQ-E-K-T-----------KE-NNVTV--E-K-V 125
O.GA.10.10Gab1190               ----V--SWG....HE--QLCA-TSV.......PI---N.-GGS.EG.....-AGE.-.......GT.EGGP..DR.....-ALPICL--IP--D--V--ARV--H-C-V----------NNIH-E-N--------------R-KE-ENVTV--REV 123
O.GA.11.11Gab6352               --QI--SGG....HE--QLCA-TSA.......PI---D.-GGGSAE.....-T-E.-.......GT.EGGP..ER.....-AL-VCL--IP--D--V--ARV--H-C-V----------NNIQ-E-E-T-----------KE-NNVTV-VE-R-V 123
O.SN.99.99SE_MP1300             -----I--SGG....HE--QLCT-TSV.......PI---D.-DGGNEG.....-T-K.-.......GT.EGGP..ER.....-L-VCL--IP--D--I-PARV--H-C-V----------NNIQ-E-K-T-----------KE-N-VPV--E-REV 123
O.US.10.LTNP                    -----QV--SGG....HE--QLCT-TSV.......PI--S-D.-GGGGKG.....-TGE.-.......GT.EREP..NR.....-II-D--II-A-V--H-C-V----------NNIQ-E------------------KE--NVAV--E-REV 123
N.CM.15.S4858                   -----E-VS-H....RET-KLPPDNNKER.......-H--AT.....---W-S........-GE.E.QAG-.-...-G-....D-GG..S-EDRELSVPTL------I-------VI-V--KEVR---------I--LQ-E-K-----------NVTVD-Q-R-- 134
N.FR.11.N1_FR_2011              ----KG-VS-....RET-KLPPDNNKEG.......TH--AT.....---W-S........-GE.E.HTG-.-...-G-....D-GG..P-EDRELSVPTL------I-------VI-V---EEEI---------I-DLQ-E-K-----------NVTV-E-R-- 134
P.CM.06.U14788                  -------SGG....QE--QLCT-TST.......TD---Y..GG........-G--D--EQA.....-G.-.KGG..RER.....-V-P-SL--I--D--VIPV---HIC----------VDNLP-E---------------KEFEEVKX--E-RQV 124
P.FR.09.RBF168                  -----K---SGG....QE--QLCT-TST.......T----Y..GG........-G--D-.PGQA.....-G.-.XGGG..X-....-GIP--L--I--D--VIPV---HIC----------VDKLP-E---N-------------T-KXFEXVKV--E-REV 123
CPZ.TZ.06.TAN5                  ----THPLVG....VQT--LCA-HP-EREE.......---G-G........---ST..DTSGT.NGP.....-TT....-DANE..PN.....-VL--IN-----MM-V-VQ---VCQ--------S-FCNIK-K-Q-A--T---------VP-SE-YN-PVQ-GNKEV 123
CPZ.US.85.US_Marilyn           ----T-VPIVERGIKET--LPGK-EG.......-H-S-N..QLCA-TST.......R--S.DTG.....-T...GKG-.K-....-AL-A---I---------LR--VA--IV-------DNIQIE-T-R---M---------K----HVN--E-R-- 129
GOR.CM.12.SIVgor_BQID2         ----V--SGG....QE--QLCA-AST.......TV---N..GGSDER.....-KGE.-.......-GS.-.RGGS..GR.....-AL-ICL--IP--D--V--H-C-V----------HDIQ-E-K-T-----------KEFNNVTVQ-E-REV 123
GOR.CM.13.SIVgor_BPID15        -----N--SGG....QE--QLCA-ASA.......-H.-GG........-GS.D.PEQT.....-G.-.R-EG........-......-L--I--D--VIPV---HIC----------VDKIP-E---V-----------KEFENVK---E-REV 122
```

HIV-1 Proteins

```
                                  protease end  p66, p51 RT start                    M41L                    D67N  K70R                          D110 catalytic site
B.FR.83.HXB2                   IGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGL.KKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPG 296
A1.CD.02.LA01AlPr                                                                                                                                                                      293
A1.CM.08.886_24                                                                                                                                                                        296
A1.ES.15.100_117                                                                                                                                                                       296
A1.KE.11.DEMA111KE002                                                                                                                                                                  295
A1.PK.15.PK034                                                                                                                                                                         296
A1.RW.11.DEMA111RW002                                                                                                                                                                  296
A1.SE.13.078SE                                                                                                                                                                         296
A1.UG.11.DEMA110UG009                                                                                                                                                                  296
A2.CM.01.01CM_1445MV                                                                                                                                                                   296
A3.SN.01.DDI579                                                                                                                                                                        295
A4.CD.97.97CD_KCC2                                                                                                                                                                     296
A6.BY.13.PV85                                                                                                                                                                          297
A6.IT.02.60000                                                                                                                                                                         296
A6.RU.11.11RU6950                                                                                                                                                                      296
A6.UA.12.DEMA112UA024                                                                                                                                                                  296
B.AR.14.DEMB14AR003                                                                                                                                                                    296
B.BR.10.10BR_MG035                                                                                                                                                                     302
B.CU.14.14CU005                                                                                                                                                                        296
B.DE.13.366396                                                                                                                                                                         297
B.DE.13.947915                                                                                                                                                                         296
B.DO.11.DEMB11DR001                                                                                                                                                                    299
B.ES.15.100600                                                                                                                                                                         295
B.FR.11.DEMB11FR001                                                                                                                                                                    295
B.JP.12.DEMB12JP001                                                                                                                                                                    296
B.PH.15.DEMB15PH003                                                                                                                                                                    299
B.RU.11.11RU21n                                                                                                                                                                        300
B.SE.15.102SE                                                                                                                                                                          296
B.TH.10.DEMB10TH002                                                                                                                                                                    300
B.US.16.2609                                                                                                                                                                           296
C.BR.11.DEMC11BR035                                                                                                                                                                    292
C.CN.10.YNFL19                                                                                                                                                                         296
C.DE.10.622166                                                                                                                                                                         299
C.ES.15.100_118                                                                                                                                                                        302
C.ET.08.ET104                                                                                                                                                                          296
C.IN.15.NIRT008                                                                                                                                                                        292
C.MW.09.703010256_CH256.w96                                                                                                                                                            294
C.NG.10.10NG020523                                                                                                                                                                     292
C.NP.11.11NP016                                                                                                                                                                        292
C.PK.14.DEMC14PK009                                                                                                                                                                    297
C.SE.15.100ET                                                                                                                                                                          292
C.TZ.08.707010457_CH457.w8                                                                                                                                                             292
C.US.14.M0118v1d14_5M04_C5                                                                                                                                                             303
C.ZA.13.DEMC13ZA152                                                                                                                                                                    299
C.ZM.11.DEMC11ZM006                                                                                                                                                                    297
D.BR.10.10BR_RJ095                                                                                                                                                                     294
D.CD.03.LA17MuBo                                                                                                                                                                       297
D.CM.10.DEMD10CM009                                                                                                                                                                    301
D.CY.06.CY163                                                                                                                                                                          295
D.KE.11.DEMD11KE003                                                                                                                                                                    295
D.SE.12.077UG                                                                                                                                                                          294
D.TZ.04.CO6405V4                                                                                                                                                                       295
D.UG.10.DEMD10UG004                                                                                                                                                                    295
D.UG.11.DEMD11UG003                                                                                                                                                                    297
D.YE.02.02YE516                                                                                                                                                                        297
F1.AO.06.AO_06_ANG32                                                                                                                                                                   300
F1.AR.02.ARE933                                                                                                                                                                        296
F1.BR.11.DEMF111BR037                                                                                                                                                                  296
F1.CY.08.CY222                                                                                                                                                                         296
F1.ES.11.VA0053_nfl                                                                                                                                                                    296
F1.FR.04.LA22LeRe                                                                                                                                                                      297
F1.RO.03.LA20DuCl                                                                                                                                                                      298
F1.RU.08.D88_845                                                                                                                                                                       296
F2.CM.10.DEMF210CM007                                                                                                                                                                  295
F2.CM.11.DEURF11CM026                                                                                                                                                                  290
G.CD.03.LA23LiEd                                                                                                                                                                       295
G.CM.08.789_10                                                                                                                                                                         293
G.CM.10.DEMG10CM008                                                                                                                                                                    295
G.CN.08.GX_2084_08                                                                                                                                                                     300
G.CN.16.224GX                                                                                                                                                                          294
G.ES.14.EUR_0033                                                                                                                                                                       296
G.GW.08.LA57LmNe                                                                                                                                                                       294
G.KE.09.DEMG09KE001                                                                                                                                                                    294
G.NG.12.12NG060409                                                                                                                                                                     291
H.CD.04.LA19KoSa                                                                                                                                                                       294
H.CF.02.LA25LeMi                                                                                                                                                                       294
J.CD.03.LA26DiAn                                                                                                                                                                       294
J.CM.04.04CMU11421                                                                                                                                                                     292
K.CD.97.97ZR_EQTB11                                                                                                                                                                    296
K.CM.96.96CM_MP535                                                                                                                                                                     296
01_AE.CM.11.1156_26                                                                                                                                                                    295
01_AE.CN.12.DE00112CN011                                                                                                                                                               294
01_AE.IR.10.10IR.THR48F                                                                                                                                                                288
01_AE.JP.11.DE00111JP003                                                                                                                                                               295
01_AE.MM.14.fKSDU26                                                                                                                                                                    296
01_AE.PH.15.DE00115PH012                                                                                                                                                               296
01_AE.SE.12.070SE                                                                                                                                                                      289
01_AE.TH.10.DE00110TH001                                                                                                                                                               296
01_AE.TH.11.40436v09_04                                                                                                                                                                292
01_AE.TH.90.CM240                                                                                                                                                                      301
02_AG.CM.10.DE00210CM013                                                                                                                                                               296
02_AG.DE.09.701114                                                                                                                                                                     296
02_AG.KR.12.12MHI11_10746                                                                                                                                                              296
02_AG.KR.12.12MHR9                                                                                                                                                                     300
02_AG.LR.x.POC44951                                                                                                                                                                    296
02_AG.NG.12.12NG060304                                                                                                                                                                 296
02_AG.NG.x.IBNG                                                                                                                                                                        296
02_AG.PK.15.PK032                                                                                                                                                                      295
02_AG.SE.14.098GN                                                                                                                                                                      296
```

Pol

HIV-1/SIVcpz Proteins

protease end p66, p51 RT start   M41L   K70R D67N   D110 catalytic site

B.FR.83.HXB2   IGTVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGL.KKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPG   296

HIV-1 Proteins

Pol

```
                                                             polymerase motif                    K219Q
                                                             M184V                    T215Y
B.FR.83.HXB2                 IRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAELELAENREILK 466
A1.CD.02.LA01AlPr            ----------------------------------A---E----V----------E----K--A---S--F------------------------------------------------A---K-------A------D-V-------- 463
A1.CM.08.886_24             ---C-------------------------S-----I------------------V---A----------------------------------------------Q----------K------------D-VA---------- 466
A1.ES.15.100_117            ----------------------------I----LK-E-I--------T---A------A--K--F----------------------------------D-E------------A------------A----DIVT------- 466
A1.KE.11.DEMA11KE002        ----------------------D---AK-EVI----------------A--K--F-----------------------------M--E--------T---A--K------A------DIVT------ 465
A1.PK.15.PK034              ----------------C----------------V--E-------A--K--F--------------------------------------A---K-------A------D-V-------- 466
A1.RW.11.DEMA111RW002       V-------------------------------------A--K--F------------------------Q----------A-Q-K------A------VT------- 466
A1.SE.13.078SE              ---------------I----Y-SK-E-I--------A-V---N--S--F-----------------------------------------G---A-------A----D-VV----- 466
A1.UG.11.DEMA110UG009       ---------------------SK-E-I--------V---KK--S--F-------------------------ET---------S---A--R------A----DIVT------ 466
A2.CM.01.01CM_1445MV        V----------------I--------TK-E-I--------A-----A---F-------------------------------K----------A-Q-K--R---A---A----DIV--R----E----- 466
A3.SN.01.DDI579             V---------------Y------S--E-I--------N-----A--K--F-----------------------------Q----------A------D-VT-------- 465
A4.CD.97.97CD_KCC2          -------------T-----SK-E------------A--K--F------------------------Q----------K--R-I-A----D-VT-------- 466
A6.BY.13.PV85               M------------S-----AK------------A--K-----D---------------D----------K------A----D-VT------ 467
A6.IT.02.60000              -----------------X-------------------X---X--X--F---------------------------M--D----------L-IK-------A----DI-T--K------- 466
A6.RU.11.11RU6950           -------------LK---------------V----A--K--F-----------------------M--D----------K------A----DIVT------- 466
A6.UA.12.DEMA112UA024       -----------------SK-E--------------A--K--F---------------D----------K------I-A----DIVT------ 466
B.AR.14.DEMB14AR003         V----------------I--------------V---H--K--F---------------------------K----------IK----A----------M-------- 466
B.BR.10.10BR_MG035          -------------------------D-----------A-V----T---------------------------------------------K--------------V--Q------- 472
B.CU.14.14CU005             -------------C----D-----L------------A--A-E--F-----------------------M---------------IK----A---------V--Q------ 466
B.DE.13.366396              ---------------A---------------------A--K--F----------------------------------------------A----------P---IV--S-- 467
B.DE.13.947915              -------------C-----G---------------E--I-D----K--F------------------------S--K------P-----V--R--I------L-- 466
B.DO.11.DEMB11DR001         ------------------E---------------E--K--F---------------------------E----------A---A---V--------- 469
B.ES.15.100600              ------------------E------------D---K--F-------------------------------A---A----IVQ------ 465
B.FR.11.DEMB11FR001         ----------------M-----------------A--K--F-------------------------------A---K----A---S----Q------ 465
B.JP.12.DEMB12JP001         -------------C----------Y---------A----R--G--F-------------------------------A-G---A----R 469
B.PH.15.DEMB15PH003         --------------------EK---I---------LEK---V----F------------------------------N-------S--K------------- 463
B.RU.11.11RU21n             -----------------D----E-I------------------K-----------------------------------A-L--------------A------ 470
B.SE.15.102SE               ---------------H-----KRN-E-I-------E----V-Q----T--F--------------------------------IK----I--T--V--- 466
B.TH.10.DEMB10TH002         V----------------D----L------------A--K--F---------------------------------S---K----I---S--I--- 470
B.US.16.2609                -----------------D---------------E--I-D--N--K--F--------------------D-E----------K---I-------I----- 466
C.BR.11.DEMC11BR035         ---------------R----T---EL----------A---E--------------------------------Q----------DIVT--D---- 462
C.CN.10.YNFL19              -------------I----A--E------------A---E------------------------------------------V---- 466
C.DE.10.622166              --------------------K-D----------A---G--K--F---------------------------------V-G--DIV----- 462
C.ES.15.100_118             --------------C----K--A--EM---------A--K--K--F-------------------------Q----------V--DIV--------R 472
C.ET.08.ET104               ----------------T--E--------------A---E--KG-F---A--LR-----------------Q-----------I--A--DIV----- 466
C.IN.15.NIRT008             -------------D---A--E---C---------A---E-----------------------------Q-----------I----A--DIV---- 462
C.MW.09.703010256_CH256.w96 --------------------AK--E---------A--K--K--F-----------------------------S-Q----------K-M----A----DIV----- 464
C.NG.10.10NG020523          ------------------A--E------------A---E--K--F--------------------------Q----------A--A----DIV---- 462
C.NP.11.11NP016             ------------------AK------------A-V---A-----------------------------K------A------IV--- 462
C.PK.14.DEMC14PK009         ------------------AK------------A-V---A--------------------------------K------A------DIVQ---- 462
C.SE.15.100ET               T--------S--------MK--------------A-V---K--F----------------------------Q----------DIV-------I- 467
C.TZ.08.707010457_CH457.w8  --------------------VR-------------A--K--F--------------------------K--S---K------DIV--------E----R 462
C.US.14.M0118v1d14_5M04_C5  -------------C----TK--E------------A-V---K--K--F----------------H-E-D-E--------K------A--DIV----- 473
C.ZA.13.DEMC13ZA152         -----------------------EL----------A--K--E-------------------------L--------S---------A-S--DIV---- 469
C.ZM.11.DEMC11ZM006         ------------------A--E------------A---E--K--E--F------------------A-Q--Q----------A----VT----- 467
D.BR.10.10BR_RJ095          ----------------I--D---------------E------------K-----------------------Q--L-A----S--------- 464
D.CD.03.LA17MuBo            -------------------EV-------------A--K--E-------------------------T--D----------K------D---A----- 467
D.CM.10.DEMD10CM009         -------------C----K--EL-------------A--K--E------------------------T----K---A-Q-KH----A------IV---- 471
D.CY.06.CY163               -------------C----K--EL------------A--K--E------------------------T---EN----------A------IV---- 465
D.KE.11.DEMD11KE003         -----------------EM-------------A---E-----F----------------------K--D-ED-------------RCI--A-T---IV---- 465
D.SE.12.077UG               ---------------Y----EM-C--L-------------E--F-----Y---------------E--E------------------IV--- 464
D.TZ.04.CO6405V4            V----------------------------------A---K--F-----------------------Q-A----------X---K------------ 465
D.UG.10.DEMD10UG004         ------------------EM-------------I----G--K--F----------------------K-----------K--RC-K-A-T---I------------R 465
D.UG.11.DEMD11UG003         ------------------EM-------------A---E------------------------T-T---E--------------CI---T---V---- 467
D.YE.02.02YE516             -------------C----EL------------------K--E-------------------------T--Q-EN---------K--R---A---D--V--R------ 467
F1.AO.06.AO_06_ANG32        -------------C----TR--EL----------A---G--K--F----------------------Q-D----------A-G--DIV--A--- 470
F1.AR.02.ARE933             -------------C----TK--M------------------K--T--F---------------------K---------K------A--DIV--A-- 466
F1.BR.11.DEMF111BR037       -------------C--D--AK-EL-------------A--K--F----------------------S-Q------------IKE-----A---DIVT-K---- 466
F1.CY.08.CY222              ------------------VK--------------I----A--K--F------------------Q-D-----------I------A--DIV--A---K--- 466
F1.ES.11.VA0053_nfl         ------------CX-----D--AK--GL-------------E--K--F------------------Q-D----------A------A--DIV--A--- 466
F1.FR.04.LA22LeRe           ------------------AR--E-----------------K--A--F----------------------Q-D----------A------DIV--A--- 467
F1.RO.03.LA20DuCl           V--------H---R--------T-------L---A-----K--F----------------------Q-N----------A------A--DIVT--A--- 468
F1.RU.08.D88_845            --------------K----TK------------K-------V---D--K---------------------Q-D-E--------------A----DIVT--A--- 466
F2.CM.10.DEMF210CM007       -----------------H----AK--E-------------A--K--F-------------------------Q-D--S----------R-KHI----A-T--D-VS--A----M----- 465
F2.CM.11.DEURF11CM026       -------------C----TK--E------------A------K--F----------------------A-Q-D----------R-KH--R--A---D-V--A----M----- 460
G.CD.03.LA23LiEd            ------------------A----RR--EM---------A----G--K--F-------------------------Q-D-E----------I------DIV--A--- 465
G.CM.08.789_10             ------------------NK--EM------------A-V---A--K--F-----------------------Q-N-E----------K---I------DIV--A--- 463
G.CM.10.DEMG10CM008         -------------C--R------A--EM-------------A---D----F--------------------K-D-D-E----------IK------I------DIV--A--- 465
G.CN.08.GX_2084_08         ------------------X----A--EM-------T------A--D----F-------------------Q-D-E----------IK----------DIV----T----- 470
G.CN.16.224GX              -------------R-------IK--E-----------------A---F---------------------Q-D-E----------IK------A---DIV------T---- 464
G.ES.14.EUR_0033           ------------------IK--E------------------A---E----F--------------------Q-N-E----------X------A---DIV--X--M--I--- 466
G.GW.08.LA57LmNe           ------------------IK-E-E----V--------A-----K--F-------------------------Q---------X------A---DIV--A--- 464
G.KE.09.DEMG09KE001         K-----------------IK---M----------------A--K--F--------------------------------K------A---DIV--A--- 464
G.NG.12.12NG060409          -------------------TK--E------------D---A---D----F--------------------Q-D-E--------S---K--R--A--DIVT--D---- 461
H.CD.04.LA19KoSa           ------------------E--E------------E----A--E--F--------------------SVK-----------K------I------IVT-A--- 464
H.CF.02.LA25LeMi           ------------------T------------------A----A--E--F---------------------VK-----------R--H----A----DIV--A-- 464
J.CD.03.LA26DiAn            -------------C------MK-E-I---------E--M----A--E-----------------------------K--------A--DIV--R--A---- 464
J.CM.04.04CMU11421          X----------------TK--XL----------E--A--X--E--K---------------------------Q----EX--------X--R-K-M----A-S--DIV--K--------K---X 462
K.CD.97.97ZR_EQTB11         -------------C------RK--M-L----------A----E--------------------------Q-D----------F----K------V---DIV--A-- 466
K.CM.96.96CM_MP535         V-----------------IK--EM------------P--------A--K--F-------------------------Q-D----------K------V---DIV--A-- 466
01_AE.CM.11.1156_26        ------------------N-E------------A---A--S--F----------------------------VK-----------A---------DIVT--- 465
01_AE.CN.12.DE00112CN011    ------------------KIK--E----------N-----I-----A--N--------------------R-----K------H-----------A---DIV----- 464
01_AE.IR.10.10IR.THR48F     ------------------KIK--EMI--H-------A------A--N--------------------------E--------K------H--------------A---DIV---- 458
01_AE.JP.11.DE00111JP003    ------------------TK--ELI----------A------A--MS--F---------------------------E-----------K--------A---DIV----- 464
01_AE.MM.14.fKSDU26        -------------C---------IK-E-I--------E--A---A--S--F---------------------------E---------IK--------A---DIV---- 466
01_AE.PH.15.DE00115PH012    -------------C------IK--EM------------A--Q--A--S--F----------------R---VE-----------------------A---DIV----- 466
01_AE.SE.12.070SE          -------------C------TK--EM------------A--A--S--F------------------R---E----------IK----------A---DIV----- 459
01_AE.TH.10.DE00110TH001    ------------------IK--E------------I--------A--S--F----------------R---E-----------G---A---------A---DIV----- 466
01_AE.TH.11.40436v09_04    ------------------RK--E------------------A--S--F----------------R---E-----------G---A---------A---DIV------R 466
01_AE.TH.90.CM240          ------------------IK--EM-----K---------A--S--F----------------R---E-----------------A---------A---DIV----- 471
02_AG.CM.10.DE00210CM013    ------------------A----AK--E-------------A---E--F-----------------------H-Q-----------A---K--A---DIVT---- 466
02_AG.DE.09.701114         ------------------A----IK--E-------------A---E--F----------------------Q-----------A---K--A--------VV---- 466
02_AG.KR.12.12MHI11_10746  ------------------A----TK--EM----------A---E--F------------------------------A--IK-----------A---DIV----- 466
02_AG.KR.12.12MHR9         ------------------AE--EV-----------A--DK--E-------------------------Q-----------K------------DIV----- 470
02_AG.LR.x.POC44951        ------------------A----TK--EM----------A--Q-D-K--F-------------------------A-R-K-----------A---DIV----- 466
02_AG.NG.12.12NG060304      ------------------A----T---EM-----------A---E--F---------------------H-E---R----------A-R-K-------DIV----- 466
02_AG.NG.x.IBNG            ------------------A----TK--E--------------A---G--K--F--------------------VE-----------IK--R--A---DIVA----- 466
02_AG.PK.15.PK032          ------------------A----AK--E--------------A---D-----F-------------------------A--IK-S----V--DIVTM----- 465
02_AG.SE.14.098GN          ------------------A--I------MK-E-----------I--------A-V---E--K----------------------Q-N---------V---DIVT------------ 466
```

Pol

HIV-1/SIVcpz Proteins

polymerase motif
M184V
K219Q
T215Y

```
B.FR.83.HXB2                 IRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAELELAENREILK 466
02_AG.SN.13.9580             -------------------A-----TK---I-C--I------------E--A-----E------F-----------------------------------Q-Q---------------A---K--------A---DIVT--------------- 474
03_AB.RU.97.KAL153_2         ---------------------------E-----------------A-----E--------------E-------------------------------------------------------A---------A-------------A-------- 466
04_cpx.CY.94.94CY032_3       --------------------C----FK--E-----------A----------E---F-------------------------------------QPA---------------------K----------A---DIV---T---------- 465
05_DF.BE.x.VI1310            --------------------C------EV--------------A-K--E---A-F--------------------------------------------------------------K-----------A---DIV------------------- 466
06_cpx.AU.96.BF90            --------------I----IK--E-----------A---E--F--------------------------------------------Q-D-E-------------------------------A---DIV--A----------- 468
07_BC.CN.98.98CN009          ---------------I---IK--E---------A-----E--F----------------------------------------------Q---------------------------A---DIV------------------ 466
08_BC.CN.97.97CNGX_6F        --------------------C----------------K--------------------Q-----F----------------------------------------------A---DIV------------ 462
09_cpx.GH.96.96GH2911        ------------------IK--E------------A------E--F-----------------------------------A------------------AR----DIVT------- 465
10_CD.TZ.96.96TZ_BF061       -------------------QM---------Q-H-TI---G--K--F-A-----------------L--Q--K-N-----------------A-V--DIV--------- 465
11_cpx.CM.95.95CM_1816       -------------------T--E-----------E-V--K-----F----------------------------E-D-EC--------------------------A---DIV--A---- 464
12_BF.AR.99.ARMA159          T-----------------------------------------A------F----------------------------------K---R-------A------K---------- 466
13_cpx.CM.96.96CM_1849       ---------------------------------E-V--A--F----------------------------A-Q-D-E----------------------A--------- 473
14_BG.ES.05.X1870            --------------------RK--EM--------A-----A--K--N--F-----------------------------------K-------A---DIVS--T----- 465
15_01B.TH.99.99TH_MU2079     --------------------C--K--EM-----------A--S--F----------------R-------------------------K-----------A---------- 465
16_A2D.KR.97.97KR004         V---------------------E-----R---------N---FY------------E------------------AE--K---A---DIV--K----E------- 466
17_BF.AR.99.ARMA038          ----------------------L--------------------K--F--------------------------I----------------IV--K----- 466
18_cpx.CU.99.CU76            -------------------E-E---------------------K-F------------------------Q--E-----------K----I---DIV--V--- 464
19_cpx.CU.99.CU7             ----------------D----E----------------K-F----------------S--E------------------K------------A------------- 462
20_BG.CU.99.CU103            --------------------G--F--------------------WK--FY----------------Q-D--------------------A---V--T---- 462
21_A2D.KE.99.KER2003         ------------------EM------------------E-----------------K--E----------------A----------------R--- 465
22_01A1.CM.01.01CM_0001BBY   ---------R-D--AK--E------------E--I-D--A------Q-N----------------A------------------------R 466
23_BG.CU.03.CB118            --------------------C----EL-----------A-----E--F------------K----------A-------------- 462
24_BG.ES.08.X2456_2          --------------------C---X----EL-----------A--F------------------Q----------A------- 462
25_cpx.CM.02.1918LE          ---------------T--GT--AK--E--------E--S--K--F-----------------A-Q--------A---DIV--V--- 465
26_A5U.CD.02.02CD_MBTB047    V-------------------A--E------------E--S--F--------EN-----------------K--------A---DIVT--K----- 464
27_cpx.FR.04.04CD_FR_KZS     --------------------C----T--E-----------A--A--K--F---------Q--E--------K--R---A---DIVS------ 464
28_BF.BR.99.BREPM12609       ------------------------------------V------E--F----------------I------------------V--K------ 466
29_BF.BR.01.BREPM16704       ------------EK------------------------F---------------------A---------V--K---------I- 466
31_BC.BR.04.04BR142          --------------------------------------K--FD-----------------E-D-E-----------------DIV------- 472
32_06A6.EE.01.EE0369         ------------------EM-------------V----E--E--F----------------E-D-E---------------AR---A---- 466
33_01B.ID.07.JKT189_C        --------------R---IK-----------D----------A-----F----------------E--------A---DIV---------K-- 466
34_01B.TH.99.QUR1969P        --------------------C---------M------------D---F---------------E--------------A---DIV---------X-- 460
35_AD.AF.07.169H             --------------I---------E-I-----------I----A--F---------T--E-----------------A---DIVS------------ 465
36_cpx.CM.00.00CMNYU830      ------------------A------TK--E-------A--K--N--F----------------Q-D--I--------A---A---A------ 466
37_cpx.CM.00.00CMNYU926      ------------------C---A--E-----------A--K--N--F----------------Q-D-I--------R-------D-VTK--K----IL-V--- 464
38_BF1.UY.03.UY03_3389       ------------------Q-----------------E--------------------------------------V------------R 470
39_BF.BR.04.04BRRJ179        ----C---------------E----V---A----K--A----W--FY----H----------M----------K-------V--A------ 469
40_BF.BR.05.05BRRJ055        ------------C----A--EL--------A----E--K--FF--EQ-------------M----------K------DIV--T-A---- 466
41_CD.TZ.05.CO6577V5         -------------------R-E-----------A-----E--K--F----------K---ET---------R-H--C------VS------ 473
42_BF.LU.03.luBF_01_03       ------------------------------------I---------A------------------------------V------- 470
43_02G.SA.03.J11223          ---C--------------TK--E---------L--A--E--F--------------Q-D-E----------K--R---A---D-VT----- 465
44_BF.CL.00.CH80             V------------------------------------K----------------Q--------------K--R---P---R--I---- 466
45_cpx.FR.04.04FR_AUK        --------------------TR--EM-------A--E--K--F--Q-------------------Q-D-E-------S---A---DIVL--A-----E------A- 464
46_BF.BR.07.07BR_FPS625      --------------------C----AK--E-------N--A--K--F---------Q-N-E------------R-K----A---DIVS--A------ 466
47_BF.ES.08.P1942            -----------------------------------------K-------------------Q--N-------R---------DIV--K-- 465
48_01B.MY.07.07MYKT021       ----------------------E-I----------A--N--F------------------Q--A--Q-K----------A---DIV--A----- 466
49_cpx.GM.03.N26677          --------------------A------SK--E-I------A--S--F----------------Q-D-EA----------K-----A---DIV---- 464
50_A1D.GB.10.12792           --------------------A-----T#--E-I---------A--S--F--------------------------K----I--A---IVT--A- 464
51_01B.SG.11.11SG_HM021      --------------D--------E-I-------E--V-------F------------R--E------------K-------A---DIV--V-- 466
52_01B.MY.03.03MYKL018_1     -----------------------------------------F---------------R-----E-----------X---A--XK-------A---D-V--- 466
53_01B.MY.11.11FIR164        --------------------C---IK--E-----------A-----F-----------------R--E-------------A---DIVQX--X-- 466
54_01B.MY.09.09MYSB023       --------------------C----------V-----A-----S--FF--Q----L-----R------------A--V-----DIV-A------ 466
55_01B.CN.10.HNCS102056      --------------------C----IK--E------V-----A-----F---------------R--E----------A--V---DIV-A---- 466
56_cpx.FR.10.URF5_patient_A  ------------I------H---------------R----V------------Q-D----------------------A------- 468
57_BC.CN.09.09YNLX19sg       ----------H--R-----A-----------------A-----K--F-------M----------------------D-V-V------ 462
58_01B.MY.09.09MYPR37        --------------------IK--E------------V-----A--S--F-----------M---N-------------A------DIV---- 466
59_01B.CN.09.09LNA423        --------------------IK--E----------I-V---A--S--F----------VE----------K---A---RV---A------ 463
60_BC.IT.11.BAV499           --------------V--G-----I-----------A--K--Q--A-------T-----Q-D--------------A---DIV----- 461
61_BC.CN.10.JL100010         --------------------------------A--E--K--F-----------Q----------------A---DIV------------R 462
62_BC.CN.10.YNFL13           --------------------C------------------A--A--K--F-----Q-----------------A---DIV--T---E-----R 462
63_02A.RU.10.10RU6637        --------------------RK--E----------I-N--F-------------Q-------------------A---DIVT---- 466
64_BC.CN.09.YNFL31           --------------------A--E-----------A--E--K--F---------Q---------------------A---DIV------ 466
65_cpx.CN.10.YNFL01          --------------------C------D-----------A--E--K--F------------N----------I-S--DIV----- 462
67_01B.CN.11.ANHUI_HF115     --------------------NK------------I-----A--K--F----------R--E-------------K-------V------- 464
68_01B.CN.11.ANHUI_WH73      ------------R------------------------I-----K-----------R--E-----------------V------ 464
69_01B.JP.05.05JPMYC113SP420 -----------------------------------------V---------------------------------------- 464
70_BF1.BR.10.10BR_PE004      --------------------C----AK--M-----------A--E--K--F-----------Q-N------------A---DIVT--T----E----- 468
71_BF1.BR.10.10BR_PE008      --------------------C----D--A-----------A--E--K--F-----------Q----------------A---DIV--A----- 466
72_BF1.BR.10.10BR_MG002      --------------------C----AR----------A--E--K--F-----------Q-XD-E-----------G--K--R------DI---A---- 467
73_BG.DE.01.9196_01          --------------------IK--E-----------A--M--K--F------------Q-XD-E----------K------A---DIV------- 465
74_01B.MY.10.10MYPR268       --------------------IK--EM--------------I---S--F---------M--X--------------K--------A---DIV------- 466
77_cpx.MY.14.14MYNBB090      ------------------------------------I------KX--F--------------------------------------IV------ 452
78_cpx.CN.13.YNTC19          --------------------IK--E-T---------A--S--F-------------Q------------------------K-----A---DI------- 475
79_0107.CN.15.SX15DT013      -------------Y-----TK--E--------I-----A--E--K--F--------------E----------------------K----A---DIV-------X------- 466
82_cpx.MM.13.mSSDU12         --------------------C-----------------A-----E--K--F-------------------------A---DIV--A------R 462
83_cpx.MM.13.mSSDU94         --------------------C----AR--EL-----------A--K--K--F--------------------T-----------A-T--DIV------ 462
85_BC.CN.14.14CN_SCYB2       -----------------------------------V--G--K--A--F-----------------------------A-T--DIV---------- 462
86_BC.CN.13.15YNHS18         ----------------------L-----------V------E--K--F--------------------------A---DIV--------I- 462
87_cpx.CN.12.DH32            --------------------------------------V------E--F---------------------A---E------------R 462
88_BC.CN.05.05YNRL25sg       --------------------T--E---------A--E--K--F----------------------------V---DIV------- 462
90_BF1.BR.11.BRG06043        --------------------R-----------------A---E--F------------Q------------------A------- 466
O.CM.96.LA51YBF35            -------------D----D--EME--------PLAE--KRV-L--E--YQ-----Q----------K--D-EV------Q-R-E---I---S-----V--SK-----E----E--- 467
O.CM.96.LA52YBF39            V------------D----E--E--------PLAE--KRV-L--E--YQ--F------------Q-R-KE----IK--S---V--SK-----E---K-- 462
O.CM.99.99CMU4122            -------------V--D----D--ELE-C--------PLTE--KRV-L--E--YQ-----------Q-N-EI------Q-R-E---I---S-IV--SK-----E----K-- 462
O.CM.x.pCMO2_3               T------------I--D----N--ELE-C--------PLTE--K-V-L--E--YQ-----------Q-N-EV------Q-R-KE---I---S--IV--SK------E---- 460
O.ES.01.Read25_HIV_GroupO    V------------D----D--ELE-C--------PLAE--KRV-L--E--YQ-----------K--N--V--------I------Q-R-E---I---S-----V--SK---M--E--K-- 461
O.FR.06.RBF206               V-------T--------N--EME--------PLAE--K-V-L--E--YK--F-----------I--Q-N--V----------Q-R-KE---I---S-IV--SK------E---- 464
O.GA.10.10Gab1190            -------------D----D--ELE-C-----H-----PLTE--KRV-L--E--YQ--F---Q-----R----Q-N--T--------I------Q-R-E---I---S-----V--SK-----E----K-- 461
O.GA.11.11Gab6352            V------------D----D--ELE-C--V--------PLTE--KRV-L--E--YK--FY---------Q-N-EE------Q-R-KE---I---S----V--SK-----E----K-- 462
O.SN.99.99SE_MP1300          -------------D----D--ELE-C--------PLTE--KRV-L--E--YQ--F-----------Q-N-EE------Q-R-KE---I---S-----V--SK------E----K-- 462
O.US.10.LTNP                 -------------D----D--ELE-C--------PLTE--KRV-S--E--YQ--F-----------Q-R-KE---I---S--IV--S-----V--K------ 462
N.CM.15.S4858                -------------D---KH-E-I------------LA--EVV----R--D--F-----------K---EI---------K---R-I------V-DF-K------ 473
N.FR.11.N1_FR_2011           V------------T------KHSE-I------------LT--KEV---G--K--F-----------K---KH--R-I-----VNF-K---- 473
P.CM.06.U14788               L--V---------H----------S--EVE--------PLSE--QRV-K--E--YV--F-----------K-----------I---S-----VAF-R-----E--K---- 463
P.FR.09.RBF168               -------------H----ES--EVE--------PLSE--QRV-K--E--YV--F-----------EA------------Q-R-KE---I---S-----VAF-R-----E--K---- 463
CPZ.TZ.06.TAN5               K-F----------D---EH--VD--------SEEE--KL-KK------L---E-----Y-EK---M-------N------N----N-T---PEQ----------G---TKE----I--V-G--DPVEM-R------E--KQ-I- 462
CPZ.US.85.US_Marilyn         -------------D---H--VI--------NLEK--E-V-L--Y--T--F-------------E-----------IK---IK-A----VNF-H--M--E------ 468
GOR.CM.12.SIVgor_BQID2       V------------D----D--E-E--------PLAE--R-V-G--E--YQ-------------K--D-EV-----X-------Q-R-KE---I---S-----V--H------R--- 462
GOR.CM.13.SIVgor_BPID15      V------------H----D----N--E-E----------PLSE--QRV-K--E--YA--F-----------K---------G--------RIKE---I---S----VTF-R-----E--K---- 461
```

HIV-1 Proteins

p51 RT end p15 RNase H start

```
B.FR.83.HXB2                    EPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPF.KNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTNRGRQKVVTLTDTTNQKTELQ  635
A1.CD.02.LA01AlPr               N------------------EH---------.---G-R-S----R--V-V--VG--G----------------A--M-H-------D----------------D--M-------------M---I-D----S-AE----A--H  632
A1.CM.08.886_24                 D-------T--V----D---------.-----K-S----R--A-V---VV-------------L--D-----------------------S------D-----I-S-E----H  635
A1.ES.15.100_117                -----A-------------D---------.-----K-S----R--A-V-VVM----------------K-D----------------D--------E-------H  635
A1.KE.11.DEMA111KE002           ------N-S--A--V---D---------.-----K-S----A-V--V-M-----------------------------D--I-----------IS-E----H  634
A1.PK.15.PK034                  D-------------------.-----KR-A--------V--AR-------R------I--MD-------I----S------DK-----H  635
A1.RW.11.DEMA111RW002           --------------------D---------.-----K-S----V--V-M--------R------------MD-------D-----------S-E----H  635
A1.SE.13.078SE                  -----A-------------D---------.-----KK-S----R--A-VI-RVV--------------------MD------D--A--------DK--S--E----H  635
A1.UG.11.DEMA110UG009           D------------------D----L--.---D-S-----A-V---VM-----------------MD-------D----------K---------D--I-S-E----H  635
A2.CM.01.01CM_1445MV            N-----------T----D--------.-----KK-TT-----------VM-------------------------------Q----------D----I-S--E----R--H  635
A3.SN.01.DDI579                 D--A------------N--------.-----K-SS-----R--V-V---VV------------------R---------M----Y-----------D-I----------S-ME--------H  634
A4.CD.97.97CD_KCC2              ------------------D-------.-----KK-S----R--S-V--VA-------------------R------A--MD------------D---------S-E-----H  635
A6.BY.13.PV85                   ------------------D-------.---KRGS----AV---VA---L---R---------R------A--M-------------------------T------D-----P-SE----Q--H  636
A6.IT.02.60000                  D-------V--------D-------.---KRGS----AV--V----I--R---R---A--M----------------------D--------P-E------H  635
A6.RU.11.11RU6950               -----L-------V----D-------.---KKGS----AV--VA------R------------A--M---------------------RI------D----------P-NE-------H  635
A6.UA.12.DEMA112UA024           --------V---------D-------.---KKGS----AV---VA-----R------A--MD------D--M------N-T------D--------P-E--------H  635
B.AR.14.DEMB14AR003             -----A---T-------N----V--.--------A-----------A---V---R---A-----------------S----------D-----------S-------H  635
B.BR.10.10BR_MG035              -T--A------------EM--------.-----K---------R--------AS-------------E------H----------D------P----------641
B.CU.14.14CU005                 -----------V--L----------.----------T---------V------------I--R-----------------------D-------SIN  635
B.DE.13.366396                  -T-------E----EL--------.-----T---------S-------I-----D--L-------------A----------S-----R--PI------636
B.DE.13.947915                  -------V----E----------Y-.-----S----------V------------I--R---------------------S--------------635
B.DO.11.DEMB11DR001             ------------------D------.-----K----------VAM------S-----------I---------------DK---IS--------638
B.ES.15.100600                  ----------------------D-.-----T-------V---AQ-G------R-R----A--M-----------------D-----S-N------634
B.FR.11.DEMB11FR001             ---------------EH--------.-----K---------AQ-----------R----A--M-----A----------DK---IS-------634
B.JP.12.DEMB12JP001             ------V----L----------.-----K---I-----------A--R-----------R----A--M-------------D--K---S----638
B.PH.15.DEMB15PH003             K------------------------.-----K----I-----------A--R-----------A--M-----E------DK---S--------632
B.RU.11.11RU21n                 --------V--L----------.----------------------V------------I-DK-----------------------I-DK-----639
B.SE.15.102SE                   -----A-----------L---.----K--T----R------S----------R-----DA----------------DK---S--------635
B.TH.10.DEMB10TH002             T--A------------------.-----K----------L---A------R-----------------------DK---S-N---------639
B.US.16.2609                    ----------T----L---------.-----K------R---V--------M---R----M--------------------A-----IS-------635
C.BR.11.DEMC11BR035             ------------------D------.-----K-T-N--------AM-------R-----------------G--M----D--I-S-E-------631
C.CN.10.YNFL19                  -T--A------------LD--------.-----K-T--------AM---------R-------AD------------D-A-V----K------D---K-I-S-E--------635
C.DE.10.622166                  -------L--------ND-------.---KR-A----------AQ------R-----------D-------------LT-V-----D-I-R------D---I------638
C.ES.15.100_118                 -------------D-------D-.-----K-A----A-------AQ---------R---A--D------------D--M-------D--A-P-E--------641
C.ET.08.ET104                   ------------------ND------.---KR-T--------AQ-----R---I--A--D-------------D-------S-E-------635
C.IN.15.NIRT008                 ------------------D------.---KV-T---------AQ---------R---A--D------------T--------V-----DK--I-SV-E-------631
C.MW.09.703010256_CH256.w96     ------------------DD------.---KQ-T-----CM----------I--R-------D-------------I----------K------D---I-S-E----633
C.NG.10.10NG020523              ------S--E-------DD-------.---KR-T---R----C-----------I--R-----D-------------A-------Q----D---IIP-DE-----631
C.NP.11.11NP016                 ------E-----------D------.-----K-T---------M-----R-------SD-------------------------IS-E-------631
C.PK.14.DEMC14PK009             ------E-----------D------.---KR-A-------AQ---------A--D-------R--------------M----D-------SISE-------631
C.SE.15.100ET                   ------E-----------HE------.----R---R----AQ------------R----D-P-------------I---------D--I-S-E------636
C.TZ.08.707010457_CH457.w8      --A-----------------D-----.---K--S------------M---A-----------D-------------------I----------D--I-S-E------631
C.US.14.M0118v1d14_5M04_C5      ------------------ND------.---KIKA------R------M---A--R---D-------------I------------I-------D--S-NE-------642
C.ZA.13.DEMC13ZA152             -------L---D---------.-----K-T---------AL-------R----A--D-------------I------------D------D--I-S-NE----A---638
C.ZM.11.DEMC11ZM006             ------------------HD------.-----K-T---------AM-C------R----A--D-------------------------D--I-S-SE-------636
D.BR.10.10BR_RJ095              -----------L----NE-------.----------T------SL-C----RI---R----DL--V------------I--------D---IIP------633
D.CD.03.LA17MuBo                ------V---H--------.----------T--------------A-----------S--ID---------L---------R-----D---SF-------636
D.CM.10.DEMD10CM009             ----------S---W------.-----K----------AM--------R----A--M---------------------D---A-SF--------640
D.CY.06.CY163                   -------K-----V----------.-----K------------AI--------R-----A--M----------------I-------S-I-------D---A-SF------634
D.KE.11.DEMD11KE003             -----Q----------------QY.-----K--T---------H-C--------R-----A--M----------------D----------S---------634
D.SE.12.077UG                   -----A------------E------.-----KK----------AL-G----------I--R-----A---------------I-------P--------633
D.TZ.04.CO6405V4                ----A---L---V----N------.---------S----------V--VC------R----D--V-------------I--------D---PF------634
D.UG.10.DEMD10UG004             -------------------QY-----.-----K--T------AQ-C----A--------------------I-----------D---PF-------634
D.UG.11.DEMD11UG003             ------V---PD--------QY.-----KL--T------AA----SQ-C--------R-----A--M-----------------D-----S-A------636
D.YE.02.02YE516                 ------V--------D------.----------V--AQ---------R----A--M---------------I----------D---PF-------636
F1.AO.06.AO_06_ANG32            ------------------HE------.---KR-S------VAM---------------------T------A---K--K-------D--KE-A-P--E----A--H  639
F1.AR.02.ARE933                 ------------------L------.---K-S-------AQ---A--R--L--D---ISI-------T--A----------S---K------D---A-P--E----A--K  635
F1.BR.11.DEMF111BR037           -----A------------------.-----K-S-------AQ-C--A--R--LR--D----------T--T-----------S--E-------D--A-S--E--A---635
F1.CY.08.CY222                  --------V---V----D------.-----K-S------VQ--------I----K--A-*ID--------T--T-----------S---I------D---S--E--A---634
F1.ES.11.VA0053_nfl             ------Q--E--------F------.-----K-S------AL---I--R--L--D----------T--A----------S---E------D---IP--E--A--H  635
F1.FR.04.LA22LeRe               ----------V----------C-.-----K-S------AV---------R---------T--I--------S---E------D---S--E--A---636
F1.RO.03.LA20DuCl               ----------T----------.-----K-S------AM-C--------------A--------T--T-------------Q-------D--A-S--E--A---637
F1.RU.08.D88_845                ---------V----------.----Y-------K-S----R----AV---------A-----------T--PE------------S---K-----DK--A--E--A---634
F2.CM.10.DEMF210CM007           ------------------HD----.---H.---KR-S-------DV---A--G----RV--R--------------T-----------S-----------D--S--E-------634
F2.CM.11.DEURF11CM026           -----A------------HD---V--.---H.---RKS-------A-G--V--R----------T-------------S-------D--S--E--------629
G.CD.03.LA23LiEd                ------E---V----LD-------.---Y.---KRGS------V--A-G-I----I-----R----DV--------D-------R--T--P--Y----D-K--II--E----A---634
G.CM.08.789_10                  ------E---V----------.---Y.---KRGS------V--A-G-I-----I--R--R-------DV--------R--T-I-P-V--Y----D-K--II-E----A---632
G.CM.10.DEMG10CM008             D---------V----VE------.----KRGS------V---A--I----I--R--R--D-------R--T-I-X---Y----D-K--IIS--E----A---634
G.CN.08.GX_2084_08             ------------V---PD------.---Y.---KKGS------A---I--V----R---V--D-------T--I---Y---------DK----II-IQE---A--H  639
G.CN.16.224GX                   ------A---------SD-----.---QY.---KKGS------V--------R--V--D----------T---Q---Y----DK----II-IQE----A--H  633
G.ES.14.EUR_0033               ------E---E----V---Q-----.---KKGS------V---A-----------R---------------T--A-D----DK-K--II---E----A---633
G.GW.08.LA57LmNe                ------E---E------V---Q-----.---KRGS------V--------R--------D-------R----------V----I-DK-K--II---E--A---633
G.KE.09.DEMG09KE001             ------------V---D-------.---Y.---KRGS-------A---V--R--R------A--P--P---Y----V----I-DK-K-II--E----S--H  633
G.NG.12.12NG060409              ------------DE-------.---Y.---KRGS--------A--V--R--R----------D-------P--P---Y----D-K--II-IQE----H  630
H.CD.04.LA19KoSa                ------------------D------.---KV-S-------VAV------------I--R------------H------T-------------S---I--------D----IIS--E----H  633
H.CF.02.LA25LeMi               T-----------A--------V---.-----K-T--I-----------A-----I--R-------------H------T--A-D------S---I-------D---IIS--E---H  633
J.CD.03.LA26DiAn                -----I---A-E----LD------.---KR----------SM-C-------R----------------M--------S-----I--DK--------633
J.CM.04.04CMU11421              -----X---A-E----V--ED-----.---KR-----------AL------R---------D---------M----------XK----I-DK------Q----631
K.CD.97.97ZR_EQTB11             ------------H------Y----.-----I-S--------V--VAM------R-----G-----------T---------------Q------DK-----ISI-E---------635
K.CM.96.96CM_MP535              ------------------ND----.---H.-----S-------A-G--VAM------R-----------H----K-R---D------SI-E----A---635
01_AE.CM.11.1156_26             N-------------V-------.---KGS-----R--V--VA-G-----I--R-----------M----------------D----S-AE-------634
01_AE.CN.12.DE00112CN011        T-----V-----------D------.-----K-S----R--V--VA---------I--R--------------D----SK-----D-----S--E------633
01_AE.IR.10.10IR.THR48F         T--------V--V---D------.-----K-S----R--V--VA--------I--------M--------------D----S------D--KR--S--E-----627
01_AE.JP.11.DE00111JP003        T-----V-----V---D------.-----K-S----R--V--VA----------------M--------------D----SK--E----D--S--E------635
01_AE.MM.14.fKSDU26             T----------V----D------.-----K-S----R--A-V---S----I--R-------------------D---------------IS-E-------634
01_AE.PH.15.DE00115PH012        T--------V---------D----.-----K-S----R--A-V----A---R------R-------------------SK-----D------S--E------635
01_AE.SE.12.070SE               T-----V---------D------.-----K-S----R--V--VA----I--R----------------------------D-----S-E-------628
01_AE.TH.10.DE00110TH001        T------------V---D------.-----K-S----R--V--VA-----I--R-------M-------------------S------D--S--E-----635
01_AE.TH.11.40436v09_04         T-M----------V----D----.-----K-S----R--A-V---VA------LR-------------------N-------------IS-E-----631
01_AE.TH.90.CM240               --------------D-------.---RGS-----R--V--VA----------------------D-------S-----------D---IS-E----640
02_AG.CM.10.DE00210CM013        ----------T---V--V---D--.---F.---KK-S------V---VA---------R----A--M-------------D-----RIIS-E---H  635
02_AG.DE.09.701114              ----------T---V--V---D--.-----KK-S------V--VA--------R------A--M--------------------D-----I-S--E----635
02_AG.KR.12.12MHI11_10746       A---------T---V-------.---Y.---KK-SS------DV---VA---------R----A--M----------------D----H-I-S--E----635
02_AG.KR.12.12MHR9              ----------T---V-------.-----KK-S------V---VA--------R----A--M-----------------D-----I-S-E----639
02_AG.LR.x.POC44951             K---------T-----D-----.-----KK-S----I----V--VA---------R----DA--M----------------DK-----S--E----635
02_AG.NG.12.12NG060304          ----------T---D-----.-----KR-S------V---VS---------I--R-------------------S-----------I------D-----S--E----635
02_AG.NG.x.IBNG                 ----------T---V----D-----.-----KR-S------V--VAM--------R----R----A--M----------------D-----S-E----635
02_AG.PK.15.PK032               ----------T---V-------.-----KR-S------A---VAM--------R----DA--M--------------------D-----S-E----634
02_AG.SE.14.098GN               ----------T---------D---.---Y.---KR-S--I-----V----AKK-------R----R--A--M-------D---------I----DK----IIS--E-------635
```

Pol

p51 RT end _p15 RNase H start

B.FR.83.HXB2 EPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPF.KNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTPKFKLPIQKETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTNRGRQKVVTLTDTTNQKTELQ 635

HIV-1/SIVcpz Proteins

HIV-1 Proteins

Pol

p66 RT, p15 Rnase H end p31 Integrase start

```
B.FR.83.HXB2                   AIYLALQDSGLEVNIVTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQD.EHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVI 804
A1.CD.02.LA01AlPr              --Q---------------------------R---I-----K-E---I-S-----------------------S---------------E.D--R----T-V------I----------------------------------------------------------- 801
A1.CM.08.886_24                --H--------P------------------K-D----D---S----K-I---------------------V----------E.D-------T------I------------------------I------------------------------------L----- 804
A1.ES.15.100_117               --H--------S------------------R--G-D--S----S-K-I-----------------------E---R-------------------------------------M------------------------------------------------- 804
A1.KE.11.DEMA111KE002          --H--------S------------------R--A----D---S----------------------------E---R---------------N------------------------------------------------------------------------- 803
A1.PK.15.PK034                 -------------T--R----------------K-R-DR--S------------------E.D--R-----K----------I----------------------------------------------------------I---------------------- 804
A1.RW.11.DEMA111RW002          --L--------S------------------R-DA-V--D--S---------------------------E---R---------------------------------------------------------------------------------------- 804
A1.SE.13.078SE                 ----------------V----R-D------K-ME-D---S-----N-----------------------E.--R---T-----------I------------------T---------------------------M----- 804
A1.UG.11.DEMA110UG009          --H--------S------------------R--G-D--S--------------------------------E.--R------------------------------------------------------------------------------------ 804
A2.CM.01.01CM_1445MV           -------------------------------R-----E---RI--S-----------------------E.--R----H---I------------------------------------------------------------------- 804
A3.SN.01.DDI579                --H--------S---------------D-I--K-E----S------------------------------E.--R----T---------------------------I-------------- 803
A4.CD.97.97CD_KCC2             --------------------------I--I--K-E----S-----------------------------E.--R----T----------------------------- 804
A6.BY.13.PV85                  --H--------S------------------R--I-K--K-E-R--S------------------------N------E.--------K----------------------I----------- 805
A6.IT.02.60000                 ----------X------------------R--A-V----K-E---S----------------------E.D---------K----------------------------X------X 804
A6.RU.11.11RU6950              --H--------S------------------R--I-K--K-E-R--S----------------------N--R----E------K----------------------------------I------- 804
A6.UA.12.DEMA112UA024          --H------S-D-I--K--K-E---S--------------------------Q----R----E-----K------------------I-I------- 804
B.AR.14.DEMB14AR003            ---I-----S-----------------K----S-----N-----------------I------------E.D------I-------------I------- 804
B.BR.10.10BR_MG035             --H------------------------------K---------S------------------------E.D------------------------------M----- 810
B.CU.14.14CU005                --Q-------------------H-H--K------------------------------E.D---------E----I----C----Q----- 804
B.DE.13.366396                 --H--------P------------R-----S--------I-------------------E.D----------------------I----------- 805
B.DE.13.947915                 --H------ST----------------R---S--------------------------E.D-------------------------I---------- 804
B.DO.11.DEMB11DR001            ---------S---V-----------K------N--I------K--R----------E.--------K-------I----------------- 807
B.ES.15.100600                 --H-----------------------K------S------I---------------E.----------I---------------- 803
B.FR.11.DEMB11FR001            --H-------L---------------K------S-----------------T----E-------------------I------- 804
B.JP.12.DEMB12JP001            --H--------------------------K--------------R--T-V------E.--------------C---------------- 807
B.PH.15.DEMB15PH003            --H-------------------K------------N----------R---N--R--------E.D-----------------L----------------- 801
B.RU.11.11RU21n                --H--------------------------K--------I-----I-N--------E.--------------------------I------ 808
B.SE.15.102SE                  ----------S------------------K----N--I----N------------E.-----------------Q-----------------I------ 804
B.TH.10.DEMB10TH002            --H-------------------------K----S------------------T--E.--------N-------I------------- 808
B.US.16.2609                   --Q---------------------------K---------------NE-------E.D-------------------I------ 804
C.BR.11.DEMC11BR035            --Q--------S---------------K-T---S--L-N-R--S-----------------NE------T----------I----M---- 800
C.CN.10.YNFL19                 --Q--------S------------------K-----R--S--------------------E.-----Q----------I--------- 804
C.DE.10.622166                 --Q--------P------------------K---------S-V---------E.D-------E----I-------------I------ 807
C.ES.15.100_118               --Q---------------V-----K----R-R--------------P------E-E.D-D--------Q--T-----------I------- 809
C.ET.08.ET104                 --Q--------S-----------------K-D-I----Q-R--S----------N-------E.--N----DE--------I-------- 804
C.IN.15.NIRT008                --------S-------------K------R--S----------------------E.-----------SQ-Q-------------M-- 800
C.MW.09.703010256_CH256.w96   --Q--------S------------RI-S------S-----E.D-R-----E-----I--------V-------I------- 802
C.NG.10.10NG020523            --Q--------P------------------K------Q----S------K-----E-E.D--N------E------I-------I-I----- 800
C.NP.11.11NP016               --H--------S------------------K------S--R-------S-R------E.------------N----------I----------- 800
C.PK.14.DEMC14PK009           --H--------S------------------K-----R-R-S----------------E.-----------------T------------I-M---- 800
C.SE.15.100ET                 --H------S----------K--I------N-R--S-------------------E.-----------DE-GI-S---H----I------A-- 805
C.TZ.08.707010457_CH457.w8    --Q--------S----------H--K--I----N-R-R--S--------N------E.------------DE-GI-S-----H------Q--I------I-------- 805
C.US.14.M0118v1d14_5M04_C5    --Q--------S------------------K-R--R--S-----------------I-----------I-Q--I------------------I------- 811
C.ZA.13.DEMC13ZA152           --Q--------S-----------------N--I-S--------------------E.D---NE----I-R-------------- 807
C.ZM.11.DEMC11ZM006           --Q--------S------------------R-R--S-----------------E.--R--N----E----I-------------- 805
D.BR.10.10BR_RJ095            --N-----------------------K------Q----------Q---I-------GE--I--------I-----------IV-- 802
D.CD.03.LA17MuBo              --N-------------------------K----S----Q-R--S-----------SE-E.-D----------------------I------- 805
D.CM.10.DEMD10CM009           --N--------P------------------K------N--I---S-----DR----------I-------------------I-------- 809
D.CY.06.CY163                 --N------------SL--------------R----N--I----S--D--N------------I---------W------------K-- 809
D.KE.11.DEMD11KE003           --NM-------------------------K----S---------R------------------------I---------------------- 803
D.SE.12.077UG                 --N-----------------------I---------Q--------S------------------------I------------------- 802
D.TZ.04.CO6405V4              --N--------------------------K----------Q---------------E-------------I------------- 803
D.UG.10.DEMD10UG004           --N--------S-----------D----------------R--S-----------E.----------------------L--------I------ 803
D.UG.11.DEMD11UG003           --N--------------------------K----Q--S-----------------E.----------------L--------------- 805
D.YE.02.02YE516               --N--------S-----------K--V-S-----------S--------NQ--E.D-D--------I--------Q------------I----- 805
F1.AO.06.AO_06_ANG32          --N----------------------------K-----Q--R-S----------E.-------------------------------M---- 808
F1.AR.02.ARE933               -VQ--------A-----------------K----------------I-----E.--------------------------------L--- 804
F1.BR.11.DEMF111BR037         --Q--------S------------------K--V----N--S------T-----I-----N---------I-----Q------N--------- 804
F1.CY.08.CY222                --H--------S------------------K---S---R-Q-S------R-----S----------I-K----R--------------E-I--- 803
F1.ES.11.VA0053_nfl           --QI----A-S------------------K----------Q--S------------E.----------------------------I------- 804
F1.FR.04.LA22LeRe             -VH--------------------------K-------Q--I-S------------I--E.D--R------N----------------I----- 805
F1.RO.03.LA20DuCl             --------S------------------K--H----Q-RI-S-----------------I---E.--N-K------------------- 806
F1.RU.08.D88_845              --------S------------------K----------Q---S----------------I---E.----N------------------------I-------- 803
F2.CM.10.DEMF210CM007         --Q--------S-------------L--K--------Q----S--------N-------E.---F-N-------------Q---------- 803
F2.CM.11.DEURF11CM026         --Q--------S-------------H-K---------Q----S-----N--------E.--N-K--I------------------- 798
G.CD.03.LA23LiEd              --N--------P------------------R---V----------S----------E.D--R-----N--C--I----------------I----- 803
G.CM.08.789_10               --NV-------------------------K------------R--S---------I--N---------E.D--R---------------Y-----------I-----M--- 801
G.CM.10.DEMG10CM008           --N--------------------------K--I----R---S------------I----N-----E.--R--N--------I--------------------QI---- 803
G.CN.08.GX_2084_08           --Q--------S-------X---------X--I----------------------E.D--R--N---------I------------------- 802
G.CN.16.224GX                 --W--------S------------------R----------I-------------S-----E.--R-N-------------------------- 802
G.ES.14.EUR_0033             --QI-------SA----------------R-------------R--S-----------E.D--R--T------------------- 804
G.GW.08.LA57LmNe             --QI-------S------------------R---R--S----------------E.--R---------------I------------ 802
G.KE.09.DEMG09KE001           --Q--------S------------------R-D----R--S----------------E.--R-----K---------------M--- 802
G.NG.12.12NG060409            --Q--------S------------------R-------R--T----------------------E.--R--K--------------I---------------- 799
H.CD.04.LA19KoSa             -----------E-----------------K--I-----E--------------------------E.--R--N-------------I------------- 802
H.CF.02.LA25LeMi             ----------------------------K------S---Q----------------E.--R--N------N---I-I---- 802
J.CD.03.LA26DiAn             --------------V--------------K----S--------E------------E.--------------------N----I--------------I----- 802
J.CM.04.04CMU11421            --L---X---X-A--------------K--------E-X-R--S-------------E.-------------------------I-------- 800
K.CD.97.97ZR_EQTB11          --H--------------------------K-D---------DR--S-----------E#-----N-------------I---------- 804
K.CM.96.96CM_MP535           --C--------S------------------K--D----RI-S-------------E.--R--N-------------I----------- 804
01_AE.CM.11.1156_26          --H--------P------------------R---I----K-E--------------E.--R----T------I-----C--------I------- 803
01_AE.CN.12.DE00112CN011     --H--------S------------------R--I----E-V--S---------I---SE----E.--R----T------------------- 802
01_AE.IR.10.10IR.THR48F      --H--------------L------R--V----S------------------------E.--R----T--------------------I-------------- 796
01_AE.JP.11.DE00111JP003     --H--------S-------------R--I-K--------S-I-----------KK-----E.D--R----T--------------I---K--Q-------------- 802
01_AE.MM.14.fKSDU26          --H--------S------------------R--I----E---S---------------E.--R----T------I----------------- 804
01_AE.PH.15.DE00115PH012     --H--------P------------------K--I----S--------I-S-------E.--R----I---------II----------------- 804
01_AE.SE.12.070SE            --H-----------------GR---V----------------------------E.D--R----T--II---N-N----------------------- 797
01_AE.TH.10.DE00110TH001     --H--------S------------------R----E--------------------E.--R----T------T----------------------------I 804
01_AE.TH.11.40436v09_04      --H--------------------------R--V----E----N------------E.--R----T------II-------N-N------------------- 800
01_AE.TH.90.CM240            --H-----------------------R--V----S---------------------E.--R----T------TN----------------------- 809
02_AE.CM.10.DE00210CM013     --H--------S------------------R--I----K-E--------------E.--R----T------------C---------------------- 804
02_AG.DE.09.701114           -------E---S----------------R-D---K--K-E-D--S-----------I--S--R------E.D--R-------G------------------- 804
02_AG.KR.12.12MHI11_10746    --H--------S------------------N---------------I--S------E.D--R---------G-------------I------------- 804
02_AG.KR.12.12MHR9_          --H--------------------------R----L--K-D---------I------S--------E.D--R-----------G----------------I--- 808
02_AG.LR.x.POC44951          --H--------S------------------R----------N--------------E.D--R---------------I------------------------- 804
02_AG.NG.12.12NG060304       -----------SQ----------------R----S---K-E-I-S------------E.D--RF------E----I------------------- 804
02_AG.NG.x.IBNG              --H--------S------------------R-------N-----------------E.--R--K----------------M---------G------------- 804
02_AG.PK.15.PK032            --H--------S------------------R---K-G-D--S-----------I---E.--R--------------Q--------G----------I---- 803
02_AG.SE.14.098GN            --H--------S------------------R-----R-E-D--S----------I--E.--R--------Y---I------Q-K---------------*K------- 803
```

Pol

HIV-1/SIVcpz Proteins

p66 RT, p15 Rnase H end p31 Integrase start

B.FR.83.HXB2   AIYLALQDSGLEVNIVTDSQYALGIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKVLFLDGIDKAQD.EHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLDCTHLEGKVILVAVHVASGYIEAEVI 804

HIV-1 Proteins

```
B.FR.83.HXB2                     PAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTGATVRAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRNPLWKGPAKLLWKGEGAVVIQDNSDIKV 974
A1.CD.02.LA01AlPr                ----------I------V----S-A-K----NVQ----------------------------E-------------------------I----------R---------------D-I--------------------- 971
A1.CM.08.886_24                  -------------------VV----STA-K-----Q------------------------------------------R---H---M----------D-I-----------------N--- 974
A1.ES.15.100_117                 -------------------VV----P---S-AFK----SVQ-----------R----------L----------I------R---H---M----------D-I--- 974
A1.KE.11.DEMA111KE002            ----------I-R------VV----P-IS-A-K----N-----------------------E-------------------------I----------N----------D-I-------------G-- 973
A1.PK.15.PK034                   -------------------VV----S-AFK----SV-----------------------E-------------------I--L-S--R-------K---F---D-I----------H-N--- 974
A1.RW.11.DEMA111RW002            -------------------VV----S-A-K----D---------------------------E-------------------------I------------------D-I--- 974
A1.SE.13.078SE                   -------------------VV----S-AFK----NVQ----------------------------E-------------------------------K-------D-V--- 974
A1.UG.11.DEMA110UG009            ---------A---------V----G--IS-AFK----DVQ-------------------------E-----------------------N-K------------------K--- 974
A2.CM.01.01CM_1445MV             ----------I------V---P---S--K-----VQ----------------------------D------------R-----------E-I--- 974
A3.SN.01.DDI579                  ---------A---------V----S-AMK----N------------------------------E---------------------------------------- 973
A4.CD.97.97CD_KCC2               ----------I--------V----STA-K----S-Q------------------------------L-----------------------D-I--- 974
A6.BY.13.PV85                    ------------------IV----P-SSA-K----D-Q-------------------------E-------------------------I----------D-I-----------------N--- 975
A6.IT.02.60000                   -------------------VV----SSA-K----N-Q-X--X---X--------V--E---------------------------------------D-I--- 974
A6.RU.11.11RU6950                -------------------V----P--SNA-K----N-Q------------------------E--------------------I----------D-I-----------------N--- 974
A6.UA.12.DEMA112UA024            -------------------VV----P--S-A-K----D-Q-----------------------T------------------------I----------D-I--- 974
B.AR.14.DEMB14AR003              -------------------P--SN--K----------------------------I--S-----------------------D-----------EDK--- 974
B.BR.10.10BR_MG035               ----------------L----P--SNV-K----------------D-T-------------------------I----------------------------------N--- 980
B.CU.14.14CU005                  -------------------P--ST--K--Y-----------------------------------I-----------F---D-I-----------------E-- 974
B.DE.13.366396                   ----------I--------V------S-------E--------R----------------------------------------I---S-L---------N-D-I----R--- 975
B.DE.13.947915                   ----------I--------L----P-IS--K----------------------------MI------TQ---H-------------N-D-I----------L---E-- 974
B.DO.11.DEMB11DR001              ------D---------A----V----SN--K----------Q--Q-------------N------------X----R---I---------------D--- 977
B.ES.15.100600                   -------------------V----SNA-K----------------N------------------------I----------------D--- 973
B.FR.11.DEMB11FR001              -------------------ST--K---------D---------------------------------------R-----I---------------D--- 973
B.JP.12.DEMB12JP001              ----------I--------ST--K------------------N----------------------M-----------------------D--- 977
B.PH.15.DEMB15PH003              -------------------P--ST--K-----------------------------------------T-----------S----------N-D--- 971
B.RU.11.11RU21n                  -------------------STA-K----------Q--Q---------------------Q------------------D-I--- 978
B.SE.15.102SE                    ------------------V----P--STA-K----------------------------------------I----------D-I-----------------E-- 974
B.TH.10.DEMB10TH002              ----------I--------V----------------L-----V--------------------E----------------------D--- 978
B.US.16.2609                     ----------I--------P--IST--K--------E------------------------------------------------D--- 974
C.BR.11.DEMC11BR035              -------------------SN--K----------Q-------------------------------------I--------------L--- 970
C.CN.10.YNFL19                   -------------------V----S-A-K----------------L--DR--------------------------I---R---------D-I--- 974
C.DE.10.622166                   -------------------V----ST--K----------------------------------------------I---------------D-I--- 976
C.ES.15.100_118                  ----------I--------VV----SS--K----------Q----------------------------I------------N--I---D-I-----------------L--- 979
C.ET.08.ET104                    ------------YI-----V----P--S-A-K----------------E------------R-----------------------D-I--- 974
C.IN.15.NIRT008                  ----------I--------V----ISNA-K----------Q-------------R--------------C----I-M--------VL-N------------D-I-----------------E-- 970
C.MW.09.703010256_CH256.w96      -T-------YI-----V----P--S-A-K----------Q---------------------------T-------I-------Q---I--------D-I--- 972
C.NG.10.10NG020523               -T--------YI-----V----S-A-K----------Q------------E---E---------------------------------D-I--- 970
C.NP.11.11NP016                  -S---------------V----SGA-K------------N----I----------------------------------------D-I--- 970
C.PK.14.DEMC14PK009              -------------------V----SNATK----D-Q-D----------------S-----------------------R----------D-I--- 970
C.SE.15.100ET                    -------------------V----S-A-K----------Q-D----------------------------------I---------------D-I--- 970
C.TZ.08.707010457_CH457.w8       -E-------YI-----V---R---SN--K-----------------------------------D-T---------------------D-I-----------------K--- 975
C.US.14.M0118v1d14_5M04_C5       ----------YI-----NV----S-A-K-------------------------------------------------------------D-I--- 981
C.ZA.13.DEMC13ZA152              -------------------V----S--K----------Q--------------------------------I----R-----I---------D-I--- 977
C.ZM.11.DEMC11ZM006              ----------YI-----V----S-A-K----------Q--------------------------------------I---------------E-- 975
D.BR.10.10BR_RJ095               -E---------------VV----R--ISSA-K-------------------------------------------I---------------E-- 972
D.CD.03.LA17MuBo                 -------------------VV----S-A-K------------------------------------I----R----I--------------------E--- 975
D.CM.10.DEMD10CM009              ----------I--------VV----S-A-K------------------------------------------I-----------------D-I--- 979
D.CY.06.CY163                    -------------------VV----S-A-K----------V------------R-------L----------R------------------- 973
D.KE.11.DEMD11KE003              -------------------VV----S--K---------------Q----------------------------R------------------- 973
D.SE.12.077UG                    -------------------VV----P--S--K---------------------------------------I-S-L-S----R--Q-----F---D-I-----------R-E--- 972
D.TZ.04.CO6405V4                 -------------------S-ALK------------------------------------------------------D-I--- 973
D.UG.10.DEMD10UG004              -------------------VV----R--STA-K------------------------R----------------------V---------D-I-----------------E--- 973
D.UG.11.DEMD11UG003              ----------I--------VV----IS-A-K----------------E------------------------------------------D-I--- 975
D.YE.02.02YE516                  -------------------VV----SGA-K------------------------------------------------------------D-I--- 975
F1.AO.06.AO_06_ANG32             ----------I--------S-A-K------Q---------------------------------------------S-----------N-D-V--------------E-- 978
F1.AR.02.ARE933                  ----------I------A----P--S-A-K-S----Q-------------------------------------S-----R---------D-V--- 974
F1.BR.11.DEMF111BR037            -------------------T--S--S---Q------------------I----------------------------I---------D-V--- 974
F1.CY.08.CY222                   ----K-----I--------I------SGA-K----Q--E--------K-I-K-------K----------I-----------------------D-I-----------------E-- 973
F1.ES.11.VA0053_nfl              ----------I--------P--S-A-K----VQ--------------I--------------------------I----------------D-I-----------------E-- 974
F1.FR.04.LA22LeRe                ------D---------M------S-A-K------------------------------------------------------------D-V-----------------E--- 975
F1.RO.03.LA20DuCl                ----------I------R------SGA-K----Q-----------------------------------I----------------------D-V-----------------E-- 976
F1.RU.08.D88_845                 ----------I--------I------ST--K------Q---------------------R----------I-S--------------D-V-----------------E--E--- 973
F2.CM.10.DEMF210CM007            ----------I--------S-A-K------Q---------------------------------------I-S------------------D-I-----------------NE--- 974
F2.CM.11.DEURF11CM026            ---------------R------I------S-A-K----Q-----------------I----------------------I-S------F---D-I-----------------K--- 968
G.CD.03.LA23LiEd                 ----------I--------V----S-A-K----N-T-------------------------------------I-S------------F---D-I-----------------NE--- 973
G.CM.08.789_10                   ----------I--------V----P--IS-A-K----N-T-------------------------------I-S------------------D-I-----------------NE--- 971
G.CM.10.DEMG10CM008              ---------A--------V----R--SNA-K----N-T-------------E------------------------------------D-I-----------------NE--- 973
G.CN.08.GX_2084_08               ----------I--------V----SNA-K----S-T------------------------------------I-S-----SR------D-I-----------------NE--- 978
G.CN.16.224GX                    ----------I--------V----SNA-K----S-T-------------------------------------I-S-M-------------D-I-----------------NE--- 972
G.ES.14.EUR_0033                 ---------F-I-------M------S-A-K----N-TH------------------------------------I----------F---D-I-----------------NE--- 974
G.GW.08.LA57LmNe                 ----------I--------V----P--SNA-K----S-T------------------------------------I-SE-----------D-I-----------------NE--- 972
G.KE.09.DEMG09KE001              ---------A--I------V----IS-A-K----N-T------------------R----------------MI--S------------D-I-----------------NE--- 972
G.NG.12.12NG060409               ----------I--------V----SNA-K----S-T------------------------------------I-S-M-----VS-----D-I-----------------NE--- 969
H.CD.04.LA19KoSa                 -------------V-----M------S-A-K----Q-------------------------------------I----N-----L-V-K--------D-I-----------------E--- 972
H.CF.02.LA25LeMi                 -------------I-----STV-K----N-Q-------------------------R-----------------------LN--K-----D-I-----------------E--- 972
J.CD.03.LA26DiAn                 ---------F-I-----I------G--ISNA-K----D------------------------------------I----------------D-I-----------------E--- 972
J.CM.04.04CMU11421               ---------F-I-------V----S----K-------X-----------E---R---------D-T-----------R-----------D-I--- 970
K.CD.97.97ZR_EQTB11              ----------I-------RV----S-V-K----D-----------------------------E-------------------------I-------------------E-I-----------#.Y--E--- 972
K.CM.96.96CM_MP535               ----------I--------V----T--STV-K----V-----------------------------------------LN--K-----E-I-----------------E--- 974
01_AE.CM.11.1156_26              -----S-------------RVV----S-A-K----N-T-----------------------------------------------------D-I--- 973
01_AE.CN.12.DE00112CN011         -------------------V----S-A-K----NVR-----------------E------------------------------------D-I--- 972
01_AE.IR.10.10IR.THR48F          -------------------V----S-A-K----NVR-----------------E------------------------------------D-I--- 966
01_AE.JP.11.DE00111JP003         ------D----I-------V----S-A-K----NVQ--------I-------I-E-------------------------------D-I--- 972
01_AE.MM.14.fKSDU26              -------------------V----S-A-K----NVR-----------------E------------------------------------D-I--- 974
01_AE.PH.15.DE00115PH012         ---T--------------V----S-A-K----NVR-----------------E------------------------------------D-I--- 974
01_AE.SE.12.070SE                -------------------V----P--STA-K----NVQ-----------------------------L-------------------D-I--- 967
01_AE.TH.10.DE00110TH001         -------------------V----STAMK----VNVR-----------------E--R--------------I-M-----------------D-I--- 974
01_AE.TH.11.40436v09_04          -------------------V----S-A-K----NVR-----------------I-E------------------------------D-I--- 970
01_AE.TH.90.CM240                -------------------V----S-A-K----NVQ-----------------E------------------------------------D-I--- 979
02_AG.CM.10.DE00210CM013         ---------F-I-------V----S-A-K----NVT-----------------------------------I-S--I------------D-I--- 974
02_AG.DE.09.701114               ------D---I--------V----SGA-K----NVT-----------------------------------I-S-----------------D-I--- 974
02_AG.KR.12.12MHI11_10746        ------P----I-------V----S-A-K----NVT-----------------------------------I-S-----------------------------I 974
02_AG.KR.12.12MHR9               -------------------V----S-A-K----NVT----------------------N----------------MI--L-S--------H-I-----------G---H-I--- 978
02_AG.LR.x.POC44951              -------------------V----S-A----NVTH-----------------------------------I-S--------Q---------D-I--- 974
02_AG.NG.12.12NG060304           ----------I--------V----P--S-A-K----NVT--------------------R--------------I-S-----------------D-I--- 974
02_AG.NG.x.IBNG                  -------------------V----S-A-K----NVT-----------------------------------I-S-----------------D-I--- 974
02_AG.PK.15.PK032                ----------I--------V----S-A-K----N-T-------------------R---------------I-S--I------------D-I--- 973
02_AG.SE.14.098GN                ----------I--------S-A-K----NVT-----------------------------------I-S--I------------D-I--- 973
```

B.FR.83.HXB2   PAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTGATVRAACWWAGIKQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVYYRDSRNPLWKGPAKLLWKGEGAVVIQDNSDIKV 974

HIV-1 Proteins

Pol

```
                                                    p31 Integrase end
                                                      Pol end
B.FR.83.HXB2                    VPRRKAKIIRDYGKQMAGDDCVASRQDED..........*         1003
A1.CD.02.LA01AlPr               ------------------------G-------------         1001
A1.CM.08.886_24                 ----------------------------------G-----------   1004
A1.ES.15.100_117                --------------------------------G-------------   1004
A1.KE.11.DEMA111KE002           --------LKH-----------G-------------             1004
A1.PK.15.PK034                  --------K------V--------G-------------           1004
A1.RW.11.DEMA111RW002           -----V--------------GG-------------             1004
A1.SE.13.078SE                  --------------------------------G-------------   1004
A1.UG.11.DEMA110UG009           --------------------------------G-------------   1004
A2.CM.01.01CM_1445MV            -----V--------------------------G-------------   1004
A3.SN.01.DDI579                 I-------------------------------G-------------   1003
A4.CD.97.97CD_KCC2              --------------------------------G-------------   1004
A6.BY.13.PV85                   ---------------------------------------------   1005
A6.IT.02.60000                  ------X-----------------------------         1004
A6.RU.11.11RU6950               ---------------------------------------------   1004
A6.UA.12.DEMA112UA024           ---------------------------------------------   1004
B.AR.14.DEMB14AR003             --------------------------------G-------------   1004
B.BR.10.10BR_MG035              -------------------------S-------------------   1010
B.CU.14.14CU005                 ---------------------------------------------   1004
B.DE.13.366396                  ---------------------------------------------   1005
B.DE.13.947915                  ---------------------------------------------   1004
B.DQ.11.DEMB11DR001             ---------------------------------------------   1007
B.ES.15.100600                  ---------------------------------------------   1003
B.FR.11.DEMB11FR001             ----V----------------------------------------   1003
B.JP.12.DEMB12JP001             ---------------------------------------------   1007
B.PH.15.DEMB15PH003             ---------------------------------------------   1001
B.RU.11.11RU21n                 -----V---------------------------------------   1008
B.SE.15.102SE                   ---------------------------------------------   1004
B.TH.10.DEMB10TH002             ---------------------------------------------   1008
B.US.16.2609                    ---------------------------------------------   1004
C.BR.11.DEMC11BR035             ----V---K--------A--M-------------           1000
C.CN.10.YNFL19                  --------XK--------A---G-------------             1004
C.DE.10.622166                  --------K--------A---D-------------             1006
C.ES.15.100_118                 ----V---K--------A-------------               1009
C.ET.08.ET104                   ----V---K--------A--M-G-------------             1004
C.IN.15.NIRT008                 --------K--------A-------------               1000
C.MW.09.703010256_CH256.w96     --------K--------A---G-------------             1002
C.NG.10.10NG020523              ------------------A---G-------------             1000
C.NP.11.11NP016                 --------K--------A---G-------------             1000
C.PK.14.DEMC14PK009             --------K--------A---GG-N-------------           1000
C.SE.15.100ET                   ----V--------V--A---G-------------             1005
C.TZ.08.707010457_CH457.w8      ----V---KH--------A---D-----QNME.......           1004
C.US.14.M0118v1d14_5M04_C5      --------K--------A---G-----QNME.......           1015
C.ZA.13.DEMC13ZA152             ----V------------A---G-------------             1007
C.ZM.11.DEMC11ZM006             ------------------A---G-----QNME.......           1009
D.BR.10.10BR_RJ095              -----------------------G----G-------------       1002
D.CD.03.LA17MuBo                -----------------------G----G-------------       1005
D.CM.10.DEMD10CM009             -----------------------G----G-------------       1009
D.CY.06.CY163                   ----V-----------------G----G-------------       1003
D.KE.11.DEMD11KE003             ----V---------------------------------------   1003
D.SE.12.077UG                   -----------------------G-------------           1002
D.TZ.04.CO6405V4                ---------------------------------------------   1003
D.UG.10.DEMD10UG004             ---------------------------------------------   1003
D.UG.11.DEMD11UG003             ----V----------------------------------------   1005
D.YE.02.02YE516                 ----V------------------G-------------           1005
F1.AO.06.AO_06_ANG32            ---------------------------------------------   1008
F1.AR.02.ARE933                 ----------------------M-G-------------           1004
F1.BR.11.DEMF111BR037           -----------------------G-------------           1004
F1.CY.08.CY222                  I----------------------G-------------           1003
F1.ES.11.VA0053_nfl             -----------------------G-------------           1004
F1.FR.04.LA22LeRe               -----------------------G-------------           1005
F1.RO.03.LA20DuCl               -----------------------G-------------           1006
F1.RU.08.D88_845                -----------------------G-------------           1004
F2.CM.10.DEMF210CM007           ------------------------A-------------           1003
F2.CM.11.DEURF11CM026           -----------------------G-------------           998
G.CD.03.LA23LiEd                ---------------------S--G-------------           1003
G.CM.08.789_10                  ------R-L----------G---G-------------           1001
G.CM.10.DEMG10CM008             -----------------------G-------------           1003
G.CN.08.GX_2084_08              -----------------------G-------------           1008
G.CN.16.224GX                   -----------------------G-------------           1002
G.ES.14.EUR_0033                -----------------------G-------------           1004
G.GW.08.LA57LmNe                -----------------------G-------------           1002
G.KE.09.DEMG09KE001             -----------------------G-------------           1002
G.NG.12.12NG060409              --------L--------------G-------------           999
H.CD.04.LA19KoSa                ------------Q---------G-N-------------           1002
H.CF.02.LA25LeMi                -----------------------G-------------           1002
J.CD.03.LA26DiAn                -----------------------G-------------           1002
J.CM.04.04CMU11421              ----V---XX-----------G--X-------------           1000
K.CD.97.97ZR_EQTB11             -----------------------G-------------           1002
K.CM.96.96CM_MP535              -----------------------G-------------           1004
01_AE.CM.11.1156_26             -----------------------G-------------           1003
01_AE.CN.12.DE00112CN011        -----------------------G-------------           1002
01_AE.IR.10.10IR.THR48F         -----------------------G-------------           996
01_AE.JP.11.DE00111JP003        -----------------------G-------------           1002
01_AE.MM.14.fKSDU26             -----------------------GG-------------           1004
01_AE.PH.15.DE00115PH012        -----------------------G-------------           1004
01_AE.SE.12.070SE               I----------------------G-------------           997
01_AE.TH.10.DE00110TH001        ------------------------D-------------           1004
01_AE.TH.11.40436v09_04         -----------------------G-------------           1000
01_AE.TH.90.CM240               -----------------------G-------------           1009
02_AG.CM.10.DE00210CM013        -----------------------GG-T-------------         1004
02_AG.DE.09.701114              ----------E----------------------------------   1004
02_AG.KR.12.12MHI11_10746       --------K--------------G-------------           1004
02_AG.KR.12.12MHR9              ------N----------------G-------------           1008
02_AG.LR.x.POC44951             ----V-L----------------G-------------           1004
02_AG.NG.12.12NG060304          -----------------------G-------------           1004
02_AG.NG.x.IBNG                 -----------------------G-------------           1004
02_AG.PK.15.PK032               -----------------------G-N-------------         1003
02_AG.SE.14.098GN               ----V---K----------------N...........           1003
```

```
                                                p31 Integrase end
                                                   Pol end
B.FR.83.HXB2                 VPRRKAKIIRDYGKQMAGDDCVASRQDED..........*              1003
02_AG.SN.13.9580             ----------------------------G-----------             1012
03_AB.RU.97.KAL153_2         ----------------------------------------             1004
04_cpx.CY.94.94CY032_3       ---------------------N----G-------------             1003
05_DF.BE.x.VI1310            ----------------------------G-----------             1004
06_cpx.AU.96.BFP90           ----------------------------G-----------             1006
07_BC.CN.98.98CN009          --------K--------A----G-----------------             1004
08_BC.CN.97.97CNGX_6F        --------K--------A----G-----------------             1000
09_cpx.GH.96.96GH2911        ----------H----------G------------------             1003
10_CD.TZ.96.96TZ_BF061       -----V---K-------A---------QNME.......--             1007
11_cpx.CM.95.95CM_1816       I---------------------------G-----------             1002
12_BF.AR.99.ARMA159          ----------------------------G-----------             1004
13_cpx.CM.96.96CM_1849       ----------------------------G-----------             1011
14_BG.ES.05.X1870            ----------------------------G-----------             1003
15_01B.TH.99.99TH_MU2079     ----------------------------G-----------             1002
16_A2D.KR.97.97KR004         ----R-----------------------G-----------             1004
17_BF.AR.99.ARMA038          ----------------------------G-----------             1004
18_cpx.CU.99.CU76            ----------------------------G-----------             1002
19_cpx.CU.99.CU7             ----------------------------G-----------             1003
20_BG.CU.99.Cu103            ----------------------------G-----------             1000
21_A2D.KE.99.KER2003         ----V-----------------------G-----------             1003
22_01A1.CM.01.01CM_0001BBY   ------------E--------G-----N.-----------             1004
23_BG.CU.03.CB118            ----------------------------G-----------             1000
24_BG.ES.08.X2456_2          ----------H----------G------------------             1000
25_cpx.CM.02.1918LE          ----------------------------G-----------             1003
26_A5U.CD.02.02CD_MBTB047    --------------------G-------------------             1002
27_cpx.FR.04.04CD_FR_KZS     --------LKH----------G------------------             1002
28_BF.BR.99.BREPM12609       ----------------------------G-----------             1004
29_BF.BR.01.BREPM16704       ---------------------G------------------             1004
31_BC.BR.04.04BR142          -----V---K-------A--M-------QNME.......--             1014
32_06A6.EE.01.EE0369         ----------E---------E--G----------------             1004
33_01B.ID.07.JKT189_C        --------------------G-G-----------------             1004
34_01B.TH.99.OUR1969P        ----------------------------G-----------             1004
35_AD.AF.07.169H             ----------------------------G-----------             1003
36_cpx.CM.00.00CMNYU830      ----------N-----------------G-----------             1004
37_cpx.CM.00.00CMNYU926      --------------------M-G-----------------             1001
38_BF1.UY.03.UY03_3389       ----------------------------G-----------             1008
39_BF.BR.04.04BRRJ179        ----------------------------G-----------             1007
40_BF.BR.05.05BRRJ055        -----V---K------------------------------             1004
41_CD.TZ.05.CO6577V5         -----V----------------------------------             1011
42_BF.LU.03.luBF_01_03       ----------------------------G-----------             1008
43_02G.SA.03.J11223          ---------------------N----G-------------             1003
44_BF.CL.00.CH80             ----------------------------G-----------             1004
45_cpx.FR.04.04FR_AUK        -----V----------------------G-----------             1002
46_BF.BR.07.07BR_FPS625      ----------------------------G-----------             1004
47_BF.ES.08.P1942            ----------------------------G-----------             1003
48_01B.MY.07.07MYKT021       --------------------M-G-----------------             1001
49_cpx.GM.03.N26677          ----------------A-----G-----------------             1004
50_A1D.GB.10.12792           L------------------S--M-G---------------             1002
51_01B.SG.11.11SG_HM021      ----------------------------G-----------             1004
52_01B.MY.03.03MYKL018_1     ----------------------------G-----------             1004
53_01B.MY.11.11FIR164        ----------------------------G-----------             1004
54_01B.MY.09.09MYSB023       ----------------------------G-----------             1004
55_01B.CN.10.HNCS102056      ----------------------------G-----------             1004
56_cpx.FR.10.URF5_patient_A  ----------------------------G-----------             1006
57_BC.CN.09.09YNLX19sg       --------K-------------------G-----------             1000
58_01B.MY.09.09MYPR37        ----------------------------G-----------             1004
59_01B.CN.09.09LNA423        ----------------------------G-----------             1001
60_BC.IT.11.BAV499           -----V---K-------A--M-------------------             999
61_BC.CN.10.JL100010         --------K--------A----G-----------------             1000
62_BC.CN.10.YNFL13           --------K--------A----G-----------------             1000
63_02A.RU.10.10RU6637        --------K--------X----G-----------------             1004
64_BC.CN.09.YNFL31           --------K--------A----G-----------------             1004
65_cpx.CN.10.YNFL01          --------K-------------------G-----------             1000
67_01B.CN.11.ANHUI_HF115     ----------------------------G-----------             1002
68_01B.CN.11.ANHUI_WH73      ----------------------------G-----------             1002
69_01B.JP.05.05JPMYC113SP420 ----------------------------G-----------             1002
70_BF1.BR.10.10BR_PE004      ----------------------------G-----------             1006
71_BF1.BR.10.10BR_PE008      -----V----------------------G-----------             1004
72_BF1.BR.10.10BR_MG002      -------V--------------------G-----------             1005
73_BG.DE.01.9196_01          ----------------------------G-----------             1003
74_01B.MY.10.10MYPR268       ----------------------------G-----------             1004
77_cpx.MY.14.14MYNBB090      I---------------------------G-----------             990
78_cpx.CN.13.YNTC19          ----------------------------G-----------             1013
79_0107.CN.15.SX15DT013      ----------------------------G-----------             1004
82_cpx.MM.13.mSSDU12         --------K--------A----G-----------------             1000
83_cpx.MM.13.mSSDU94         --------K--------A----G-----------------             1000
85_BC.CN.14.14CN_SCYB2       --------K--------A----G-----------------             1000
86_BC.CN.13.15YNHS18         --------K--------A----G-----------------             1000
87_cpx.CN.12.DH32            ----------------------------G-----------             1000
88_BC.CN.05.05YNRL25sg       --------K--------A----G-----------------             1000
90_BF1.BR.11.BRG06043        -----V----------------------G-----------             1004
O.CM.96.LA51YBF35            --------L-H-------A-----G-T-SESMEQPGEIP-             1015
O.CM.96.LA52YBF39            -----------------T-SM--G-A-SESMEQPGEIP-             1010
O.CM.99.99CMU4122            ----------N------T-SM--G-X-SENVEQLGEIP-             1010
O.CM.x.pCMO2_3               ----------H------T-SM--G-T-SESMEQPGEIP-             1008
O.ES.01.Read25_HIV_GroupO    -----------------T-SM--T-SENVEQPSEIP-               1008
O.FR.06.RBF206               ---------Q----T--M----T-SESVEQPGEIP-                1012
O.GA.10.10Gab1190            -----------------T-SM--G-T-SESMEQPSEIP-             1009
O.GA.11.11Gab6352            -------------S-S---S-T-SENMEQPGEIS-                 1010
O.SN.99.99SE_MP1300          ----------H------T-SM--G-T-SENVEQPGEIP-             1010
O.US.10.LTNP                 -----------------T-SM--G-T-SESVEQPGEIP-             1010
N.CM.15.S4858                -------------------G--GG---NQEME.......--           1015
N.FR.11.N1_FR_2011           -------------------G----G---NQEME.......--          1015
P.CM.06.U14788               ----------N----------DT-R-SESLEQSG...--             1008
P.FR.09.RBF168               -------------------------DI-R-SESLE-                1004
CPZ.TZ.06.TAN5               ---------KE-------AGSMDD--N-T..........--           1000
CPZ.US.85.US_Marilyn         ---------K--------A-SM-G----S..........--           1006
GOR.CM.12.SIVgor_BQID2       -----------------X-SM-XG-T-GESMEQPGEIP-             1010
GOR.CM.13.SIVgor_BPID15      -----------------N----DT-G-SESLEQSGQVS-             1009
```

HIV-1 Proteins

HIV-1/SIVcpz Proteins

Vif

Vif start

```
                             MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHYESPHPRISSEVHIPLG.D.AR.LVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPELADQLIHLYYFDCFSDSAIRKALLGHIVSPRCEYQAG.HNKVGSLQYLALAALI.....T.PKKIK 160
B.FR.83.HXB2
A1.CD.02.LA01AlPr
A1.CM.08.886_24
A1.ES.15.100_117
A1.KE.11.DEMA111KE002
A1.PK.15.PK034
A1.RW.11.DEMA111RW002
A1.SE.13.078SE
A1.UG.11.DEMA110UG009
A2.CM.01.01CM_1445MV
A3.SN.01.DDI579
A4.CD.97.97CD_KCC2
A6.BY.13.PV85
A6.IT.02.60000
A6.RU.11.11RU6950
A6.UA.12.DEMA112UA024
B.AR.14.DEMB14AR003
B.BR.10.10BR_MG035
B.CU.14.14CU005
B.DE.13.366396
B.DE.13.947915
B.DO.11.DEMB11DR001
B.ES.15.100600
B.FR.11.DEMB11FR001
B.JP.12.DEMB12JP001
B.PH.15.DEMB15PH003
B.RU.11.11RU21n
B.SE.15.102SE
B.TH.10.DEMB10TH002
B.US.16.2609
C.BR.11.DEMC11BR035
C.CN.10.YNFL19
C.DE.10.622166
C.ES.15.100_118
C.ET.08.ET104
C.IN.15.NIRT008
C.MW.09.703010256_CH256.w96
C.NG.10.10NG020523
C.NP.11.11NP016
C.PK.14.DEMC14PK009
C.SE.15.100ET
C.TZ.08.707010457_CH457.w8
C.US.14.M0118v1d14_5M04_C5
C.ZA.13.DEMC13ZA152
C.ZM.11.DEMC11ZM006
D.BR.10.10BR_RJ095
D.CD.03.LA17MuBo
D.CM.10.DEMD10CM009
D.CY.06.CY163
D.KE.11.DEMD11KE003
D.SE.12.077UG
D.TZ.04.C06405V4
D.UG.10.DEMD10UG004
D.UG.11.DEMD11UG003
D.YE.02.02YE516
F1.AO.06.AO_06_ANG32
F1.AR.02.ARE933
F1.BR.11.DEMF111BR037
F1.CY.08.CY222
F1.ES.11.VA0053_nfl
F1.FR.04.LA22LeRe
F1.RO.03.LA20DuCl
F1.RU.08.D88_845
F2.CM.10.DEMF210CM007
F2.CM.11.DEURF11CM026
G.CD.03.LA23LiEd
G.CM.08.789_10
G.CM.10.DEMG10CM008
G.CN.08.GX_2084_08
G.CN.16.224GX
G.ES.14.EUR_0033
G.GW.08.LA57LmNe
G.KE.09.DEMG09KE001
G.NG.12.12NG060409
H.CD.04.LA19KoSa
H.CF.02.LA25LeMi
J.CD.03.LA26DiAn
J.CM.04.04CMU11421
K.CD.97.97ZR_EQTB11
K.CM.96.96CM_MP535
01_AE.CM.11.1156_26
01_AE.CN.12.DE00112CN011
01_AE.IR.10.10IR.THR48F
01_AE.JP.11.DE00111JP003
01_AE.MM.14.fKSDU26
01_AE.PH.15.DE00115PH012
01_AE.SE.12.070SE
01_AE.TH.10.DE00110TH001
01_AE.TH.11.40436v09_04
01_AE.TH.90.CM240
02_AG.CM.10.DE00210CM013
02_AG.DE.09.701114
02_AG.KR.12.12MHI11_10746
02_AG.KR.12.12MHR9
02_AG.LR.x.POC44951
02_AG.NG.12.12NG060304
02_AG.NG.x.IBNG
02_AG.PK.15.PK032
02_AG.SE.14.098GN
```

Vif start

B.FR.83.HXB2   MENRWQVMIVWQVDRMRIRTWKSLVKHHMYVSGKARGWFYRHHYESPHPRISSEVHIPLG.D.AR.LVITTYWGLHTGERDWHLGQGVSIEWRKKRYSTQVDPELADQLIHLYYFDCFSDSAIRKALLGHIVSPRCEYQAG.HNKVGSLQYLALAALI.....T.PKKIK 160

HIV-1 Proteins

```
                                                    Vif end
B.FR.83.HXB2              PPLPSVTKLTEDRWNKPQKTKGHRGSHTMNGH......*            192
A1.CD.02.LA01AlPr         ------R------------R-LQ----R--C......-            193
A1.CM.08.886_24           ------K------------R-RD--------C......-            193
A1.ES.15.100_117          ------R--A---------------E---------...-            193
A1.KE.11.DEMA111KE002     ----RI------------R-R--------------...-            193
A1.PK.15.PK034            ------R------------M-PQ-N-----CWSC...-            196
A1.RW.11.DEMA111RW002     ------R--A------R----P--N-I---C......-            193
A1.SE.13.078SE            ------R-----S-----RDP--N-----C......-            193
A1.UG.11.DEMA110UG009     ------R-----S-----R-P-----N--C......-            193
A2.CM.01.01CM_1445MV      ------R--V---------R----N-----C......-            193
A3.SN.01.DDI579           ------R------------R----------C......-            193
A4.CD.97.97CD_KCC2        ------R------------R----------C......-            193
A6.BY.13.PV85             ------R------------DR--N---S-C......-            193
A6.IT.02.60000            ------R------------R-R--------C......-            193
A6.RU.11.11RU6950         ------R------------R-R--N-----C......-            193
A6.UA.12.DEMA112UA024     ------R------------RDR--N-----C......-            193
B.AR.14.DEMB14AR003       ------A----------R-------N.........-            193
B.BR.10.10BR_MG035        ------R------------------K--L---...-            193
B.CU.14.14CU005           --V---A---------S----D-K-N---S--...-            193
B.DE.13.366396            ------A------------------E---------...-            193
B.DE.13.947915            ------S------------QD--N--HQ--H--...-            194
B.DQ.11.DEMB11DR001       ------R---------------------------...-            194
B.ES.15.100600            ------R---------R------------------...-            193
B.FR.11.DEMB11FR001       ------K--------V----------------Q......-            193
B.JP.12.DEMB12JP001       ----------------------------------...-            193
B.PH.15.DEMB15PH003       ----------------------R--N---C--...-            194
B.RU.11.11RU21n           ------------------R--D------IS--...-            193
B.SE.15.102SE             -----IA----------------------L......-            193
B.TH.10.DEMB10TH002       ------------------------M--K------...-            193
B.US.16.2609              ------K--------------------------S-...-            193
C.BR.11.DEMC11BR035       ------R--V--------R-IRDR--N--L---...-            193
C.CN.10.YNFL19            -----IQR-V-----N----M-R--N-----...-            194
C.DE.10.622166            ------R--V-----N----RDR--N------...-            193
C.ES.15.100_118           ------K--V--------R-R--K--LS--...-            193
C.ET.08.ET104             -----S--V--K-----R-R--N-------...-            193
C.IN.15.NIRT008           -----IQ--V-----N----R-R-------...-            193
C.MW.09.703010256_CH256.w96 ------S--V--K--------R-R--N------...-            193
C.NG.10.10NG020523        ------R--V-----N----R-R--N------...-            193
C.NP.11.11NP016           ------S--V--------N--R-R--N------...-            193
C.PK.14.DEMC14PK009       ------S--V--------RIR-R--N------...-            193
C.SE.15.100ET             -----IS--V------R--R-R--N------...-            193
C.TZ.08.707010457_CH457.w8 ------R--V-----N----R-R--N-I---...-            193
C.US.14.M0118v1d14_5M04_C5 ------Q--V-----N----R-R--N------...-            193
C.ZA.13.DEMC13ZA152       ------R--V-------------R--N-----...-            193
C.ZM.11.DEMC11ZM006       ------R--V---------------N----L...-            193
D.BR.10.10BR_RJ095        ------R------------R----N------...-            193
D.CD.03.LA17MuBo          ------R------------Q------N-----...-            193
D.CM.10.DEMD10CM009       ------R------------Q------T----...-            193
D.CY.06.CY163             ------R------------Q--K----S---...-            193
D.KE.11.DEMD11KE003       ------R------------R---E--------...-            193
D.SE.12.077UG             ------R------------HQ---R--R----...-            193
D.TZ.04.CO6405V4          ------R------------M--------------...-            193
D.UG.10.DEMD10UG004       ------R------------R--------------...-            193
D.UG.11.DEMD11UG003       ------R------------R-R-------S--...-            193
D.YE.02.02YE516           ------K------------Q--------------...-            193
F1.AO.06.AO_06_ANG32      ------R--V---------E--------------...-            193
F1.AR.02.ARE933           ------Q--V---------R--------------...-            193
F1.BR.11.DEMF111BR037     ------Q--V---------RDR-----------...-            193
F1.CY.08.CY222            -----IK--V---------E-R---N------...-            193
F1.ES.11.VA0053_nfl       ------R--V---------E-----E-------Y......-            193
F1.FR.04.LA22LeRe         ------K--V---------R-R-----------...-            193
F1.RO.03.LA20DuCl         ------K--V--------REIR------------...-            193
F1.RU.08.D88_845          ------R--V---------E--------------...-            193
F2.CM.10.DEMF210CM007     ------R--V---------N----R--------...-            194
F2.CM.11.DEURF11CM026     ------K--V-----N---R--------------...-            194
G.CD.03.LA23LiEd          ------K------------R--E-P--------...-            193
G.CM.08.789_10            ------R---------S--IR---E-P-----...-            193
G.CM.10.DEMG10CM008       ------Q---------S--IRD-KE-P--S--...-            193
G.CN.08.GX_2084_08        ------K------------------NPS--S--...-            193
G.CN.16.224GX             ------K----------------*ENPS-R--WNC...-            195
G.ES.14.EUR_0033          -----IK------------R-R-R--NP----...-            193
G.GW.08.LA57LmNe          ------R------------R------P------...-            193
G.KE.09.DEMG09KE001       ------R---------S--IR-R-ENL--T---...-            193
G.NG.12.12NG060409        ------R-----------------ENP--S-PWKC...-            196
H.CD.04.LA19KoSa          --------A-----E-----------I----...-            193
H.CF.02.LA25LeMi          ------S-----------H--R---E---S--...-            193
J.CD.03.LA26DiAn          ------Q--V------NS---R--------S--...-            193
J.CM.04.04CMU11421        ------Q--V---------N---R---------...-            193
K.CD.97.97ZR_EQTB11       --V---Q--V---------R--------S-Q......-            193
K.CM.96.96CM_MP535        --V---K--V------------R----------...-            193
01_AE.CM.11.1156_26       ------R------------R---ENP--S--WNC...-            196
01_AE.CN.12.DE00112CN011  ------K------------IR---ENPS----...-            193
01_AE.IR.10.10IR.THR48F   ------K------------I----ENL-----...-            189
01_AE.JP.11.DE00111JP003  ------K------------IR....P-------...-            193
01_AE.MM.14.fKSDU26       ------K------------IM---ENP-----...-            191
01_AE.PH.15.DE00115PH012  ------K------------R---ENP----Q......-            193
01_AE.SE.12.070SE         ------K------------------V---ENP-----...-            193
01_AE.TH.10.DE00110TH001  ------K------------IRA--ENP-----...-            193
01_AE.TH.11.40436v09_04   ------K------------NS--I---ENP----Y......-            193
01_AE.TH.90.CM240         -------.EI----------R---ENP-----...-            191
02_AG.CM.10.DE00210CM013  ------R--------KE--------------...-            193
02_AG.DE.09.701114        ------A--A---------IRD-----S-----...-            193
02_AG.KR.12.12MHI11_10746 ------K--A-----KD---RD----RS--C...-            193
02_AG.KR.12.12MHR9        ------K--A-----E----RD----S-----...-            193
02_AG.LR.x.POC44951       ------K--A-N-------IR-----RS----...-            193
02_AG.NG.12.12NG060304    ------K--A---------R-R-R--RS----...-            193
02_AG.NG.x.IBNG           ------R--A-----KE----R---RP-----...-            193
02_AG.PK.15.PK032         ------Q--A-----E----R-----M-----...-            193
02_AG.SE.14.098GN         ------R*-------E----R-----RS----...-            192
```

Vif

HIV-1/SIVcpz Proteins

Vif end

```
                                                     PPLPSVTKLTEDRWNKPQKTKGHRGSHTMNGH......*
B.FR.83.HXB2                                                                                  192
02_AG.SN.13.9580                    ------R--------E----R------S-S-........-              193
03_AB.RU.97.KAL153_2                ------------------R--D-------S-........-              193
04_cpx.CY.94.94CY032_3              ------K-V---------R-R-ENQI----........-              192
05_DF.BE.x.VI1310                   ------------------R--N-----Y........-                193
06_cpx.AU.96.BFP90                  -----Q--V---------R---E-.......-                      193
07_BC.CN.98.98CN009                 ----IK--V-----N---IR-R--N-.......-                    194
08_BC.CN.97.97CNGX_6F               -----IK--V-----N---IR-R--N-.......-                   194
09_cpx.GH.96.96GH2911               ------K--V---------R-R---ENQ----.......-              193
10_CD.TZ.96.96TZ_BF061              ------R-----------R-----------.......-                193
11_cpx.CM.95.95CM_1816              ------K-V---------R---EN----CWNC...-                  196
12_BF.AR.99.ARMA159                 ------K--V-----N---R---------.......-                 193
13_cpx.CM.96.96CM_1849              ------K--V---------RD--ENQ-----.......-              193
14_BG.ES.05.X1870                   ------K-----------R-R---NP-......-                    193
15_01B.TH.99.99TH_MU2079            ----R-K-------D----IR---ENP--D-.......-              193
16_A2D.KR.97.97KR004                ------R--------------------.......-                   193
17_BF.AR.99.ARMA038                 ------K--V-------E-R---------.......-                 193
18_cpx.CU.99.CU76                   ------R--A--------R----NQ-.......-                    193
19_cpx.CU.99.CU7                    ------R-----------R---ENP------.......-              193
20_BG.CU.99.Cu103                   ------R-----------R-R---ENP-.......-                  193
21_A2D.KE.99.KER2003                ------R----------I---E--M--.......-                   193
22_01A1.CM.01.01CM_0001BBY          ------R-----------R---EYP-----WKC...-                196
23_BG.CU.03.CB118                   ------R-----------R-R---DN-.......-                   193
24_BG.ES.08.X2456_2                 ------R-----------R-R---N-.......-                    193
25_cpx.CM.02.1918LE                 ------R--------E-----ENQ-.......-                     193
26_A5U.CD.02.02CD_MBTB047           ------R-----------------E----.......-                193
27_cpx.FR.04.04CD_FR_KZS            -S----R-----------R---ENQ----Q.......-               193
28_BF.BR.99.BREPM12609              ------R-----------I-----N----.......-                193
29_BF.BR.01.BREPM16704              ------R----------R-----------.......-                194
31_BC.BR.04.04BR142                 ------R--V-----R----RDR--NQ-.......-                 193
32_06A6.EE.01.EE0369                ------Q--V-----E----R--E------R.......-              193
33_01B.ID.07.JKT189_C               ------K-----------IR---ENP--S-.......-              193
34_01B.TH.99.OUR1969P               ------K-----------QIR---ENP-T---.......-             193
35_AD.AF.07.169H                    ------IR--A--------R-P-E-----C.......-               193
36_cpx.CM.00.00CMNYU830             ------K-----------R----RST-.......-                  193
37_cpx.CM.00.00CMNYU926             ------K-----------R---ENP---.......-                 193
38_BF1.UY.03.UY03_3389              ------K--V---------R-----------.......-              193
39_BF.BR.04.04BRRJ179               ----LK-----------T------------.......-               193
40_BF.BR.05.05BRRJ055               ----LK----NK-----R---------.......-                  193
41_CD.TZ.05.CO6577V5                ----IR-----------M-----------.......-                193
42_BF.LU.03.luBF_01_03              ------R-----------------------.......-               193
43_02G.SA.03.J11223                 ------A---------S---R-----RX----.......-             193
44_BF.CL.00.CH80                    ------Q--V--------E-----E-.......-                    193
45_cpx.FR.04.04FR_AUK               ------K--A----------RD----------C.......-            193
46_BF.BR.07.07BR_FPS625             ------K--V-------E-R----------.......-               193
47_BF.ES.08.P1942                   ------Q--VX-------------E---------.......-           194
48_01B.MY.07.07MYKT021              ------K-----------IR---ENP--S-.......-              193
49_cpx.GM.03.N26677                 -----TR-----------R-R--N-.......-                    193
50_A1D.GB.10.12792                  ------R--A------E----R-P-----I-.......-              193
51_01B.SG.11.11SG_HM021             ------K-----------I----ENPI-H-.......-              193
52_01B.MY.03.03MYKL018_1            ------K-----------I-D--EH------.......-              193
53_01B.MY.11.11FIR164               ------K-----------I--R-ENPI---.......-              192
54_01B.MY.09.09MYSB023              ------K-----------IR---ENL--S-.......-              193
55_01B.CN.10.HNCS102056             ------K-----------M---N-.......-                     193
56_cpx.FR.10.URF5_patient_A         ------A-----I-------R-----RP----.......-             193
57_BC.CN.09.09YNLX19sg              -----IKQ-V-----N---R-R--N-.......-                   194
58_01B.MY.09.09MYPR37               ------K-----------IR---ENP--S-.......-              193
59_01B.CN.09.09LNA423               ------K-----------IM---E-P-----.......-             193
60_BC.IT.11.BAV499                  ------R--V----------RDR--N---H-.......-             193
61_BC.CN.10.JL100010                ------K--V---------IR-R--N-.......-                  194
62_BC.CN.10.YNFL13                  ------K--V-----N---R-R--N--X---.......-             192
63_02A.RU.10.10RU6637               ------K----------------RS-.......-                   193
64_BC.CN.09.YNFL31                  ----IK--V----D---IM-R--N-.......-                    194
65_cpx.CN.10.YNFL01                 ----IK--V-----N---I--R---D--D-.......-               191
67_01B.CN.11.ANHUI_HF115            ------K-----------I----ENP----.......-              193
68_01B.CN.11.ANHUI_WH73             ------K-----------I----ENP-.......-                  193
69_01B.JP.05.05JPMYC113SP420        ------K-----------IR-L-ENP--H-.......-              193
70_BF1.BR.10.10BR_PE004             -----XK-----------I--K--------.......-               193
71_BF1.BR.10.10BR_PE008             ----------Q---------------KE-.......-                193
72_BF1.BR.10.10BR_MG002             --------------E-------K-N------.......-              193
73_BG.DE.01.9196_01                 ------R-----------R-R---NP----.......-               193
74_01B.MY.10.10MYPR268              ------K----X-------IR---EHP--S-.......-             193
77_cpx.MY.14.14MYNBB090             ------K-----------IR-L-ENPP---Q.......-             193
78_cpx.CN.13.YNTC19                 ------K-----------R---ENP-----.......-               193
79_0107.CN.15.SX15DT013             ------K---------Q---IR---ENPS-.......-              193
82_cpx.MM.13.mSSDU12                -----IQ--V-----N---IR-R--N-A-S-.......-             193
83_cpx.MM.13.mSSDU94                ------------------R--------S-.......-                194
85_BC.CN.14.14CN_SCYB2              -----IQ--V-----N---IR-R--N-L--.......-              194
86_BC.CN.13.15YNHS18                ------K--V-----N---IR-R--N--I--Q.......-            194
87_cpx.CN.12.DH32                   ------Q--V-----N---IR-R--N-.......-                  194
88_BC.CN.05.05YNRL25sg              -----IK--V---------IM-R--N-.......-                  194
90_BF1.BR.11.BRG06043               ------R--------D--R-----E-.......-                   193
O.CM.96.LA51YBF35                   -----IQ---------RHLRIRDQL---S-----.......-          193
O.CM.96.LA52YBF39                   ------Q---------RC--IRDQL---S-----.......-          193
O.CM.99.99CMU4122                   ------Q---------HLRIMDQL-N-S-----.......-           193
O.CM.x.pCMO2_3                      ------Q---------RHL-IRD-L-N-L--R-.......-            193
O.ES.01.Read25_HIV_Group0           ------Q---------HPRIRD.LE--S----WSS...-             195
O.FR.06.RBF206                     ------Q----K---H-RIR-QLENQS-H-.......-              193
O.GA.10.10Gab1190                  ------Q----E----HLR-RDQLEN-S--Y.......-             194
O.GA.11.11Gab6352                  ------Q-------K-S-HLMIRD-LE--S-----.......-         193
O.SN.99.99SE_MP1300                ------Q---------SRHLRIRDQL---S-----.......-         193
O.US.10.LTNP                       ------Q---------RHLR-RDQPE--S-----.......-          193
N.CM.15.S4858                      ----------------H-R-Q----NP--------.......-         193
N.FR.11.N1_FR_2011                 ----------------EH---Q----NP--------.......-        193
P.CM.06.U14788                     ------QI--Q-I-SN--R---XQE--S-S-CYKP...-             196
P.FR.09.RBF168                     ------QI--Q-I-SX--R-X-RQE--S-S-CYKP...-             196
CPZ.TZ.06.TAN5                     R--------------H-----PK-N--LC-C......-               199
CPZ.US.85.US_Marilyn               ----------------C----QPEN---S-.......-              193
GOR.CM.12.SIVgor_BQID2             ------Q--A-----RYPRIRDQLE--S----WKFWKN-             199
GOR.CM.13.SIVgor_BPID15            ------QL----K---R--I---Q-N-L-S-C......-             193
```

HIV-1 Proteins

Vpr

```
                                    amphipathic α-helix                     H(S/N)RIG motifs
                                    oligomerization                        Vpr end in HXB2
          Vpr start                                        frameshift in HXB2             Vpr end
B.FR.83.HXB2          MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIW.LHGLGQHIYETYGDTWAGVEAIIRILQQLLFIH#FRIGCRHSRIGVTRQRR....ARNGASRS*  96
A1.CD.02.LA01AlPr     ----------P---YQ--M--------H--------P-.---------N-H----L-----V---M-----V-.------Q----II-G----.-V-D--G-P-  97
A1.CM.08.886_24       -----------------Y---M--------------P-.----------N-------E---V----M-----V-.------Q----IIPG---.-V-D--SG--  97
A1.ES.15.100_117      -----------------Y-------------------P-.--Q------N-------E-----------V-.------Q----II-G---.--G---SD---  97
A1.KE.11.DEMA11KE002  -----------------Y-------------------P-.------DD----E-----K-----------Q----ILPR--.-G---S----  97
A1.PK.15.PK034        -------H-P--V--Y--L----D--H-------P-.-------------E----V-M-------V-.------Q----I------.-G------P-  97
A1.RW.11.DEMA11RW002  -----------------Y---M--D--H-------P-.---------D-H----E--------V--------V-.------Q----II-G----.-V-D-PG--  97
A1.SE.13.078SE        -----------------Y---M--------------L-.---------N-------E----------V-.------Q----II-G----.-I-D-S----  97
A1.UG.11.DEMA110UG009 ----------P---YEQ-M-D---D--H-------P-.------------D----V---KT------V-.------Q----II-G----.-D-SG----  97
A2.CM.01.01CM_1445MV  -----------------Y----------Q-------P-.--D-----DI---V---------V-.------Q----II-R----.-V-D-T--P-  97
A3.SN.01.DDI579       -----------------Y---M------Q-------P-.---------N-------E-----M----V-.------Q----II-RG-R--.-V-D----P-  98
A4.CD.97.97CD_KCC2    -----------------Y------------------V-.---------N--------E---------Y----Q----I-PS-G...V----P--  97
A6.BY.13.PV85         -----------------M--------DI-H------L-.--------N-------E---T-----V-.------Q----IVQR--.--G---S----  97
A6.IT.02.60000        -----------------Y---M--DI-H--------P-.--------Q-N-------E--------V-.------Q----ILQR---.-V---S----  97
A6.RU.11.11RU6950     -----------------Y---M--DI-H--------P-.--Q------Q-N-------E---T------V-.------Q----I-QR---.-V--SR---  97
A6.UA.12.DEMA112UA024 -----------------Y---M--D--H--NT--Q-------N-------E---T------V-.------Q----ILQR---.-V---S----  97
B.AR.14.DEMB14AR003   -----------------Y-----------R--G---P-.--N--------------------------Q----II-----.--------N---  97
B.BR.10.10BR_MG035    -------T---------Y----------R--------P-.--YV-----------------------Q----I-P-----.-------N----  97
B.CU.14.14CU005       ---V--------------Y------V---------V-.---Y--------------------------Q----IIP---.---------P-  98
B.DE.13.366396        -----------------Y------------------V-.--Y--G--------------------------Q----I------.-------S---  97
B.DE.13.947915        ---T-AG--Q--A-N-A--------A-------P-.---Y-----T---LL---HM-----Q----IIQ----.---.T------  97
B.DO.11.DEMB11DR001   -----------------Y----------R-------P-.--S---Y--Q---------T-----------Q----ILR-G--.  97
B.ES.15.100600        -----------------Y-------------------P-.--S---Y-------LL-----------Q----I-----  96
B.FR.11.DEMB11FR001   --R--------------Y---I-----R--------T-.--S---Y-Q------------------H----IN--T---.T-----  96
B.JP.12.DEMB12JP001   -----------------Y----------------P-.----L---N----------------Y-L--Q----I-------.--G---  97
B.PH.15.DEMB15PH003   --------------YDV-----------------P-.---Y-----------L--M-----------Q----I----.-------  97
B.RU.11.11RU21n       -----------------Y-Q-------------A--.--Y--------L--T--------V-.------Q----II-RS--.---H----  97
B.SE.15.102SE         -----------------Y--A-G----D--Q-------P-.--S---Y--------E----T------------Q----II------.  97
B.TH.10.DEMB10TH002   ------------H----------------P------P-.---------N------------------------Q----IIP-----.  97
B.US.16.2609          -----------------Y-------D--R---Y--T-.-Q---Y---------T---------L---T---L----I------.----T-----  97
C.BR.11.DEMC11BR035   -----------------YT--------Q-------P-.------------------------------Q----II-----  97
C.CN.10.YNFL19        ----S---H----Y--------Q-------P-.--------------LL---V-----------Q----IV----.-G-T--N---  97
C.DE.10.622166        -----P-----------Y----------Q-------P-.--M--------T------V-------Z---Q----I-----.---S----  97
C.ES.15.100_118       ----------------TYI--------Q-------P-.--N---Y---------T---------I--F----H-I---IL----.  98
C.ET.08.ET104         -----------------Y-----------Q------P-.--S---Y-G-------E-----------Q----I------.  98
C.IN.15.NIRT008       --P--------------Y------I-Q--------P-.--S---Y---------T---TL-M------------Q----ILP-----.----N---  97
C.MW.09.703010256_CH256.w96 --------------Y----------Q--------P-.--S---Y---------T----------------V-.------Q----II------.-----N--  97
C.NG.10.10NG020523    --P--------Y--A-------Q------------P-.---YV-----T---TL-M-----------Q----ILP-----.----N-P-  97
C.NP.11.11NP016       -----------------Y----------Q------P-.--Q--LYV-----------------------Q----IL-----.-T-------  97
C.PK.14.DEMC14PK009   -----------------Y-----------Q------P-.--S---Y---------T---LV-T-----M-V-.------Q----ILP----.  97
C.SE.15.100ET         --------N--------Y----------Q-------P-.--Y-D----R-----V---------V-.------Q----ILP----.  97
C.TZ.08.707010457_CH457.w8 --P--------------Y----A-I----------P-.--A--Q-N----T-------------------Q----IV-----.  97
C.US.14.M0118v1d14_5M04_C5 --P--------------Y---------I-Q------P-.--S---Y---------T------------------Q----ILQ----.-G---P---  97
C.ZA.13.DEMC13ZA152   -----------------Y----------Q------P-.--S-E--------T--V-----------H----IM-----.---.--TN-P-  97
C.ZM.11.DEMC11ZM006   -----------------Y----------I-Q------P-.--S--Y--------T---V----------------H----IL-----.-T--S---  97
D.BR.10.10BR_RJ095    -----------------Y----------S--------.--S------------T-----------T------Q----I-----.---S----  97
D.CD.03.LA17MuBo      -----A-----------Y----------S-S------V--Q----Y-N-------T-----L-----------Q----INL----.-T--S---  97
D.CM.10.DEMD10CM009   -----A-----------Y----SK----------S-.-------Y-N-------T-----------------Q----IIP-----.---S----  97
D.CY.06.CY163         -----A-----------Y----------S-----T-.-----------L-T--I---------Q----IIPP----.---S----  97
D.KE.11.DEMD11KE003   -----------------Y----------S-------A.--S--Y-----------------------Q----IL----.--I---SN--  97
D.SE.12.077UG         ----S------------Y----------S----A--S--Y-----------E---------------Q----II-H--.-T---S---  97
D.TZ.04.CO6405V4      -----A-----------Y----------S--------.--Y----------L-T----V-.------Q----IM-----.---L---S---  97
D.UG.10.DEMD10UG004   -----------------Y---A------R--------.--S----Y----------V-------------Q----IM-----.---L---S---  97
D.UG.11.DEMD11UG003   -----------------Y---A------R--------.--S--------------LT-S--V-.------Q----INP-----.-T----P-  97
D.YE.02.02YE516       -----A-----------Y----------S-------A.----------T-----L--.-----Q----ISP-----.-----S--P-  97
F1.AO.06.AO_06_ANG32  ------G----------Y-----------------P-.--S---Y-Q------L---L-M-M------V-P---I-P----.-V---N----  97
F1.AR.02.ARE933       -----------------Y---A-D----------E--K---S---Y-T------------------H----I-----.  97
F1.BR.11.DEMF111BR037 -----------------Y---------E--K----P-.--S---N--------E------------------H----IL--A---.L-H-SN-P-  97
F1.CY.08.CY222        ------G----------Y---A-------------P-.---------N-------E--------------------Q----I-M----.-V-------  97
F1.ES.11.VA0053_nfl   ------G----------Y---I-----S--------P-.----N-------E----TKT-----------I------.-V-------  97
F1.FR.04.LA22LeRe     -----------------Y---A-------P-.--N---Y-H-----E---L------------H----II-----.----  97
F1.RO.03.LA20DuCl     ------G----------Y---A------------P-.--------N-------E--------------------Q----II----.-V--T----  97
F1.RU.08.D88_845      ------G----------Y---A------------P-.--S----D--------E----------------Q----IL-----.-V-------  97
F2.CM.10.DEMF210CM007 --P--------------Y----------------E-.---M--Y--N-------E--------------------H----II----.---I--N---  97
F2.CM.11.DEURF11CM026 --P--------------Y----------H------E-.----Y-------G-Q----------T--.---Q----IN-----.-T---N-P-  97
G.CD.03.LA23LiEd      -----------------Y---A-------------P-.-----------N-------E--------H------Q----IL-R--.-V-D-P----  97
G.CM.08.789_10        ---V-------------N---------------P-.-Q--LY-N-------E-----------------Q----I-P----.-V---P----  97
G.CM.10.DEMG10CM008   --V-N------------Y---A------H-------P-.--Y-N-------E----------V-.------Q----INLR-R--.-V-D-PG--  97
G.CN.08.GX_2084_08    ----------F------------------T-----P-.--Q--L--D-----E-----------------Q----INLR--R..-V-D-SG--  98
G.CN.16.224GX         -----N---L--F-A------------T-------P-.--Q---L--D-----E------V----I------Q----INP----.-V-D-SG--  98
G.ES.14.EUR_0033      -----------------Y---------------T--.---YV-N-------E------M----M-----V-.------Q----ILP----.--R-D-S---  97
G.GW.08.LA57LmNe      -----------------Y---A------------P-.--Y---N-------E--------------------Q----I-PR--.-V-D-P----  97
G.KE.09.DEMG09KE001   ---V--------Y-D------------------P-.--Y--------------------------Q----IIP----.-V-D-PG--  97
G.NG.12.12NG060409    -----N-----------Y---M-----H-------P-.--N-------E--------------------Q----I-PR--.-V-D-SG--  97
H.CD.04.LA19KoSa      --R------------------------------Q----.----N--D---------------------Q----I-------.-V---S----  97
H.CF.02.LA25LeMi      -----------P---Y----------------P-.-Q---Y-N-------G---LV------------Q----II----.-V--T-P--  97
J.CD.03.LA26DiAn      ---L-------------Y---A------S------P-.--N-------N----------------------Q----IIP----.-G-------  97
J.CM.04.04CMU11421    --P--------------Y---------------X--.---X------X----E-RX----------------Q----IMP----.-G-------  97
K.CD.97.97ZR_EQTB11   -----------------Y----I-I-----R-----P-.--D-----NN------E----------------Q----IL-----.--GG-----  97
K.CM.96.96CM_MP535    -----------------N-----I----H--------P-.--N--------T----E-L----------------Q----IIP----.--G-------  97
01_AE.CM.11.1156_26   -----------------Y----------S------T-.----D----------------------V-.------Q----II-G---.-G---PG--  97
01_AE.CN.12.DE00112CN011 ----------------F--------S---Y-T------E--------------V-.------Q----IVPG---.-G---T--P-  97
01_AE.IR.10.10IR.THR48F -----------D-----------H-----Y-NQ----E--------------V-.------Q----IIPG---.-G-------  97
01_AE.JP.11.DE00111JP003 ------------------Y------------P-.-----------------------------Y----Q----IIPG---.-G-H-TG--  93
01_AE.MM.14.fKSDU26   -----H--------------------S-------H-.--F--NH----E--------------------Q----IIPG---.-G--G--Y  97
01_AE.PH.15.DE00115PH012 -----------------Y------------------P-.----EY--N----E----------Y-.----Q----ILPG---.-G--T----  97
01_AE.SE.12.070SE     -----G------N-----I------------P-.--L--Q--Y--N----E----------------Q----IIPG---.-G-------  97
01_AE.TH.10.DE00110TH001 -----------------Y---------GP--------V---------N-------E----------------Q----IIPG---.-G-------  97
01_AE.TH.11.40436v09_04 ---L--------Y--I---------------P-.--Q--Y--N-------E------------------Q----IIPG---.-G--G----  97
01_AE.TH.90.CM240     ------.E-----Y-------------P-.--Q--Y-NN----E--------------------Q----IMPG---.--G---TG--  96
02_AG.CM.10.DE00210CM013 --R---------F------------H---P-.----E--------Q--Q--------V-.------Q----IIPG---.-G--G----  97
02_AG.DE.09.701114    -----------------F----V----H------P-.---------------E-Q----------------Q----IIPG---.-G-H--G--  97
02_AG.KR.12.12MHI11_10746 --RP--------F---M-----H------P-.---Y-------------E------------V-.------Q----II-G---.-G---G----  97
02_AG.KR.12.12MHR9    --R---------F------V---------P-.--N-------E-G--T-H----V-.------Q----II-G---.-G---G----  97
02_AG.LR.x.POC44951   -----------------F----------Q------P-.--EY----N----E----------V-.------Q----II-G---.-G---S----  97
02_AG.NG.12.12NG060304 -----------------F------------H----P-.--S-------N----E-I--M--------------Q----II-G---.-G---G----  97
02_AG.NG.x.IBNG       --R------S------------H----------P-.------------E-K------------V-.------Q----IIQG---.-G---G-P-  97
02_AG.PK.15.PK032     --RT--------D-------------H-----P-.---Y-------E--RV-V-M----------Q----IIPG---.-G----G-P-  97
02_AG.SE.14.098GN     --R-----------F---------------P-.------HN--------CEW--------P--V-.SG----S----INPSL--.-G-T------  97
```

Vpr

HIV-1/SIVcpz Proteins

amphipathic α-helix
oligomerization

H(S/N)RIG motifs
Vpr end in HXB2

Vpr start　　　　　　　　　　　　　　　frameshift in HXB2　　　Vpr end

```
B.FR.83.HXB2                       MEQAPEDQGPQREPHNEWTLELLEELKNEAVRHFPRIW.LHGLGQHIYETYGDTWAGVEAIIRILQQLLFIH#FRIGCRHSRIGVTRQRR....ARNGASRS*    96
02_AG.SN.13.9580                   --R-----------F------------H--------P---------------------Q------V-.----Q-----IIPG-...G---SG---          97
03_AB.RU.97.KAL153_2               ------------Y--------------S-------V--.-S---Y------------------V-.-----Q-----IIQR-...-------            97
04_cpx.CY.94.94CY032_3             ------------N-------------------E-------------------------Q---I-P-...#E-GD-----                        98
05_DF.BE.x.VI1310                  ------------Y--I--------H--------P-.-S--Y--N----------E-----------Q-----II------I---S---              98
06_cpx.AU.96.BFP90                 ---------------------P----.NI--Y--S------LL-T--L--V-.----Q-----IIP-----G-----                          97
07_BC.CN.98.98CN009                ---S--------Y---------P--.-S---Y-N----------------------Q-----IL-----T------                          97
08_BC.CN.97.97CNGX_6F              ---P--------------Q------P---.-YV------T---T-L-------------Q-----IV---------                          97
09_cpx.GH.96.96GH2911              ------------N------------P-----------------E--------------Q-----II-G--R---G-D-G---                     98
10_CD.TZ.96.96TZ_BF061             ------------YL--------------N--------T---------------------P-----IV------T-T--                       98
11_cpx.CM.95.95CM_1816             ------------Y--M-------H---------P-------N------V---------Q-----IIPG-...G-------                       97
12_BF.AR.99.ARMA159                ---P--------Y---------P--------N------V---------------Q-----IM-----V-----                          97
13_cpx.CM.96.96CM_1849             ------------N---------H--MKQ----P-----------E----T-M-------Q-----V-----G--S---                      97
14_BG.ES.05.X1870                  ------------Y----------P-.----Y-N----E----XL-X-------Q-----I-P---V-D-S---                        97
15_01B.TH.99.99TH_MU2079           -G------------Y-G-P--------K.-S--P-------E----V--H---V-.----H-----AIIPG--A...R---PG---                 97
16_A2D.KR.97.97KR004               ------------Y----------D--R----P*.-S---Y-A-----T-----------M---.----Q-----IM.P-G-...---                 95
17_BF.AR.99.ARMA038                ----G--------Y-------P-.-S---Y-N----E-------------------H-----II------V--D---                      97
18_cpx.CU.99.CU76                  ------------Y---------P----------Y-N-----E-------------Q-----IV----------G-------                      97
19_cpx.CU.99.CU7                   -----N------Y--A-------D------P--.----Y-N----E--L---------Q-----I-PR--.--V-D-PG---                    97
20_BG.CU.99.CU103                  ------------Y---------P--.-V--N-N------L---------------Q-----ISLR----V--D----                      97
21_A2D.KE.99.KER2003               --------D----------S-----P-.-S---A------M------V---------H-----I-------S---                        97
22_01A1.CM.01.01CM_0001BBY         ------V-Y------V------------GP.---E--N------E-------------V-----Q-----II-G--...G---G-PG---              97
23_BG.CU.03.CB118                  --------Q-----------------P-.-V--N-N------E-------------T---.---Q-----ISPR--...G---PG---                97
24_BG.ES.08.X2456_2                ------------Y---------P--------A----N---E-----------------Q-----ISPR----R-D-SG---                   97
25_cpx.CM.02.1918LE                --RT--------N---------P--.----Y-N----E-V--------------Q-----IIP---.--G---G-P---                    97
26_A5U.CD.02.02CD_MBTB047          ------------Y-----------Q------P-----------E-V-----------V-----Q-----IF-S--...G---G-P-                  97
27_cpx.FR.04.04CD_FR_KZS           ------------N---I-----------P-.-YV-N------E-V--V----------Q-----IIP-----G---P-P-                    97
28_BF.BR.99.BREPM12609             ------------Y----------S---------------------L---M----------Q-----IN------                           97
29_BF.BR.01.BREPM16704             ------------Y----------S---------------------------------Q-----INP------                            97
31_BC.BR.04.04BR142                ------------N---------Q----P-.-Y-N------E-----------------V-----Q-----II-------                       97
32_06A6.EE.01.EE0369               --R---------S----K----P-.-S---Y-N----T------------------N-----IIP---...G---T---                     97
33_01B.ID.07.JKT189_C              -------------------K---P-.-X----N------X---------------X----Q-----I-PG---...G---G---                  97
34_01B.TH.99.QUR1969P              -----A--------------H-------P-.-Y--DN-----E-------------V----.---Q-----ILPG---...G---G---              97
35_AD.AF.07.169H                   ------------Y--M-----D--H-------P-.-Q----Y-N-----------------L-----I--G---.----V-D-SG-P-               97
36_cpx.CM.00.00CMNYU830            ----------F----------H-------P-#.---Y-------E------M--------V-----Q-----II-G--...G---G---               97
37_cpx.CM.00.00CMNYU926            ------------Y--A------------P-.-------Q--------Q---K-------V-----Q-----I-PR--.---V--P--                 98
38_BF1.UY.03.UY03_3389             ------------Y---------A------P-.-S---Y-N----E----T-T-------H-----IH------V--N---                    97
39_BF.BR.04.04BRRJ179              --------H----Y---------S------P-.-N--Y--I----T-----------Q-----II-L---.--S---S---                   97
40_BF.BR.05.05BRRJ055              ------------Y---------S--.----Y-N---------------------------Q-----II-G--...G---G---                   97
41_CD.TZ.05.CO6577V5               -----N------Y---------S--.-Q--F---------------------------Q-----IF------S---                        97
42_BF.LU.03.luBF_01_03             ------------Y------------------P-.---Y-----------V---------------INIR--------N---                    97
43_02G.SA.03.J11223                --V---------X---------H-------P-.---Y-N------E--------X-------Q-----II-G--...G-H-G---                  97
44_BF.CL.00.CH80                   ------G------Y----------S-----P-.-Y--N------E-------------------H-----II-------V-----                 97
45_cpx.FR.04.04FR_AUK              ------------Y--M-------P--.-------V--------------------Q-----IIPG-...V-D-PG---                       97
46_BF.BR.07.07BR_FPS625            ------G------Y----------P--.----Y-N-----E-------------------Q-----II------V---G---                    97
47_BF.ES.08.P1942                  ------------Y---------P-.----Y-N----E---------M-H---------Q-----N-K--II----T-------                  97
48_01B.MY.07.07MYKT021             ------------Y---------T--.-------Y-N----E-------------------Q-----IIPG----G----R---                   97
49_cpx.GM.03.N26677                ------------Y--I------Q-.----S--Y-N----E-V--K-----------H-----II----.--G-----P-                     97
50_A1D.GB.10.12792                 --R---------Y-Q-A-DM--D--H-----P-.-S---Y-N----E-V---------Q-----II-G--...G---S---                    97
51_01B.SG.11.11SG_HM021            ------------Y-A----------------P-.----Y-------E------T------V-----Q-----I-G-----G-P---                 97
52_01B.MY.03.03MYKL018_1           -------A-Y--A---------------P-.-Y--Y--TN-----E-------------Q-----I-PG---...G---G---                    97
53_01B.MY.11.11FIR164              ------------N---------H-------Q-.--X--N------E----------X-------Q-----IIPG---...G---G---                97
54_01B.MY.09.09MYSB023             -------------------S-P-.-L----N------RE-----------------Q-----IIPG---...G-----                      97
55_01B.CN.10.HNCS102056            ------------Y---------P--.-S----N-----E-----------------------Q-----ILPG---...G---G---                97
56_cpx.FR.10.URF5_patient_A        ------------S---------------P-.------------E-----------------Q-----ILP-G--...V-D-SD---                 97
57_BC.CN.09.09YNLX19sg             ---P--------------Q------P-.-YV------T------K------------------IL---------                           97
58_01B.MY.09.09MYPR37              -------------------T-----P-.-----Y-N------E-----------------Q-----I-PG---...G-----G---                97
59_01B.CN.09.09LNA423              --------H----Y---------P-.-Y--H------E------------------------Q-----IIPG---...G-----G---              97
60_BC.IT.11.BAV499                 ------------Y-A--------Q--.----Y------T-------------V-.-V-GQ---II--N--...........N--                   97
61_BC.CN.10.JL100010               ------------------Q------P-.-S---YV-DN----------------------Y---H-----IV------                         97
62_BC.CN.10.YNFL13                 --P---------YX---------T--.----Y-----------------T----T------Q-----IL------                           97
63_02A.RU.10.10RU6637              -RP--H------Y-------------P-.-YV--N----T--T-------------Q-----I--G--...G---S---                       97
64_BC.CN.09.YNFL31                 ------------Y---------Q----P-.---Y-N----T------------------------Q-----IM------                         97
65_cpx.CN.10.YNFL01                ---S---------G--------Q-.-----N------E------------------------Q-----ILPG---...G---G---                 95
67_01B.CN.11.ANHUI_HF115           -----N------Y------------P-.------Y-N----E-------V------------Q-----ILPG---...G---G---                97
68_01B.CN.11.ANHUI_WH73            ------------Y------------P-.-Y--S------E------------V---------Q-----ILPG---...G---G---                97
69_01B.JP.05.05JPMYC113SP420       ------------Y-A-A------------P-.-Y--N-----Q--------------------I-------G---                          97
70_BF1.BR.10.10BR_PE004            ------------Y--A-------S-.-------D--------------------------Q-----IIQR-...-----TG---                   97
71_BF1.BR.10.10BR_PE008            ------------Y------------P-.------R-----LVK----------R-Q---ILLR-...-T----G---                        97
72_BF1.BR.10.10BR_MG002            --R---------Y---------V--.-----Y--N------E---------------------Q-----I-PS--...                        97
73_BG.DE.01.9196_01                ------------Y---------P-.-----Y-N----E----------------------Q-----IIP---...V--PG---                   97
74_01B.MY.10.10MYPR268             ------------A-Y--A------S-.------P-----N-----E--L-M--------Q-----IIPG---...G---TG---                 97
77_cpx.MY.14.14MYNBB090            ---------------S----I-----P-.----Y-NX-----E----------------Q-----IIPG---...G---G-P-                   97
78_cpx.CN.13.YNTC19                ------------Y-A---------P-.-----Y--N------E--------------------Q-----IMPG---...G---PG---              97
79_0107.CN.15.SX15DT013            --P---------F---------D--.-------Y-N----E----------X--M---V-.----Q-----IVPG-...G---G---               97
82_cpx.MM.13.mSSDU12                --P---------R---------Q-.-------Y-N------E------------------Y---.--I-PG---...G---G---                   97
83_cpx.MM.13.mSSDU94               --P---------F----------Q-.------P-----------H---T----------------IVI------                           97
85_BC.CN.14.14CN_SCYB2             --P---------F----------P-.-YV-D-------------------------------------IVI------                          97
86_BC.CN.13.15YNHS18               --P---------Y-Q-AI------Q-.------LY-----------------------------Q-----IV----------P-                  97
87_cpx.CN.12.DH32                  --P---------Y---------Q-.-S---S--Y-D-----------V-----------------Q-----II------                       97
88_BC.CN.05.05YNRL25sg             --------H----Y---------P-.-YV------T---T----------------------IL------                               97
90_BF1.BR.11.BRG06043              --RP--------------------S-.------Y-----T--L-----V-.------Q-----                                       97
O.CM.96.LA51YBF35                  ------------A--F--------N-.-QA--Y-D-----E-M------A-------------Q-----INPSNPRRRGR-D-S---                101
O.CM.96.LA52YBF39                  --V-------A--F--------N-.-QT---Y-------I-M----I--T-------------Q-----INPSNTRGRGR---S---                101
O.CM.99.99CMU4122                  --------H-A--F--A-T-D-A------P-.--Y--N------V-I-M--------T------Q-----INPSNTRGRGR---S---                101
O.CM.x.pCMO2_3                     ----------P--F--*A--T-D-A------P-.---Y-N------V-M-M--------------Q-----INPSNTRGRGR---S---                100
O.ES.01.Read25_HIV_Group0          ------------P--F--A--------P-.----LY-------V-M-----------T-.Y----Q-----INPTNTRGRGR---SNS---             100
O.FR.06.RBF206                     ---N--V--V-A------------A-.----N--QA--Y-D--------E------I-A---.----Q-----INPSSTRGRGR---S---            101
O.GA.10.10Gab1190                  ------N---F--F-------I-T--------P-.----Y-D-----E-G------------T-.Y----Q-----INPSN-RGRGR---S---          101
O.GA.11.11Gab6352                  ---N-----P--F-----------A-.----N--QA--Y-D--------E------I-A---.----Q-----INPTNPRGRGR---S---            101
O.SN.99.99SE_MP1300                ------------A--F-------S---.-QA--Y-D-----V-M------H--A----------Q-----INPSNTRGRGR---S---              101
O.US.10.LTNP                       ----------A--F--A------T-.------H-.-------V--N------R-M------Y-.-Y----Q-----IDPRG--.-GR---S---         98
N.CM.15.S4858                      -------A----Y--A---------------Y--N------E------------M------Q-----I-P---.-R----                      96
N.FR.11.N1_FR_2011                 --R--L---F--Y-----------------P-.------Y-N------E-----------Q-----I-P---I---                         96
P.CM.06.U14788                     ---P---------P---F----M-QT----E------P-.-S---Y-N------E-T------------I---.Y----Q-----ILTPS-R.-GR-H-P---   99
P.FR.09.RBF168                     ---P---------P---F----M-QT----A---P--LQY.---S--Y-N------E-T---------I---.Y----Q-----ILPPS-R.-GR-H-P---   99
CPZ.TZ.06.TAN5                     ------------YTL-L-DT--I-Q----TV.-Q-V-NWVFTVH.-S-E--QEL-----RA--T-.Y-H--A----                          84
CPZ.US.85.US_Marilyn               --V-------A----Y---A-DV------E--L--------P-.-Q---Y-S-----------E-----------------Y---H-----ISLA--TPGGR---  101
GOR.CM.12.SIVgor_BQID2             --V-------A---F--A---V------A--LK---P-.---A---------------E-----------I-T--Y---H-----IIPSNARGRRR---P----   101
GOR.CM.13.SIVgor_BPID15            ---------R-P---F---M--T----A------A-.---P-.--S-----D------E-T--------I---.Y---P-----IAPHPQR..GR-H-PN---    99
```

Tat

```
                        Tat start          C-rich region          nuclear localization          exon 1 end exon 2 start                    Tat end
B.FR.83.HXB2            MEPVDPRLEPWKHPGSQPKTACTNCYCKKCCF.HCQVCFITKALGISYGRKKRRQR..R.RAHQNSQTHQA.........SLSKQPTS.QP.RG..DPTGPKE*KKKVERETETDPFD...............*   100
A1.CD.02.LA01AlPr       --L---N-D-N------T---NS------W.---L--LR-G--FH------R-...GTPHS-KN--N.........PIP--I-.T.QR..S---E-S-----SKAKS-R-A............           102
A1.CM.08.886_24         -D---N----N------R-T--SS-----W.---L--LR-G---------R-...GTPYS-KN--N.........P-PE-SLP-.T.Q..EW---E-S-E---SQ----R-............            102
A1.ES.15.100_117        -D-I--N----N-----R-A--N--SY.-N--L-N-G---------R-...-SP-SKND--S.........PIPE--L-.-A.Q-.-S---E-S-----SKAAA-R-............            102
A1.KE.11.DEMA111KE002   -D----NI--N------I-P-NK------Y.--L--PA--LN-G---------...GTP-S-KN--S.........PVP---IP.RA.Q-.IS----SE----------............            102
A1.PK.15.PK034          ------N----N-----T-P-NK---V--Y.-L--G-G--------SK...GTP-S-KD--N.........PIPE--I.RA.P-.I----ST--M-G----...............            102
A1.RW.11.DEMA111RW002   -D----N----N-----V----P--LS-G--V-----KR-...GTP-SND--N.........PVP--SIP-.T.K..N---E-S--E-SKA---R-...............            102
A1.SE.13.078SE          -D---N-A--N------R-E-P-N----.--L--LQR-G-...-A-PPS-KD--N.........PVPQ--LP-.A.QR..S---E-S------SK--P-R-...............            102
A1.UG.11.DEMA110UG009   -D---N----N------T-P-SK----E-Y.-I---QH-G------------W---...-T--G-KD--N.........PIP---LP-.A.P-.-S------------SKA---R-...............            102
A2.CM.01.01CM_1445MV    ------K----N-----R----N----V--W.----LN-G---------P...-PS-S-KD--N.........PV-E-SLP.HA.QR..VS--Q--S--E--SK---HQ-...............            102
A3.SN.01.DDI579         ------N----H--------S------A--W.---LN-G---------K-...-GTP-SKD--N.........PVP---LP-.T.P-.------S----SK----R-...............            102
A4.CD.97.97CD_KCC2      ------N----E------P-NR-----Y.--P---VN-G---------KL-...-TPRSNKD--N.........PIP---L-.-A.QR.------E-S--E--SKA-P-R-...............            102

A6.BY.13.PV85           -D---S-D--N---------SS----R--L----I--LK-G---------H-...-GTSH--ED--N.........HI----LP.HT.QR..-Q--E-ST----SKA-AAQL...............            102
A6.IT.02.60000          -D---N----N---------SK------W.----LN-G---------R-...GTP--NKD--N.........PIPN--LP.HT.QR.VQ--E-S------SKA---R-...............            102
A6.RU.11.11RU6950       -D---N----N---------V-----W.---L--LK-G---------R-...GTPHS-KD--N.........PI----LP.HT.QR..-Q--E-SA----SKA---R-A...............            102
A6.UA.12.DEMA112UA024   -D---N----N-----------------W.---L--LK-G---------H-...GTPH--KD--N.........PI----LP.HT.QR.NQ--E-P-----SKA----L-...............            102

B.AR.14.DEMB14AR003     ------------------R----P-------------------G---N-------...-SP-G-EN--V.........------------L.----P------S------R-............            102
B.BR.10.10BR_MG035      ---I----------------P----------PK--L-------...----PKH----D.........P----S--S.H-.-------W----T--A-DRW............            103
B.CU.14.14CU005         ----------------------S---W-Y.----M--G----------...-PP-D-E---D.........------AT-.L-.----------S-E------V--R-............            102
B.DE.13.366396          -D-----A--E-----------NT-----L.----SR----------...--S---------.........P-PN--A-.H-.---------S----------GS............            102
B.DE.13.947915          ------S---------------E-P-T--------L.---MK---------...GPPK-N---V.........-------T.-H.-.-----------K-----DHWCMDS............            107
B.DO.11.DEMB11DR001     -------K-----------K-F--R--Y.--A--L-G-------...-TSED---R-G.........-P---A--.-.-N--Q-SE-T---------VG............            102
B.ES.15.100600          -------A--E-----------R----L----------G----------...-PP-D--AN-D.........-P-R--A-.---.------S-T-------QV............            102
B.FR.11.DEMB11FR001     -----------------NK--R-Y-L----G---------...-TP-D-P----.........------------.---.------S-E---KA--H-DA............            102
B.JP.12.DEMB12JP001     ------S----N--------N-----P---TK-G---------...-PP-G-----V.........-P-P---A-.-L.-----S-E------QV-............            102
B.PH.15.DEMB15PH003     ------S---------S-P-A--------A--TK---------G-...-P-S---D.........P----SAP-.L-.-----S-E--------SH............            102
B.RU.11.11RU21n         ---------------------L----MK-G---------...-SP--N--D.........-P--A--.---.------S--------H-............            102
B.SE.15.102SE           ------------------R---------M--R-G--Y-----...-P-DN--D.........-PE---.-T-.-----P----SK-----VH............            102
B.TH.10.DEMB10TH002     ------------------------TK-G---------...-SP-D--N.........-Y-P--A-.---.-----Q----S-E------KH............            102
B.US.16.2609            ------S---------------P--P---R---.--YA--M--G---...-P-D-E---E.........A-PE-SDP.--.--------E-S--T-----A--N-............            102

C.BR.11.DEMC11BR035     ----------N----------SP---S-SY.--L---Q--G--------RA--G-PPSGED--N.........LI----L-.-A.--Q--SE-S-----SK............            104
C.CN.10.YNFL19          L--I--N----N-----D---N--S-SY.--L--QK-G---------...PI----L-.RT.K-.-S--SE-S------SK-K---A............            102
C.DE.10.622166          -D---N----N---------N--V-SY.--L--G--Y---------...GTPPSNKN--T.........PV---LP-.T-.-N--SSE-S----SK............            102
C.ES.15.100_118         ------N----N---------NK-H----Y.-P--L-G--Y---------...-SRS--KD.........-P---L-.-A-.--Q--SE-S------SK............            102
C.ET.08.ET104           ------N----N---------NK-F---SY.--L---G--G--------...S-PPS-KD--D.........PI----L-.-----S---R---V.PA-E---V............            100
C.IN.15.NIRT008         ---I--N----N-----R----N---SY.--L--QK-G---------...S-PPS-EN--D.........LV----LP.R-.H-..NQ-VSE-S-----SK............            102
C.MW.09.703010256_CH256.w96  ---I--K----N---------NK-------SY.--L--QK-G---------...N-P-P--L-.T-.N---SE-S-----SK............            103
C.NG.10.10NG020523      ---I--N----N-----S-P-NS---H-SY.--L--QK-G---------W---...S-PPSNKD--G.........-IRN--LP.RA.N-.NQ--E--S--EM-SK--R-Q............            103
C.NP.11.11NP016         ------N----N-----NQ---R-SY.--L--QK-G--Y---------...--PPS-ED--N.........LI----LP.RT.Q-.-Q--SE-S-----SK............            102
C.PK.14.DEMC14PK009     ------N----N---------P-N------SY.--YK--L-G---------G-...S-PPS-KD--N.........LIPN-SLP.RT.Q-.-QP--E-S-----TK-TAA............            102
C.SE.15.100ET           ------N----N---------NK-F-S-SY.--T--L-G---------...K-PPS-KD--N.........LI----L-.-T.-.-Q--SE-S-----SKA-----............            102
C.TZ.08.707010457_CH457.w8  --Q--K----N---------NT-F--R-SY.--L--Q--G--------K---...S-PPS-ED--N.........PI----LP-.T-.----SE-S--E--SK-T---L............            102
C.US.14.M0118v1d14_5M04_C5  -D---N-D--N--------N-P-NQ-----C-SY#-L---Q--G--------K---...S-PPS-ED--N.........-I----LP.-T.-.N---SE-S----SKA-------............            102
C.ZA.13.DEMC13ZA152     ---I--N----N---------P-N------SY.--L--Q--G--Y---------...N-PPS-KD--N.........-P--LPQ-.-.--N----SE-S-----SK--A---*............            102
C.ZM.11.DEMC11ZM006     -D---N----N---------NK---Y.--L--Q--G--Y---------...S-PPG-KD--N.........PI----SP-.A.-.N---SE-S-----SK-KA--YA............            102

D.BR.10.10BR_RJ095      -D---N----N-----E-P-NK-H---------Y.-----G---------...KPPNGD-A--D.........PVPE--S-.--.---------T-Q----SWA-------W.............            103
D.CD.03.LA17MuBo        -D----NIA--N--------S----A--W.---Y-----T----S-G-----------G-...-PPHG---I.........PVP--S-.---.H-.---------Q---------TT--A--L-W.............            103
D.CM.10.DEMD10CM009     -D---N----N--------S-----NK---Y.----R---S-G-----------G-...KPPKDNT---V.........PVPD--S-.-.H-.--------P----NPS--E-ASK--A----W.............            103
D.CY.06.CY163           -D---N----N--------S-----NS---R--Y-.--C-----G-----------G-...-PP-GNK---G.........PVPE--S-.--.H.P-.-P-P--NPS--E-ASK--A----W.............            103
D.KE.11.DEMD11KE003     -D-I--S----N-----R-P-NK-------Y.-----Q--VS-G-----------...-P-GG-V--D.........PIP---S-.--.---------------SKA-A---AW.............            103
D.SE.12.077UG           -D---N----N-----R-P-NK-------Y.-----N-G-----------...-TP-GG---N.........PIP--S-.--.L-.-.-Q---------SKA-A---AW...STDYLHLSGTT-           114
D.TZ.04.CO6405V4        -D---S----N-----NK-------SY.---P-----L-G-----------...-PP-DNKA--D.........PIPE--S-.--.---E----Q-K------S--A------W.............            104
D.UG.10.DEMD10UG004     -D----NI--N-----N-P-NK-----N-Y.-Y---------...-PP-SG-A--N.........PIP---LP-.T-.------N-----S--------Q--CWEDYRHLSGTT...-           114
D.UG.11.DEMD11UG003     --L--N-D--N-----R-P-NK---------Y.-----N-G-----------...GTP-GG-A--Y.........PVP---S-.-L.-.-----------SKA----------W.............            103
D.YE.02.02YE516         -D---N---------R-S-P-SP---------------G-----------G-...-SP-SN----V.........PVP---S-.--.------------SQAKAP---W.............            103

F1.AO.06.AO_06_ANG32    --I--N-D--N-----T-P--K------YY--G--G--------...PVPE--L-.-A.-.---------N------S--EM-SK---S--*............            101
F1.AR.02.ARE933         --L--N-D--N-----T-P------R-----YR-LAR--G-...H.-TPHGT--I--D.........LVP---L-.-A.-.---N-------S--E-SKAK---C-............            102
F1.BR.11.DEMF111BR037   -D-I--N----N-----P-P-NK---Y.----YW--A--G-...H.-TP-G--I--D.........-VP---S-.-A.-.-.NQ---------S--E-SKA----CA............            102
F1.CY.08.CY222          --------D--N-----T-P--R---------YY--A--G-...H.-TP-S--N--D.........PV---L-.-A.-.-------N------S--E-SKA---C-............            102
F1.ES.11.VA0053_nfl     ------S-D--N-----A-P-S-----W.--V--G-----------...H.--PP-S-PL--D.........PV----I-.-A.-.------N--------S--E-SKA--N-C............            102
F1.FR.04.LA22LeRe       --L--N-D--N-----K-T------K------R---Y-----------...--PP-SN--I--D.........PIPE--L-.-A.-.-----N--------S--E-SKAK----*............            101
F1.RO.03.LA20DuCl       ------N----N-----K-T-P-NK-----Y.---R--T--G-------...H.--PPS-S--D.........LVP---L-.-A.-.----N-------S-----SKA-----C............            101
F1.RU.08.D88_845        --L--N-D--N-----A-P--K--------Y.---R--TM--G-------...GTX-S--I--N.........PVPE--S-.--.----------Q-----S--SK-G--*............            101
F2.CM.10.DEMF210CM007   --I--N-D--Q-----E-P-N-------------R---Y----------...-SP-S-KV--D.........PVP---LP-.T-.--N-------------E-S-----SK--G--*............            102
F2.CM.11.DEURF11CM026   --I--E-D--N-----E-P--K----------L--TS-G--L-------...-S--G-EI--D.........PIP---L-.-T.-.-NH----EKS--E-SKAK----*............            102

G.CD.03.LA23LiEd        -D----NI--N------Q-P-NK----R--W.----LN-G---------KH-...-GPS-S-KD--N.........PVP---L-.LP.IA.-.--N--------S--E-AS---A-............            102
G.CM.08.789_10          -D---N----N------N-P-NK----A--W.----LN-G---------KH-...-GPS-G-KD--N.........PVPE--IS.---.N---Q--S--E-ASKA-A-Q............            102
G.CM.10.DEMG10CM008     -D---N----N------N-P-NK----A--W.----LN-G---------KH-...-GSSHG-KD--N.........PVP---IP.TT.-.--N--Q--S--E-ASKA--Q............            102
G.CN.08.GX_2084_08      -D---N----N-----K----NK----A--W.----LK-G---------KH-...ATPSSNKD--N.........PVP---IP.TT.-.--N-------------SK-K------............            102
G.CN.16.224GX           -D---N----N------N-P-NK----A--W.--L--LK-G---------KH-...ATPPS-KD--N.........PVP---IP.TT.-.--N----------SK-............            102
G.ES.14.EUR_0033        -D---N----N------N-P-NK----V--W.----LH-G---------KH-...-GPP-S-ED--N.........PVP---LP.TT.-.--N--------S------SK............            102
G.GW.08.LA57LmNe        -D---N----N------N-P-NK----V--W.----LK-G---------KH-...-GPP-G-KD--N.........PVP---LP.TT.-.--N-A--E--S--SK-K---L-............            102
G.KE.09.DEMG09KE001     -D---N----N------N---R----N--V-W.----LK-G---------KH-...--SPG-KD--N.........PVPE--LP.TT.-.--N-A-----S--E-SK----............            102
G.NG.12.12NG060409      -DL---NI--N------N---------N---V-SW.----LK-G---------...TPSGNKD--N.........PVP---VP.TT.-.------KQ--E--SQ-K---*............            100

H.CD.04.LA19KoSa        -D----NR--N------A-P-NK----A--Y.--YL--S--G---------...ETPRGN--D-.........HI-E--L-.RI.H-.---K---------SKA-A----W.............            103
H.CF.02.LA25LeMi        ------NQ---N------N-P-N---------LK-G---------...-GTPASL-NN--N.........LV----L-.-A..-LL-.RA.---A----E-ASK-----H.............            102

J.CD.03.LA26DiAn        ------NK---N---------NQ-----------Y.--M--LK-G---------...-T--PPG--N--D.........LIP---LP-.T.QR..KQ--E--S--E--SK--P-R-............            102
J.CM.04.04CMU11421      ------NI--N------R----Q-----------LK-G------N---...XVXE--LP.LT.SR..KQ----Q--E-SKA-PXR-............            102

K.CD.97.97ZR_EQTB11     ------NI---NQ--------NQ------Y.--L--LQ-G---C---E----...T.TTPYA-KN--KD.........PIP---LP-.A.--.----------S--E-SK-K---*............            101
K.CM.96.96CM_MP535      -D---NI---NQ--------NK-----Y.---L--LK-G---------N-------P-...TTPY--EN--D.........P-R---L-.-.---EQ-D---S-----SK-K--Q............            102

01_AE.CM.11.1156_26     -D---N----N--------T--SK---V--W.----H--L-G---------KH-...-GTPPG--N--N.........PVP---LP.II.-.--N----E-SK------ASKA............            102
01_AE.CN.12.DE00112CN011  ------N----N--------T--S--------W.----I--LK-G---------KH-...-GT--S-KD--N.........PVPE--LP.TT.-.--NS------S--E-ASK----C-............            102
01_AE.IR.10.10IR.THR48F   ------N----N-----P-V-SK----I--W.--L--LK-G---------KH-...--TP-SNKN--D.........PIPE-SLP.TS.-.--N--D---S--E-ASK---G-CA............            102
01_AE.JP.11.DE00011JP003  ------N----N-----T-D-SK---------W.-L--LK-G---------KH-...-GTP-S-KD--N.........PIPE--FP.TT.-.--N--D---S--E-ASKA----CA............            102
01_AE.MM.14.fKSDU26       ------T--T-N-----T-V-SK---V--W.-L--LK-G---------KH-...-GTP-S-KD--N.........PVP---RT.XS.-.--N--DQ--S--E-ASKA----*............            102
01_AE.PH.15.DE00115PH012  ------N----N--------T--SK-----W.----LK-G---------KH-...GPP-SRKD--N.........LIPE-SPP.II.-.--D---S----ASKA----+H............            102
01_AE.SE.12.070SE         ------N----N-----I---S-----V--W.----LK-G---------KH-...-GPS-S-KD--N.........-VP--S-T.H.-.--N----S--E-ASKA---------*............            101
01_AE.TH.10.DE00110TH001  ------NI---N--------T-P-S-------W.----LK-G---------KH-...-GN-PS-KD--N.........HIPE-SLP.II.-.--K-------S--E-ASKAK--*............            102
01_AE.TH.11.40436v09_04   ------N----N--------S-----SW.---L--LK-G---------KH-...-GTP-SRKD--N.........PVPE--LP.II.-.--NQ--SE-----ASTA-A--RA............            102
01_AE.TH.90.CM240         --L--N----N--------T--SK---V--W.----LK-G---------KH-...-GTP-S-KD--N.........PIP---LP.II.-.R.--N--D---S--E-ASKA----QC-............            102

02_AG.CM.10.DE00210CM013  --L--N----N--------T--SK---V--W.--L--LK-G---------KH-...-GTP--RRD--N.........PVP---LP.TT.-.--N--E-P--E--SK-----Y-............            102
02_AG.DE.09.701114        ------N----N--------R-A--S-----I--W.--L--LN-G---------...-GTP--D--N.........PVP---P-.IT.S-.--N--QE-S-----SK-----Y-LRTAS............            107
02_AG.KR.12.12MHI11_10746  --L--S----N--------T--SK---M--W.-L--LN-G---------RG-...-GTP-SH-D--N.........PVP---LP.TT.-.--N--N---S--E-ASKA--NQC............            102
02_AG.KR.12.12MHR9_        --L--S----N--------T--SK---V--W.-L--LK-G---------RG-...-GTP-SH-D--N.........PVP---LP.TT.-.--N---S--E-ASKA--NQC............            102
02_AG.LR.x.POC44951       -D---SI---N--------T--SK---M--W.--L--LK-G---------...GSPRSN-DR-N.........PVP---LP.TA.-.--N--S--E-ASK----NLC-Y-............            103
02_AG.NG.12.12NG060304    --L--S-D--N--------T--N----I--W.---Y--LK-G--------Q-R-...GPP-SREDR-N.........-VP---LP.TT.-.--N---S--S-TE-ASK-N-*............            101
02_AG.NG.x.IBNG           --L--S----N--------T--SK---M--#.--L--N-G---------R-...-GTP-SR-D--N.........PVP---LP.TT.-.--N-----S--E-SK-K---C-............            101
02_AG.PK.15.PK032         --L--N-S-N--------T--NT----I--W.--L--LK-G---------R-...-GTP-SR-D--N.........LV---SLP.TS.-.--S-DQ--S--E-ASK----CA............            102
02_AG.SE.14.098GN         L-L--N----N--------T--SK----M--W.--L--LN-G---------R---S-...GTPPSR-D--D.........PVP---LP.TT.-.--D---S--E-ASK-----*............            101
```

Tat

Tat start   C-rich region   nuclear localization   exon 1 end exon 2 start   Tat end

```
B.FR.83.HXB2                 MEPVDPRLEPWKHPGSQPKTACTNCYCKKCCF.HCQVCFITKALGISYGRKKRRQR..R.RAHQNSQTHQA.........SLSKQPTS.QP.RG..DPTGPKE*KKKVERETETDPFD...............*  100
02_AG.SN.13.9580             -DL---S----N------T--SK------W----L--LN-G-------KH-...GTP-SR-D--N----------PVP---E-S--E-ASK--K--S-----------------------------  102
03_AB.RU.97.KAL153_2         ------------------------------L----MK-G----------------P-DN--D-V------------P--A---------------M----M------------------------  102
04_cpx.CY.94.94CY032_3       -D-----N----------T-D-NK-F----W---------LK-G-------KH-...GSL-G-KG--N----------LIP---L-Q--N---S---E-Q----ASK--A---A-------------  103
05_DF.BE.x.VI1310            -D----N----N------R--NQ---M--Y---N-------------P---PP-G--A--D.--PVPE-P-----.-----KQ--E-S--A-Q--W.--------------------------  103
06_cpx.AU.96.BFP90           ------KI---NQ---K------K------Y---LN-G---------------Q-PPG-KN--D----------PV---L-.T.QR.EQ---EKS--E--KA-P-R--------------------  102
07_BC.CN.98.98CN009          ------K--S----------------------MK-G---F---------------S-P-S-ED--N----------LI---LP.RT.Q.-------S---SK-K----------------------  102
08_BC.CN.97.97CNGX_6F        ------N----------E--N------R-SY---L--QK-G-------------S-P-S-ED--N----------LI---SLP.RT.Q.-QA-SE-S----SK-KPA----------------------  102
09_cpx.GH.96.96GH2911        --L---N----N------N--Y----Y--H--LN-G-----------A-SCG-KSN-D-------PIP---L-.T.--.N------S--E--SK-KS-Q---------------------------  102
10_CD.TZ.96.96TZ_BF061       ----H--N----N------NK-H--A-SY---L--Q--S---------S-PPS-ED--N----------LI---S-.----.------------A-A---WWTDS----------------------  107
11_cpx.CM.95.95CM_1816       ------NI---NQ--------Y---M--LK-G---------------P-SHS-KN--D-------HIPE--L-.-A.TR.-K--V-E-S--E-SK--P-R--------------------------  102
12_BF.AR.99.ARMA159          --L---N----N---K-S-P--K----R----YW--V--G-----------P-G--N----------I-.-A.----.----S--E-SKAKS--C-------------------------------  102
13_cpx.CM.96.96CM_1849       -D----KI---NQ-------P-K------------G----------------LIP---LP.I.----.NQ------S--E--SK-------WSADS-----------------------------  107
14_BG.ES.05.X1870            -D----N----N--------K--R-W------LN-G-----------KH-...GPP-SXKD--T----------PVP--LP.TT.--.N---E-S------SK-K------------------------  102
15_01B.TH.99.99TH_MU2079     -D---LN----N---R-T--N----I--W----I--LK-G-------KH-...GTP-S-KD--N----------PAP---L-.IS.---.-K--E-S--E-ASKA----*-------------------  101
16_A2D.KR.97.97KR004         -D----N----N------R--NK----Y--P--LN-G----------P-...-PP-GGTG--N----------PIPE-SLP.R-.QR.-I---TE-S-TE--SKA---C--------------------  102
17_BF.AR.99.ARMA038          --I---N-D--N-------T-P--R-F--E---YW--T--G--N----------PP-GNPP--...-AP--A.--.N----S--E--SKA--Q*-W--------------------------------  102
18_cpx.CU.99.CU76            ------N----NQ-------N------R--Y---LK-G----------------TTPYD--KH--D...LIPE--L-.-RT.SW.-------Q-E--ASKA-N-QC------------------------  102
19_cpx.CU.99.CU7             -D----N----NT---------N------Y---LN-G-----------------GPP-SH-D--N----------PVP---IP.DT.-E.-G--T--S-E--SK-K------------------------  102
20_BG.CU.99.CU103            -D----N----N------NK---V--W----LN-G-------KH-...GPPPG-KD--N----------PVP---IP.TT.--.N------S-EE--SK-T---CA------------------------  102
21_A2D.KE.99.KER2003         -D----N----N------R--NK-H--R-Y---IK--L--G--R--------PP-GD-A--V----------PVP--S-.----.---------S-A-----C-------------------------  103
22_01A1.CM.01.01CM_0001BBY   --L---S----N------T--N--F--V--W---L--LK-G-------KH-...GTP-G-KD--N----------PIP---LP.T.KR.--S---E-S--E--SK---NR-A--------------------  102
23_BG.CU.03.CB118            -D----N-D--N------N----N------W---L--LN-G-------KH-...GTP-D-KD--I----------PVP--LP.TT.--.N------EKS-EE-ASK-KA----------------------  102
24_BG.ES.08.X2456_2          -D----N----N------NK-F--V--W----LN-G-------KH-...GSP-S-KD--N----------PVP---IP.TT.--.N------EKS-EE-ASK-KA----------------------  102
25_cpx.CM.02.1918LE          ------N----N---G-P-N--N------------K-G-----------S--GTPPS-KD--N----------PVP--IP.TT.--.NQ---E-S-----YA-------------------------  102
26_A5U.CD.02.02CD_MBTB047    ---RNID----N----P-NK----Q--Y--L----G-----------P-...GTSRSN--D----------P-P--S--.A.QR.ESP--E-S-------SK--Q-R--------------------  102
27_cpx.FR.04.04CD_FR_KZS     -D----N----NQ-------NQ--R-Y---H--LK-G--H-------------T-PYGNKN--D...PIQE-SIP.-T.QR.--S--SE-S-----SKARA-R----------------------  102
28_BF.BR.99.BREPM12609       ------Q----R--NK------------L-------MK-G--------------P--D--NN-V-------A-.-S.--.N----E-S--------T--A-----------------------------  102
29_BF.BR.01.BREPM16704       ------N----N------R--NK------------G------------------T--D---D-V-------A-.-F.----.---------S--T---A-----------------------------  102
31_BC.BR.04.04BR142          -D----K----N------N--R-SY-------LK-G----------------S-PPS-ED--N----------LI----VP.-S.--.N---E-S-----SK----C-------------------  102
32_06A6.EE.01.EE0369         -----HKI---NQ---R--P--P------Y--YM--LN-G--Y----------T-PPG-KN--N----------PE----V-.-T.QR.-E---EK----K--SKATPNR----------------  102
33_01B.ID.07.JKT189_C        ------N----N----AT-SK---I--W----L--LK-G-----------KH-...-PP-S-ED--N----------PX-E--L-.IS.---.N--D-E-S---ASK----RC----------------  102
34_01B.TH.99.QUR1969P        ------N----N------R-SK---V--W----L--LK-G---H----KH-...GTP-S-KD--N----------PIPE--LP.II.---.NQ----S---E-SKAA-------------------  102
35_AD.AF.07.169H             -D----N-D--N------P-P-S---I-SY----LN-G----------R-...TPP-RD--D--N----------PIP---LP.RA.K---LSA-QE-S----SK--P-R----------------  102
36_cpx.CM.00.00CMNYU830      --L--NI----N------R---N--F------W---LN-G-----------R-...G-P-SR-D--D----------PV----LP.AI.-K.N------S-E--ASK---N-C---------------  103
37_cpx.CM.00.00CMNYU926      -D----N----N----------------TK-G-----------------S---T-A-SR-DN-H-------VP---LP.TT.--.N----E-S-----ASKA-----CA------------------  102
38_BF1.UY.03.UY03_3389       --I---K----N------T-P-NK---R----.----YW--T--G-----N---------TP-SN-I--D...PVP---V-.-A.----.N----S--E--SK-K----CA---------------  102
39_BF.BR.04.04BRRJ179        -D----------------K--------Y-------T--G--------------TP----I--D----------PVPE--I-.-A.HR.N------S----SKAK----*-----------------  101
40_BF.BR.05.05BRRJ055        -D----------------N----------------G-----------------SP-GR-S--GSLSSSK*LQV...SAP.-Q.--------E-S-------CA--------------------------  111
41_CD.TZ.05.CO6577V5         -D----N----N------R--NK-H-----Y--P--LN-G--X------------PP-SNK--D...PIPE--L-.-T.--.Q--SE-S-----SK----C--------------------------  102
42_BF.LU.03.luBF_01_03       ---I--K----E-------------------R--G---------------------NPPPD--Z--------------AT.-----------------Q---------DL-------------------  102
43_02G.SA.03.J11223          --L---S----N------T--SK------W----L--LN-G--------...GTP-SR-D--N----------PV---LP.IT.SE.-N-------Q-E--SK------------------------  102
44_BF.CL.00.CH80             --LI--N-D--N------T-P--R---R-Y---YW--V--G-----------PP-SD-I--D----------PVP---L-.T.--.Q--SE-S-----SKA---CA---------------------  102
45_cpx.FR.04.04FR_AUK        -D----N----N------T--NK-H---Y---YS--LN-G-------KH-...GTPYS-KN--N----------PIP---L-.-T.--.N------E-S-----ASK-K-----------------  102
46_BF.BR.07.07BR_FPS625      --L---N-D--N------T-P--K---R--Y------V--G--------...GPP-S--I--D----------LVP---I-.-A.QR.N-----E-S--E--SKA-G-W-------------------  102
47_BF.ES.08.P1942            ------------------R------R--Y------------G------------XPP-SG---N----------T-.L.----E-S--------V-RRADS----------------------------  107
48_01B.MY.07.07MYKT021       ------N----K-T--SK------W----RL--LK-G-----------KH-...GTSHRRED--N...PIPE--L-.IS.---.N--D--S-----ASKA-G-C------------------------  102
49_cpx.GM.03.N26677          ------KI---N------------Q------Y---L--LK-G--C--------K----T-PPG-KN--D...LVQ----L-.T.QR.-K-.--E-S--E--SKAGP-R--------------------  102
50_A1D.GB.10.12792           -D----N----N------T-P--K------Y-V-PA--LN-G-----------G--PPKGD-A--V----------PIP--S-.---.-------K----------SKA-A-Q--W----------  102
51_01B.SG.11.11SG_HM021      ------N----N----G-P--A---R-S------A--K-------------------T--D-E--------A-.--.-S.--.N------S-E------HSSRQRLDS--------------------  107
52_01B.MY.03.03MYKL018_1     ------N----N------T--SS----I--W----L--LK-G--H-----KH-...GTP-S-KD--N----------PVPE--FP.II.----.N------S---SE-ASKAA--------------  102
53_01B.MY.11.11FIR164        ------S----N----X-T--N------W----L--LK-G--H----KH-...GTP-S-KD--N----------PIP---LP.IX.---.NNS-D--S-E-ASKA--X-C------------------  102
54_01B.MY.09.09MYSB023       ------N----N------T--SK------W----L--LK-G-------KH-...GPP-S-KD--N----------PIPE--LP.TT.---.N------S--E-ASKA-G-*-------------------  101
55_01B.CN.10.HNCS102056      ------N----N------T--SK------W----L--LK-G-------KH-...GNP-SR-D--N----------PVPE--LP.II.---.N------S--E--SK---CA--------------------  102
56_cpx.FR.10.URF5_patient_A  -D-I--N----N---------NK------W---LN-G-X---------SH-...-TPPG-KD--N----------PVP---LP.TI.N..---.N---QE-S-----ASK------*-----------  101
57_BC.CN.09.09YNLX19sg       ------N----N--------N------R-SY---L--QK-G-------------STPPS-ED--N----------LI---LP.RT.Q.-.-Q-SE-S------SK-KP-R-------------------  102
58_01B.MY.09.09MYPR37        ------N----N---T-D-SK---I--W----L--LK-G-------------PVPE--LP.TT.---.NQ-N----S-E--ASQ------CA----------------------------------  102
59_01B.CN.09.09LNA423        ------N----N------T--SK------W----L--LK-G-------KH-...PVPE-SSP.II.---.N------S---SKA-----YA------------------------------------  102
60_BC.IT.11.BAV499           ---LETN----NQ-E-----P-N----R-SY.------L--G--Y--------S-PSSNKD--N----------PV---L-.-A.--.-Q-DLEKS----KSK-K--------------------  102
61_BC.CN.10.JL100010         ------N----N------E--N--F--R-SY.---L--Q--G---------------P-S-ED--S----------LI---LP.RT.Q.--.T--SE-S------SK--P-------------------  102
62_BC.CN.10.YNFL13           ------N----N--------E--N--R-S-.---L--QK-G----------------PIP---LP.-I.H.---.----SE-S-----SK-K---A------------------------------  102
63_02A.RU.10.10RU6637        -D----N----N------E--N--F----Y---L--LK-G-------GR-...GTS-SR-D--N----------PVP---LP.TT.--.N------S--E--SK--QC--------------------  102
64_BC.CN.09.YNFL31           ------N----N------E--N--F--SY.--L--HR-G--V--------------PPSGED--N----------PI---SLP.KA.Q.-------SE-S------RA-Q--------------------  102
65_cpx.CN.10.YNFL01          ------N----Q--N--F----------X--TK--------------------TP-S-EDY-N----------P----L-.II.---.NN--DQ---S-E--ASKA----S--------------  103
67_01B.CN.11.ANHUI_HF115     ------N----T-D-SP-----VW-.---L--LK-G-----H-----KH-...-SPRS-KD--N----------PIP---L-.IS.---.N--D-E-S-----ASKA----C----------------  102
68_01B.CN.11.ANHUI_WH73      ------N----T---SK---AW-.---L--LQ-G-------KH-...-SPRS-KGY--N----------P-PX---L-.II.---.N------EKS------ASK----QG---------------  102
69_01B.JP.05.05JPMYC113SP420 ------N----N-----------W----L--LK-G-------KH-...-RGTP---KD--N...PIPE--LP.II.---.NNQ-D--KS----ASKA----C----------------------  104
70_BF1.BR.10.10BR_PE004      --L---K----------R----K----Y---------------------W------SHS--N--V----------P-------------PE----------QDN------------------------  102
71_BF1.BR.10.10BR_PE008      ------Q----R----T--SY---S--L-G----------------------G-P-SG---------------A-.-L.----.------E-S-------G---HLGNQWMDS--------------  107
72_BF1.BR.10.10BR_MG002      ------K----------N--F---Y--R-------KPP-D-----------------P---.-T.-L.----.N----E-S------A--QKH---------------------------------  107
73_BG.DE.01.9196_01          -D----N----N---------------W-------LN-G-------KH-...GPP-S-ED--N----------PVP---LP.TT.--.N---E-S-----SK-K-N-------------------  102
74_01B.MY.10.10MYPR268       --L---N----N------T--SK------W----L--LK-G-------KH-...GTP-S-KD--N----------PIPE--L-.IS.-R.N-D--S----ASKA---L*-------------------  101
77_cpx.MY.14.14MYNBB090      ------N----N-----T-P-S----------W----L--LK-G------K----.-TP-S-ED--N----------LTP-SP-.FS.--.N---Q--Q--ASA----*-------------------  101
78_cpx.CN.13.YNTC19          -D----N----N------K-T--NK------W-----I--LK-G-----H-----------------GTP-S-KDR-N----------PVP--SLP.FS.--.NQ--------SKA---QCA-------  102
79_0107.CN.15.SX15DT013      ------N----N------K-SK------W----L--LK-G------H---KH-...GTP-S-KS--N----------P----L--.R-.Q.---------E-S----ASK-----C------------  102
82_cpx.MM.13.mSSDU12         ------N-----M-T--SK----N-W----L--LK-G---------R--KH-...-PPC-KD--X----------PVP--LP.RT.Q.--.-Q-SE-S-----SK-KA-R-------------------  102
83_cpx.MM.13.mSSDU94         ------N----N--------N------------L-------------MK-G---------TP-S-KD--N----------PTPE--LP.RT.Q.-.-Q-SE-S----ASKA-Q-QYX------------  103
85_BC.CN.14.14CN_SCYB2       ------N----N------K------------MK-G---F---------------S---R-ED--N----------LI---LP.RT.Q.D-Q--SE-S------SK-K----------------------  103
86_BC.CN.13.15YNHS18         ------K----N------K-------R--------K-G------------------P--ED--D----------LI---SLP.RT.Q.-----SE-S-------SK----------------------  102
87_cpx.CN.12.DH32            ------N----N------------------MK-G---F--RT--P--.---K-PPD--AN-D----------------LP.RT.Q.-K--SE-S-----SK-KA--C--------------------  102
88_BC.CN.05.05YNRL25sg       ------N----N--------N------R-SY---LA--QK-G---------------STPPS-ED--N----------PI---L-.RT.Q.-N--SE-S-----SK------A--------------  102
90_BF1.BR.11.BRG06043        ------N----------R--------------TR------N----------------G-PPDGSAR---------P-------S.--S--V-P-DG----------------------------------  101
O.CM.96.LA51YBF35            -D----EVP-H--------P-N----R--Y.-P---TK-G---H-------R-P.A-AASHSANKD...PVPE-SL-.ST.SR.KQKRQE-QEE---T-AG-G*YHCLQDSCNCCTRTSGQ-  115
O.CM.96.LA52YBF39            -D----EVP-H------Q-P-NT-----Y.-YL--TR-G---H------R-P.A-ATAASCPDNKD...LVPE--L-.IT.NR.-K*KRQE-QE-E--K--DPS*-HRRKGSCHSCTRTSGQ-  114
O.CM.99.99CMU4122            -D----EMP-H--------P-NK------Y.--Y---TS-G---------R-P.A-AAHPDNKD...LVPE-SPT.IT.NR.KQERQE-QE---K--GP-*YPCRQVSCICCTRTSGQ-  114
O.CM.x.pCMO2_3               -D----EMP-H------Q-P-NK------Y.--Y--TS-G---------R-P.T-AARHPDNKD...PVPE-SP-.IT.NR.-KQERQE-QE-E--K--GPN*HACHKDSCYSCTRTSGQ-  115
O.ES.01.Read25_HIV_Group0    -D-IA-EVP-H------Q-P-N----N-Y.--I---TK-G---------R-P.A-ATSHPD-KD...PVP--SL-.IT.KW.-KQKRQE-QE-E--KKAGPSEPR.CKVSCHCCTRTSEQ-  114
O.FR.06.RBF206               -D----EMP--Q-----Q-P-N----C--Y.--L--TK-G---H----R-P.A-AASH-NNKD...PVPE-SP-.NT.IR.-KQKR*E-QE---T-AG-G*RLGHQDSCNCCTRT*GQ-  113
O.GA.10.10Gab1190            -D----EVP-H--------QSP-NA-F----SY.--YL--TR-G---------R-P.A-QPAASHPD--KD...PVP-SP-.IT.RE.-EQKRQ--QE-E--KKAGPG*QHRRKGSCHCCTRTSGQ-  114
O.GA.11.11Gab6352            -D----EMP--H--------N------Y.--TK-G---------R-P.A-AVGASRPDNK-...FVPE---A.IT.DR.-KKAGP-*QPRRKDSCHCCTRTSGQ-  116
O.SN.99.99SE_MP1300          -D----EMP-H------Q-P-NK------A--Y.--Y--AS-G--L------R-P.A-AARNPDN-D...IVPE---PP.ITNNR.KHKRQE-QE-E-K-GPS-RYPRRKDSCHCCTRTSGQ-  117
O.US.10.LTNP                 -D----EVP-H------Q-P-NT-----Y.--Y--TK-G---------R-P.A-TASHPD-KD...PVP--SL-.IT.KW.-KKAGPGGDHCSKGSWHCCTRTSEQ-  115
N.CM.15.S4858                ------I----N------R--------N---------------V----TK-G---------S---...-TPKS-KS--D...PIPE--L-.RQ.Q.--.-Q--Q--Q-AL-SK--A----------------  102
N.FR.11.N1_FR_2011           ------Y----N------N------------Y--VC--MK-G-----X----S---...-PKS-KN--D...IIPE--L-.-Q.-D.Q.--.-Q--Q--Q-AL-SK--CA---------------------  102
P.CM.06.U14788               -D----D-P--QQ------SSP----A--Y.-----TK-G---------RP.A.--DN-N-N--D...PV--*-FA.CT.GA.-QQE*HQ-QT----KQATA-RS-CWKDYCLSSGTT---  112
P.FR.09.RBF168               -DQ---D-P--QQ------N------A--Y.-----TK-G---------RP.A.--DN-N-N--D...PV--*-*LA.*T.GI.-QQE*HQ-QT----KQATA-RS-CWTDSCLSSGTT---  111
CPZ.TZ.06.TAN5               -N-I--QVA------AA-E-P-S--W------.----PL--TK-G----------R-G-...KFAA-NTSN-D...LVRQ--LP.KR.SN.-IQSSQE-SQ-E--A-ASSGGRPRQQDSSLSSGWTSETS  115
CPZ.US.85.US_Marilyn         -D----NI---Q-----N------------V----TK-G---------GR-R...TTAKS-ASN-D...P-R--F-.-Q.L.-NQ--Q--Q-AL-KQ-TS-QR---------------------  103
GOR.CM.12.SIVgor_BQID2       -D----EVP-H------Q-P-N--X--X-AY.-YL--TK-G---------RP.--AVAASH-GDKD...PVPE--LP.HT.PR.-NQERQ-KQEE---K--S-SGLHCHQDSCISYTWTSGQ-  116
GOR.CM.13.SIVgor_BPID15      -D-T--D-P--QQ------SSP-N--F--A--Y.--Y--LVK-G------------GRA.A.--PNN-N-N--D...PV--*SLA.*T.GA.-*QE*HQ-QT------*-TA-RS-CWTDYYQSSGTI----  109
```

Rev start   exon 1 end exon 2 start   NLS   Leu-rich effector domain   Rev end

```
B.FR.83.HXB2              MAGRSGDSD..EELIRTVRLIKL.........LYQSNPPPN.PEG.TRQARRNRRRRWRERQRQIHSISERILGTYLGRSAEPVPLQLPPLERLTLDCNEDCGTSGTQ.......GVGSPQILVESPTVLESGTKE......*      116

A1.CD.02.LA01AlPr         -------A...A-LTA--I--I...........Y-K.-R-.....Q--------A-N--D-L---V-S-C---PE--------S----R--S-------QSQGTET--R--VSG--SVI-GT-AEN-.....-      124
A1.CM.08.886_24           ----------G-LTA--I--I...........Y-K.-K-.SG---------A-Q--D--R-V-S-C---PE-S--F----I--C--H-------QSQRAET--R--VSG---DI-D--A-N......-      124
A1.ES.15.100_117          ----------D-L-AI-I--I...........Y-K.-R-.....Q--------A-Q--D--R-V-S-C---PE-S--F----I--C--H-------QSQRAET--R--VSG---DI-D--A-N......-      123
A1.KE.11.DEMA111KE002     ----------L-A-I--I...........Y-E..Q-.S---K------A-----D--Q--NSC----------------HI--C---------QSQGAET--R--VFG----I--RI-T......-      124
A1.PK.15.PK034            ----------A---D-L-A--I--I...........Y-E..Q.S---K------A-----D--Q--NSC----------------------------QSQGVET--RT-VS----V--GQ--NQ......-      124
A1.RW.11.DEMA111RW002     -------SG------L--AI-I--I...........Y-K.-R-.....Q--------A-----D--Q--S-H--G-E---HF------N-S-----------QSQGAAT--RS-VSG--D-I-D---EN.......-      124
A1.SE.13.078SE            ----------Q-LQA--I--I...........-R---Y-K.-R-.....-K------A-N--D--Q---S----F-----------S----R--GR----QSQGVET...........-AV.......-      117
A1.UG.11.DEMA110UG009     -------P--D-LKA--I--I...........Y-T.-R-.S---K-Q----A--H-S----S-C---------HI--S---P-----QSQGVET--GT-VS----A-N......-      124
A2.CM.01.01CM_1445MV      -------P--D-LKA--T--I...........Y-T.-R-.S---K-Q----A--H-S----S-C---------HI--S---P---QSQGVET--GH--S---SVI-G---E-......-      124

A3.SN.01.DDI579           ------N-----L-AI--I--I...........Y-K.-Q-.....-K------A----D-L---C-----E-------------H--S---------TGS*-S--CSAI-GE--KQCYLFA-      118
A4.CD.97.97CD_KCC2        ------S---A-LAAI-I--I...........Y-K.-R-.....-K------A-N--D----R-C----E----FC----------H--S-------QSQGTET--G--S---SG------EN......-      124

A6.BY.13.PV85             -------T---A-LT---I--I...........I----Y-T.-R-.....-K-Q----A--LN--------AC---XT-----I--N---S-G---D--QPQGTET--G--VS---SGI-G----......-      124
A6.IT.02.60000            -------A---L-A--I-I--I...........T--Y-E.-R-.S---K------A--------S----------R--S--------QPQGTET--G--SM--SSI-GP-A-......-      124
A6.RU.11.11RU6950         -------T---A-LTA--I--I...........-Y-T.-R-.....-K-Q----A----D-L---SSC------S-SF---I--H-S---------QSQGTET--G--SG-PSGI-G-......-      124
A6.UA.12.DEMA112UA024     -------T---A-LT---I--I...........-Y-T.-R-.....-K-Q----A---------S-C------------------N--S-------QSQGTET--G--SG--SG-GA-......-      124

B.AR.14.DEMB14AR003       ----R-------D-LKA--T--Y...........S.S----------------A--R-QE--NW--S-H---PE-----------------S---A----------G-Y-AI------      117
B.BR.10.10BR_MG035        ------------LN-----RI...........-T.........-Q-----QTVG-W--RAR---PV---QF-------------------S---DS-V-X--A--------ECH....-      120
B.CU.14.14CU005           ----------D-LK-----I...........S.S---------------W-RE--G---SNH---PE---E---------S-----------EL-GT---L----I------      117
B.DE.13.366396            ----------E-----L-A--R...........-T----S.T---------------RDL-GW--S---PE-----P---N------------DS-V-X--A-------      117
B.DE.13.947915            ----------D-LK-I---F...........S.T-----------R----QT-GAW--S-H---PT----FH---D----------------T---R--PS------E-ECY....-      120
B.DO.11.DEMB11DR001       ----R-------LKI--V-V...........-G............---------R---D-L-ASC------------------------------V-----A-------      117
B.ES.15.100600            ----------D-L-----T...........-E----S.............---------R---D-L-ASC----------------S----A---------V-----A----Q------      117
B.FR.11.DEMB11FR001       ----------E---LK-----R...........-E----S.............-K--H-RML-GW--DN---PE-----------------S-------A------G--V----A--------      117
B.JP.12.DEMB12JP001       ----N----D-LKA-----F...........-L-S.S.............-------------PE-----------------D---N-------N--------AI------ECH...      120
B.PH.15.DEMB15PH003       ----------LKA-----I...........S.S.............----RA----DSF-D-P----------T-S--I----------N-------T---------I-----ED.....-      117
B.RU.11.11RU21n           ----------D-LK--I...........................---H-R---IC----PE-----P--S---------------------------A--D-----      117
B.SE.15.102SE             ----------D--L--I...........-K............----A----Q--G---SNL--P-V-LQ---------------S------------V---S-I---IE.....-      117
B.TH.10.DEMB10TH002       ----------D-LK--K-----..........-I----L.S............---RN--DW--S---PE---F--------------------N--V--P-A------      117
B.US.16.2609              ------------LK----K...........-T-S.............----R-RT---W--D----PEG----------D-----------N--------I-----Q-----      117

C.BR.11.DEMC11BR035       ------EEL-A-L-V--I--I...........----------------A--------G--S-C---PE--------I------IN-S---S-----QPQGTTA---N-*V-GKPCA--G-R-KECY.....      128
C.CN.10.YNFL19            ----------A-LKA--I--I...........Y-E.-R-.....-K------A--A-K-R-L--------SSC---PT---F----I------HIGDS-SG-A--A--QSQGTTD--E-----*-SGR-CA--G---A-KECY....      126
C.DE.10.622166            ----------LQAI--I...........Y-K............----A--------A-CV--PE---F----I---HI--G-TG----QSQGTTE---N--VFRQ-CAI-G-A-KECY.....      126
C.ES.15.100_118           ----------AA-LEA--I--I...........Y-K............-----A-K-------S-C------L-F----I----IN-S--G------QSQGATE---N--VFRKPCA--GARA-KECY....      127
C.ET.08.ET104             ----------A-LQA-KI--I...........Y--.--..PP----G---.CQPKKT-LF-*-G-F.RS--GF--L--F--S----N-S--SK----K-----------N----*PF-VLG*GRK--CFIAL....      114
C.IN.15.NIRT008           ----------A-LQA--T--I...........Y-D.-T-.--S-K-----------A-------S-C---PE-------F----I------HISGS-SG-----A--QSQGTTE-----*-SGK-CA--G-A-KECY....      126
C.MW.09.703010256_CH256.w96 ----------A-LQA--I--I...........Y-K.-...--S-K-----------A-K----G----SS----P-------------I-----HIGDS-SG-------QLQGTTE-----*VSGKPCA--G-RA-KECY....      126
C.NG.10.10NG020523        ----------A-LQAI-I--A...........-ET--Y-E.-T-.....-K------A--E-S--D---SAC---PT--------I------HIGDS-SS--------Q-----*VSGK-CAI-GPRI-KECY....      119
C.NP.11.11NP016           ----------D--LQA--T--RI...........Y-K............-----A-----R--V-SAC---P------------I-NISDS-RG------Q......L----FGK-CA--G---A-.      126
C.PK.14.DEMC14PK009       ----E-----A-LQA-KI--A...........---T--Y-E.-R-.-S.......-QK-Q-L---Q--SAC---P------P--I--AIVDS-SS-------QSQGTTE----S*-SGK-CA---GTRD-KECY....      126
C.SE.15.100ET             ----------K-LQA--I--I...........Y-K............------A----------S-C---PT---F----I---I------HI--S-SG-------QSHPGTTE----S*-SGK-CA--GTRD-KECY....      126
C.TZ.08.707010457_CH457.w8 ----E-----A-LQA--I--I...........Y-K............-K------A---------S-C---------------I--HI--S-GG-----QPQGTTE----N--*VSGK-CA--G--R--KECY....      126
C.US.14.M0118v1d14_5M04_C5 ----E-----A-LQA--I--I...........Y-K............-K------A----S-V--G-T---F---I------HI--S-SS-----QSQGATE----S*-SEK-CA--G-----KECYWFA-      129
C.ZA.13.DEMC13ZA152       ----------T-LQA--I--I...........Y--NH---.....-A------A--------RE---S-C-D-PT----F-----I---HI--S-SS------QSQGTTE--N-*VSGRTCSI-GFRV-KECY....      127
C.ZM.11.DEMC11ZM006       ----------A-LQA--I--I...........-H-K.L------------------A----TL-----S-C---P-----F---I---HI--S-SS------QSQGTTE----SGK-Y-I-G---KECY....      127

D.BR.10.10BR_RJ095        ---------N-LTAI--I--I...........-S.L.............-AG-----GN---ES----------------SA-------Q------T-VS--P-AI----E-      117
D.CD.03.LA17MuBo          ---RE-----D-LTA----F...........S.-T.............-Q---------G---IS-----------------S-N-S-------Q----Q-----P----E-......      117
D.CM.10.DEMD10CM009       ---RE-----N-LK-IQ----I...........-I----S.T..........-------A-------G---------------I--N-N-S--EP------Q---T-S---A--------      117
D.CY.06.CY163             ---RE-----D-LKAI----V...........S.IQ.AP---TH----------------A------G---------------I--N-N-S------Q---T-S----A---T--E-      117
D.KE.11.DEMD11KE003       ----------LKA--F--I...........-S.-A.............-------A--LG--F-S----PE--------------------H------S-D-D------      117
D.SE.12.077UG             -------R----LKV-----I...........S.S.............----A---------LGQ--IC------PE--------------N-NS------------P-P-A--D-      117
D.TZ.04.CO6405V4          ----------D-LK-I----I...........-S.S.............-------A-------S-C---------------I--N-N-S------Q--P--P-A-D--S--...-      118
D.UG.10.DEMD10UG004       ----------DD-LKA----I...........S.............----------A--N--AG--ID----PE--D---------NIN-I--L-----Q--P-A-D-S-.....-      117
D.UG.11.DEMD11UG003       ----------LKA----I...........S.S.............----A---------RT-G---S----------F---------INS-N---A----Q--GH--S---P-D-......-      117
D.YE.02.02YE516           ---RE-----D-LKAI----F...........-N----S.............-------A--HL--G---S-L--G--------I--N-N-S---P------S--PS----P--ED.......      117

F1.AO.06.AO_06_ANG32      ---------N-----L---I...........Y-K............------A---VRE-----ASC---P-------------HIN-S--EQGAEE-----R-----S--RAI------      117
F1.AR.02.ARE933           -------------T-LTAL-Y--I...........Y-K............------A-----RA--D---SSC---P----------HIN-S---PGTEE-----L-G--AG--HA----I-N......      117
F1.BR.11.DEMF111BR037     -------------T-LKA-KY--I...........Y-K............------A-----RAL-----SSC---PE----------HIN-S--EQGTE----------SG--HA--V-------      117
F1.CY.08.CY222            -------------T-L-A--T--I...........Y-K............------A-----L-LSC---PE-------S----HIN-SK--EQGTEE-----N--S--H---D----      117
F1.ES.11.VA0053_nfl       -------------TD-LKA--I...........Y-K............------A-----ISC---PE----------------HIN-S-G-EQGTEE-----R-----TS----R------      117
F1.FR.04.LA22LeRe         -------------AD-LKAI-Y--I...........Y-K............------RE--Q--DSCV--PE---------------HIN-S---QGAEA-----S--S---NA--G------      117
F1.RO.03.LA20DuCl         -------------T--L-A--I...........Y-K............------A-----RA--D---SSC---P-----------HI--S--EQGAER-----S--H-----S--HA------      117
F1.RU.08.D88_845          -------------A--LKA-KY--I...........Y-K............------A-----RE--Q--SSC---P----------HIN-S---QGTEK-----S--P---S--I-G------      117
F2.CM.10.DEMF210CM007     ----------D-LKA--Y--I...........Y-K............------A-----VRE-----RAC---PE----------HIN-S--EQGTEG-----E-----S--PRAA------      117
F2.CM.11.DEURF11CM026     ----------D--KA--Y--I...........Y-K.L--T..........------A-----SSC---PT----------HIT------QGTEG-----E-----S--HA--G------      118

G.CD.03.LA23LiEd          ------NT---D-L-A--T--I...........Y-P............------A--------SAC--FE--------HI--S--S---QPQGTETR--GS--P-A-G-----      124
G.CM.08.789_10            ------ST---D-L-A--T--I...........-Y-P............-------K------A----N---------SAC---LE----------CI--S--S----QSQGTEP---K--TS----VI-G------      124
G.CM.10.DEMG10CM008       ------ST---A-LTA--I--I...........-H-S.A............----------A----N--------SSC---P-----V-----H--S--SNF--QSQGTEQ---------SV--G------      124
G.CN.08.GX_2084_08        ------ST---Q-LQAI-I--I...........Y-P............-----T-K------A----R------S-CV--LE----P--I--H---S--SR-PE--QSPGTET--G---S----V--G------      124
G.CN.16.224GX             ------ST---D-L-A--I--I...........-Y-P............------A----R---S-F--LE-----P--ID--H---S--NR-PE--QSPGTET--G---S----V--G------      124
G.ES.14.EUR_0033          ------ST---D-L-A--T--I...........-F-P............-------K------A----H-C--LE--------------H--S--P----QSQGTET--G---S----V--G------      124
G.GW.08.LA57LmNe          ------ST---D-L-A--I--I...........Y-K............------A----S-CV--LE---S-L-------H---SK-NR-P---QSQGTET---E----SA--G-----      124
G.KE.09.DEMG09KE001       ------ST---G--LQA-I--I...........Y-K............------A------SAC---P------P---HIA-S--S--GTET-------EL-G--S---G--GA-A------      117
G.NG.12.12NG060409        ------------LLAI-T--I...........-S-P............------A----------S-C---P-----I--N-H-S--S--P--QSQGTET---G--S---H-V--G-----      123

H.CD.04.LA19KoSa          -----------K-LEAIKI--I...........I----Y-E.ST.--K----------A----G--V-T-L---HQ-------------S----NK--------------S--SA--GA-A-.-      117
H.CF.02.LA25LeMi          -------N----D-LQVC-I-RI...........Y-E.-A-.--V--K---------A----E-CM--P-------------SK-------EK-----------VP---A--GT-V--ECH.....-      120

J.CD.03.LA26DiAn          ----------Q-LLA--I--I...........Y-K.-N-.S..........-A-N--D---V-S-C---P-----------I--R--S---D-------------DS--SGGTGML-GA-----      117
J.CM.04.04CMU11421        ----------Q-LLA-XX--I...........-R---Y-S.-R-.S..........-A-N-XD-------C-C--------------I--HI--S---D-------------SG--PCM--GA-----      117

K.CD.97.97ZR_EQTB11       ---R---E--QQ-LTP--I--I...........Y-K............------A-----RE--Q-V-SSC----T---------S-N-D--S-QGTEG-------EL----P--PD---G--D-.....-      117
K.CM.96.96CM_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------      117

01_AE.CM.11.1156_26       ------ST---LQA--T--I...........-Y-S.S.............-------K------A----------RDC-----------------------QSQGTET---R--SG--SVI-G-----      124
01_AE.CN.12.DE00112CN011  ------ST---KA--T--I...........-Y-S.S.............-------K------A-------RA------------------H--S--S-------QSQGTET---R--SG-PSVI-G---N.      124
01_AE.IR.10.10IR.THR48F   ------ST---D-L-AI-I--I...........-Y-P.A.............--T-K------A----------VRAL-D---SDC----T-------I--S--S-------QSQGTET-L-R---SG-PSGI--P--EN.......      124
01_AE.JP.11.DE00111JP003  ------ST---L-A-KI--I...........-S-P.............---T-K------A----------S-C---P-----------------H--S-----HSQGTET--R--SG-PSDI-GP--EN.......      124
01_AE.MM.14.fKSDU26       ------XT---LGA--I--I...........-E-X.A.............-T-K------A----------RA------S-C---PT----------------H--S--DX---QSQGTET---R--SG--SGX-GP---N......      124
01_AE.PH.15.DE00115PH012  ------ST---D-L-A--T--I...........-L.S.S.............-T-K------A----------RD-----SACV--*-T-----------I--H--S------QSQGTET---R--S-PSDI-GP--N.......      124
01_AE.SE.12.0705E         ------ST---D-L-A--I--I...........-L.S.S.............------A----------RE---V-SSC---PT------------I--H--S------QSQGTET---R--SG---SA---S.......      124
01_AE.TH.10.DE00110TH001  ------ST---I-QAAKI--I...........I----Y-S.S..N.........--------S-C----T-----F-----------H--S--------QSPGTET---R--SG--SV--GP--N.......      124
01_AE.TH.11.40436v09_04   ------ST---L-V--I--I...........-Y-S.SG.............-K------AQ--R--AL--S-C---T----------------S--G------QSQGTET---R--SG-GT--QN......      124
01_AE.TH.90.CM240         ------ST---L-A--I--I...........-F-S.S.............-T-------A-----SA-------S-C---T----V----N--S---G------QSQGTET---R--SG--SVI-GP--N......      124
02_AG.CM.10.DE00210CM013  -------T---L--EI--I...........-Y-K............------A-K--RT-----SAF--P-----HF------I--C-------QSQGTET---S---S---SVI-G------      124
02_AG.DE.09.701114        -------A---L-I--T--I...........-H-S--V..........-------K------A-------T-C--L-S-C-------------P--I--C---------SR---QSQGTET-----S--S-PSIL-G------      124
02_AG.KR.12.12MHI11_10746 -------A---L-A--I--I...........-Y-P............-------T-K------A----RA------S---G-------------I----S---A--R--QSQGTET---G---S-SG------      124
02_AG.KR.12.12MHR9        -------A---L-VI-I--I...........-Y-P............------A----------SA------S---P---------I----S--SK------RQSQGPETR--G--TP---SSI-G------      124
02_AG.LR.x.POC44951       ------NA---G-LGVI-IVRI...........-Y-K............------A----CA-I--V-S-C---PE---------------I---C---S---------QSQGTET---S---YII-G------      124
02_AG.NG.12.12NG060304    ----R-E---L-A--IV-I...........-Y-P............------A----H-RE---V-SAC---PT---------I--C-N-S------QSQGTET---S--PG-H-FI-G--D-......      124
02_AG.NG.x.IBNG           -------A---L-A--I--I...........-Y-P............-------T-K------A----RA------S-C---P---------------I--C--S------LSQGTET---S----YII-G------      124
02_AG.PK.15.PK032         -------T---L-A--I--I...........-Y-P.A............--T-K------A----RAL-D---S-C---PT----------C--S--------I-QSQGTET---S----II-G------      124
02_AG.SE.14.098GN         ---G--EA---L-A--T--I...........-Y-P............-------T-K------A----E-KQ---RA----T----------I--S--SK-------QS*GTKT------S--FLII-G----K......-      123
```

Rev start exon 1 end exon 2 start NLS Leu-rich effector domain Rev end

Rev

HIV-1/SIVcpz Proteins

```
B.FR.83.HXB2                MAGRSGDSD..EELIRTVRLIKL.........LYQSNPPPN.PEG.TRQARRNRRRWRERQRQIHSISERILGTYLGRSAEPVPLQLPPLERLTLDCNEDCGTSGTQ.......GVGSPQILVESPTVLESGTKE......*    116
02_AG.SN.13.9580            -------NT-.....L-A--I--I.........-Y-P.-----K------A---K-RQ----SAC-E-PT---HF-----K-CI--S----Y----EPVSRD*DL----F---SII-G---R-......    123
03_AB.RU.97.KAL153_2        --------.....LK-I----F...........--S.----------------H-----Q--S---PE---L-H---------S--------------------------D-------......    117
04_cpx.CY.94.94CY032_3      -------NI--.-D-FKAA-A--I.........-Y--N-T.--------A-K-----A----------P----------K--N-S----DK-------VS--L-A--GT-A-......    118
05_DF.BE.x.VI1310           -------R-.-D-LKA-----I...........-S.---------------A--------E--------------N-N-S--------------------S--P-AI-----......    117
06_cpx.AU.96.BFP90          -------.-DR-LLA--I...........-Y-K.-N-.------A----N--D----V-S-C------------I--R---S---N-------N--SG-PDML-GT--T-.    117
07_BC.CN.98.98CN009         ------E-----A-LKA--T--I.........-Y-E.-R-.----K------A------------SAC--------I--HISGS-SG------QSQGTTE-----*-SGK-CA--G--A-KEYY...    126
08_BC.CN.97.97CNGX_6F       --------.....-A-LQA--I...........-Y-E.-R-.----K------A---N-L------SAC--PT------HITDS-RG------QPQGTTE-----*-SGK-CA--G--A-KEYY...    126
09_cpx.GH.96.96GH2911       -------N-.....-A-LQA--V--I.......-Y-K.---------------A--S--N-------RAC--PE---------H-N-S----N-----------AD--SG--CAI-G-----.    117
10_CD.TZ.96.96TZ_BF061      ------N-.....-A-LQA--I...........-Y-K.---------------------GG---DA---PE---------N-N----------------H---GH--S----I-D-----.    117
11_cpx.CM.95.95CM_1816      -------N-.DQ-LTA--I--I.........I----Y-K.-Q--S--S-----A--D--Q---SSC--P----------N-----S-N--------------DX--PG--Y---PR-E-.    117
12_BF.AR.99.ARMA159         --------.....-LKAA...........---Y-K.----------------A--S--RA---ID---PE---F-----G------------------------------......    117
13_cpx.CM.96.96CM_1849      --------.-V-LT----I--I.........-N--Y-S.Q--------K-------A----GQ--S--------------H----------QSQGTET----G--P---GI-G----N...    124
14_BG.ES.05.X1870           -----ST-.-D-L-AX--I...........-Y-P.------------K------A-----SACV--E---------H-G-S--S--V---QSQGTET----GS--S---SVI-G-----.    124
15_01B.TH.99.99TH_MU2079    -----ST-......L-A--I--I...........-Y-S.A--S--T------A--KR-RE----SPC-----------------S-----------QSQGTET---R---SG--SV--GP--N-.    124
16_A2D.KR.97.97KR004        ------P-.-D-L-A--A-RI..........-Y-D.-R--S--------A----SAC--PT------S-----------QPQGTET---R---S---PSV--G-----.    124
17_BF.AR.99.ARMA038         --------.-D-LKAI...........-----------A---NK-G-----P------L------SK--------------E----V---PS-I------.    117
18_cpx.CU.99.CU76           --------.-Q-LT--SI--I.........-Y-E.-A-.-----------A----SA---N-C-GP---------S----F---QSQGTET----GS-VP---SG--GT-D--.    124
19_cpx.CU.99.CU7            --------.....-D-L-VI--I...........-Y-T.Q.RA-----K----A----------S--C-----F--------H-N-S-----------QSQGTET--R--G---SV--G--A-N.    124
20_BG.CU.99.CU103           -----ST-.....-D-L-A--I...........-Y-P.------------K------RAL----T-C--PP------------S--------------QSQGTET---R--VS----G--G---.    124
21_A2D.KE.99.KER2003        ------R-----DD-LKAI----F.........-----------------A---------A--VSYH----E-----I----N-N-R----A----.--Q-G--R---S---AI-G---E-.    117
22_01A1.CM.01.01CM_0001BBY  -----ST-.....L-A--I--I...........-Y-K.-R-.----K------A-K-D--L-----Q--ADC--PP------------S--I------QSQGTETR--R---S----DI-GA--K..    124
23_BG.CU.03.CB118           -----ST-.....L-A--I--I...........-Y-P.------------A------Q--ADC--PP------------S--I------QSQGTET----R---S----DI-GA--K..    124
24_BG.ES.08.X2456_2         -----ST-.....-D-L-A--I--I.........-Y-P.------------K------A----NSC--P------S-GTE----------------GS-VS--SV--G---.    113
25_cpx.CM.02.1918LE         ------VN-.....-LQAA--I...........-S-P.----------------TL---A--ASC--LE---------H-----S--------LSQGTET--R--VS-----A--GT--.    124
26_A5U.CD.02.02CD_MBTB047   ------P-.-A-LGAI--I...........-L-K.-R--S.----------A--N-K-D----V-S-C--GPP----------S-----S--F-N----QPQGTET---R---S--SGI--GA--E-.    124
27_cpx.FR.04.04CD_FR_KZS    ----NE----Q-LTAI--I...........-K---Y-K.-R-.----K------A---K-D--K-V-S-C---P--------P-----C--S---------QPQGAET---R---S--SA--GT-A-.    124
28_BF.BR.99.BREPM12609      --------.-D-K-I-I--F...........-S.--------------------L-RW--SNH--PE---F-----I----------------V------A--G---.    117
29_BF.BR.01.BREPM16704      --------.....-D-K----I...........-S.S.-----------------RW--SNH--PT-------F---------S------------------V-----A---.    117
31_BC.BR.04.04BR142         --------.....-A-LQA--I--I.........-Y-.-------------A------RA---S-C--P----FL--I--NIN-S-SG------HPQGTAE---N--*V-KKPCA--G--RA-KECY...    126
32_06A6.EE.01.EE0369        -----N-.-Q-LLA--I--I...........-N---Y-K.-N-.S.---------A--H-D---S-C--P------I--R---S-----E-----------D---PEKPGVL-GT---.    117
33_01B.ID.07.JKT189_C       -----ST-.....DD-L-A--I...........-X----Y-S.A.------K------A---GA----SAC---P--S------S--S----------QSQGTET---R---SG--SVI-D---N..    124
34_01B.TH.99.QUR1969P       -----ST-.....-LA-I--I...........-Y-S.S.------K------A--K--RE--D--ISC---T-------------SK--------------Q------G-----N.    117
35_AD.AF.07.169H            --------.....-LKGI--I--I.........-Y-E.-R--S.------K------A--N-D--Q---S--C--P--------Q--Q---------QSQGVET---R--VF----V--G----N.    125
36_cpx.CM.00.00CMNYU830     -------A-.....-L-V--T--I.........-Y-P.S--R.----K-Q------A------S-C---FE------------H--S--S---------QPQGTEA--Q--V----GI-GT--N..    125
37_cpx.CM.00.00CMNYU926     --------.-Q-L-A--NI...........-Y-P.----------K------RAL------S-C---------------H-N-S---RD--------QPQGTET--L-G--S--P-G--G----N.    117
38_BF1.UY.03.UY03_3389      ------G-.....-LKAI-Y--I.........-F-K.----------K------RAL------S-C------------------HIN-S----SAC---------N--PV---P-AI-D---.    117
39_BF.BR.04.04BRRJ179       --------.....-LK---Y--I...........-Y-L.-T-.------------RE----F-SAC--RE-----------------HIS--S---EQGAEE.........-TSG--LA--.    117
40_BF.BR.05.05BRRJ055       --------.-D-L-A--A--VLYQAVSSFKF...-S.S.-----------RK---W--ISNH--PT---F---------------S--F.........-Q-----N.    117
41_CD.TZ.05.C06577V5        --------.....-D-LKAI--V...........-Y-K.L.------K------A------S--PP--FL--HIGDS-GG------QSQGTTE---T--S----AI----E-ECY..    127
42_BF.LU.03.luBF_01_03      --------.-T-LQ-----R...........-S.------------------QXF-GW--N---PE-------H----------S--------------V-----A--T-.    117
43_02G.SA.03.J11223         --------A-.....-L-X--I--I.........-Y-S.-A.R------------RAL------SAC--------------H--S--SR-P---QSPGTET---G--P---V-------.    124
44_BF.CL.00.CH80            --------.VD-LKVI-Y--I...........-----------------RAL------SSC--LE---------H-N-S----QGTEG.........-SG----A--.    117
45_cpx.FR.04.04FR_AUK       ------NT-.....-LTA--I...........-Y-K.------------A------S-C--P----------------H-N-S----QGTEE.........-AD--TSG--CAI-G--.    117
46_BF.BR.07.07BR_FPS625     --------.-D-LKA--Y--I...........-Y-K.-R-.------Q-----A---VRG-----RSC--PE---------H-N-S----QGTEE.........-TSG--RV--G--.    117
47_BF.ES.08.P1942           --------.-X-LKA--I...........-S.S.------------------R--G----N--D-PT-----------------S-------------T-----P--.    117
48_01B.MY.07.07MYKT021      -----ST-.....-LTA--I...........-Y-K.------------A------R--S-C--P----------------------S-----------P---QSHGTET---R---SG--SDI-GA--N.    124
49_cpx.GM.03.N2667          ------N-.-Q-LLA--I--I...........-K---Y-K.-N-.S.---------A-D--D--Q--IDC--P---------S---C--S-S----N--------D--VSG--CPI-GE---.    117
50_A1D.GB.10.12792          ------S-.-D-LKAI----F.........-------------K-----------N--G--G---P--PE--G-----N-N-S---N------------Q-----AI-D-----.    117
51_01B.SG.11.11SG_HM021     --------.....-S.------------K------H-RVV-DW--S---T------------------S-K--------------Q-------V-----P-..    117
52_01B.MY.03.03MYKL018_1    -----ST-.....L-A--I--I...........-F-S.S.------K------A------C-CV--PT----P------S-------G---P-----A-.    117
53_01B.MY.11.11FIR164       ------T-.....-L-A--I--I...........-Y-S.X--T--T---T-K------XRA------SXCV--I-----T------N--S-G------QSQGET---R---SG-PSVI-G---N..    125
54_01B.MY.09.09MYSB023      ------NT-.-D-L-A--I--I...........-Y-P.-------------------VRE------RD---G--T-----I----N--S-G---------QSQGTET-L-G--SG--SG--G--X-N..    124
55_01B.CN.10.HNCS102056     ------ST-.....-IL-A--I--I.........-Y-P.------------K------A-----AACV--T---P---------HI--SK--E----G---H--S-----SG--SGI-G-V--EN..    120
56_cpx.FR.10.URF5_patient_A ------VT-.....-LQA--X--I.........-Y-P.ST.------K------A------RE------S-C---T--H-P---I--C------GG-----LSQGTETR-----TS----IL-G-----ECY..    127
57_BC.CN.09.09YNLX19sg      --------.-A-L-A--I--I...........-C-E.-R-.------K------A--N-D---A--S-C--P------F-----G-HISDS-SG-A---QSQGTTE-----*-SGKYCA-----E--KEYC..    126
58_01B.MY.09.09MYPR37       -----ST-.....-L-A--I--I...........-Y-P.------T---------A------RAL------S-C--P------S--------HQSQGTET---R---SG--SVI-GQ---N.    124
59_01B.CN.09.09LNA423       ------ST-.....-LA-A--I--I.........-Y-S.S.------K------A------TL---V--S-CV--PT----------------H--S--S----SG--SGI-G--P---N.    124
60_BC.IT.11.BAV499          --------.-A-LQAI--I--I.........-Y-R.------T*---K--*-A------S-C--P---------------I--NINHS-SS---R--QPQRTAEK--N-*VFRKPCA-------.    121
61_BC.CN.10.JL100010        --------.....-L-A--I--V...........-Y-E.-R-.------QK------A--N-R------SSC-----------I-G-HISDS-SG------QSQGTTE-----S-*-SGKYCA--G----A-KEHY..    126
62_BC.CN.10.YNFL13          --------.....-L-A--I--I...........-Y-K.ST.------------A------SAC--P----------------H--HISGS-SGE------QSQGTTE-----*-SGKHCA--G----A-KERY..    126
63_02A.RU.10.10RU0637       --------.-V-.....-L-A--T--I.......-Y-P.------------------A------SA----SSF--------------S--------------AA----QSQGTAT----S-SG--G-.    124
64_BC.CN.09.YNFL31          --------.....-L-V--T--I...........-Y-A.-----K------A-E--N-------R-C--PE----S-------HISDS-SG-P----QSQGTID--R--*-SGK-CA--G--A-ECY..    126
65_cpx.CN.10.YNFL01         --------.....-L-A--T--I...........-F-S.S.--T--T---T---------Q--D---RAC---PE-----I---YISD--S------QSQGAET---R--XG-PSV--GP---N..    125
67_01B.CN.11.ANHUI_HF115    ------NT-.....-D-LAA--I...........-X-S.A.------------RA------XPF---T------------H--------------QSQGTET---G--SG--SVI-G----EN..    124
68_01B.CN.11.ANHUI_WH73     ------NT-.....-D-LAA--A--I.........-Y-S.S.------------KE------A-X--T----------D--S--------------QSQGTET---G--SGK-SVI-G----EN..    124
69_01B.JP.05.05JPMYC113SP420 ------ST-D.A--L---A--I.........-Y-S.S--T--T-K------A------RE------A-S---T---------H--SD------------QSQGTET---G--SG--SVI-G--N..    126
70_BF1.BR.10.10BR_PE004     --------.....-LTA--I--F...........-L-S.-----------------RT--DW--SNH--PT---F---I----------S--------------P--.    117
71_BF1.BR.10.10BR_PE008     --------.....-L-A-...........-S.S.------R--------------KH-WA--GW--SNH-----E--H--Q-------S---------------S-I-.    117
72_BF1.BR.10.10BR_MG002     ------E-.-N-LK...........-S.S.------------------RN--GW--N--D-PE------------------------S-----------D--R--HT-.    117
73_BG.DE.01.9196_01         ------SA-.-D-LKA--I--I...........-Y-P.------------A------SACV--XE---------H--S--S--L-RQSQGTET---G--S--SVI-G---N..    124
74_01B.MY.10.10MYPR268      ------ST-.....-L-A--I--I...........-Y-S.A.----T-K------A------CE------C-G-T---------S---Q--------QSPGTET---R---SG--SVI-G--N..    124
77_cpx.MY.14.14MYNBB090     ------NN-.....-L-A--I--I...........-Y-S.A.------------RE--Q---SACV-QLE---------N-GDS--------QSQGTEA--R---P---SV--G--N..    124
78_cpx.CN.13.YNTC19         ------ST-.....-L-A--IV--I.........-Y-S.A.-------------SAL------SAC------------------S--S-----QSQGTET---R---SG---V--DP--N..    124
79_0107.CN.15.SX15DT013     ------ST-.....-LKA--A--I.........-Y-D.-R-.------------A--S--F----Q--------CIGDS-SG------QPQGTTE---R---SE-PSV--G----N..    124
82_cpx.MM.13.mSSDU12        --G--ST-.....-L-A--X--X...........-X---Y-E.-R-.------K------A--K-D---T---SAC-----P-------------K--------Q-----G-PS---D---.    117
83_cpx.MM.13.mSSDU94        --------.....-L-A--I--I...........-H----Y-E.-R-.------Q------A--NK-NTX--A---S-C------P----------S-------E-HQSQETEP----R---PG-SV----G----N..    117
85_BC.CN.14.14CN_SCYB2      ------E-.....-A--SA--I--I.........-Y-E.-L-R-T.--K------A------S--PP--------HIADS-SG------QSQGTTE-----*-SGK-CA--GAR--KECY..    127
86_BC.CN.13.15YNHS18        --------.....-Y-E.-R-.------K------A------S-C--P------------H-GDS--------------Q-L-R---SG--SV--G---.    117
87_cpx.CN.12.DH32           --GRE-P-.-K-LQ--K---I...........-Y-E.-R-.------K------A------RA------ISC--P------H-GDS.........-Q-----EI-G----N.    117
88_BC.CN.05.05YNRL25sg      ------S-.....-A-LQA--T--I.........-Y-E.-R-.------K------A--K--L---S---T-------------SK--------Q-----EI-G----N.    117
90_BF1.BR.11.BRG06043       --------.....-D-LA--Q--V-R.........-L-S.------K--GW--SA---GW--PE------------------S------------------D--V---V-A---.    117
O.CM.96.LA51YBF35           ------ED.....-QQ-LQAIQ---I...........-C-A.-A-.S-S------R--A-VDT-AS---A-VVHGPQDNNLVE----Q-S-RDP-VDQPPE-WTMDSRAEDN*-L---WSCD-I-DTRIAN....    122
O.CM.96.LA52YBF39           -----DAD.....-QQ-L-A-QI---I...........-C-S.-I-.S-N------R--T-VD--AA-V-A-VVHGPQDNNLVE----Q-SISDP-ADQ-P--WTVDPRAENN*-L*NW*SCN-I-ATRIA....    120
O.CM.99.99CMU4122           -----D-V.....P-LH-I--I...........-Q-S.-I-.S-N------R--A--DTLAA-F-ASVVHGPQDNNLVE---Q-SIRDPDGDQP---GTVDPRTEDN*CL-N-CSCS-I-ATRIA....    121
O.CM.x.pCMO2_3              -----D-V.....QP-LHAI--I...........-Y-S.-I-.S-N------R--A--DTLAA-I--A-VVHGPQNNNIVE---I-Q-SIRDPDGDQP--VSVDPRAEDN*CL-N-YSCS-I-ATRIA....    122
O.ES.01.Read25_HIV_GroupO   --------A.....-Q-LQVIQI--I...........-H-S.-S-.S-N------R--A-VNLAA--F-A-VVHGPQNNNIVE---Q-SIRDP-GDRL-E-WTVDPRTKDN.YL-N-CSCN-I-ATRIT-....    120
O.FR.06.RBF206             -----EED.....-QQ-LQAI-I-RI...........-H-T.-S-.S-N-K----R--A-VDALAT---A-VVHGPEDSN-VD--S--H-SIRDP-ADQL-E-WTVDSGTKDN*RL-TCWGCN-I-ATRIA....    122
O.GA.10.10Gab1190          -----D-D.....-N-L-AIQI--I...........-Y-S.-S-.S-N-K----R--A-VDN-AA-V-A-VVHGPQDNNIVD--S--H-SIRDP-ADQL-E--WTVDPGTEDN-SL-S-CSFN-I-A-RIEK....    121
O.GA.11.11Gab6352          -----DAD.....-QQ-VQA--I...........-Q-S.-I-.S-N------R--V--D-LAT---A-VVHGPQDNSCVE--S--H-SIGDS-GDQP-E-WTVDPRREDN*CL-N-YSCS-I-VTRIAK....    121
O.SN.99.99SE_MP1300         -----DGD.....QP-L-AIQI--I...........-H-SPTT.-STS------K--A--DTLAA--A-VVHGPQDNNLVE---Q-SIRDPDGDQP-E-WTVDPRTEDN*CL-N-HSCN-I-ATRVA....    123
O.US.10.LTNP               -----D-D.....R-LQAIQI--I...........-Y-S.-S-.S-N------R--A-VET-AA-V--VVHGPQNNNIVE---Q-SIRDS-GDQP-E-WTVDLGTEDN*CL-N-YSCN-I-ATRIA....    121
N.CM.15.S4858             ------V-.....-L-A--V-RI...........-Y-D.NK.-----------RA------IS-C--GPQ---D-P---ID-----PK--SE.Q.......ATTE*TQN---GNNSI-GKRV-R......    116
N.FR.11.N1_FR_2011         X-----VN-.....-L-A--I--RI...........-Y--.ST--S-----------RA------S-C--GPQ---D-P---TK--S--P-ESQQGTATTE*IQN---GNTSI-GK-V-R......    123
P.CM.06.U14788            -----DEDL.R---T-I-I--I...........-S-A..R-PS-NST------S------DQ-AG--IASH---PE-LGGAD--DISQ-HIEDQGPPDN-VDTPGPTTE..Q*L-AR-LCCVWL-D-RI-T.......X    121
P.FR.09.RBF168            -----DEDL.R---T-I-I--I...........-W-E..R-SS-NST--K-----S------DQ-AG--ASR---PE--GGAD--D-S--HIGDQGPPNNPVDTPGPTADQA*L-AR-LCCVWL-D-RI-TKCN...Q    126
CPZ.TZ.06.TAN5             ----EE-.....AN-LQ-I-V--I...........-D---Y-S.GA-.S-A----------R--A-VDALAS---HYR--GPQK-P--DI-D-SK-H--PV-HPASPE-GDNQPSRQPSNNT*KHSSRS*TSP-VEWQYLC....    122
CPZ.US.85.US_Marilyn       -------GDAD.-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......    124
GOR.CM.12.SIVgor_BQID2     -----EEDQ.-Q-LQAIKV--I...........-F-T..H---N--K------P--R--A*-D--AT--ASLIPGPQDNN-VE--S--Q-SIQP--TKPP--GTADPRTKD-*-L-NT*GYS-I-ATRIAKKXY....    124
GOR.CM.13.SIVgor_BPID15    -----EEEL..QD--T-I-I-RI...........-W-E..Q-PS-NST------D-----DQ-AG--ITNR----E-LGGVD--DISH-HIGDQRPPEHPVDTPETPTEQ-N*LLER-LGCVWL-GTR-TTECN....    126
```

HIV-1 Proteins

```
                                                                    phos phos
                    transmembrane domain                       cytoplasmic domain
                  Vpu start                           α-helix              α-helix         Vpu end
B.FR.83.HXB2              TQPIP.......IVAIVALVVAIIIAIVVWSIVIIEYRKI....LRQRKIDRLIDRLIERAE.DSGNESE.GEISAL.VEMGVEMGHHAPWD.VDDL...*         82
A1.CD.02.LA01AlPr        MT-LE.......-G--AG-I--L-L--I--T-G-----L....--K--R----LEKIR----.------D.-DTEE-.SKL.----DRDF-A.GNN-...-           82
A1.CM.08.886_24          MLSLGIWE....-W---G-I-VL-------T--V--CKRV....-K-K----KI-A-IR----.------D.-DTEE-.ARL.----NYNLG-.DN--...-           85
A1.ES.15.100_117         MS-LQ.......-W---G---A-L-----T--G--QCKRV....-K-------EKIR----.------D.-DTEE-.SKL.ID--DYDLGI.DNIV...-            82
A1.KE.11.DEMA11KE002     MT-LE.......-S---G----L--A--I-LG-IK-L....-N------I-E-IR----.------D.-DTEE-.SKL.ID--DYDLG.DNI-...-             82
A1.PK.15.PK034           LDSLQ.......-C---G-I--L----T---G-----L....-------E-IR-------.------.-NTEE-.STL.-D--NYNFGF.DN--...-              82
A1.RW.11.DEMA111RW002    M-LLV.......VW---IG-I--L-L--I-T--GL--K-L....------LE-IR-----.------D.-DTEE-.WAT.----NYDLGN.DNN-...-           82
A1.SE.13.078SE           ML-LE.......-C---G-I--L-L-----T--G---K-L....-K-K----LE-IR----.------D.-DTEE-.STL.I---NYNLGF.DN--...-          82
A1.UG.11.DEMA110UG009    MR-LE.......-W---G-I--L-L----T--G-----K-L....RK------E--IR--E-.------D.-DTEE-.AGL.IN--NYDLG-.DNN-...-        82
A2.CM.01.01CM_1445MV     ML-LA.......-LS--G-I--F-L----T--F-Q-KEL....KK-----W--K-IS----.------D.-DTEE-.STL.----NLDF--.ANN-...-          82
A3.SN.01.DDI579          MTSLE.......-W---IG-I--L-L----T--G---K-V....-K-KR-----E-IR----.------D.-DTEE-.STL.--V-DDI-LY.DNN-...-        82
A4.CD.97.97CD_KCC2       MN-L........-WS-I--I--F-L----G-I-AF---K-L....RK-----W--E-IR----.------D.-DTEE-.STL.M---NLDFL-.NNN-...-          81

A6.BY.13.PV85            MT-LV.......-C--A----VS----I--T--GL---L....--K-------E-IR----.------.-DEEE-.STL.M-G-NL-LL-.NNNV...-            82
A6.IT.02.60000           MS-LQ.......-C--I----VG-------T--G----RL....------N--E-IR----.------.-DAEE-.SKL.--V-DYDLL-.DNNV...-           82
A6.RU.11.11RU6950        MT-LE.......-Y--A----VF-------T--G----RL....--K-------E-IR----.------.-DEEE-.STL.M---NL-LL-.DNNV...-          82
A6.UA.12.DEMA112UA024    MS-LE.......-Y--I----VF-------T--G----RL....RK------E-IR----.------.-DTEE-.STL.MGV-ND-FV-.D-NV...-        82

B.AR.14.DEMB14AR003      M--L........--------IVA--------------....-----------IA------.------.-DQEE-.SAL.M-------...-                    82
B.BR.10.10BR_MG035       M--VV.......-L--G----A--------T--L-------....---S--L--I--------.------.-DQEE-.SAL.----N-----...-                82
B.CU.14.14CU005          M--LV.......--G----I-A---------------....-----E-IRD----------G-SGRII.QPL.------G----.----EC-                  85
B.DE.13.366396           MP-LV.......-A---G--TG-L--A---M-F-------....---------------IR----.------.-DQEE-.SAL.-D---L-----N----...-      82
B.DE.13.947915           M--LV.......-F--------TA-----------------....------L-IR------.------.-DQEE-.SAL.----D-------...-               82
B.DO.11.DEMB11DR001      M-SVQ.......-S--------GV------I--F---KRL....R--K---------IR------.------.-DEE--S-L-L-----N----...-            82
B.ES.15.100600           M-SLT.......-L-------L----T--F----------....---R-N----IR----.------.-DQDE-.SK.--DR--L------...-                82
B.FR.11.DEMB11FR001      M-SLT.......-L-------L----T--F----------....----------S---IR----.------.EDQ----.EAL.M---L---.I---...-          79
B.JP.12.DEMB12JP001      M--LQ.......-L-------L----T--G---------....--K------N-IS----.------.-DQEE-.SAL.M-------.N--M--...-              82
B.PH.15.DEMB15PH003      MT-LMTPLQ...-A-------TA----I----A--------....-----------IA------.------.-DQEE-.AAL.--R--D---N.I---...-        86
B.RU.11.11RU21n          M--LV.......-AT--F--VG---------L--------....--------A-I--------....-DQ.-EAL.M----L----.---M--.-                79
B.SE.15.102SE            MW-VQ.......-L--T------V-L----T--F----R-....VK-----IL--IR----.------D.-DHEE-.SVL.-----D---I.-N---...-         82
B.TH.10.DEMB10TH002      MSSLY.......-LT--F-LVT----------F----R-....------------I---IR----.------.-DQEE-.SAL.M-------GN.I----...-      82
B.US.16.2609             MFSLK.......VA--------A--------L-------....------------IR----.------D.-DQEE-.SGL.--R--L----.I-----...-        81

C.BR.11.DEMC11BR035      MVNLLEKID..YKIGVG--A--LV---I--T--Y-----L....V--R--W--K-IR----.------.-DTEE-.QT--.D---LRLL-.GN--...-          87
C.CN.10.YNFL19           MVDLD.......YK-TVG-FI--L-L--I--T-AY------L....-TK-----W--K-IG----.------.-DTEE-.ST--.D---RLRL-.-N--...-        83
C.DE.10.622166           MINLAAIVD..YRIGVA-FII-L-----I--T-AY------L....VK----NW--K-IR----.------.-DTEE-.ET--.D---LRLLA-.DN----...-     85
C.ES.15.100_118          MLKFIERID..YRLGVG---I-A-----I--V-AY------L....V--KR--W--E-IR----.------.-DTEE-.ET--.D---LRLLA-.AN----...-     87
C.ET.08.ET104            MTNFLEQVD..YRI-VG-FI--L-----T-AY-------L....--------W--K-IR----.------.-DTEE-.ST--.D---NLRLM-.AV----...-       87
C.IN.15.NIRT008          MVDLG.......YKLTVG-II--L-----T--Y---K-L....----------W--E-IR----.------.-DMED-.ST--.D---NLRLL-.-NE---...-      83
C.MW.09.703010256_CH256.w96 MLDLVARVD..YRIGVG--I--L------ILAY---KRW....-K-K---W--K-IR----.------.-DTEE-.S--.-D-E-LGLL--.NG---...-    86
C.NG.10.10NG020523       MLESVD.....YRIG-A--LI-L----I--T-AY-----W....V--T---W--K-IR----.------D.-DHEE-.ST--.D--QLRLL-.ANG---...-       84
C.NP.11.11NP016          MIDLD.......YRLGVG-FI--L--V----T--Y-----L....V--------W--K-IR----.------.-DTEE-.ST--.D---LRLL-.GN--...-        83
C.PK.14.DEMC14PK009      MLNLLERVD..YRIGVG--GI-L------T---YL--K-L....VK--------W--E-IR----.------.-D-EE-.ATE.-D-DNLRLLV.GN-...-      87
C.SE.15.100ET            MLDLLEKVD..YRIGV--F--L-----I--T-AY------L....V--KR--W--E-IR----.------.-DTEE-.ST--.D---NLRLL-.AN--...-      87
C.TZ.08.707010457_CH457.w8 MLDLLARVD..YRIGVA-LI-L------T--Y--K-L....----KR--W--K-IR----.------.-D-EEV.ETL.-D---NLRLL--.IN.......-      86
C.US.14.M0118v1d14_5M04_C5 MLQ-D.......YKLGVG----I-I--F----L-----....------R--W--K-IR----.------.-DTEE-.ST--.D---NLRLL-.AN--...-     83
C.ZA.13.DEMC13ZA152      MLNLLARVD..YRLGVG-FII-L-L--I--T--Y----L....VK--------W--K-IR--V-.------.-DVEE-.ST--.D---LRLL-.-N--...-       87
C.ZM.11.DEMC11ZM006      MISLIERVD..YRLGVG--II--L------T-AYL----L....----K---W--E-VR----.------.-D-EE-.ST--.D---LRLL-.GNE-...-        87

D.BR.10.10BR_RJ095       M--LV.......-LG---IL--L-----T--F--C--L....RK-----Y---IR-------.------D.-DTEE-.SAL.-D---L---N.-------           82
D.CD.03.LA17MuBo         MK-LE.......-A---------S-L-----T--Y----KR-....-K-------IQ----.------.-DREE-.STL.---------VVA---.-          82
D.CM.10.DEMD10CM009      M--LV.......-IS------TL------T--Y------....Q--K----W----IR--E-.------.-DKEE-.STL.-D----IVA----.-            83
D.CY.06.CY163            M-SLE.......-L--A-----L-L----T--Y------....K------W--IR----.------.-DKEE-.SQL.M-------NVA--I.-             83
D.KE.11.DEMD11KE003      M-SLQ.......-A---------L-L-----T--F-C-RL....R--#---W----IR----.------.-DTEE-.SAF.-G--Y-----.IN-I...-         81
D.SE.12.077UG            M--LV.......-L--------L-----T--F--CKRL....R----------IIKK-E-.N----.-DTEE-.STL.MD--------N.---I...-             81
D.TZ.04.CO6405V4         M--LE.......-I---------A-L-----T--G------....R------L---IR----.------.-DEEE-.SKL.M-G--L---N.I---...-          82
D.UG.10.DEMD10UG004      M--LV.......-L---------LLL------T--F--C--L....R--KR--W---IR----.------.-DEEE-.SKF.M-------N.I--M-...-        82
D.UG.11.DEMD11UG003      MNLL.......VIG--------L----A--F--W-RL....-K------K-N-IA------.------.-DREE-.SAL.-G--A---L----.IG-I...-         81
D.YE.02.02YE516          M-TLE.......-LS-----I-A----I--T--Y------....R------Q---IR----.------.-DEEE-.STL.M----A---NVA---.-           83

F1.AO.06.AO_06_ANG32     MSDLL.......AI---------L------T-AYL----V....VK-KR-N--YE-IR-------.------D.-DAEE-.AAL.G-V-PFI-GG.INN-...-       82
F1.AR.02.ARE933          MSYLL.......AIG-T--I--L-----T-AY---K-L....V--R-N--YK-IR------.------.-DAEE-.AAL.G---PFI-R.I-N-...-            82
F1.BR.11.DEMF111BR037    MISLL.......AIS-A--I--L--V----T--F---K-L....RK-----N--YE--R-----.------.-DAEE-.AAL.G---PFL-G-.INN-...-      82
F1.CY.08.CY222           MTNLL.......AI--A--IL-L-V----T-AY----K-V....I----N--YKKIR------.------.-DAEE-.AAL.G-V-PFI-G-.INN-...-        82
F1.ES.11.VA0053_nfl      ML..Y.......-IS----IL-L-V----T--Y-----L....------N--YE-IR------.------.-DAEE-.ADL.G-V-PFI-G-.INN-...-          80
F1.FR.04.LA22LeRe        MSYLL.......AI--T--I--L-V----T--Y-----K-L....V------N--YK-IR----.------.-DAEE-.AAL.G----PFI-GN.INN-...-      82
F1.RO.03.LA20DuCl        MSY-L.......AIS-------I--V----T--Y-----L....V------N--YR-I-----.------.-DAEE-.AAL.G----PFV-G-.-NN-...-        82
F1.RU.08.D88_845         MTE-I.......IG----I--L-V-----T--F------....--------K--R-------.------D.-DAEE-.AAL.G----PFV-G-.-NN-...-        82
F2.CM.10.DEMF210CM007    MSLSI.......--V-A-FI-VL-----I--T--Y-----K-L....----KR-NK-YE-IR-----.------.-DAEE-.AAL.G---PFI-GN.INN-...-    82
F2.CM.11.DEURF11CM026    MSLVL.......A-G-----I--L-----T--T--Y----K-L....V--KR-NW-FE--S-----.------.-DAEE-.AAL.G-V-PFV-GN.LNI-...-   82

G.CD.03.LA23LiEd         M--LE.......-A---G-I--FLA-----T--F-L--E-....RK----EK----IR----.------.-DTEE-.ATL.M---DFD--V.G-N-...-           82
G.CM.08.789_10           M-SLA.......--V-G-I--FLA-LI----YVL--N-....RK-KQ-EN-L--IR----.------.-DTEE-.AKL.----DFDA-V.G-NV...-             82
G.CM.10.DEMG10CM008      M-ALV.......-A---G-I--FLA--T--T-F-Q--E-....RK-N---K-L--IR--I.----.-DTEE-.AAL.-D-VEFD--V.G-NM...-              82
G.CN.08.GX_2084_08       M--LE.......-A---G-I--F-A-----T--Y------....R--KR--K-L--IR----.------.-DTEE-.ATL.-DVVEFDL-V.G-N-...-          82
G.CN.16.224GX            MKALE.......-A---G-I--F-A-----T--Y------....R--KR--K-L--IR----.------.-DTEE-.ATL.-DVVEFDL-V.G-N-...-          82
G.ES.14.EUR_0033         M-SLE.......-A---G-I--S-A-----I--F-Q-KE-....RK-K--Q-IL--IR----.------.-DTDE-.ATL.----DFD--V.G-N-...-           82
G.GW.08.LA57LmNe         M-SLE.......VA---G---G-A--I--I--F-Q-KE-....KK-K--Q-IL--IR----.------.-DTEE-.ATL.----DFD--V.G-N-...-           82
G.KE.09.DEMG09KE001      M--LE.......-AS--G-I--FLA-TI---T--F-Q--E-....RK--------Y---IR----.------.-DEEE-.ATL.-GR-AFD--V.G-N-...-       82
G.NG.12.12NG060409       MS-LE.......-A---G-I--V-A--A--T-L----K-....RK-KR-----LE-IQ----.------.-DTEE-.ATL.-D-VDFE--V.GNN-...-         82

H.CD.04.LA19KoSa         #MSLY.......ALGVG-----S-LV-I--T--Y-----L....VK-K-----I-E-IR------.------.-DAEE-.SKL.--P-DL-N-A.-           81
H.CF.02.LA25LeMi         MYIL........TIG--G---F--V-I--T--Y------L....VK--R----LE-IR-----.------D.-DTEE-.STL.--R-QLNLGY.-A--...-        81

J.CD.03.LA26DiAn         ML-LQ.......-I----I--SLL--G--TV-Y------L....------N--K---IR----.------D.-DTDE-.SKL.----P-DL-N-...-            82
J.CM.04.04CMU11421       MNSLQ.......-AS-------FFL-VXA--XVAY------....V---R--K-X--IR--E-.----D-D.-DTEE-.AKL.---P-DL-N-.------...-       82

K.CD.97.97ZR_EQTB11      MV-L........T-G-I---A-L-L--I--T-AYL----V....V--KR-NW-F--IR-------.------.-DTEE-.AAL.G-T--LILG-.INN-...-       81
K.CM.96.96CM_MP535       MVSL.......AIS-------L-L--I--T--Y-----L....VK-KR-NW---IR-------.------.-DAEE-.ADI.G-L--LILGN.I-N-...-         81

01_AE.CM.11.1156_26      MNALQ.......-W---G---GV-L-V---A-------K-....-K--R--K-VK-I------.-DTDE-.ATL.----DFDH-V.G-N-...-                 82
01_AE.CN.12.DE00112CN011 MT-LE.......-S---G-T--L-L-----TV-AL-I-I-....----------K-IR--E-.------.-DTDK-.AKL.----NFD--V.G-NM...-         82
01_AE.IR.10.10IR.THR48F  MT-LE.......-S--AG---L--L-----T--A------....---N-------K-IS---X.----.-DTEE-.AKL.----DFD--V.G-N-...-          82
01_AE.JP.11.DE00111JP003 MT-LQ.......-S--#.......................-RF....-K---------K-IR--E-.------.-DTDE-.AKL.----NFD---V.G-X-...-       61
01_AE.MM.14.fKSDU26      MX-LE.......-S-X-G-I--L-L-----T-A-VL-----....-----VK-IR--E-.------.-DTEE-.AKL.----DFD--V-G-X-...-             82
01_AE.PH.15.DE00115PH012 MT-LD.......-C--AG-IIEL-L----T--G--V--V-....-K--------K-IR-------.------.-DTDD-.AKL.----DFD--V.GN---...-   82
01_AE.SE.12.070SE        MT-LD.......-S---G-I--L-A-----T--G--FK-....----------VK-IS----.------.-DTDE-.AKL.----DFD--V.G-NV...-          82
01_AE.TH.10.DE00110TH001 MTHLE.......-S---G-I--L-L-----T--A--F----....-K--------K-IR----.------.-DTEE-.AKL.----DFD--V.G-N-...-      82
01_AE.TH.11.40436v09_04  MS-LV.......-S---G-I--L-L-----T--V--L----....-K-------K--IS----.------.-DTEE-.AKL.----DFD--V.G-NV...-       82
01_AE.TH.90.CM240        MT-LE.......-S---G-I--L-L-----*T--A--VK--....------VK-IR----.------.-DTEE-.AKL.----DFD--V.G-N-...-             81

02_AG.CM.10.DE00210CM013 M--LE.......-S---G----F-A-----T--Y------....KK-K------K-IR----.------D.-DEEE-.STL.M-I-YDNLL-.NNN-...-        82
02_AG.DE.09.701114       M-SLE.......-A---G----F-A-V---T--F----R-....-RK-K----W-L-KIR----.------.-DTEE-.STL.M---YDNIL-.N--V...-        82
02_AG.KR.12.12MHI11_10746 M--LE.......-A---G----F-A-----T--Y----A-....-RK-K------L--IR----.------.-DTEE-.STL.M---YDNIL-.NN-...-      82
02_AG.KR.12.12MHR9       M--LE.......-S---G-I--F-A-----T--F------....RKE---SK-L--IR----.------.-DTEE-.STL.M---DYNIL-.N-N-...-          82
02_AG.LR.x.POC44951      M-SLE.......-A---G----F-A-V---T--Y------....-RK-K--H--L--IR----.------.-DTEE-.STL.M-I-Y.DIL-.NNN-...-          81
02_AG.NG.12.12NG060304   M-SLQ.......-A---G----F-A-----T--Y------....-KK-K------K-IR----.------.-DTEE-.STL.L---PDDILG.I-N-...-         82
02_AG.NG.x.IBNG          M--LT.......-S---G----F-A-----T--Y------....-RK-K------L--IR----.------.-DTEE-.STL.M----YEYIL-.N-N-...-      82
02_AG.PK.15.PK032        MHSLE.......-A---G----F-A-----T--LV------....-RK-K------L--IR----.------D.-DTEE-.YTL.M-R-NDNVGN.D-I-DNL-   85
02_AG.SE.14.098GN        MT-LC.......-----G----F-A-----T--Y---KR-....-RK-D--N--L--IR------.------.-DTEE-.STL.M---DDNIL-.YNN-...-      82
```

Vpu

HIV-1/SIVcpz Proteins

```
                                                                                        phos phos
                                        transmembrane domain                   cytoplasmic domain
                                       Vpu start                              α-helix        α-helix           Vpu end
B.FR.83.HXB2                           TQPIP.......IVAIVALVVAIIIAIVVWSIVIIEYRKI....LRQRKIDRLIDRLIERAE.DSGNESE.GEISAL.VEMGVEMGHHAPWD.VDDL...*            82
02_AG.SN.13.9580                       M-SLT.......-A---G----F-A----*I---L----R....TK-SA-G--L-IR----.------D.-DTEE-.TTL.M----SN.LA.F-----.-          80
03_AB.RU.97.KAL153_2                   M-SLA.......-A-------VG------G----L-------....------IR----.------.-DQ....EAL.M----LV---A-----.-              79
04_cpx.CY.94.94CY032_3                 MLFWE.......-W--G---L--V----TL-F---K-L....R---R--S-YN-IR----.------D.-DAEE-.STL.-G--NFD-V.G-N----.-         82
05_DF.BE.x.VI1310                      MSDLL.......TI-V---I--L------I-Y--K-L....V-----N-YK-IR----.------.-DAEE-.AAL.G-V-PFI-G-.INN----.-            82
06_cpx.AU.96.BFP90                     M-ALE.......-A---G---FLA----T-F-Q-E-....RK-K--EK-L-IR----.------.-DTDE-.ATL.M--DFD-V.G-N----.-               82
07_BC.CN.98.98CN009                    M-TLT.......-------A-------T-L-------....---KR-------IR----.------D.-DQEE-.SAF.M----------.-                 82
08_BC.CN.97.97CNGX_6F                  MLDLE.......L-VG--I--L------T--Y---RL....VK-----W--K-IR----.------.-DTEE-.ST-.-D--RLRLL.-N----.-             81
09_cpx.GH.96.96GH2911                  M-SLQ.......-AS--GII--F-A--IA-T--F--V-....RK-KR------IRD----.------.-DTEE-.AKL.----P-DS-N.-----.-            82
10_CD.TZ.96.96TZ_BF061                 MLSLLARVY..YKLRVG--I--L------IR-Y-Q---L....VEPT----W--K-IR----.---R-HD-D.-DTEE-.ST-.-D--NLRLL-.-N----.-      87
11_cpx.CM.95.95CM_1816                 MISLQ.......-AS-I-----F--VV---T--GL----L.....-------IR----.------.-DTEE-.AKL.----P-DL-N.-----.-             82
12_BF.AR.99.ARMA159                    M-SLV.......TL--------A-L--A--L-------....--------IR----.------.-DAEE-.AAF.G-L-LLI-GN.I-N----.-              82
13_cpx.CM.96.96CM_1849                 MNAL-.......-IS----I--F-L-V---T-AY-----L....VH-K--EK-L--IR---.------.-DTDD-.AKL.---NFD--V.G-N----.-          82
14_BG.ES.05.X1870                      M-SLE.......-A---G-I--G-A--X--T-F-Q-KE-....RK-K--QGIL--IR----.------.-DTDE-.AVL.M---GFD--V.G-N----.-         81
15_01B.TH.99.99TH_MU2079               MS-LE.......-S---G-I----T-A--I-N--....-K------VK-IR--E.------.-DTDE-.AKL.----DLD--V.G-N----.-                82
16_A2D.KR.97.97KR004                   MNSLQ.......-LS--G-I--F-L-V---T--F----K-....KK-N---W--K-IS----.------.-DTEE-.STL.M---NLDFG-.AN----.-         81
17_BF.AR.99.ARMA038                    M-ALE.......-L-------T---------L----------.....------L.---------.-DAEEI.SAL.M---------.IN---.-                82
18_cpx.CU.99.CU76                      MXXFE.......--G-I--I--F-----VC--T--Y-----....-K--G-FE-IRD----.----Y-.-DTEE-.SKL.M---N-N-GY.-A-V-.-            82
19_cpx.CU.99.CU7                       M-ALQ.......-A---G-I--F-A----T-Y-----....RKE----EK-L--IR----.------D.-DTDE-.ATL.M-L-DFD--V.G-N----.-         82
20_BG.CU.99.Cu103                      M-SLE.......VA---G-I--G-A--I--A-F-L-KE-....RK-K----ILE-IR----.------.-DTDE-.AVL.-D--DFDH-V.G-N----.-         82
21_A2D.KE.99.KER2003                   M-ALQ.......-I--A----T-------T-F---R-.....K--R--C----IR----.------.-DREE-.SKL.----Y---N.I----.-              82
22_01A1.CM.01.01CM_0001BBY             M--LE.......-CT--G-I--L-L-----T--A--IK-V....R--K-----IE-IR----.------.-DTDD-.AKL.---NFD--V.GNN----.-         82
23_BG.CU.03.CB118                      MHSLE.......-A-V-G-I--G-A--I--A--F-L-KQ-....RK-K----ILE-IR----.------.-DTNE-.AAL.-D--DFDH-V.G-N----.-        82
24_BG.ES.08.X2456_2                    M-SLE.......-A---G-I----T-A--I-T--W-Q-KE-....RK--------ILE-IR----.------.-DTNE-.ATL.-D--DFDH-V.G-N----.-     82
25_cpx.CM.02.1918LE                    M-SLE.......-F---G-I--F-V-VI--T--YL-I----....RK-K---Q-LE-IRT----.------.-DTEE-.STL.M----DFDA-V.G-N----.-     82
26_A5U.CD.02.02CD_MBTB047              MT-LQ.......-Y---G-I--L-L-------LV--K-----....-K-------L--IRD----.------D.-DTEE-.SAL.----D-YG-.YNV---.-      82
27_cpx.FR.04.04CD_FR_KZS               MISLK.......TIS-A--I--FL-VV----T-Y-------....---KR----LE-IR----.------.-DTEE-.ARL.----P-DLGN.DN----.-        82
28_BF.BR.99.BREPM12609                 M-ALI.......-A-------VA-------T--F-------....----------IR----.------.-DAEE-.AAL.GD--PLVLGN.INN----.-          82
29_BF.BR.01.BREPM16704                 M-SLQ.......--I-------VA-------T--F-----....----R---I-N-IR----.------.-DHED-.LAL..V--D-----.-N----.-         80
31_BC.BR.04.04BR142                    MFTL-ERID..YR-G-G--LI-L-----I--T-Y------L....V--R--W-VK-IR----.------.-DTEE-.ET-.-D--NLRLL-.-P----.-         87
32_06A6.EE.01.EE0369                   M-ALE.......-S---G-II-SLA-----T--L-Q--E-....RK----EK-L--IR----.------.-DTDE-.-TL.M---NFD--V.G-N----.-        82
33_01B.ID.07.JKT189_C                  MS-LE.......-S--AG-II-L-L-----T--G--CQ--....--Q-------VK-IR----.------.-DTEE-.LKF.----NLD--V.G-N----.-       82
34_01B.TH.99.OUR1969P                  MS-LE.......-S---G-I--L-L-----T--A------....-E--------VK-IQ----.------.-DTEE-.AKL.----MFD--I.G-N----.-        82
35_AD.AF.07.169H                       MTSLY.......-W--IG----L-L-----T-Y-G-I-RV....------VK-IS--T.------.-DTDE-.AKL.-----NYDLG-.DNN---.-             82
36_cpx.CM.00.00CMNYU830                M-SLQ.......-A--AG----F-A--GG-T--L----K-....-I--K------LE-IK--K....------D.-NTEE-.STL.----DFDH-V.G-N-.-       82
37_cpx.CM.00.00CMNYU926                M--LE.......-YS--G-I--F-A-----T--L--K-....RK--------N-IR-----.------D.-DTEE-.STL.M-L-DFD--V.G-N----.-         82
38_BF1.UY.03.UY03_3389                 MSNLL.......VLG----I--L-------T-Y-----....V----R-N--YE-IR----.------.-DAEE-.AAL.G-V-QLVI-GN.INN---.-          82
39_BF.BR.04.04BRRJ179                  MSDLL.......AIG-T--I--L-----T-AYL----F....V----R-N--YE-IR----.------.-DAEE-.AAL.G-V-PFILG-.INN----.-          82
40_BF.BR.05.05BRRJ055                  M-SLQ.......-A---G---VA-------T--LL----L....--------E-IA----.------.-DQEE-.STL.M----------.-                  82
41_CD.TZ.05.CO6577V5                   M-SLQ.......-T-------VA-------T--F----R-....R------Y-L--IR----.------.-DEEE-.SKL.----------N.I----.-           82
42_BF.LU.03.luBF_01_03                 M-SLV.......-A-------A-V--------L--------....-K---------L-KIS----.------.-DHEE-.SAL.----------N----.-Q         82
43_02G.SA.03.J11223                    M-XLE.......-A---G-I--FLA-----T--F-------....R--KX----LX-IR----.------.-DTEE-.AIL.-DNVXXD-.V.G-N----.-         82
44_BF.CL.00.CH80                       MTDLL.......AIS-A--I--A--V----I--Y------L....V-----N--YK-IR----.------.-DTEE-.AKL.G-V-PFI-GN.I-N-.-            82
45_cpx.FR.04.04FR_AUK                  MNYL........--G-A--I--ALL---I--T--Y-----V....KK-ER-NW--YN-IR----.------.-DTEE-.AAL.G----LILG-.IAN---.-        81
46_BF.BR.07.07BR_FPS625                MSDLL.......AISLA--I--L-------T-AYL--K-L....V-----N--YE-IR----.------.-DAEE-.AAL.G----PFI-GN.INN----.-         82
47_BF.ES.08.P1942                      M-SLE.......-L-------A-L------L-------....-E--VYK----IR----.------.-DQEE-.SA-.--R--L-----.IN---.-             82
48_01B.MY.07.07MYKT021                 MT-LQ.......VS---G----T-------T--G--F-....-------K-IR--T.------.-DTDE-.AKL.----DFD--V.G-N----.-               82
49_cpx.GM.03.N26677                    M-SLQ.......-A-----I--FFLV-I--T-Y-A--QL....V----E------IR----.------D.-DMDD-.#KL.--R-SLDFRN.AN--.-            81
50_A1D.GB.10.12792                     M--LV.......-LS-A---IVS-------T-F--KRL....R----------IR----.------.-DREE-.SA-.L---D------M.-.-               81
51_01B.SG.11.11SG_HM021                MSSLD.......---------A-L-I----T--F----....--K-K------IR----.------.-DQEE-.AAL.--R--LV-G--I.--.-               82
52_01B.MY.03.03MYKL018_1               M--LE.......-S--AG--I--L-L-----T--A--FK--....----R----VK-IR----.------.-DTDE-.ATL.----DFD--V.G-N----.-       82
53_01B.MY.11.11FIR164                  MT-LQ.......-S---G-I--L-L-----T--VL------....----------VN-I----.------.-DTDE-.AKL.---NFD--V.G-X--.-           82
54_01B.MY.09.09MYSB023                 MT-LQ.......-S--AG-I--L-L-----T--F------....----------K------.------.-DTDD-.AKL.-D--DFDD-V.G-N----.-          82
55_01B.CN.10.HNCS102056                MT-LE.......-S---G-I--L-L-----T--A--K---....QKE---EK-LE-IR----.------.-DTDE-.AQL.-D--DFDH-V.G-NV.-            82
56_cpx.FR.10.URF5_patient_A            M--LE.......-A---G-I--F-A--I--T--Y------....RK-QR-----L-KIR----.------.-DKEE-.ATL.MD--VDFDS-.V.G-N----.-      82
57_BC.CN.09.09YNLX19sg                 MLN-N.......YRLGVG-----L-----T--Y-K---L....V-------C--K-KIR----.------.-D-EE-.ST-.-D-E-LRLL-.-NE---.-         83
58_01B.MY.09.09MYPR37                  MS--LE.......-S---G-I--L-L-----T--A---LL--....-RK------K-VE-IR----.------.-DTDE-.AKL.---NFD--V.GGN----.-      82
59_01B.CN.09.09LNA423                  MLSLET.......-GV-I-F-IVG-------T--YL------....--------K-IR----.------D.-DQEE-.SAL.M----R---N.-----.-          83
60_BC.IT.11.BAV499                     MLGSID.....YRLGVG--I-RL----I--I--Y-----L....V----R--W-VK-IK----....N------.-EDAEE-.DN-....V-YLRLL-.-N-V-.-    82
61_BC.CN.10.JL100010                   MLDLE.......SAL--VG-------T--V----T-Y-----L....VK--------K-IR----.------.-DTEE-.ST-.-D-NLRLL-.-N-W.-          83
62_BC.CN.10.YNFL13                     M*SLT.......--LS------A-------T--F-------....R-----N---IR--S.------.-DQED-.SAL.M----------N.IN--.-            81
63_02A.RU.10.10RU0637                  M-TLE.......-A---G----F-A-----T--Y-----....RK-K---K-L--IR----.------D.-DTEE-.STL.M-V-NY-LL-.DNNV.-            82
64_BC.CN.09.YNFL31                     MLDLN.......YTITVG-FI--L-----T--Y---RL....KK-----W--K-IR----.------D.-DTEE-.ST-.-DL-NLRLL-.ANN----.-          83
65_cpx.CN.10.YNFL01                    M--LV.......-S---G-I--L-L-----T--L-------....VK------LE-IQ----.------.-DTDE-.AKL.----DFD--V.G-N----.-         82
67_01B.CN.11.ANHUI_HF115               MT-LE.......-S---G-I--L-L--X--T--G--F-....----------K-IR----.------.-DTDX-.AKL.----DFDH-V.G-N----.-           82
68_01B.CN.11.ANHUI_WH73                MT-LE.......-S---G-I--L-L-----T--F-------....----------K-IR--E.------.-DTDE-.AKL.----DFDH-V.G-N----.-         82
69_01B.JP.05.05JPMYC113SP420           MT-LE.......-W--AG-I--L-L-----T--L-----....--------K-IR----.------.-DTEE-.AKL.----DFD--V.G-N----.-            82
70_BF1.BR.10.10BR_PE004                M-SLL.......-YS-------VA-------....----R-N------IR----.------.-DAEE-.AAL.GD--ELLI-G-.IGN----.-                82
71_BF1.BR.10.10BR_PE008                MP-LY.......-L-----I--G-------........------------IR----.------D.-DAEE-AAAI.G---PLVLE-.IIN---.-               83
72_BF1.BR.10.10BR_MG002                MT-LY.......-S---------G-L-----T--L-------....RKEKR-E---N-IA----.------.-DQEE-.SAL.--R--L----I-Q.-             82
73_BG.DE.01.9196_01                    MK-LE.......-S---G-I--S-A--I--I--F-Q-KE-....RK-K--Q-IL--IRD----.------.-DTEE-.STL.----RD-H--.-N-.-           82
74_01B.MY.10.10MYPR268                 MS-LQ.......-S---G-I--L-L-----T--A------....-K-K----K-VN-IR----.------.-DTDE-.AKL.----DFD--V.G-N----.-        82
77_cpx.MY.14.14MYNBB090                MT-LE.......-S---G-I--L-L-----T--A--LK-L....----------VK-IQ----.------.-DTEE-.AKL.----DFD--V.G-NV.-           82
78_cpx.CN.13.YNTC19                    MNSLE.......-S--AG-I--L-L-----T--A--A-....---Q-------VK-IR----.------.-DTDE-.AKL.----DFDH-V.G-NV.-            82
79_0107.CN.15.SX15DT013                MT-LE.......-S--AG-I-L-L-----T--A-QI-R-....----------VK-IR--E.------.-DTDE-.AKL.----DFD--V.G-NV.-             82
82_cpx.MM.13.mSSDU12                   MS--LE.......-G---G-I--L-L-----T--L---Q--....-K--------IR-IQ----.------.-DTDE-.AKL.----DFDH-V.G-NV.-          82
83_cpx.MM.13.mSDU94                    MT-LE.......-S---AG----L-----T--L---Q-....-K--------IR----.------.-DTDE-.LAF.M-----AG--.---.-                82
85_BC.CN.14.14CN_SCYB2                 M-TLI.......FS--------VA-------L-------L....-K--------I-IR----.------.-DQEE-.LAF.M-----AG--.---.-             81
86_BC.CN.13.15YNHS18                   M-SLYIS....LGL--------------T--L-------....-K-------E-IR----.------.-D-EE-.ST-.-D--NLRLL-.--NV---.-           83
87_cpx.CN.12.DH32                      M-ALT.......-----VA----I--T-L-------....-K------VK-IR----.------.-DTEE-.STL.-D--NLRLL-.-NV.-                 81
88_BC.CN.05.05YNRL25sg                 MLDVD.......YRLGVG--I--L-V---T--Y-----L....VK-----W--K-IR----.------.-DTEE-.ST-.-D--LRLL-.-N--.-             83
90_BF1.BR.11.BRG06043                  M-ALT.......-IS-------T-------T-------....----KI-X-IR----.------.-DQEE-.SGL.------------.I----.-              82
O.CM.96.LA51YBF35                      MHHWE.......LITLI-FS--CLLCVL--ML-FRL-LEHKKQDRKE-ELLKRLK----IRD...-DY--N.--EE....R-V.MDLV.-SHGF.DNPMFEL-       86
O.CM.96.LA52YBF39                      MPHKG.......LLL-II..ALLL-NVI--MFNLRK-LEQKKQER-E-E-IERLR-IR-VRD...-DY--N.--DE...Q-V.RDLI.-SHGF.DNPMFEL-        84
O.CM.99.99CMU4122                      M-HKD.......LLILIITSALLL-NVIL-LF-LKQCLEQKKQTK-E-E-I-RLR-IR-IED...-DY--N.--EE....QTV.RDLI.-SHGF.DNPMFEL-       86
O.CM.x.pCMO2_3                         M-HRD.......LITLIIISALLL-N-IL-MFILKQCIEQKKQSR-E-E-IERLR-IRKIQD...-DY--N.--EE....Q-V.RDLV.-SHGF.DNPMFEP-       86
O.ES.01.Read25_HIV_Group0              MHYRG.......LLTLI-IIALL--NVI--MFILRK-LEQKEQDR-E-E-LKRLR-IR-IRD...-DY--N.--EE....Q-V.RDLV.-SHGF.DNPMFEL-       86
O.FR.06.RBF206                         MHQRD.......VLVL-VFIALS-VSLLI-LFSLKY-LEKRNQER-EEE-LKRLR-I--IKD...-DYGTT.E-EE....Q-V.MDLM.-NHGF.DNPMFEL-       86
O.GA.10.10Gab1190                      MYLKDLLLYHKDLILVLIIIALLL-NVIL-MYILKK-LEQKEQEK-E-E-LERLR-IR-IRD...-DY--N.--EE....Q-V.MDLV.-SHGF.DNPMFEL-      93
O.GA.11.11Gab6352                      M-HKD.......LIIVIISSALLFLNVLI-TL-LKQCLDHKKQDR-E-E-LERVR--R-IKD...-DY--N.--EE....Q-V.MDLV.-SHGF.YNPMFEL-       86
O.SN.99.99SE_MP1300                    MHHRD.......LLTLIIISALLTNVIL-AFILRQ-LRQKKQDR-E-E-LERLR-IRQIED...-DY--D.-KEE...Q-V.RDLV.-GYGF.DNPMFEP-         86
O.US.10.LTNP                           M-HKD.......L.LIIVSALLL-NVI--MY-LRK-LEQKRQDR-E-E-LKRLR-IR-IRD...-DY--N.--EE....Q-V.RDLV.-SHGF.DNPMFDL-        84
N.CM.15.S4858                          LL..........PLGLT--G--L-L-V---VLLYK--K-L....KL-EQ-SYIT..-RD-E-....-----D.-DTEL-.-TL.LSCNKLEQG.-SWV.-          75
N.FR.11.N1_FR_2011                     L...........LLGFI-VG--F---V-I-VLLYR--K-L....KV-E--RHIRQ-IRD-....-----D.-DEEL-.-TL.LPPDKLDQG.-NWV.-           76
P.CM.06.U14788                         MH-RD.......EAVLIIAG-LLLCC--X-GK-LLLVL-E....RERD-FVQRLA-WR-GQ-..-E-Y--N.E-EE....EQL.R-L-.NLLGF.DHV-.-         79
P.FR.09.RBF168                         MH-RD.......EAVLIIAG-LL-CI----GK-LLFVL-E....RERD-FVQRLV-WR-GQ-..-E-Y--N.E-EE....EQL.R-L-.NLLGF.GNI-.-         79
CPZ.TZ.06.TAN5                         VGSLL.......TNV-GIFCILL-LIGGSLL----IK-EL....R-T-QQQ-VLE---R-LSI---I-ED.E-.........LNWNNFD-HN.YNPRDWI-         79
CPZ.US.85.US_Marilyn                   MLNWF.......EIGLI--GIEG-LVVII-GL-ARLW-Q-KIKENT-QE-QN-LE-IRI-E-.........-D.--EEETLAKLL.--LSSLEL.DNPRIV.-      85
GOR.CM.12.SIVgor_BQID2                 MHYRG.......QIV-II-CALLL-XVLI-MLILKL-LDNKRQDR-E-E-ISRLR-IRXIRD...-DY--N.E-EE....R-X.MDLX.-SYGF.YNPMFEL-       86
GOR.CM.13.SIVgor_BPID15                MHYRD.......EIVLLIAG-L--GGCLII-GYLLLYWL-E....RKRDIFVQRVA--RK-Q-..-E-Y--N.E-EE....EKL.R-L-.AQLGFAYYM-.-        80
```

HIV-1 Proteins

glycosylation NDT

V1

Env start   signal peptide end gp120 start   glycosylation NVT   CD4 binding   glycosylation NSS

B.FR.83.HXB2	MRVK.EKYQHL.WRWGWRWGTMLLGM.LMIC....SAT..EKLWVTVYYGVPVWKEATTTLFC̈ASDAKAYDTEVHNVWATHAC̈VPTDPNPQEVVL.VNVTENFNMWKNDMVEQMHEDIISLWDQSLKPC̈VKLT.PLC̈VSLKC̈TDLKNDTNT................NS	142
A1.CD.02.LA01AlPr	---M.GTQKSY.QNL.----IVIF-L.VI-----.N--..GQ---------A--D-V-----------K---------IK-.------E-----P-----T--G-------Q-----T-D-H-NITSD-ITNN........GNT.TN	146
A1.CM.08.886_24	---M.GTOMNW.QGL.----II---I---I------.-VA..GN---------RD-E----------M---------D--IN-.E---S------T---------T-N-SSIITNS...........	137
A1.ES.15.100_117	--A-.GIQRNF.YNL.LT---TI--L.IIL----.-A..-N---------D-E--------------IP.E------------S-----T-I-------T-D-SHNVTLKG-....	141
A1.KE.11.DEMA111KE002	---M.GTERNC.QNL.LT---I--I.IIF---.N-A..-N---------D-E--------E-K--------ID..K----S-----T--------G---------T-H-EKYIKNITNQNE........T-	144
A1.PK.15.PK034	--AR.GIQRNE.Q--.LT---II--L.I-----.NTA..-----.-N---------N-E----------K---------ID.K--K-D-----N-----T---I---------T-H--VN--NVTV........	139
A1.RW.11.DEMA111RW002	-K-M.GIQRNY.GQL.-----I---I.----.-.-N---------N-E----------------Y-.-E---E-------N-----T-----------Q-----T-N-NNITVNGN.......	139
A1.SE.13.078SE	-K-M.-TQKNC.Q-L.L----II--L.I---.---.-TA..-N---------RD-K----------K--A--I-----LD-.E---E------T---N----------T-N--NNVTSI-S....	139
A1.UG.11.DEMA110UG009	---R.GTQRNC.QGL.LT----II--I.-N-A..-N---------D-E--------E-K---------E--------ID.K-----S-----T---------T-S--NA-SGS........TG	139
A2.CM.01.01CM_1445MV	-K-M.GTQRNC.QHL.-----ILIF---I------.K--..-N---------RD-E----------R----------D-.K-----------M--------T-N-SEV-DN-....	136
A3.SN.01.DDI579	---M.GIQRNY.LPL.----MIS-C-.-I-----.K--..-N---------D-E----------K---------LD.-E----E------T----------Q-----T-N---IPKNFTNTE....	141
A4.CD.97.97CD_KCC2	-K-R.GIQRNS.QHL.----LIFWT.II-----.--A..---------D-E--------H-N--EK-Q-I--------L-MD-.D---E-------N---I-D--R--Q-R--------T-D-QPVNSTN-....	139
A6.BY.13.PV85	-KAR.-MKRNY.QHL.----G-IL-FWI.IIM----.K--..-N---------RD-E----------K----------D-K-II-.E---T-D---Q----QT---------T-N-SEPRPEPTSA....	139
A6.IT.02.60000	-KAR.GMKRNY.QNL.-----IFW-.IIM----.K--..-N---------RD-E----------------D--IM-.D---E-----Q----QT---------T-N-QEPNVTNSNDTS........YNNS.--V	148
A6.RU.11.11RU6950	-KAR.GMKRNY.QHL.------W-FW-.IIM----.K-A..-H---------D-E----------K---------D--IM-.E---D----N-----QT---------Q-----T-H-AGSNSTSSNNST........SYN.---	147
A6.UA.12.DEMA112UA024	-KAR.GMKRNY.QHL.-G--M---W-.MIM----.K-A..DM---------RD-K----------H----------IS.E---D--Q----QT---------E-----Q-----T-N--P-PNSTISNS.........S-	143
B.AR.14.DEMB14AR003	----.GIRKNC.Q-L.-K------------------.---.--N---------S-I----------S----K-.E---D-D----N-------K-------E-------T-D-I-VPTNCTNT.............GN	142
B.BR.10.BR_MG035	----.GIRKNC.Q--.-----AL------------.--.-.-N---------G-E---------------------E----------------T-E-IKVNTTGTNRT....	140
B.CU.14.14CU005	----#GIRKNY#SH-.----I-V----W...-K..Q---------T--GLPVWKNATTTLFC.---L--RDKC----------M.G--Y------------IF-----I-T-N-N----C	136
B.DE.13.366396	-K-.-GIRKNY.QHL.-T--IS---------.-A..GN---------S---------------D----IV.E---D-D----D------------T-N-S-GNYT....	143
B.DE.13.947915	----.GIRKNY.Q-L.----M---------N----.--.-GP---------------------A-----------D---------------T-N--VNTNNTNSN.........T-	143
B.DO.11.DEMB11DR001	---T.-MRRNC.QNF.-K--L------------.-TA..-D------------K----------------------S--IE.E---D---------A-------T-N--NWNATAAKNIT.........SD	144
B.ES.15.100600	---.-IKMRNC.Q--.-IG--L---I.-T---.-K..-Q--A-I----------K--------------P---I--------S--I-.E----D-----Q--N----------T-N---VG-S--DSH.......NST.--G	146
B.FR.11.DEMB11FR001	-K-.-KTR..-KHL.-----IL---I.-----.-.-A..-Q---------N-E---------------E----------D------N---------L-----T-N--G-N-A--I....	136
B.JP.12.DEMB12JP001	-K-.GTRKNC.QHL.-K--I------------.-----VEKDN---------A----------R-E-----I-E-N-------V----E--------T-N--AG-T-EAGNT....T-	146
B.PH.15.DEMB15PH003	----.GIRKNW.QHL.-GV----I.-----.--.---------D-K---I-----E--------V----------D--.--------N----------T-N-NATINST-P....	140
B.RU.11.11RU21n	--M.-IR...KHL.-----W---------------.-QN---------------SK-A---Y-----S--IP-.K------G--N--D-----Q----------T-N--NVASTN-S....	136
B.SE.15.102SE	-K-M.GTMRNC.QYL.-----L------------.--A..NN---------R--NV------------------IA.E---------P-----V-N--------T---------R----T-N---V-EGS-S....	139
B.TH.10.DEMB10TH002	-K--.-IRKNC.RHL.----L--I.----.-.--.-E---------V----------K----------IE.E---D--N-N--D------------V--T-H-CK-NDTKCHDLNVTNST...TPEG...-N	152
B.US.16.2609	---M.GIRKNY.QD-.--KG-IL--I.----S.--.-.-E---------G-E-A---------S--IKM.E---H------S-----------T-N-INTTSSGVN....	139
C.BR.11.DEMC11BR035	----.GIQRNC.RQ-.-I--ILGFW.-V--Y....NVM..GA---------K----------E----------S--I-.E---D-D----D--------Q----T-R-S-A-SNHSNINSTQ............GA	146
C.CN.10.YNFL19	---R.-ILRNY.QQ-.-I--VLGFW-------...-VG..GNR---------D-K----------E----------M-.E---E--------N---Q----------T-N--NVT-VNSNS.............TN	142
C.DE.10.622166	---R.GTLRNY.PQ-.-T--ILGFWL-------..IISG..NVV..GN---------K---------EK--I--------LL.E--Y------N---------Q----------T-N-SEV-GSNTSK....	140
C.ES.15.100_118	----.GIQRNW.KQ-.-T--ILGFWL.---Y.....NVR..GN---------RD-K----------E----------E---LK-.E---------D-----Q-----T-N-SNVNVNGTG....	139
C.ET.08.ET104	----.GIQRNW.Q--.-T--ILG-W-.-LTY....NVM..GN---------NPP---------G-E----------S---I-.E---------N---------Q---------T-N-SNTRISN....	139
C.IN.15.NIRT008	----.R.GIWKTY.QQ-.-I--ILGFW-------...-NVK..GN---------E----IK----------IH-.E------------N-------T-E--NVTSTE-G....	139
C.MW.09.703010256_CH256.w96	---R.GIQRNC.QQ-.-I-SILAFW.---V----.-VR..GN---------K---------K----------IK.E----------D------------T-N-SHNITVNG-MGNGT............RG	145
C.NG.10.10NG020523	---M.GIMRNC.LQ-.-I-VSLGFW-.---AY....NVI..GN---------K--------HER---I----------I--.G------------E-----D--------T-E-RKMNISNSNT....	142
C.NP.11.11NP016	----.-TLRNY.PQ-.-T--ILGFW.-V--Y....NGG..GN---------K----------K--------R---M--G--------T-Y----Q--V--------T-E-SNVSRNVSS....	139
C.PK.14.DEMC14PK009	-R.GILRNW.PQR.-I-IILGFWL.-V-----.NVG..G---------R--K----------G-ND--I-------------T-D-RNVSENV-ST....	144
C.SE.15.100ET	-K-M.GTQRNC.QQ-.-I--ILGFW--.-TY.....-VM..GN---------K---------------ML.E-I--------N---Q-V---M---------M-----T-N--NVNITS-MTRE............MG	144
C.TZ.08.707010457_CH457.w8	-K-M.-ILKRW.KH-.-I--ILGFW.-L-S....NVM..GN---------R--K----------ER---------M--.E---------D--N-----D--F------M--.--R-ENATYNDTN....	139
C.US.14.M0118v1d14_5M04_C5	----.R.GILRSC.QQ-.-I--ILGFW-.-Y.....NVG..GN---------K----------K--------S--F-.K-.D----D--N-----F-------M--.--T-N--SVNVTNSND.............TN	142
C.ZA.13.DEMC13ZA152	---M.GMWRNC.QQ-.-I-V-LGFW-.----....GNM---------K----------------MI-.A---------N---------D--------T-N--AYNSTN-VG.........I-	140
C.ZM.11.DEMC11ZM006	---R.GILRNY.PQ-.-I--ILGFW.-V-S....-VR..GN---------K----------IW-.E---D--N----N---------T-S--NFTTNSTII............TI	142
D.BR.10.BR_RJ095	----.GIQRNY.QHL.-T--IL-----....NGK..KD---------S----------E-A--I--------ME-.E---S--N----D------------T-N--....	142
D.CD.03.LA17MuBo	---R.GTGRNY.QHL.-K------LL--TY....-TA..-DN---------KAI--------E-A--I----------IK-.E---------N-------T-N-S-V-VNNTSTNAATNS.........S-	141
D.CM.10.DEMD10CM009	--AR.GIKRNY.QHL.-T--I----LL--TY....-A.DN---------K----------TKA-A--I----------K-.E-------D-------T-N-ETHLNSTQS....	141
D.CY.06.CY163	---R.GIKRNY.HNL.-K--------L--TY....-A..DN---------VTKA-A--I----------I--GE------D-------T-D-I-VETNNK-T....	142
D.KE.11.DEMD11KE003	--A-.GMKRNY.QHL.-G--I----I.-TY....-VA..-Q---------S--KE-A--I--------------I---------P-------T-N--EW-DNSTNSNST....	145
D.SE.12.077UG	-K--.GIQRNY.QHL.-T--I-----W-TY....-VA..-Q---------R--N------S--EPQA---I----------L---ME-.L--------N--H--K-------S.--P-H--ICEGN--TEV...NTTTEGNSAT-G	154
D.TZ.04.CO6405V4	---R.-MKRNC.QNL.-KG-IL---I.-----.-.-A..-Q---------R--N------S--E-A--I------------S-----------Q-SEM--G-RGASNS.............TG	145
D.UG.10.DEMD10UG004	---R.-MKKNY.QNS.----I---I.-T.....-V-..-GQ---------K-P-A--I-----------IY-.E---------N-----------T-N--KWTSN-TS....	139
D.UG.11.DEMD11UG003	---R.GIEKNY.RHL.-G---L----L-VTY....-VA..-Q---------K----------S-R---------IH-.E------N-----------T-N--EWEPNATNA.............-N	142
D.YE.02.02YE516	---R.GIKRNY.QHL.-K---L----L--TY....-A..D---------KA----------SVK-A--I--S----------IE-.K-------------T-S---I--V-VD....	140
F1.AO.06.AO_06_ANG32	---M.GMQRNW.QNL.GK--LLF--V.-I-----.N-A..-N---------G-ER---------------IL-.K--------Q-N-----T---------T-H-K-ANTSS....	137
F1.AR.02.ARE933	---R.GMQRNW.QHL.GK--PLF--I.-I------.N---..-N-L---------S-ER----------D-R---K------E---D---N---T-----------XT-X--AIRN-....	137
F1.BR.11.DEMF111BR037	---R.GMQRNW.QHL.GK--LFF--I.-I-----.-.-A..-DN---------H----------S-E-A--------S--LD.----D-S----N---------T-M--NANITNANITN.....ANITNANI.TQ	151
F1.CY.08.CY222	---R.VTQRNW.QHL.G---LLF--I.-I-----.-A..DN---------G-EK----------IH-.E-I-D------N---------T-N--AKNSTEQS....	138
F1.ES.11.VA0053_nfl	---R.-MQRNW.PI-.G---LLF--I.-I------.K--..-N---------E---EE----------N---2--I----K------D---------T-N-VNYNTTNA-S.............SD	141
F1.FR.04.LA22LeRe	---R.GMQRNW.QHL.GK--LLF--I.-I------.N-K..DN---------ER----------------D-.T--A--D--N-------T-N-SAT-----....	137
F1.RO.03.LA20DuCl	---R.-MQRNW.PHL.GK--LLF--I.-I------.N-A..DN---------G-EQ---------------IR-.K-----D--S-------T-N--FNGNST-QNSTYPN.STVSPTNSTSA	157
F1.RU.08.D88_845	---M.GMQRNW.QHL.GK--LLF--I.-I------.N-A..-N---------R---------E--A---------F-.E---D-D--N-S----------T-N--NATIPNGNNN............ST	143
F2.CM.10.DEMF210CM007	---R.GMQRNW.QHL.GK--LLF--I.-I------.N-A..-N---------V---E--A--Y------------L--E----------N---------T-S---ANDTSAVN....	145
F2.CM.11.DEURF11CM026	---R.GMQRNW.QHL.GK--LLF--I.-IF---.N.A..-N---------EK----A--Y------------MR.I------N-S-D----N-----------I------.T-N-NAQKVN-SICT.........NN..TD	145
G.CD.03.LA23LiEd	----.GIQRNW.QHL.-K---LI--L.VI-----.-.-S..DN---------E-N-P--------S-S------------IT.E-----------D----------T-N--E-....	132
G.CM.08.789_10	----.GIQRSW.QHL.-K---LM--L.VIM----.--S..-N---------ED-N-P--------S-S----------------II-.E---------N---------------T-N--NSN-SYSNSS............TQ	142
G.CM.10.DEMG10CM008	--R.GIQRNW.QHL.-K--LIF-L.VI-----.-.-S..-N---------D-E----------S-R---------D-LF-.E-------A--N-------------IT-H-NNWLNYTDANG....	142
G.CM.08.GX_2084_08	--R.GIQRNW.QHL.-I-WNLI--L.VI-----.-TS..-N---------E--D-P---------R--S-K-------IKM.D--P--A--N---------T-D-AVTCANA--....	141
G.CN.16.224GX	-K--.GIQRNW.QHL.-M-WSLIF-L.VI-----.-TS..-N---------ED-E-S---------S--K----------IE.E---P----------S------------T-D--NATYTNV-NN.........A-	143
G.ES.14.EUR_0033	-K--.GIQMNW.QPL.-K---ILI--L.GI-----.--S..RD---------ED-D-L---------S-K----------IP.E-----------N---------T-N--NVTVA--SSN.............XE	143
G.GW.08.LA57LmNe	-K--.GIQRNW.PPL.-K--ILI--L.VI-----.--S..NN---I-----ED-D----------S---------------IT-.E--------N-----------T-N--TVN-TRTNTN....	141
G.KE.09.DEMG09KE001	--A-.-MKRNW.Q-L.-EG-LLI--L.VI-----.--S..IN--A-----ED-D-P--------S-R------------LP-.T------------N----------IT-N--NVTSTGIENN.............TE	140
G.NG.12.12NG060409	---R.GIQRNW.QHL.-T-W-LS--L.VI-----.-.-A..NN---------ED-D-P---------S-R------------MP-.E---T--------D--------T-N--NYNHTN-ATK.........NNETV	145
H.CD.04.LA19KoSa	---R.-MQRNY.PHL.-K---L-IF---.-L.V--.N-E..-N---------E--K-----------S---IR-.D----E-------D----------T-N--KVS-N-HA.............-V	141
H.CF.02.LA25LeMi	T--M.-TQRNC.P-L.--G-SLI----.-L-Y....NVV..GN---------K-------R--EP-M--------M-.E-------------T-N-S-VIRNNTSE.............IN	142
J.CD.03.LA26DiAn	---M.-TQMNC.QNL.-----L-IF------------.--..-N---------D-N---------S-K-I--------IE-.K---E------Q------V--E---------M--.T-T-SNNVTIKG-SND........ASN..TE	147
J.CM.04.04CMU11421	-T-M.TQRSW.QHL.-K--L-IF---------.K-A..D-M---------S-S---------S-S------------MII.A---E----------N--------M--.T-N-SHNISTNS-NSR.........TTP..A-	147
K.CD.97.97ZR_EQTB11	--AR.-IQRNW.QHL.GKR-ILF--I.-I------.--A..NN---------E----------------------E----------------N-----T--------T-T-NVT-NRTNANKN.............DT	145
K.CM.96.96CM_MP535	--R.GIQRNW.QTL.GN--ILF--I.-I------.NA..D---------N-E---------------EM.E-------------N---------T-N--Y-GTNS-N....	140
01_AE.CM.11.1156_26	-K--.GTQMNW.LHL.-K---LII--L.VI------.--N---------QES---------THM.E----------N-----Q--V--E------T-N-NATSTNSS....	139
01_AE.CN.12.DE00112CN011	----.-TQINW.PNL.-K---LI--L.VI------.--S..NN---------RD-----------HTE------------MT-.E--------N--D-Q--V--E--------T-N--EA-PNA-LTE.........YNK..TN	146
01_AE.IR.10.10IR.THR48F	-K--.-TQRNW.LN-.-K---LI--L.VT------.--S..DN---------RD-E----------QE-A----------IY-.E-------Y----P-A--Q--V--------T-N--KANLTKLANA.........SNI..TN	146
01_AE.JP.11.DE00111JP003	-K--.-TQRNW.PNL.-K---LI--L.VI------.--S..DN---------H---------RD-V--------------IH-.E---Y----------D--------TMN--NVNATNVNAT.........NAT..IN	147
01_AE.MM.14.fKSDU26	----.-TXMSW.PNL.-K---LI--L.GIX-----.--S..D---------D-X----------HE---------MF-.E---D----------D--Q--V--------I--------T-T-SNAHLNNTNTTN.........VNT..TN	147
01_AE.PH.15.DE00115PH012	----.-TQMIW.PNL.-K---LI--L.V------.--S..-N---------RD-D-----------GHE----------IH-.E--Y----I-E-K-----Q--V--------T-N-SANFTKTNNSTG.............I	145
01_AE.SE.12.070SE	----.GTQMSW.PN-.-K---LI--L.GIM-----.N-S..DN---------HT-----------R-IH-.E---Y-----K-------Q--V--------T-N--RANLMF--TNN.............A-	144
01_AE.TH.10.DE00110TH001	----.-R.-TQMNW.PNL.-K---LI--L.VI------.--S..DN---------RD-K----------MH-.K--Y----------S--------T-H---F-....	143
01_AE.TH.11.40436v09_04	-K--.GTQMNW.PNL.-K---LI--L.VI------.--S..DN---------RD-K----RGHE------------IY-.E----D----N--S-----------T-N--SATVTNY-KV....	133
01_AE.TH.90.CM240	----.-TQMNW.PNL.-K---LI--L.VI------.--S..DN---------RD-K----------MH-.-E-----------N------------T-N---NANLTNGGSK.............TN	143
02_AG.CM.10.DE00210CM013	-K-M.GMKKNY.PLL.---L-MIIFWI.II-----.K--..-N---------E-----------------THM.E----------------------Q---------T-N-S-IIFRSSSINNS.....TSSTPTTTI-	152
02_AG.DE.09.701114	---M.GIQRNC.PLL.---V-IIFWI.M-M---.-.-N---..KQM---------EP-A--I-------H-TH-.E---M----------D-------T-N-S-PT-S-S-NSS........SDSP..SN	147
02_AG.KR.12.12MHI11_10746	---M.-IQRNC.LLL.----MIIFWI.I------.---.-N---------E---------------IH-.E----S----N--D-----------T-D-QTFNGSS....	136
02_AG.KR.12.12MHR9_	---M.GIQRNY.PHF.----IIFWI.MI------.--N---------D-E----------------IQ-.E---D--------N---------T-D-HNFNSSSNSSTFI....	141
02_AG.LR.x.POC44951	---M.-IQRNC.PLL.-L---IFWI.MI---.---.N.-.-NNM---------RD-E----------N--T---I--------IH-.E----------N--D--------T-D-HNINSS....	134
02_AG.NG.12.12NG060304	---M.GIQRNY.PHF.----L-IFWV.-I-----.--N---------D-E---------------MR-.A----S------N----------T-N-R-APINQTG....	138
02_AG.NG.x.IBNG	---M.GIQRNY.PLL.----IIFWI.MI-----.N---..-D---------E------------------IH-.E---Y----------N--------T-D-HNYN-SYSN....	138
02_AG.PK.15.PK032	-K-M.GIQKNY.TPL.--G-MIWW---.MIFW..IICNA..-Q---------QD-K----------------------D------------Q----------T-E-HGNITINS-NSA........NSS..SE	149
02_AG.SE.14.098GN	---R.GIQKSC.PLL.-----MIIFWI.II-----.K--..-N---I-----RD-E----------A--------S---IH-.E----------D-----------G-T-E-W-KN-TWDK....	138

Env

HIV-1/SIVcpz Proteins

glycosylation NDT

V1

Env start   signal peptide end gp120 start   glycosylation NVT   CD4 binding   glycosylation NSS

```
B.FR.83.HXB2               MRVK.EKYQHL.WRWGWRWGTMLLGM.LMIC....SAT..EKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVVL.VNVTENFNMWKNDMVEQMHEDIISLWDQSLKPCVKLT.PLCVSLKCTDLKNDTNT................NS 142
02_AG.SN.13.9580           ---M.GTQRN*.PLL.----IVI.WL.-------N-....------------------------K--------------------IE-.K---Y----N----Q-------G----P----T-N---ANQWN...................... 134
03_AB.RU.97.KAL153_2       -----IR...KHL.----LF.-------------N-....----------------------SK-------Y------S--IP-.K------G-N------------------T-N---KEVTS.................TN 138
04_cpx.CY.94.94CY032_3     ---M.GMQRNW.PHL.GK--LL-I.-VI--S..NN-......-D-F-----------R-K-C---G-K--------------A-.I----------------NEG---A---S---TFT-INATTTNS--.............. 139
05_DF.BE.x.VI310           ---R.GMQRNW.PHL.-K--LI-I.-VI----S..NN--.....-D-F---------R-E---E----EK--I------------A-.I--------------G-----------------T-N---F-ANSTANS.................TT 143
06_cpx.AU.96.BFP90         ---R.GIQTSW.QHL.-K--LI--.-LVI----..KNM------A-ED-D-I---------SPDK------------IS-.K---------N-D----------E--------T-T--NATLGNK-LGNNST......NS..TL 149
07_BC.CN.98.98CN009        ---T.GIKKNY.RHL.-----------V.GN--...........---------------------------MF-.E---------E-N---V----------------T-E-KNVSSNS-GT................... 140
08_BC.CN.97.97CNGX_6F      ---R.GTRRNY.QQ--I--VLGFW--..NVE..GN--......---------------S--I-M.E-----N---N---VT--------------T-E-RNVSSNG-GT................ 140
09_cpx.GH.96.96GH2911      ---M.GIQRNC.QNL.----I.------N------...........-------------Q--SK-K--------IG-.E---------N---G---GH-71-----T-D-HRVN-SKPNT................ 140
10_CD.TZ.96.96TZ_BF061     -T-T.GIQRNC.RQ--I--ILGFW--...N-R...........---------T----------KA-A--I---R-I-.E--------G--D---Q----GH-27--T-N-NGINATND-.................. 139
11_cpx.CM.95.95CM_1816     ---R.TQRSW.HNL.---L-IF--..N--....------------RD-D-P---N----S--IH-.E---------D--E--N---------Q------T-N--NV-KNCTNVEK......NNTTLGNS..-C 152
12_BF.AR.99.ARMA159        ---R.GMQRNW.QHL.GN--FLF--I.-I----N--...........--------S-ER----------E---D--N---G---T-N-SNANATA-NAT.......AN...GT 145
13_cpx.CM.96.96CM_1849     --AR.GIQMTW.QNL.----ALI--L.VI-S..-G-S.ND.....-RD-E----------------S------IN-.G---------E-N---D--Q-----------T-N--NNITS-SSI.................. 139
14_BG.ES.05.X1870          -KA-.GIQRNY.QSL.----ALI--L.VI-S..-G-S.ND......---------------------N-------E-N-D-Q------T-----FNGTNTNNT............DV-.K 145
15_01B.TH.99.99TH_MU2079   ----.TQMNW.PNL.-K--LI--L.VI----..-A..-N-....--D-S---------A-G-------IE-.I.G-----------E--------Q------T-N---R----I--................. 139
16_A2D.KR.97.97KR004       ---R.GMQRKS.QHL.------------I------S..K----.-D..........---I------G-E-----------MP-.E------D---N------T-V---------------T-N-N-VNVTATN.................. 139
17_BF.AR.99.ARMA038        ---R.GMQRNW.PSL.-K--ILI----..S..K---D......-RD-E-------------I-T-G-----------D----N------T-N-NVXXSG-NTX................ST 141
18_cpx.CU.99.CU76          ---R.GTRRSY.PPL.-N--LI--L.VIM----..-F..DH-...........---D-V--------S-S--S------------T-D--I................... 141
19_cpx.CU.99.CU7           ---R.GIQTNW.QCL.-T--LII--L.GI---..-S..-D-......-ED-N-P-------S---S------------I-P-.E---------N---------A-----------T-D-..-...................... 130
20_BG.CU.99.CU103          ----.GIQTNW.QCL.-T--LII--L.GI---..-S..-D-......-ED-N-P-------S---S------------I-P-.E---------N---------A-----------T-D-..-...................... 137
21_A2D.KE.99.KER2003       ---R.-IERNC.QNL.-K--I----I.----..-A..-DN------...........---------S-K-I---------ME-.G---------N---------------T-N-N-QVRNS..................... 137
22_01A1.CM.01.01CM_0001BBY ---R.GTQMTW.QNL.----I-I.---L.VI---..-S..-DN.....-RR-E------------P-----------S----I--------P--T-N---E-ARFNITN.................... 138
23_BG.CU.03.CB118          ---R.GIQKNW.Q-L.-T---LII--L.GI---..-S..-DN-......-ED-Y-P--------S-----IP-.X-------------H--------Q-----T-N--KVDCE-RN.................CT 141
24_BG.ES.08.X2456_2        --AR.GIQRNW.QLL.-S--LII--L.GI---..-S..-DN......-ED-DAP-------S-S---S------------IS-.Q------N----D---------T-N--NVNSTR-NHV................-- 143
25_cpx.CM.02.1918LE        -K-R.GIQKNY.PLL.----ILM--L.VI--..-S..-D.....-D-K------------K-A--I----------IDI.E---------E----------Q------T-N-SNININSTGN................TT 142
26_A5U.CD.02.02CD_MBTB047  ---M.GTRKNC.QHL.LK---LI--V.I-F----.K-A..-D.......------------------MN-.T---------E-N---Q-----------T-N---ITTNRTSR.................TN 140
27_cpx.FR.04.04CD_FR_KZS   -K-M.GIQRNW.QGL.----L-I---I.-------.VQ..GN.....-S--R--D-------------HE--A------------IL-.D-------R-N-A----------F--T-N---VNTTRIHD................... 140
28_BF.BR.99.BREPM12609     ---R.GMQRNW.KHL.GT--LLF--I.-I---..N--....-NQS.......---------N-VK--A--I----------I----K---------------E-I----------T-N---NNVTV--MN........ATN..TT 145
29_BF.BR.01.BREPM16704     ---A-.TMKTY.-HS.----M------.---------..-Q----.........---------------S--------V--A---------L-.E---------------T---IN-NATNLNDAS......TQAPT..TK 147
31_BC.BR.04.04BR142        ----.GTQMNW.KQ.--I--ILGFW--.FL-Y..NVR..GN--P-......-RD-K------------------I-E-.Q------E---D--Q--------T-H-S-ANRTH-G.................. 139
32_06A6.EE.01.EE0369       ---R.GTQMNW.SHL.-K--LI--L.VI---..-S..-KD.....-A-ED-D-I------HS--R--------------IH-.Q-K-I--------H--------------K-KNATFKNA-ING.............TS..EG 146
33_01B.ID.07.JKT189_C      ----.TQRNC.SNL.-K--LI--L.VI---..-S..-DN....-RD-E---------HE-------------MYM.E------K-AD-Q--V--------Q------T-N-AYANLTKSNNT.................VN 143
34_01B.TH.99.QUR1969P      ----.TQMNW.LNL.-K--LI--L.VI---..-S..-DN.....-RD-E---------HE--------------Y-----------D-------------T-N-NVTTNVNI...............TN 141
35_AD.AF.07.169H           ---M.GTQTNW.QNL.----I--I.VI-S..-N--....------------------------R-IH-.Q--E-----N------------T-T--NANVN--NDSN...............TN 144
36_cpx.CM.00.00CMNYU830    --AR.GTQRNC.LHL.-K--LII-L.VI---..-S..ND------.........-RD-N-----------------------E--------------T-E-QNFNGSN-H.................. 139
37_cpx.CM.00.00CMNYU926    ---M.GIQRNY.PPL.-N---LI--L.VI---..-N..NL------.........-D-D-----T--VP---------LP#---#--KINV.G--KD----------N----QI-----GPK----#--#---TF-#SSH-VNP#...........KT-G 138
38_BF1.UY.03.UY03_3389     ---R.GMQRNW.QHL.EKCSLLF--I.-I---..N--....-------------------E-------------------F--K--F-D--N--------------T-N-S-ASAPKT-.................. 138
39_BF.BR.04.04BRRJ179      ---R.GIRKNW.QPL.---I--I.-----..N--.....-A..DN.......---------------------S-H-EF-.E--------------T-N----AVRA--GT................. 140
40_BF.BR.05.BRRJ055        ---R.GIRKNW.QPL.----I-I.-----..N--....-A..DN.......------------------------------T-N--AAFNGTQ-STQGG..............DQ 145
41_CD.TZ.05.CO6577V5       --AR-.MKRNC.QN--K--I----T.----..-A..DN.......-------K------E-RS-KK-A--I------IE-.G---------E----------T-N--AIS-S-A-.................-N 141
42_BF.LU.03.luBF_01_03     -T-R.GIMKNY.QHL.-K--I----T.-------..-N--....-------XX--------X--------X--------T-N--NAAAHNINS..................TD 142
43_02G.SA.03.J11223        -----.IQRNW.QYL.-TMWXLI--L.XI---..-S..-N------.....-ED-X-P---N----S---K-----------------MT-.E---------E-------------M----T-T-NATCNKTN.................. 139
44_BF.CL.00.CH80           ---R.GMQRNW.QHL.GK--LLF--I.-I-L..--..-DN....--------------------ID-.P--------N---T-V-V-----------T-E--NATATNHSD................AT 139
45_cpx.FR.04.04FR_AUK      ---R.-TQRNW.QHL.----ILF--I.-L-----..-A..DN.......---D-Q------------------ID-.-----R----N-------P------Q--------T-N-N--IT-RNGS................. 136
46_BF.BR.07.07BR_FPS625    ---R.GMQRNW.QHL.GK--LLF--I.-I---..-R--......-----------------K-A-------------I-M.G---------D-P--------T-E--NAA-TND.................. 142
47_BF.ES.08.P1942          --AR.-IKKNY.QQ.--KG-IL---L.T-------.N-EKS--......-------------------A--------------ML-.E-------N---------T-N--NNTVTTNASM.................. 142
48_01B.MY.07.07MYKT021     ----.TQGNW.LN.--LI--L.VV---..-S..-DN......-RD-E-A------QVE--------------------S---Q-V-----------T-N-SEKITISGNGN...............ID 143
49_cpx.GM.03.N26677        ----.TWMI#.QN.--G-LI--L.-------..-GA..GN.....-------------------S-K-A---------------V-------------T-N-NV-TNNS-ANN..................TN 142
50_A1D.GB.10.12792         ----.IERNY.QQC.----M------T.-------..-S..-DN......-------------------S-KA-A--I------------ID-.E---------------T-N-G-SNIN-A-.................... 137
51_01B.SG.11.11SG_HM021    ----.GIRKNW.QHL.---G-ILF--I.----..N.A.-------..........-----R----------------LP-.E---Y-A--Q-V---------T-N--TA-FE--STT..............Y- 142
52_01B.MY.03.03MYKL018_1   ----.TQMNW.PHL.-K--LLF--L.VI---..-S..-DN.....-D-D-----------HE-A---------------IH-.E---------K-A--Q--V----------T-N--NATVADHNKNK................TI 144
53_01B.MY.11.11FIR164      ----.GTQMTW.PSL.-T--ILMI-L.VI---..-S..NN........-RD-D----------HE--M--I-------------IL-.E---------K-AD-Q--V-D-----T-H-ASANVTGKNIT..............TD 143
54_01B.MY.09.09MYSB023     ----.TQMNW.PNL.-K--LI--L.VA---..--L..DN.......-RD-D---------------E---------------S-R---M.E-----------N-------------T-N-KANFVKDNTT.........TNSPNNTTT 150
55_01B.CN.10.HNCS102056    -K-R.VTQMSW.LNL.-I--LIT-L.VIM--...-TS..DN........------------HV--M-----------------IY-.E-------D---N----------RM-.H--T--E-KNVTYCK-DT................TN 139
56_cpx.FR.10.URF5_patient_A ----.GIKRNW.PHL.-T-W-LI--L.VI---..--TS..TN.......-D-E-----------R--------------S--K-E--------------T-N-NNVTTSISN.................... 140
57_BC.CN.09.09YNLX19sg     ---R.GTLRNC.QQ--I--SILGFW--.-S..NVG..DN.......-R-K-----------E-------------------D---D---Q--V--------RM-.H---T-E-KNVTYCK-DT................TN 141
58_01B.MY.09.09MYPR37      ----.TQMNW.PNL.-K--LI--L.VA--...--L..DN.......-RD-D-------------E------------------S-----R----M.E---------------T-N--KANFVKDNTT.........TNSPNNTTT 150
59_01B.CN.09.09LNA423      --AM.GIKRNW.QHS.-K---V-----.----..-TA...-Q........-R-N-----------E-K---------------A-------------E----T-N--NYN-TK-V-...............MK 140
60_BC.IT.11.BAV499         ----.RMQRNW.IT--ILGFW---MY..-N.........-R-K--------E-------------E---D--Q-V--------T-N-SNVNRTE--NA.................TA 140
61_BC.CN.10.JL100010       ---R.GILKNY.PQ--I--ILGFW---G..NGG..KG.......-RD-K-------------------------S-----------N-Q-V----------T-N---V--SS-.................... 137
62_BC.CN.10.10RU0637       ----.M.GTRKTC.QHL.---------I-------..-A..-N--.........----------------------S-L.-------------E-----K---V------T---SS-N--N--V................. 139
63_02A.RU.10.YNFL13        ----.GIVKNY.QHL.----LI--FW-.IIM..K-A..-N-.........-D-K------I-E--H--M--T---------------R-----------D----T-D-YECRSNVSN.................. 139
64_BC.CN.09.YNFL31         ----.GILKNC.QQ--I--ILGFW---I.-.NVG..-DN......-D-K-----------------HK----------------R-----B----N---Q---V---T-N--Q-V-YNGTAGNY................T- 142
65_cpx.CN.10.YNFL01        ---R.GTQMNW.PNL.----LI--L.VI---..-TS..DN......-D-K------------R-IN-.E---------S--A--------O--V-----G---T-N-GNVTVRKYNESV...............V 144
67_01B.CN.11.ANHUI_HF115   ----.GTQMXW.LNL.-K--LII-L.VIT----.R-LD..NN------...........--RD-E-------------R-HE-------------IP-.E------S-A---Q--V------X-----T-H-SNAXLNXT-N.................XN 142
68_01B.CN.11.ANHUI_WH73    ----.GTXMNW.LNL.-K--ILII-L.VIT----..R-LD..NN------...........--RD-E------------R-HE--------------IP-.E--------S-A--Q--V------------T-H-SNA-LNVTANN.................. 142
69_01B.JP.05.05JPMYC113SP420 ----.GTQMNW.PDL.-K--LI--I.VIT----..-S..DN-----.........--RD-K--------QEA-A--I--------IP-.Q---------------N--------I--I--KNAILNG-VS.................DN 142
70_BF1.BR.10.10BR_PE004    ---R.GMQRNW.QH--GT-SFLF--I.-V--....NTA..GN-----.........--------X--S-----------XX-----------------J--Q-----------N----------T-H--NNXTKSGTNS.................T- 142
71_BF1.BR.10.10BR_PE008    ---M.GMQRNWQQQL.GK--LWF-RT.-L-----........----A--A---------S--K---------I----D------------T-N---WG-T-SN.................... 140
72_BF1.BR.10.10BR_MG002    -KAM.GIKRNY.QH.--KG-IL---I.-S...-D.......------------------------ME-.K------------A-----------T-N-SNHG-A-.................... 137
73_BG.DE.01.9196_01        -KA-.GIQRNC.Q-L.-K--VT-I-----..-N-V..-Q......-------------T---------------S-L.---E----------------------T-D--N-NETESGNS...............TT 143
74_01B.MY.10.10MYPR268     ----.TQMNW.PN.-K--LI--L.VI---..-S..NN-......-RD-D-------------HE--------------IP-.K---------N---D--------------T-N--NVT-VF-ATI................FNT..TK 146
77_cpx.MY.14.14MYNBB090    ---R.-TQMNW.PNL.-K--LI--L.VIM--...-S.D-.......-D-V-----------HI------Y----P-----I-P-.K---I-E-R--D--Q--V---------T-N--KI-VA-A-...............TF 141
78_cpx.CN.13.YNTC19        ----.TQMNW.PNL.-K--LI--L.VIM--...-K-S..DN.......-D-D------------HS-----------------IP-.K---------N------------Q--------T-N-SNKVEPH--TFN..............-T 144
79_0107.CN.15.SX15DT013    ----.TQMNW.LNL.-K--LI--L.VIM--...-S..DN.......-D-D------------HE--------------IF-.T------D---K--D--Q--V---------T-H--NVS-NGSNSS.........NS..S- 146
82_cpx.MM.13.mSSDU12       -K---.GTRMNW.PN.-K--LII-L.GIM--...-S..DN.......-------RDVK---------EK------------MP-.M-G----------------------T-H--NVSSSSTDT................... 143
83_cpx.MM.13.mSSDU94       ----.GIKKSY.-HL.-----I-------.------..-N........-RD-N-----------HS--G----------------------------T-N-VENVTVK-ETM................. 140
85_BC.CN.14.14CN_SCYB2     ----.GIKKSY.-HL.-----I-------.------..VG..GN-----.........-R--KAP--------------------MF-.G---------N-----V--------T-E-NVTSNSS-V................TP 141
86_BC.CN.13.15YNHS18       ---R.GTLRNC.QQ--I--ILGFW--.IM---.NVG..GN-----.........-----D--------------------R---------R-D--K-V--------T-N--NANRT-DN................... 137
87_cpx.CN.12.DH32          ---R.GTRRSY.QH--I--ILGFW--.IMY...NVG..-D.......-------------ER-------------------MS-.G---V---R-D--K-V---------T-D-RNVTAAYHG................... 138
88_BC.CN.05.05YNRL25sg     ---R.GTRRNY.PQ--I--ILGFW--.NGG..-D.......-------D-V----------VE--K---------------I--.E-----------N-----------------T-D-KNVS-----SNS..............T- 143
90_BF1.BR.11.BRG06043      ----.GIRKSC.QGL.----I--I.-----I.-R-A..-D.......-----D-V-----VE--K------------------------------E--------T-N---W-HANL-SAT........QTTA..SQ 148
O.CM.96.LA51YBF35          -KAM.-KN...EKS.-I-C.IVMAL.IIP....LSC.NM-YA---S-----ED-APV------NMTS--K--I--SQ--------YL-.N--DS-I-E-Y------EM-.F--QMN--VTDI-KM-............ 136
O.CM.96.LA52YBF39          -K-M.-MN...RKS.-ILY.IVMAL.IIP....LSS.SQ-YA---A-----ED-SPI------NLTS-K--I--SQS-------T-H-YP-.H--DY--I-Y--D-----QM-.F--QMN--EEQ-V-............... 134
O.CM.99.99CMU4122          -AM.-MN...RQL.GILY.IVMAL.IIP....LSY.DQ-YA---A-----E-KP--------VNLTS-Q--I-SQ--------T-Y-YP-.T---D--IV-Y--D--------QM-.F--QMN--NVNDT-............ 134
O.CM.x.pCMO2_3             -K-M.-KN.RKL.GTLY.IVMAL.IIP....LSH.DQ-YA--FA-----ED-APV------VNLTS-Q--I-SQ--------T-Y-YP-.T---D--IV-Y--D---------QM-.F--QMN--NVNGTNANG........... 137
O.ES.01.Read25_HIV_GroupO  -KAM.-RKN...KKL.-ILY.IVMAL.TTP....LSS.KP-YA---A-----ED-IPV------NLTS-R--I--S--------T-Y-YP-.H--DK-DI--Y------QR----D--QM-.F--QMN----PTD-............ 132
O.FR.06.RBF206             -EPQ.RRKN...KKL.-ILC.ITMAL.IIP....LNS.K--YA---S-----ED-EPV------TNMTSQ-Q--I--S--------S-H-YF-.N--D--TI--Y-------Q--QMN--NYT-N-TP............. 136
O.GA.10.10Gab1190          -KAM.-KN...KKL.-I-Y.-VMAL.IIP....LSC.K-SYA---A-----ED-KAI------VNLTS--K--I--SQ--------S-F-YP-.N--DS--I-E-Y--D--Q-----M--FM-QMN--N-NETSKGN.......... 137
O.GA.11.11Gab6352          -KAM.-PN...KKL.-I-C.IVM-F.IIP....LSSN.QQ-YA---A-----ED-SPV------NLTSP-K-E--AQS-------T---YP-.K--DK-DI--Y----------TQ-S-F-QMN--NE-N-SKS.............. 138
O.SN.99.99SE_MP1300        -K-M.-KN...KKL.GILC.-VMAL.IIP....LSH.NQHYA---A-----ED-SPV------VNLTS-Q--I--S--------S-Y-YP-.K--DK-DI--Y--------QM-.F--QMN--NVNDE-S........... 136
O.US.10.LTNP               -KAM.-RKN...R-S.-ILC.IVMAL.IIP....LTS.KQ-YA-I-A-----ED-APV------NLTSQ-Q--AQ------S-N-YT-.H---DK-DI------Y------QM-.FM-QMN--N--S................ 130
N.CM.15.S4858              ----.M.GIQSYW.SLF.CLVISLSK........VVGS..NQH.......-RD-E-Y-------HS-----I--Q---------------L.---------N-A---Q-----------E----------TML-NNSNGNMSVEGTT.........QRTT 142
N.FR.11.N1_FR_2011         ----.M.GMKSYW.SLF.YLLXSLIK........VIGS..-QY.......-------D-E-I---X---HS--A--I---Q-----------TMN--SD-SS-V............ST 136
P.CM.06.U14788             -K-M.K-KK...NSL.GNLE.IC-AL.I-YFNAISYTS..GTRY.........-RD-KI-------A--SLTSX-Q--I--Q------R-I--RI-D------S--I-D-Y--T--Q-----V---TMX-S-CSTVDC-NNSS....R...VNNS..TD 151
P.FR.09.RBF168             -AM.K-KK...NSL.GNL-.IC-AL.VIYFNAISC-S..GIHY.........-RN-EV-------A--SLTSX-Q--I--Q------T-I-IPI-------S--I-------T--Q------I----TMN-SECHREASC........NNNS..S- 145
CPZ.TZ.06.TAN5            *KKM.RN.........LIGI-LT-II.TTLG....IGF..SQYYA--F----I--D-KPS-------DVTSRDE--I----N---L----Y--P--A-M-VD---EE-Y--QE-K--L---FQ--F------F----MS-NLTVPS-TVSPATKVSST.........TK 142
CPZ.US.85.US_Marilyn       -K-M.--KKR-WLSYCLLSSLIIP-L........SS--A--------RDVE----------H--P---D-E--N-A--Q-------TMT-LNPDSNSSA................ 134
GOR.CM.12.SIVgor_BQID2     -K-M.K-KK...EKS.-ILX.IVMAF.IIP....-SS.RP-YA-----------RD-Q-----A----MASS-M-----Q---------Q-I-LK--T--T--I-SG-------Q------VM--TMN-XRXTPN-T-HSLNMTGPAPRTSXPTTS-.... 155
GOR.CM.13.SIVgor_BPID15    ---M.R-KK...RSL.GN--.LS-ALLIICFNTGSYVM..GTRY.........-R--K------A--NLASK-Q--I--Q------T-IK-K-...IS-PY-I-E-YI-K--QKN----LN---------VI--TIN-SKF-PA-TNTTSA.......PINA..SL 150
```

HIV-1 Proteins

Env

| glycosylation NIS |
| glycosylation NCS |

CD4 binding   glycosylation NVS   CD4 binding

V1   V2   CD4 binding   glycosylation NGT   glycosylation NFT

glycosylation NSS   glycosylation NDT   glycosylation NTS   glycosylation NKT   glycosylation NGS

```
B.FR.83.HXB2              .....SSGRMIMEKGEIKNCSFNISTSIRGK.VQKEYAFFYKLDIIPIDNDT....................TSY..KLTSCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNNKTFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEE.EVVIRSVNFTDNAKT 283
A1.CD.02.LA01AlPr         .....-PNITNEMG................-........GQ-..R-IN----A--------T-----------------TE-R-----R---------K-------------------N-T--E-I------ 292
A1.CM.08.886_24           .....TVNETRQGMR---------TT-AL-D-.R-RV-SL---R---VE-G-SSSS--------NNH...........NT-..M-IN-----A--------T-----------T----D-------R----S----S---------------------Q.K-M-E-I------T-- 277
A1.ES.15.100_117          .....TNTSTPNM-E-VR------VT-EL-D-K--RK-VHSL--R---VE-NG-NST.......SS...........SQ-..R-IN----A-K----------------RD-E-----R----------------A-----------------G.-K--E-IS----N 282
A1.KE.11.DEMA111KE002     .....NNQSE-SGE--L-----VT-EL-D-K--V-SL--R--VK-NEGNSS...........DN..SQ...........NE-..R-IN---A--------T-----------------------------------G.-K--E-IS----N 289
A1.PK.15.PK034            .....NNNSNVQ-M----STT-G--D-.KR-A--S---LVSMN-NNNS.................NE-..R-IN---A---------------KD-N-T--L-K-------------------------------I--E-IN---- 279
A1.RW.11.DEMA111RW002     .....TKGT-IEGMT----AT-EL-D-.N--M-SL---V-NE-KNN.......SSKN.NSS...SLN..R-IN----------T-----RDPK---------S----K---------------K---I--I-N--- 286
A1.SE.13.078SE            .....TATVNNSAIE-------VT-E-WD-.R--V-SL--H--VQ--ENS..........-E-..I-IN---A-K----------------KD-GY-----S----S--------------IR--E-S----- 280
A1.UG.11.DEMA110UG009     .....N-SSTVEGE-------VT-EL-DR.KRRM--L---V--NETNN.........SY-..S-IN---AL--------------RDEE-------K--------S----------------VKG.-R-K-E-ISE-H-N 281
A2.CM.01.01CM_1445MV      .........TK-M----YLMP-ELKD-.T-EV-SL----VQ-NSTDNVT......QQN...NDT...QP-..R-IR--------T--------------KD-K-----S-R--S--------------M--K------N 278
A3.SN.01.DDI579           .......NVTKAYE-M-------T-EL-D-.K--I-SL--R--VV--E-GSNST.......ENS...-E-..R-IN------------------KE-K-------K--------------------M--E-L-N---- 285
A4.CD.97.97CD_KCC2        .....KVEVP--MT-----MT-ELSD-K--VRSL--RI-LVQ-G-N-ND......SS...NRS...LQ-..R-IH--------T-----------V-----------KDQE------K----------------G.GIK---A-ISY---N 283
A6.BY.13.PV85             .....SKNSSSTV-RD--SMT-EL-D-.TKNV-SL-D---V--GENRNS...........SNAN.NSRGQYEQ..R-IN----AM--------T---------------EP--------X-------------------R-I--E-I---- 289
A6.IT.02.60000            .....TQYNNSNG-E-M--------T-EL-D-.TK-V-SL----V-TGSNSSN.......SS...EE...R-IN-----------------------KD-N--------------------------------K-----M--E-I------ 293
A6.RU.11.11RU6950         .......-NSSDSWVEVMR-----VT-EL-D-.KK-V-SL--R--VSTGSNDS.........EQ-..R-IN----AV--------T----------KDTN-T-K-6----------------------K-.-M--E-I---G-I 289
A6.UA.12.DEMA112UA024     .....N-YSNDSWVKDM-----YY-T-EL-D-.KKNVHSL---VVSLGTSNNS.......SQY..RQ-..R-IH----T-K------T---------------KDPM-----K------------------K--.-M--Q-I---S-- 289
B.AR.14.DEMB14AR003       .....CTNWK-K-G--L-------T-YRTD-.M--GN-L--N---K---EGN............N-...R-L---------I-W------T-----L--------------------------------G.DII--S-- 284
B.BR.10.DEMB_MG035        .....GTYNTSHNWE-M-------VT-N-GDN.KKT-------L-V--EPNS...........R-..M-KN-----K------------------------------------------------E---------N 282
B.CU.14.14CU005           .......--G-V--LD-S----VS---F-T-H.-VIADEVMDTD--LLVSI-S......INN...----L----LM----RSERCK-------------------------------------E------T--- 281
B.DE.13.366396            ............-------------KTNPNLKD-.S-RA--S----V-LEENS............-T-..R-I----S--------------R-D---------D------------------------KK.-I--E----- 267
B.DE.13.947915            .....KTNSTASSPI-M-------T-DRLD-.RK---L-S--VVQ-EEGNN...........N-..R-I----K---------I--------------DQK---S---D----------------------I--E-I----- 285
B.DO.11.DEMB11DR001       .....RNSSLTI-D--V------T--D--D-.R---R-D---V-V--ETNDK.......NNS...S-..R-IN-----------------------------T-G-----S-------------------------I--HL-N 289
B.ES.15.100600            .....NY-NWER--P----K-----T-NMQD-.R--V--L-----VV----NDN............S-..R-I------------------------------K-I---------------------A--------I 272
B.FR.11.DEMB11FR001       .......T..SSNGMT-------T--L-D-.W--V--L---LVK--D-NN...........R-I-----------------------------D-K-----K--------------------------A-------I 272
B.JP.12.DEMB12JP001       .....TNSSK-E---------VT-T-DN-.AR--S-L--R--VV---D-NNS........DS...STS...A--..R-IN---------T--------------L-------S-K---------------H--------K--L--K---N-E- 294
B.PH.15.DEMB15PH003       .....ETE-LRTL-E-VR-----VTP--KN-.M-----LL---LVA-NKTDN.........R-I-----TL-------I------------D-K-S-K-S--------------------D.-I--E-I------ 294
B.RU.11.11RU21n           .........S..I-GS-M-----T-----T--MKKG.-K---L-N--VVQ---NDT..........S--..R-I----V---L------------I------------I---------------------------I--E-I----- 273
B.SE.15.102SE             .....TENSTMTGEAVMT-----VTAG-KD-.M-------GG--ST-N...........R-K--P---N-G--KT----------RD----------K--------K---------------------YQ-IS--- 280
B.TH.10.DEMB10TH002       .....TTNITNNGGDDLR-----MT-IMTE-.M--N-L--TS--V---TDNT.........SD-..R-I----IL-------------I-------------------K------------------DI--K-S--- 295
B.US.16.2609              .......GTI-D-DM------MT-DMKD-.--R-----N----V-M-SSGS............TNE...-N-..R-I----A--------I------------------------------------------DII---E-L-N--- 280
C.BR.11.DEMC11BR035       .........K-HEINITQ-----T--MT-E-D-.KR-VH-L--R---V--LEERKNT......SN..NSS...RD-..R-VN-----A------------D--------Y-------D----------------G.-II--K-LK--T-N 294
C.CN.10.YNFL19            .NSSGTNISNSTDSGM-------AT-EL-D-.RK-----L--R---V--EGENSS......EG..KNF...-E-..R-IN----A-------N-D------T--Y------------------------G.-II--E-L-N--- 293
C.DE.10.622166            .......-NNTHDAKDSGMQ---------T-D-.TR-----L---KQMGVNS..........SN-..R-IN--S-T-A-------D---------------------I-.R-.II--E-LN-S--- 281
C.ES.15.100_118           .....GNATVE----ML------TTSEL-D-.KE----L-----VHLGK-N..........-N-..-G-KN---S----------D----------------K-----------------II--KLS---- 279
C.ET.08.ET104             .............---------TT-DL-D-.K-SA-L----V-LG-NSN.......YSN..GD-..R-LN--T-----------D---------------N-----------------------D.-I--E-I---- 274
C.IN.15.NIRT008           ..........TRVEDMR--T--AT-E-DR.K--TH-L--R---V-L-GEG............NK-..R-IN----A--------T-----T--Y---------------------------II--E-L-N--- 274
C.MW.09.703010256_CH256.w96  ....NDSTIGKMGD-MT--AT-E-KD-.KKL---L----VVTLEENS............SE-..R-IN----V--------H-----------NKI-------------K.-IIV--E-IQ--V-- 286
C.NG.10.10NG020523        .....NDATYNETIEG-R-----MT-EL-D-.KK-V--L-----V-LKENSNS......SE..NTS...EE-..R-IN-----------H----------------------------------I--G.-IM--K-L---I 290
C.NP.11.11NP016           .....YNTYNGSVE--------TT--MT-E-D-.R--M--G--V-LNKKNSS..........NN-..R-IN----T-D------------Y-----------S-----------------K-------G.-II--E-L-N-V- 284
C.PK.14.DEMC14PK009       .......-STSNDM-------TT-ELKD-.KKSV--L--R---VLEDSANS......SH..XNS...NH-..R-IN----AV--------T-D--------YV-----------------------K.G-I--E-L-N-V-- 282
C.SE.15.100ET             .....NATNLNTM-E-M------TT-E--D-.KKQ---L--E-.-E-LHGKN........SNS...SD-..I-IN----A--------T-D------------------S-----------------K--L-N--N- 288
C.TZ.08.707010457_CH457.w8   ......YKNITD--VM-----T-E--D-.KR--S-L--RV--V-L--NS..........AE-..R-IN-K--SL----I-D----------------------------------DII--E-L-----I 277
C.US.14.M0118v1d14_5M04_C5   ..I-NGTNSYLE-MR------TT-EVKD-.KR-A--L-----V-LNEKIAIN.......RT..NDN...RT-..R-IN--T-S----T-----------Y--------------------------G.-II--E-I--V-- 291
C.ZA.13.DEMC13ZA152       .....N-SNVELTERD-------MT-EL-D-.SKQ-H-L--RD--V-LNENS..........NNS...NE-..R-IN--S--K--T------------Y----------N------------------K------M---I-- 284
C.ZM.11.DEMC11ZM006       .....NNANVTNSTD---R----MT-EL-D-.SKQ-H-L--RD--T-LGGNDS...........-N-..R-IN----A-K----------------N-----------------------------G.DII--E-M--I-N 284
D.BR.10.DEMB_RJ095        .........E-VR-------T-V-ED-.KE--V-L--R----P--VQM-...........-T-..R-IN--I-A--------------RD-K---------------------------M--E-L-N--I-- 280
D.CD.03.LA17MuBo          NSTAAPMTTVASSSDM-------T-AV-DE.QK-VHSL---VNMEGKNN........TGD...GT-..R-IN---TT------------------D-Q------K--------------------KIM--E-L-S-S-- 298
D.CM.10.DEMD10CM009       .......VNTTDTDMR-------T-MLQD-.K--V-L--R--M-LVE-GKNS..........-K-..R-IN--A--------T--------------D-Q-----K--S---------------D.-II--E-I--V-N 278
D.CY.06.CY163             .........VDNSSI-DL-----VT-IL-N-.KKQVH-L--R--VVK--SQNNR.......AN...QSN...GT-..R-IN---------------------DPR-E-K-T-K------------------M---D.-II--E-L-N-I-N 287
D.KE.11.DEMD11KE003       .....T-IKDESKEIGM-----VT-EL-D-.K----L-----VVQ--GSNS.......TDN...-R-IN----T-------------------DI--E-ISS---- 290
D.SE.12.077UG             .....NCTAVHN-TIDM-----AT-ELKD-.-A-VN-L--R--VV----ND...........R-IN-----T----IT--------------D-N------K------------------------II-K-K-IS---- 295
D.TZ.04.C06405V4          .....TNN.GTISSTDRM-----VT-VVND-.QK-VH-LS-R--E-G-NS..........-E-..M-IN---A-A-S--IT--------------L--RD-Q-------S----------------II--E-I----- 285
D.UG.10.DEMD10UG004       .....NSTE-GLGV-------VT-EV-D-.KKQVQ-L--SH-VVQ-R-.............NT-..R-I-----A------IT---------------D-K--Q-R---------------------M--TE-L-NHL-- 275
D.UG.11.DEMD11UG003       .....ATIAHNVTEKIMT----KVT-E-ND-.KK-VH-L--VVK--DNDT...........SA-..R-IN----AL---------T----------D-N-T-----S---------------IV-TK-L---T-- 284
D.YE.02.02YE516           .......NNSTRE-M-------VT-VV-D-.QKQ---L---VAM............ST-..R-IN----T---------G-----E-NYT-K-------------------K---------------F---------I--E-I--I-- 275
F1.AO.06.AO_06_ANG32      .....TENSTLPE-A-A-------MT-E--D-.RL-VN-L--R---V--NSSDS..........NS...NE-..R-IN----T-K--------WD---------------------K------------------DII--TQ-IS--T-- 281
F1.AR.02.ARE933           .....TNNTLAD-P-A-Q-----VT-EV-D-.KL-VH-I--R---V--SGSND..........NS...ST-..R-IN----T-------T-----Y------XD-R-----D----------------DII--H-IS----- 283
F1.BR.11.DEMF111BR037     .....NDDTLREDP-A-Q-----TT-E-ID-.QK-VR-L----LXDGNSGSNG.........AM-..R-IN----T--------------------KEEK----------------------TK.-II--Q-IS--T-- 294
F1.CY.08.CY222            .....-QSSLEE-S-A-Q-----MT-EV-D-.QL-VH-L--R---V--NSSDN.........ES...SE-..R-IN----T-----------------*D--------Y--R--Q-D-------H----------------DII--Q-IS--T-- 289
F1.ES.11.VA0053_nfl       .....NNSTLKE-S-A-Q-----TT-E-D-.QKQVS-L-----V--NSNNS..........-F-..R-IN--TM----WD---------------Y-----------------------K---------------V-K.DI--S-Q-IS----I 302
F1.FR.04.LA22LeRe         .....GPNSTLKEDP-A-Q-----MTIEV-D-.KL-VH-L--R---V--NNNS.......SNS...-E-..R-IN----T----------------WD-----------KD-R----T-K------------------K---------SII--Q-ISN---- 284
F1.RO.03.LA20DuCl         .....NQTTLKEDT-V-Q-----DVT-ETNDR.KS-VH-L--R---V-LKDNSK.......NSS...RE-..R-IH--T-K--------------------D-K-----KD---------------------DII--FQ-IS----- 302
F1.RU.08.D88_845          .....QNDSL-EDQ-A-Q--T--M--EVKD-.TKQMRTL--S---EL-NKQSKX.......N...RTX..KE-..I-R-------T--I-D---X-T--Y---X-------X-K---G--X--K-X--------G.DI--PX-IKN--G- 290
F2.CM.10.DEMF210CM007     .....-TEAPPI-APPM-------AT-EVKD-.KK-----L--R---V-LNANANS......E-N-T-K-L-R--I-----------------------------K---------------D.-II--E-I-S-S-N 283
F2.CM.11.DEURF11CM026     .....NKYNISEDP-A-------T-E--D-.KE-Q--L--R---V-LEGSNS..........ST-..R-IN----TV--T--------D------------X-----E-N----------N------------S-G.NII--K-IS--T-N 287
G.CD.03.LA23LiEd          ...........KDMR-------T-E--DR.IKE----L--I--V--N-SN............NT-..R-IN--V-A-K-----T-D--------------KD-N-------K-----------------II--E-L----- 264
G.CM.08.789_10            .......SSDMNV--Q------TT-E--D-.RKT----L-----VQMGTND...........ST-..R-IN--V-T-K-------------RD-K----T-K-------------------------KIM--A-E-ISKSE-V 279
G.CN.10.DEMG10CM008       .....EKINANASGM-M-----MT-E--D-.KETK--L--R--VV---NGNN..........ST-..R-IN--V-T-VK----------------RD-N---Q-N--S------------------S-G.-IM--AE-ISKSE-V 283
G.CN.08.GX_2084_08        .....TNVTCNNATSDM------T--TT-E--DR.TK-----L---VVSLNQKLDN......NST..SN-..R-IN--V-T-K-----------T-----L-RD-N------Q-K-----------S-----------GI--TE-I-S-T-- 287
G.CN.16.224GX             .....TADTSVKDNIKML-----TT-E--DR.TK----L-----VVSLNQKLDN......KSS...GE-..R-IN--V-T-K----------------RD-N-------Q-K----S----------------GI--TE-I-S-T-- 289
G.ES.14.EUR_0033          .....-YSGTVDPXX-L------T-E--D-.KKQ-----L----TQ--DNSN..........MS...-N-..R-IN--V-T-K----------T-----Y------RDSK-----K---------------K.DIM--E-I-N----N 282
G.GW.08.LA57LmNe          .......ITANPR--LR------T-E-KD-.KKQ---L-----VQ-NDEGN...........ST-..DN-..R-IN--V-T-K----------T-------------GD-E---A-R----------------K--I----G.DIM--E-I----N 282
G.KE.07.DEMG09KE001       .....TN.STV-AK-MT------T-T-E--D-.QKN----L----RD-VTQ-RDEGN.......S...TNS...-L-..R-IN--V-T-K----------T--------D-N-T--S-Q--I-------------------K---M--E-I-----N 288
G.NG.12.12NG060409        TINN-TTSNLTGENPFV-----T-EL-D-.QT----L--RT-VV-L--N-DL......NI...NSS...GD-..R-IN--V-T-K------------D-------V------RE-N-T--E-K--S-----------------D.-IA--TE-I-N---I 297
H.CD.04.LA19KoSa          .....-S.I-SDVNDRL-----VT-I--D-.Q-RVT-I---V--VV-MEGEDNN...........-M-..R-IN----A--------N-D------------------K---------------------------K-I--K-IS-ST-N 283
H.CF.02.LA25LeMi          .....-NATYTDGEQ-V-----VT-V--D-.K--VH-V--R-----ENSA.............ND...-K-..R-IN-----------------------------M------------------------------I--E-I----N 285
J.CD.03.LA26DiAn          .....T-INGNNSSLGMR-----T-E--D-.QKR---L--RQ--V-LNR-NSS......D...NSS...DD-..-I-IN-----K----------N-D----------------K----------N-------------IS-.GII--E-IS----N 294
J.CM.04.04CMU11421        .....-PT-DAGRVNISEIM-----KMT-E--D-.TRR----L--RQ--V--SGRPN.........D-..I-IN------K--------N----------------------------DRN------------------KS-G.DII--E-I-N---N 291
K.CD.97.97ZR_EQTB11       .....NINATVTSTD---------T-ELKD-.KKRVS-L-----VQ-KQSEIN.......QSE...SED..R-IN---TV------------------------N-C--------------------------II---EDI-K-T-N 291
K.CM.96.96CM_MP535        .....NATSTVVSPA--------T-E-KD-.KK--S-L--R--VL-NGSS.........-E-..R-IN--T---T--T--------Y---------RD-R-------------------K--------------II--E-I--N- 285
01_AE.CM.11.1156_26       .......NVENI---R-------T-E--D-.QK-VH-L--RS-TV--NEGNS..........ST-..-I-IN------K----------D---------Y--R--D-------------K----------R--K.-II--E-I------D 277
01_AE.CN.12.DE00112CN011  .....VPIGIGNITD-VR--T-VLLAD-.K--V--L-----V-LNDNHNN...........N...ESS...GE-..R-IN------K-----I-D--------Y----------------------------I-----II--E-L-N--- 293
01_AE.IR.10.10IR.THR48F   .....A-DILGNVTE-VR--T--MT-VLQD-.K-RVH-L------LVQMNGSN..........NT-..R-IN------K-----I-D-------T--Y-X---D-R--------------------------II--E-L-N--- 287
01_AE.JP.11.DE00111JP003  .....-IEIGNIAD-VR--T--T-E-ID-.K-QIH-L-----V--K-NNN...........GE-..R-IN------K-----I-D-----T-Y-L----D-N---S---------------------------D.-I----E-I------ 289
01_AE.MM.14.fKSDU26       .ANCTGL-LENITN-Q--------T-E--D-.QK-V--L-C----VK-N--KIN......NDSNNSNSTTSS......NE-..-I-IN------K-----I-D------------Y---------T-N--S--------------II--E-I--T-- 305
01_AE.PH.15.DE00115PH012  .....TIENIGNITD-VR--T--MT-E--D-.Q--V--L-----V--NANR...........ST-..-I-IN--S--K------I-D--------Y-M----D-N------------------------K---I-----K-I----- 286
01_AE.SE.12.070SE         .....RTD-IGNMTD-VR--T--MT-E--DR.KKQVR-L----VQMNSS...........EK-..R-IN----I-K------I-D-------------Y-----------S---R-K------------K.-I----E-I----- 284
01_AE.TH.10.DE00110TH001  .....T-DIIGNITDDVR-----MT-EL-D-.Q--V--L-----V---DSSNN.......GS...SNF...-N-..-M-IN------K------------Y--R--D-R----------------GII--E-L-N--- 291
01_AE.TH.11.40436v09_04   .....INSTVNITD-VR--T--MT-EL-D-.K-QV--L-----LVS-NKNSS.......ENS...SE-..-M-IN------K------------Y----E-N--------------------------II--E-I----- 277
01_AE.TH.90.CM240         .....V-NIIGNITD-VR--T--MT-ELTD-.K--VH-L-----VQ-EDKKTS..........SE-..R-IN------K------------Y---------------------------------II--EDL-N-- 286
02_AG.CM.10.DE00210CM013  NY-.TRDGQEEM-R-MR-----AT-ML-D-.EKR-F-L--RT--VQ-NETKSN......NRC..STN...SC-..R-IN--S-T---------------------------------KD-K------Y----------------IM--E-IS---N 303
02_AG.DE.09.701114        .....-TISLTTNSTDM------T-ELKD-.KK-F--L--R---VQ-NETVDN......NNT..SSN...NQ-..M-IN--S-A--------T-----T--------RD-R-------K--------------IM--E-IS---N 296
02_AG.KR.12.MHI11_10746   .....NITDEMR--------VT-EL-DT.KKNVF-L--R--VQ-NENNSA......QYD..STK...SQ-..R-IN--S-A------I------------------------R---DEE-T--L---------------N---G.K-A--E-I-N--- 285
02_AG.KR.12.12MHR9_       .....NGSISSDMRE--------T--MT-EV-D-.K--VS-L--RY-VVQ-NENKN.......SS...SQ-..R-IN--S-A------I------T---------------KDEE--------------------------G.K-A--E-I-N-V-N 285
02_AG.LR.x.POC44951       .....LNDTQE---VR--------T-E--D-.KK-F--L--R---VVQ-N-SKGS.......KND..SQ-..R-VN--A----T------------SDED-K---Q-K--------------------K.KI-----E-I-N--- 276
02_AG.NG.12.12NG060304    .....NRSI...TTD-M-------KMT-EL-D-.E-RMH-L-----VVL-NGT-NST.......NSS...GE-..R-IN------------------------------R------------------------II--E-IS---N 282
02_AG.NG.x.IBNG           .........NLTSDMN-------T-EV-D-.KK-MH-L--R--VVQ-NENNG...........SQ-..R-IN------------------------------KD-G---------------K-------------II--E-I----- 280
02_AG.PK.15.PK032         .....GKNISAAETE--LT----TT-F--D-.Q-RVT-L----VVQ-N-NGNN..........-H-..R-IN------T------------D--------------------------K-.-I----E-I----- 291
02_AG.SE.14.098GN         .......NNTDME--V--T--MT-EL-NR.K--VK-L--RS--VVQ-KEPNND..........-Y-..R-INW--------------T--------------RDIE-K-------GK----------K--A--------N.RTM-T-E-I-N--- 278
```

Env

HIV-1/SIVcpz Proteins

glycosylation NIS
glycosylation NCS
CD4 binding
glycosylation NVS
CD4 binding
V1
V2
glycosylation NGT
glycosylation NFT
glycosylation NSS
glycosylation NDT
glycosylation NTS
glycosylation NKT
glycosylation NGS

```
B.FR.83.HXB2                     ....SSGRMIMEKGEIKNČSFNISTSIRGK.VQKEYAFFYKLDIIPIDNDT.....................TSY..KLTSČNTSVITQAČPKVSFEPIPIHYČAPAGFAILKČNNKTFNGTGPČTNVSTVQČTHGIRPVVSTQLLLNGSLAEE.EVVIRSVNFTDNAKT 283
02_AG.SN.13.9580                 ......MTRTVFPDM.........MT-VL-D-.KE-VTS#--RQ-#VQ---A-TC......DNTTCDNR...-Q-#AR-IN---A-------I------------D-N------K----------------K-----------N.-TMV--E-ISN--R- 285
03_AB.RU.97.KAL153_2             .......TSSIKMM-M---------T-DL-D-.-K---L---VVQ-------------......D--.R-I-----A-------I---------------D-K--------------------------------------------T-- 274
04_cpx.CY.94.94CY032_3           .......NGTVI-EG---------D-T-E-D-.KK---L---RI--V--NARVPIN....GSNRN..NST...EE--.M-IN--A-T-K---------E-N-T-L-----S-R-------------------K----------T------K-I---T--N 288
05_DF.BE.x.VI1310                .......N-TTLKE-T-AVQ------MT-EVND-.KL-VH-L--R--V--VVVR------......SN---S---.-R-IN---T-K-------WD------------Y-----E-K-S-----K----------------K-.GII--Q-IS----N 287
06_cpx.AU.96.BFP90               .GNN-TIVVDISK-<---------T-E---D-.TK---L---RP--V--GD-SNN-......SD--.M-IN--V-T-K-----------------RD-D-----K-----------------------------K-.NII-KTE-I--T-N 294
07_BC.CN.98.98CN009              ......CNETCRESMK-M-------AT--VV-D-.K-TV--L--R--V--FTKKNSS......G..NSS...EH--.R-IN---A-------T-D-------DRR----------H----------------K-.II--E-L-N-V-- 287
08_BC.CN.97.97CNGX_6F            .......YNETYNESVK--------AT-LL-DR.KKTV--L--R--V--LNDENSG......E..NSS...EY--.R-IN---A-------T-D------T--Y-----D-I---Q-H-----------------R.-II--E-L-N-V-- 287
09_cpx.GH.96.96GH2911            .........TSDGSEVRMQ------MT-E--D-.H--V-SL--RS--VS-NASN........NE--.R-IN-------------------T----------RD-E----I-K--S--------------------DII--E-L-N-V-V 278
10_CD.TZ.96.96TZ_BF061           ........VGDPRMT--------TT-EVKD-.KKQ---L-----VVQ-GDSN.........-R-IN------A-------T-D----------DRK----K-I------K-----------------II--E-L------N 274
11_cpx.CM.95.95CM_1816           .......T-NSTVEGQA-M------MT-E--D-.QKQ---L-----VV-LNDNDN.........R-IN-XV-TVK-----T---------------D-K-------------------------G.--R--K-I----- 294
12_BF.AR.99.ARMA159              .......QNNTLEE--P-A-Q----MT-E--D-.QL-VH-L--R--V--G-GS..........ET---R-IN-------T---*----WD------T--Y-L---D-K-------------------DII--Q-IS--T-- 285
13_cpx.CM.96.96CM_1849           .......NSST-QGRE---------Y-MT-GVGDR.R--V-SL--RV--E--EGNSNS......GN...SSS...SE--.R-I-------N---------T------------DEE----------K--------------G.-IK--AK--S--VGN 286
14_BG.ES.05.X1870               TNTNNDDIGG--R-----------AL-D-.IK---L--S--TQM----------......R----------------T-----------V---V----K----------------------I----K------T-- 288
15_01B.TH.99.99TH_MU2079         .......NRTI-G---------T-.V-D--L--VEA----------......R-RK--I-----------T--------------------------------------------------I--K----S---N--I 274
16_A2D.KR.97.97KR004             .......TQSPDSNNTM-------ETT-EL-D-.K--VQ-L--E---VQLNSSDSN....D...TLN...RQ--.R-IH-D--T-----------------KDPE----------K-----S---------A------GNKTI--A-I---T-N 284
17_BF.AR.99.ARMA038              .......QNSTLKE-A-A-Q-----MT-EV-D-.QL-VH-L--R--V--NGSNG.........SG...GE--.R-IN---T-------------WD---------Y-----D-K----------------------II--K-IS--I--N 285
18_cpx.CU.99.CU76               ......A-NSS-#ATE-M-------F-T-EV-X-.QKQ---L--X---VS-#G-XN.........SNY..SD--.R-IN--V-T-K-------------QP-----------N-------------------T--K-M--I--I--N 284
19_cpx.CU.99.CU7                ...........HM------------MT-E--D-.K-QI-SL-----V--GD...........-----D-----C-----K-R-------------------------------------K------ 256
20_BG.CU.99.CU103              ......TVISSED-NI---Y-T-E-KDR.KKQ---L--R--VV--NDKNSN.....NS..NNY..SD--.R-IN--V-A-K-------D----------------------------------------K.-IM---E-L---V 281
21_A2D.KE.99.KER2003            .......EG-TQN-NQTVDM-----T-VV-D-.RKQV--L-----N--DN.........TT...ST--.R-IN------A-------T-------Y-------RD-R------------------------------K.-II--E-L---V 282
22_01A1.CM.01.01CM_0001BBY      ......SSVEEL-T-VT-------MT-EL-D-.RREV-SL--R---VQ-NGNNS.........F.-R-IN----------------------------KE-----------K-----------------K.-IM---E-L-N-V 276
23_BG.CU.03.CB118               ......TTNGTVNNSAD--------T-E--D-.SK-----L-------M--DKNNG.......SN..KNY..SD--.R-IN--V-T-K-------D--------------RD-E----------------------------------K.-IM---E-L-N--V 289
24_BG.ES.08.X2456_2            AGDGTVRIDSNVTDDMTX--XMT-DLQD-.QK---XL-----X--SDNSN.......N..NTX..SU--.R-IX--V-T-X---X-T-D-X-------Y------KDEE--X-----R------H---X-K--X---K.K-M---E--AN---V 293
25_cpx.CM.02.1918LE            ......NNNSTVNNRE---------MT-E--D-.KK-----L-----V-L-D-N..........S--.R-L----T-K-----------------KD-N---A----------------------------I---E-I-N--N 285
26_A5U.CD.02.02CD_MBTB047       .......NHTDINR-MR--------MT-HL-D-.QE-VH-L--R--VLMNGNDSK......NSS...GE--.R-IN-----------------D----------RD-D-----------------------------G.G-KV--E--N---V 281
27_cpx.FR.04.04CD_FR_KZS         .......-TNTS-HTTE-L------VT-A-KD-.KL--LQ-L-ES---V-MNKSS........NT--.R-IN-----K---T-------F-------------------L--------------------II--K-EL--T-- 281
28_BF.BR.99.BREPM12609          .......GNSSVEAME--M------VT---GD-.M-----L-------VV--A--N........-T-IN---------------------------------------L-------------------II--E-L--N--- 286
29_BF.BR.01.BREPM16704          .......-NSSW-N-E--M------VT--RKD-.MK-----L--HT---VSTGEGDN.........-----I-IT-------------------T----------K-----------------T-K-------DI--K-E----- 289
31_BC.BR.04.04BR142             ............TIV------T-MT-EL-D-.QR-VH-L--R---V-LESESY..........ED--.R-IN----------------D--------YV----------K-----------------K.-II--K-LS--V-- 275
32_06A6.EE.01.EE0369            ......TKNITVNN-E-MT------VT-EL-ED-.KKT-----L------V--EGKNNN......GNS...SA--.I-IN--ATN-E---------------------RD-N-----K-----------------G.-SII--E-I---T-N 292
33_01B.ID.07.JKT189_C          .VSSEIANLGNITE-VR-------VT-EL-D-.KE-A-L------VQLG-SNS.........SK--.R-IN------A---------D-------T------------K-----------------II--E-L---N 288
34_01B.TH.99.OUR1969P          .GPNL-NIENITE--VR-------VT-E--V-L-----VVQ-NGNG...........GK--.R-IN-----K---I-D------T-----Y-D-K------------------------II--E-I---N 285
35_AD.AF.07.169H               .NDSKIEDMV-----------MT-EL-D-.K--VQSL-----VV--D-N...........S--.R-IN------A---------T----------------E-E-----S-K--S----------------G.-I---E-IS----N 285
36_cpx.CM.00.00CMNYU830         ......NVSKDMQ----------TT-VL-D-.I--VHSL--R--V--ASNN........DS...SR--.R-IN--A-A-----------T------------KD-Q----L-K--S--------------KN.-TI--E-I-N--N 281
37_cpx.CM.00.00CMNYU926         ......-KSLDGG--EI--------Y-MT-EL-D-.K--I-SL-----VV--GDNN.........ET--.R-IN-------------T------Y--------D-EY-----K--------------------G.K-M--E-I-N--N 279
38_BF1.UY.03.UY03_3389          .......DSTL.E---A------T--MT-E--D-.QL-VQSL-----V--R--V-........RY--.R-IN---T-K-------WD------------D-K-----------------------------II--Q-IS---- 279
39_BF.BR.04.04BRRJ179           .HNSTLKEDP-A-Q----FMT-EV-D-.-K-----L--R----V--S-SSN........SSS...GE--.R-IN-S--A---------T------Y-------------K-----------------KII--Q-IS---- 285
40_BF.BR.05.05BRRJ055           .N-TVLKE-P-A-Q----MT-EV-D-.QK-VH-L--R----V--NGEGN........DN...GT--.R-IN-----K-------------T-----Y---------------K-----------------II--K-L-G-TN- 289
41_CD.TZ.05.CO6577V5            .......TNST-MGNAETRM----MT-VV-D-.KKQV--L-----VS--S...........GD--.R-IN------A-------T-------------RD-Q--------------------------------II--K-L-G-TN- 274
42_BF.LU.03.luBF_01_03         .......KLNSTEEI-N--------VTLG-KS-.M-----L--RS--VV---#...........SS...-M-I----T-----------------D-K---X-K--------------------------DII--Q-IS----- 283
43_02G.SA.03.J11223            ......-VKCEDV-NL--------VT-EL-D-.KK-----I------VV---GNSN.........NES...GD--.R-IN--V-T---------T-D----------KE-N----Q-K--S-----------------I---K--S---QV 281
44_BF.CL.00.CH80               .......AASTLKE--P-A-Q----MT-EV-DR.RL-VN-L--R---V--S-NSS..........RE--.R-IN------A-------T-K-----WD----------Y-----KD-R----L-Q--S------------I--KII--Q-I--T-- 284
45_cpx.FR.04.04FR_AUK          .......AVE--MR------MT-EL-D-.Q-ER-SL----VVQ-SENND.........S---.R-IN---T-K--V---T--------------------KD-R------L-Q--S---------------I------KIM--E-L-N-V-- 278
46_BF.BR.07.07BR_FPS625         ......TQK-ILRE-Q-A-Q----MT-E-D-.QL-VH-L--R---V--S-NSS........-----.R-IN-S--T-------WD------------D-K--------------------------------DII--Q-IS----S 278
47_BF.ES.08.P1942              .......NNS---------T-QT----.RKR---L--N--VVQLES-KTS.........-R-I-----------I----------------------------------------------------DII--Q-IS----S 297
48_01B.MY.07.07MYKT021          .......TT-KIGNMTD-------MT-VL-D-.Q--VH-L------V-E-NSNR......NNSNSSNSSSEYSE....-D-.I-VH---T--K----------Q----------Q--DIM-----I----K.GII--E-IS--T-N 286
49_cpx.GM.03.N26677            DTTPT-PTATSSPI-M-------T-.GD--.KK------T--R---NS-.........-----.M-IN------A---------------------------------------------K.DII--E-I---N 282
50_A1D.GB.10.12792             .......NNNTELED--EG-----Y--T-DL-D-.K--I-SL-----T--NENHS........SDS...SE--.I-I-----T--K----------Q---------------------------------K.DII--E-L--N 282
51_01B.SG.11.11SG_HM021         .NVSGIGNITD-VR--------T--MT-E--D-.-QK-SH-L------V--GTQSNT.........-----.R-IN------A---------I-D-------T--X-----D-N-----K-----------------II--E-L---N 285
52_01B.MY.03.03MYKL018_1       .-T-KTIGNLTD-VR--------SMT-ELQDR.K--IQ-L-------VQMNSNSNS......NS..SNSSKYRE...-R-IN------A-----------T--------T-X---------D-N-----K-----------------II--E-I---T-N 295
53_01B.MY.11.11FIR164          .TTASIGNITD-V---T--MT-ALKD-.X--A--L-------VE-NKNSNS......SN..SSS..SX--.M-XN-X-----T---------------------------------------------------K.DII--E-L---N 292
54_01B.MY.09.09MYSB023         .TTKVIGNITD-V---T--MT-EL-D-.K--VQ-L---I--V--NNNN.........NS...SD--.R-IN---------------T-------Y----------------------------------------DII--E-L-N--RN 288
55_01B.CN.10.HNCS102056         ......GPD.TPNITD-VR--T--MT-E-D-.K-QVH-L--R--V--SG-ENN.........-----.R-IN------A------------------T-------------K-------------------DII--E-L-N---N 280
56_cpx.FR.10.URF5_patient_A     ......STVDSGTD------TPTV-D-.RK------NN---VQ-NDGDN.........S--.R-V------------------F----------------------------K.-IM---E-I-N---I 279
57_BC.CN.09.09YNLX19sg          .ICNHINGRTLDENVR-----TT-LVKD-.K--VH-L--R--V-L-GKDSSD.......-----.R-IN--S-T-------N----------------------------S-D.GII--QDLKN--- 288
58_01B.MY.09.09MYPR37          ASPLGNITLGNITD-V---TT-LVKD-.K--V-L--RV----E-NNKS..........SE--.M-IN------A-------T-----------XT---Y-------D-N-T----KK--S------------K-.-II--E-L-N---N 297
59_01B.CN.09.09LNA423          ......DTMN-T---------TS--KD-.IR-----L---------T--I.........-----.-T-I------------------------------F--S---K--------------------IV--R--N---N 282
60_BC.IT.11.BAV499             .......YAPSMQ----------MT-EL-D-.KK-VH-L--R--V--LE-ESSDN......GNN...NTF..GE--.R-IN----AM--------------------------E-----------N------------G.-II--K-L----- 287
61_BC.CN.10.JL100010           .......KGQ-M-E-----TT-VL-D-.K-TG--L--RI-LV-LGENSSES......YRL...NSS...KF--.R-IN------A-------N-D-------T--Y---Q--------------------------K.-II--E-LA--V-- 280
62_BC.CN.10.YNFL13             .......TYE-RQ----------MT-E--D-.QK---L---R--V--LNDKNSS......D...KNS..GE--.R-IN------A-----------------T--Y---------SP--------N--------K.-II--S-E-L----- 287
63_02A.RU.10.10RU0637          ......ST-LL--F--MT-V-D-.R-RMS--L--R---V--TEGQNN..........SSK..GE--.R-IN------A-------T-D----------------------------------------KD-E.-II--K-L----V 277
64_BC.CN.09.YNFL31             .TDNGTMEDNANV-S-----V-MT-V-D-.K--VN-L-----VLNES..........SSK..RE--.R-IN------A---------------------------------------------------II--K-L----I 285
65_cpx.CN.10.YNFL01            SNESVAKVSNESVEHM------VT-IV-DR.Q-TAF-L--RS--V-LEGNSNN......SS..NSS...GA--.R-IN------T----------N-D-------------------------------------DII--E-L-N-V-- 296
67_01B.CN.11.ANHUI_HF115        .......XTXXIGNITD-VR----T--MT-E-D-.K--XVH-L------VEMEKNK..........SE--.R-IN------A---------T-D----------D-K---------N-----------------II--E-LXN---I 287
68_01B.CN.11.ANHUI_WH73        .NIDNITD-VR----T--MT-E--D-.K--VH-L------VQMEGNN..........SE--.R-IN------A-------T-D------T--Y-D-K----------N-----------------II--E-L---N 280
69_01B.JP.05.05JPMYC113SP420    .GPKEIGNMTE-V----A-EL-D-.K-QA--L-----V--RD-NNS......SDNNSSAS....KE--.M-IN-X-----T---------T-D----------Y----------L--X-----------------K.-IM---E-L---N 287
70_BF1.BR.10.10BR_PE004         .NRTLGM---E--L--------T-N-GD-.-K---GI--N---V---------......-R-I----XL-------X----------------------------------II----S-L---T-I 283
71_BF1.BR.10.10BR_PE008         .N-SEGK--R--M---------VH-DM-N-.E----------VTQ-KE-NN.........ASN...S--.R-I-----V------------------G--K-K--K-N---------------------E------N 289
72_BF1.BR.10.10BR_MG002         .-Q-NNTLGE-P-A-------MT-E--D-.QR-VH-L--R--V--N-NENNN......SDN...NSS...GE--.R-IN----T---T--I-WD-------------Y--------------K-----------------II--Q-IS----- 288
73_BG.DE.01.9196_01            .......NSSSW--M-------VTSIT-D-.K-Q--Y--N---VQ--SESNS..........-K--.R-I---------I-------T-----------------------------------------K.-II--E-L---N 282
74_01B.MY.10.10MYPR268         .DLDKIGNITD---------T--MT-EL-D-.K-QV--L------VQ--NVTSNN.........S...GE--.R-IN------A--------IT-D-------T--X-X---XDXX------------------G.-II--E-L----N 290
77_cpx.MY.14.14MYNBB090        .TNNTIGNITDDVR-------MT--H--A--L----T---EENSKN......STY...RE--.R-IN------A---------T-D-------T--Y-----E-N-----K--------------------I---K-ISN-VN- 287
78_cpx.CN.13.YNTC19           .ENRLIGNITEDL--------MT-EV-D-.KKNV--L------V--NEKKDN.........N---.R-IN------A---------T-------T--Y---------E-N--S------------------D.-II--E-L---N 287
79_0107.CN.15.SX15DT013        .-NSTLDNRQ-M----------MR---TT-VL-D-.K--V--L--R--VE--KNS..........SE--.R-IN------A--------T-D-----------------------------------------G.DII--E-MS--V-- 287
82_cpx.MM.13.mSSDU12           .VNNTIGKITD--S--T--MT-E--D-.KKQVQ--L--R--LVQ-G-NNST.........KE--.M-RN------------------M--------T--Y-L---D-K-------N--------------------G.DII--E-I----N 274
83_cpx.MM.13.mSSDU04          .VNNTIGKITD--S--T--MT-E--D-.KKQVQ--L--R--LVQ-G-NNST.........KE--.M-RN------------------M--------T--Y-L---D-K-------N--------------------G.DII--E-I----N 286
85_BC.CN.14.14CN_SCYB2        .TIV-NSINGTL-EQM-----AT-EVKDR.K--V--L------V-L-EKNSS......G...NHS...EY--.R-IN------A-K--Y--T-D-------T--Y-L-------------N------------------G.-II--E-I---N 291
86_BC.CN.13.15YNHS18          ......EAH---Q-----AT-EVKN-.K-SV--L------V-L-P-..........KND..SY--.R-IN--A-K--Y--T-D-------T--Y-L-----------F-----N------------------G.-II--E-L-N-V-- 287
87_cpx.CN.12.DH32             .TYNGTQ-M--------VT-ELKD-.KKN-F-L--R---V-L--NKTG..........QY...NET..RP--.R-IN------------------------T--Y------L------------------S-G.-I-V--E-L----- 285
88_BC.CN.05.05YNRL25sg        .YRTYN--TDKEAYK--------AT-ELKD-.QKSV--L--R---V-LNDSKPLNESKPRNDSN.....HSS...EH--.R-IN------A-------T-D----------Y--------------------------DII--E-L-N-V-- 302
90_BF1.BR.11.BRG06043         .TNSTKEDMN-DM--------A---GN--M-D-.-------V---V---GNDN.........-----.F..I-IG-----------------------------------------K.-I---A--S---S 291
O.CM.96.LA51YBF35             .N-TKTTDTTNPM-A-D--VT-VLKD-.LE-KQ-L--VS-LMKVEG..........SNS...-T--.I-IN--STT-R----------------Y-F-VDED-----S-N-I-V-T-------K-A------I---T-S--.KIR-MGK-IS-S--N 278
O.CM.96.LA52YBF39             .DLLTNALK-D--VT-ILKD-.KE-KQ-L--S-L-K-NS..........TEG...NT--.I-IN--STT-K-------------------M-S-ITV-T---------K-A------I---T-S--.KIR-MGK-ISNTG-N 270
O.CM.99.99CMU4122             .NATAASSAV-K-E--VT-VLKD-.KE-QQ-L--RS-LNK-E-S........TTS...-M--.T-IN--STT-R-------------------G-Y-F----Q-R-ITV-T-------K-T-----V---T-S-K-.KIR-MGK-I--SG-N 274
O.CM.x.pCMO2_3               .TRKEQEQLM-Q-E--VT-VLKD-.KE-A-L--VS-LSKLEGNN........TSD...-M--.T-IN--STT-K----T--------------Y-FQ--SPE-----E-S-ITV-T-L---T-----V---T-SK-.KIR-MGK-IS-S--N 278
O.ES.01.Read25_HIV_Group0     .NGTTIENPV-K-E--TT-VVKD-.KE-KQ-L--VS-LMKLNDNS.........TNT...-M--.T-IN--STT----------------------T--------E-T---T-K-ITV-T-------T------I---T-SKG.K-R-IGT-IS-SG-N 273
O.FR.06.RBF206               .GLVGETPM-K-E--TT-VVKD-.KE-KQ-L--VS-LAKVNTSN.........VND..TNT...TMIN--STT-------------------------T------E---L-H-I-V-T-------K-T-----I----S-K-.KIR-MGK-I-K-E-N 277
O.GA.10.10Gab1190            .SSSENETPV-K-E--AT-VVKD-.QE-KQ-L--VS-LMKPNET.........DNA...-M--.T-IN--STT-------------------------G-Y-F----E---T-S-ITV-T-------TI------I---T-SKG.KLR-MGK-IS-S--Q 277
O.GA.11.11Gab6352            .NITENDGETPM-V-E--VT-VLKE-KEE-KQ-L--VS-LVKL--NG.........TNQ...-I--.TIIN--STT-K-------------------G-Y------P-Q-T---TV-T-------K-T-----I----S-Q-.KIR-MGT-ISAS-QN 282
O.SN.99.99SE_MP1300          .SVKNDTSSSENLM-K-E--VT-VLKE-.KE-QQ-L--VS-LMKVNE.........NND...-M--.T-IN--STT-K-T-------------------G-Y------P---TQ-----T-V-T-----------H-I--TISKG-KIR-MGT-IS-TG-N 279
O.US.10.LTNP                 .NTTDM----E--AT-VVKD-.KE-KQ-L--VS-LT-L-DKKNS.........SNS...-M--.T-IN--STT--------------------T--------K-T-ITV-T-------K-T-----H-I--TISKG.KIRMMAK-I---N 283
N.CM.15.S4858                ...T-STG-.GNKLM-------AT-ERKD-.KKTV-SL--AE-VVK-S--.........ST--.R-I----T-V-----I-----------P---L------EGN-S-N-T-K---------------K--I--I------STG.DI---N....SSRG- 277
N.FR.11.N1_FR_2011           .INNTS-IQVKAM-------AT-ERKD-.KKTV-SL--NE--VQ-NDTVNN......RTE...V..R-IK-D-T-V--V---T----------P---M--GNX-SAK-S---------------K--I--X----GT.-I-XXN....SSR- 277
P.CM.06.U14788              .TNTSKTNPL-LFQ-----TT-IVKD-.Q-TQQ-L--RA-LTKL---...........ST--.R-IN-------T---N---L--Q------Y-LM--QTG--------NKTVITH------K-T-----I---T--KG.-PLVITQ-VS-TG-V 292
P.FR.09.RBF168             .VNXTADNPLDLFK-----TT-VLKD-.K-GQQ-L--RT-L-ALNET.........ANN...-L--.R-IN-------T---N----L--Q------Y-FR--Q--------NQTVITH------K-T-----I-------.-PLVITQ-VS-TG-V 289
CPZ.TZ.06.TAN5             EPEIKRLGKNITNY-VYK-----QT-EF-D-.KKQI-SL---E--MKNN-TNNN..........N--.Y-HN-----A---E-A-ST-------M---KDQN-T-K--S---V-H----Y-MIA-A-H---T-E-N.KTTAYFADS-H--QP 288
CPZ.US.85.US_Marilyn          .VNTTDIMR----------T-EL-D-.KKQV-SL--VD-LAH-N..........P----.R-IN--T----------------X------------X--TK.N-TV--K--A---I 265
GOR.CM.12.SIVgor_BQID2        H..NMTSTPNIPEL-VY--T--VT-VLKD-.KSQQQ-L--RE-LSQ-GDXN.........ST--.R-IN--T-S-------L--Q-------Y-LM----QSK----T-NETLITH-------T----FIF--T--K.-LLML-K-IS-SG-- 297
GOR.CM.13.SIVgor_BPID15       ......Q-STTAQNNLNFY-----VT-VLKD-.KE-KQ-L---ENLV-LNQT.........SKE...KL--.R-IN--TT-------K-L-Q-----KY-LI--KQ-G--X-XL-NXTVITH--------T-----IF-------.-ALVM-KDV-ATG-N 293
```

HIV-1 Proteins

HIV-1/SIVcpz Proteins

Env

```
                              glycosylation NNT                                                          glycosylation NST
            glycosylation NCT                    V3 tip                                         glycosylation NST
                                             V3                                               glycosylation NST
        glycosylation NTS                                glycosylation NNT                         glycosylation NKT   CD4 binding   glycosylation NST      V4
B.FR.83.HXB2          IIVQLNTSVEINCTRPNNN.TRKRI.RIQRGPGRAFVTI.....GK.IGNMRQAHCNISRAKWNNTLKQIASKL..REQFG.........NNK.....T..IIFK..QSSGGDPEIVTHSFNCGGE.FFYCNSTQLFNSTWFN.....ST.......WST.......EGS 405
A1.CD.02.LA01AlPr     L----VKP-N-T-I-------G-.G-...--QT-YAA...DAI--I-H-Y-VVNKTE--K-Q-VVH-----HW...................A.....P----I-T----------TSG----HV..NN....T-......... 407
A1.CM.08.886_24       -----Q--V---S------.--RS-..---QS-YAT...-EI--DI-E-R--VNGTE--KA-QKVVIQ-----H-..................G----K-N..S-----L-T----A---TSG----HS..NG....TANG....... 395
A1.ES.15.100_117      -----KEP-N-T---G--..--TS-..---Q-YAT...-DI--DI-K-Y--V-K-A-K--QGV-QQ---KY-.......-E-N..........N..S-----L-T--------TSS----ET..NN....A......... 399
A1.KE.11.DEMA111KE002 -----REP-R--------.--SV-...--Q-FAT...-EI--NI-K-Y--V-GVE--K--QKVVRQ-.-K-Y-.....................--S-N..N-A--M-T-------DTSG----NG..TA....SNV.......SI 407
A1.PK.15.PK034        -----TKA-Q--------..--QS-..---Q-YAT...-AVT-DI---V-KSA--A--Q-V-A------QGYW...............-V-A..N-----I-T---------TS-----..........A 396
A1.RW.11.DEMA111RW002 -----VKPIG--S--.-.--SV--..---QT-YAT...-EI--DI---V--ED-FEA--N-SKQ--GKH----......G...........-VYD..K---L-VT----------T-G-----SR..N-.......TDNN.TWFRNNA- 412
A1.SE.13.078SE        -----VDP-I------.QK.KSRSM..---V-YAT...-E-IT--.K--S-TVNGSA--E--QKV-AQ----KHW.........T...RP---L--TM------TSS-----DL..NA....TRE....... 393
A1.UG.11.DEMA110UG009 --H-TKP-S-R-I------.--TSV-..---QE-F-...-S-I---I-----Y--V-KSE-E----Q-V-TQ-----KYW.................-K-T.GPV---V--T--------TSS----SG..G-.....NN......... 397
A2.CM.01.01CM_1445MV  ----FTEP-Q---------..--S--.F-...-Q--Y-N...-N-I--DI-----K-D--V--QKVVKQ-.-HSH----...........P-AT.......N..SP----L-T-------F--T-G-----TK..EG....N--W...KKNDSV 403
A3.SN.01.DDI579       -----VNP-Q-S-------.--TSV-..---Q--YAT...-DI--DI-------V--EN-TE--Q-V-KQ------KRH--...........T..G----L-T--------TSG-----ND..TA....IND.......TI 403
A4.CD.97.97CD_KCC2    -----DIP-K---S------..--QT-YAT...-DI---I-------L--TA--D--YNVSKA--.--H-.....................P----------N..K----L-VTQ-M-I-------TS-----S-AP....NG....SWI.......NN 402
A6.BY.13.PV85         -----KEP-N-T-I------..--TS-...--Q-YA-...-EV--DI---I-R-Y--V-EKA-EK---QN-STQ----.-Y-.........G......T---T-A..G----V-VT--------DTSK----NH..XK........NI 402
A6.IT.02.60000        -----AKP-N-T-I------..--TS-..---QT-YAT...-DI--DI-X-Y-FVN-TE-----QK-SK--.-GAY-.........K----------R..N----L-T--------T-A-----NE..NN....T.......I 405
A6.RU.11.11RU6950     -----TE--S-T---G--..--TS-..---QT-YAT...-DVX-DI---V-Y--V--TE--S--QK-STQ----.-RY-.........K----------R..N----V-VT--------T-E-----DG..NG....TA......... 402
A6.UA.12.DEMA112UA024 -----KEP-N-T-IK-GT-..--T--HH-..---P-I---V-S----K--QG-RRQ---.-NY-.........T---------R-P..N----L-T--------T-D-----DR..ND....TDAN...STQV 412
B.AR.14.DEMB14AR003   -----KT-S-T-Q------..--S---N----IYAT...-QI---I---Y---N-T-L---Q-V-N-..................--V-N............--------T--------NS..TW....NW.......NST 402
B.BR.10.10BR_MG035    -----Q--A-T-I------..--S-.HT..F--TIYAT...-EI--DI---Y---KE--T--L-VE-----Y-............-K-D..KP-----IM------TS----T-YTGL..ND..........TTN 401
B.CU.14.14CU005       -----E--V---S------..--S-.-H---Q-FAT...---I--E---K-A---K--SK--.-GK--.........E------------S.Q..P----L-E--------T-P---S-YK..N-........-KS...QRT 383
B.DE.13.366396        -----K--N-T-V--G--..--S-.-YAT...-EI---I------QA-EH--NET-K--G--L......................KIT.......-Q..P----L-E--------T-E-----V-N..-G......... RV 383
B.DE.13.947915        -----KEP--------GV.P-..--KVIYAT...-QI--DI-------E---I-EK--K---Q--...................D----S--..N-------M-I-------F--T-N-----I-DI..NS....TWN...STER 407
B.DO.11.DEMB11DR001   --H--K--Q-------S---.-N-.-----Y-T...-EI--DI-------L--TD-----RS--I--K-----............R...........................-N..T-K---S-DM..-N....LN....STE 407
B.ES.15.100600        --H--K-S-T-V-------.--SV.HM..--SVWYAT...-DI--DI-------END-Q-A---V-T-.............-N-A..KHA---VEL........-T-K---S-DM..N-....-N........SN 407
B.FR.11.DEMB11FR001   --H--K--Q-------S---.-N-.-----Y-T...-EI--DI-------L--TD-----RS--I--K-----.........-AT.........-V-N...--TM-M--G----T-P-D--NG..GE......PV....SN 389
B.JP.12.DEMB12JP001   --H--Q------------..--RS-.HL..--QTLYAT...-DI---Q-A-----R-VVA--.-K--.......-AT..........-V-N...--P-----T-K-----N-........PV....WPN 411
B.PH.15.DEMB15PH003   --H-SEA-Q-------S---.-T-..-----YAT...-EVT-DI-------GTE-----V-VSE--.-DK--.........K---------V--.P---L-EM------T-K-----N-.....-N......RNN 398
B.RU.11.11RU21n       -----KEP--------G--.HM..-K--YAT...-DI--DI-------V-KI-----I-VVG-----.......K----LN-T..K------.-V-N..--T-K-----DK..NI....TE....FNKT 393
B.SE.15.102SE         -----KEP-Q-------G--.HF..-KT-YAT...-DI--DI------KET-----VE----R-.......K---------.-V-N...--T-K-----RI..NG....TL....N- 397
B.TH.10.DEMB10TH002   ---E-IP-V-V--------.--S----TWYAT...-DI--DI---Y--TVGN-S---QR-DK--.-K-----...........R-----N-N...--------T----N--LT..WK....AN....ASD 414
B.US.16.2609          -----KDP---T-I------.--S-.TM..-----Y-T...-DI---I-------N-TD--K-----.......K----------D..P----L--M-----------T--------DC..TG....TAC..TLNGTV 401
C.BR.11.DEMC11BR035   -----N----------..---ES-..---QV-YAT...-DI--DI-S-Y-T-NGTL--YT--REVERT-.-AKH--...........NDT-NT...--K-A..NH----L-T-------TSD--DIDTS........ 409
C.CN.10.YNFL19        --H--Q--V---S------.--S-.----QT-YAT...-DI--DI-----AS---K--QGVSK---A-L-.........P------R-A..S-----I-T--------TS---G-YNGTYM.SNG.......TES....... 414
C.DE.10.622166        --H--E------------.--SV.TL..-Q-IYAT...-DI--DI-----V-EG--T-DV-K-Y---L-H-----.......P----R..N-----L-T--------TSK----FNSTHLEDNA..TNN......... 403
C.ES.15.100_118       --H-K-A-K----------.--S-.-G-..---QT-YAT...-DI---I------KKE-VE--TNVSK---A-H-.........N-T..KPA----L-T--------TS---S--G-YN........GDL......... 396
C.ET.08.ET104         --H--K---M-I-N-----..--S-..---QT-YAT...-DI--DI-K-----RN-TEA-EKV-A---KVQQ.........GEY-I-----D.--T-D..--------L-TM--I-R---L--TSG--NSLLA..N-......... 392
C.IN.15.NIRT008       --H--Q---V---S-----.--SV.----QT-FAT...-EV--DI-K-----E--QRV-K--.-K-H-.........P----E-E..PP----L-T--------T-N-----YLA..NS....TLS......... 389
C.MW.09.703010256_CH256.w96 --H--E------------.--SV-..---QT-FAT...-DI--DI-Y-T-NGTQ--D--NVRRET-.-QRH--...........R-..RH----L-T--------TSKI-D-DN..........AN......... 398
C.NG.10.10NG020523    --HF-R--P-V--G--..--S--..---QT-YAT...-GI--DI-----T-SQ--E--QWVSK--.-E-H-.........P----Q-N..-A--L-T--------TS---G-YS-I..NG....TYS.......NI 410
C.NP.11.11NP016       --H--Q--IK-V-------.--S-..---QT-YAT...-DI--DI----E---E--QVGK---S-H-....................H--K-A..S-----L-VT--------T-G--G-YMPTER.PNG.......TES.......N 405
C.PK.14.DEMC14PK009   --H----P---------.--SV-..--Q-YAT...-DI--DI------EKA-DT--QEVGK-----H-.........P----S-N..S--------T--------S-----YNG........TEX.......N 398
C.SE.15.100ET         -----KEP-Q----GT-.R-RS.--Q-IY-T...NEVT-DI------VTNT---Q-Q--VRN---ERL.........QT.......--L-E..-PV--L-VTM-I--R---TSN--RSIL-V..TL....TN......... 403
C.TZ.08.707010457_CH457.w8 --H--EP---R---G--.--SV-..--QT-YAT...-EI--DI-K-----S--E-ER-SN---E-H-.........P.......-2-R-N..S-----L-T--------TSK--NNI........-T......... 389
C.US.14.M0118v1d14_5M04_C5 --H--D---V---S-----.--SV-..--QT-YAT...-AV--DI-----EKQ--KA-QGVSE---K-H-.........K------R-N..S-----L-T--------TSG--G-YNFT..ND....TGS......... 409
C.ZA.13.DEMC13ZA152   --H--E----E-I--G--.-RSV..---QT-YAT...-DI--DI-----V-EK--DE--HRV-R-----Y-.........PRR----N..-T-Q..PP----L-VT-------F--TSG--ENRTYTL..ND....TES.......SN 414
C.ZM.11.DEMC11ZM006   --H--E----E-V------..--SL.---QT-YAT...-DI--DI-----V-EK-DE--HRV-R-----Y-.........PR----N-T..SPA----L-VT-------DTSG--ENRTYTL..ND....TGS......... 402
D.BR.10.10BR_RJ095    --I-F-K-------Y--R.QKT-S..-Q-Q-LY-S...-Q-IT-DVK--Y--V-GKD-QEMIQEV-K--GNLY.........TT......-R--E--.SP----L-TR--------TSG---N--NV..TA....KEW....NN 400
D.CD.03.LA17MuBo      --E-LS-T-I-------.--QGV-Q..---QM-F-T...-GI--DI------GT---S--QKV-EE--KDRL-.........IT......K-------.PH----L-T--------TSK-----EW..VN....GTL....TNF 417
D.CM.10.DEMD10CM009   --FKET-N-T-E-------.--SLY-T...-DI--DI-N-Y--V--TN------H-V-E--GKLL-.........-KT......-K-E..PAA-----TA--------TSY---G--I-G..-RN.....D 393
D.CY.06.CY163         --KT-----I-T------.--DS-.H..-Q-LYAT...-TNI---I-E-----SGD-----Q-V-K--GKLL-.........-V--.P-----T--------TKL---T--SS..E-....-EN....STG 406
D.KE.11.DEMD11KE003   --DKT-V-------S----.--SV.HM.-----Y-T...-NIV-SI-----L--IE--R-DRVIE--KNLT-.........IT......K--S-A..S-----T-T--------TSK--TTNITS..N-......... 403
D.SE.12.077UG         ----F-ET-R-T------.--QST.----Q-IY-T...--I-DI---Y-T-NATD-YSM--QKV-I-----KKLY-.........-KT......K-D..S-----VT--------TSN---T-LG..E-......... 408
D.TZ.04.CO6405V4      --K-IA-------Y----.--QV-H.----Q-YAT...-DI--DI-----V--DE--MTQWV-T---GSLY.........-RS......S..NA-----T---------TSG---DI..TN....T-S......... 406
D.UG.10.DEMD10UG004   ---F-K--T-----Y----.--QST.H.---A-IY-T....-T--DI-----VD-DT--Q-V-A---DLL-.........-KT......KIE-V-G..P-----L-VT--------TSG----QE..NS....TLL.......NKT 396
D.UG.11.DEMD11UG003   --I--KE--N-T-I--Y---.-S-.NM.-----Y-S...-NDI--DI-----V--TA-DK--Q-V-A---GNLY-.........-ET......-T..P----L-T--------TSE---I-T..N-....GNWT...GNTTN 406
D.YE.02.02YE516       ------ETIN-T-I------.--QG-.H..-A--Q-LY-T...-NV--DI-------GTQ-----QK--K--GKLY-.........-KT......-V--.P-----T--------TST---TR-INT..TN....DN-.......VN- 394
F1.AO.06.AO_06_ANG32  --H--E--Q----------.--RS-.HL..-Q--YAT...-DI--DI-K-----GDQ-----QKVRA--KSLL-.........P-A..-K-N..S-----L-TM------TSG--NSD........ 391
F1.AR.02.ARE933       --H--E--Q----------.--S-.HL..-Q--YAT...-DI--DI-K-----V-GMQ--T--ERVKA--KPL-.........P-A..-K-N..S-----L-TM------TSG--D-D........ 391
F1.BR.11.DEMF111BR037 --H--V-IQ----------.--PL.----YAT...-DI--DI-K-----GGT--K--EKVKKE--ESH-.........L----------H..PHA-----TM------T-K--N-RY........ 405
F1.CY.08.CY222        --H--E--Q----------.--S-.HM.-----Y-T...-AI--DI-K-----GTQ--K-MERVRE---SY-.........-T-N..S-----L-T--------TSE--NSN........ 392
F1.ES.11.VA0053_nfl   --HF-Q--Q----------.--S-.HM.-----Y-T...-EI--DI-K-----KTQ--H--S-VKK-----H-.........P-..S-----L-VTM------T-K--N-RY........ 396
F1.FR.04.LA22LeRe     --H--E--Q----------.--S-.----QT-YAT...-EI--DI-K-----V-GQQ--K--E-VKVE---SH-.........-P-R..-Q-N..S-----L-T--------TSG--N-G........ 394
F1.RO.03.LA20UcL      --H--E--Q----------.--SV.----YAT...-DRI--DI-R-----V--G-S---E--V-QRVIA--ASH-.........P-A..KLN..S-----L-TM------TSE--KTNI........ 412
F1.RU.08.D88_845      --H-KEAIQ----------.--GV.QL.A--A-LYAT...-DI--DI-K-----V-GS--EE----VNKG--DHV.........VGNI.......-A.SPPP----I-T--M-------TSE--KTNI........V 410
F2.CM.10.DEMF210CM007 -----E--A-----LG---.-S-.H.--QV-YAT...-EV--DI-K-F-I-NGTQ-K--TK-KE-F--QNLKI.........Q-I..-N-T-S..P-A--L-T--------DTD--ISQSE..KA......... 399
F2.CM.11.DEURF11CM026 ------R-I---------.--S-.H..--QV-YAT...-DI--DI-K-Y-V-NKTL--D--QKV-KEF--KNKSN.........L-G..-T-N..T-A--L-VT-------T-X--MANMX..DN....DXM.......X 406
G.CD.03.LA23LiEd      ------S---S-----.--S-.-F..-Q--YAT...-EI--EI-K-Y-TVNKT--RDM-QKVKT--GKL-.........-N-T..S-A--L-VT-L---Q---TSG--MSL........S-E........ 381
G.CM.08.789_10        ---F-K----Y----S----.-RS-.SF..-Q-VLYAT...-DI--DI-K-----TVNQTL--E--QKVKGE--Q-H-.........PDR....N..-S-N..S-----L-VT-L---Q---TSG--DRPEN........ 390
G.CM.10.DEMG10CM008   -----ETI----------.--S-.TSM.-----QV-YAT...-AI--DI-E-----V-EK--YEI-QKVKIQ--GKV-R-VL.......KNANISFT.......N..-A-N..S-----L-T--------TSR--SPLNS........ 403
G.CN.08.GX_2084_08    -----ETIQ----------.--S-.----QT-YAT...-GI--DI-E-----V-SKE-ST--QRVEAQ--KKV-.........-KT.......N..VT-A..SPI----L-TM--I-------TSG--NSKL........T 403
G.CN.16.224GX         ---F-ETIQ----------.--S-.----QT-YAT...-DI---I-----V-SKE-ST--QRVEAQ--KKV-.........-KT.......N..VT-A..SPI----L-TM--I-------TSG--NSLL........-T 405
G.ES.14.EUR_0033      ---ETIT-T----------.--S-.HL.-----QT-YAT...-DI--DI-----E-AKMINNVTE-----AL-.........-KTDG..N..-T--.P----L-T--------TSE---SL........NS......S....S 404
G.GW.08.LA57LmNe      --K-ID-T----------.-RS-.H.-A--Q--YAT...-DI--DI-----G-D-A-MIGNVTK---G-I-.........-KTHN..Q..-T--.S-----L-T--------TSN---SK-.....NS......S........ 399
G.KE.09.DEMG09KE001   --Q-I----------G--.--Q--YAT...-DI--DI-----V-GTA-R-MMGVKA---EGI-.........-TSKT..N..-T-N..TPA--L-TA------T-SR--M-ELD........X 403
G.NG.12.12NG060409    -----K-I-----------.-RS-.--QE-YAT...-AV--DI------KVNKT--TE--QGVKEQ--KKV-.........-K-I..NV--N..S-A--L-T--I-R-------ESL-T........TN......... 412
H.CD.04.LA19KoSa      --I--KEP-----I-TG--.-S-.-Q--YAS...-DDI--DI-R------GER--E--DS-VKQ--KLL-.........-Q..-V-N..S-------I-T--------TS-----I-YV........NG....TNGT...VLNGTI 406
H.CF.02.LA25LeMi      ----F-E--Q---S-----.--QSV.H.-----Q-IY-T...-DI--DI-R------ERD-EKA-SKVVQQ--KKH-.........R..P----I-T--------TSK-----DA..NK....T-S....... 402
J.CD.03.LA26DiAn      -----DKP-N-T-------.G-.HM.---QVLYAT...-EIV--I-A------GKD--D--SKV-T-----Y-.........P-..N-D.SPP----I-------TS-----NS-NG..-S......... W 410
J.CM.04.04CMU11421    -L--F-ET-N-T-F-----.--S-.HM.---QVLFAT...ERVV-DI-R-Y--V-S.N--RV-GTV-A-----GY-.........-EST.......N-T..SP----L--M--T-------TSS---S-SS..KN....DTN......... 408
K.CD.97.97ZR_EQTB11   ------EA----------S----.-S-.H.-----YAT...-DI--DI------GGQ--K-VN-VKKE--GKH-.........--------Q..P-------QVTR-I--R----S--DT-DTVDD-EE........... 401
K.CM.96.96CM_MP535    ------ET-Q----------.--S-.HM.-K--Y-T...-DI--DI------GE--M--SRVKE--K-H-.........K-G..-T--.PPNP---L--M--A-----T-K--E-G........ 396
01_AE.CM.11.1156_26   --H--E-T----S----.-I-SE..--QV-YRT...-EI---I-K-Y-Q-NGT---E--QKV-R--.--H-.........K-Q..P------VTM--------T-D-----YI..NG....T-N......... 394
01_AE.CN.12.DE00112CN011 --H--E-------S----.-TS-.H.-Q--YRT...-DI--Y-Q-NGTR-----R-VVK--K-H-.........Q..PP----L-TM--H-R------T-K-----E..NE....TMK......... 409
01_AE.IR.10.10IR.THR48F --H--K--A-----S----.-T--.TM.-KV-YRT...-EIT-DI-K-----E--NGTE--X-----V-E--K-R-.........-KA..-K-Q..PH----L-VTM--H--R-------T-K----G-RTE.NNG........TI......... 407
01_AE.JP.11.DE00111JP003 --H--D------I--ST.-I.TS.----QVLYRT...-DI--DI-K-Y-T-NKT-----GEV-T-VKK--.-A-H-.........S--------QPQPP---L-I--I--H-----T-K-----ND........TM......... 422
01_AE.MM.14.fKSDU26   ---H-D-------I---ST-.-I.TS.T...--QV-YRT...-DI--DI-K-Y-T-NET--E--NLVSR--EGH-.........PG--.....N..-T----AM-H-I-R------T-G--WSKEH........ND....TGG......... 399
01_AE.PH.15.DE00115PH012 --H--E---------S----.-TGV.T..-QV-YRT...-DI--DI-K-Y-EVNGT---KEV-T-VKK--.-KGL-.........-----G..P----L--T--H-I-R-------I-K--C-D........ND......... 399
01_AE.SE.12.070SE     --H--E------SHK.K-TSK.-----QV-Y-...EGID-DI-I-----K-NGTE-TRV---VVN--A-H-.........P.......-G..PP----L-TM-H-I-R-------T-G--N-CI........ND....T......... 399
01_AE.TH.10.DE00110TH001 --HF-E--------S----.-TG.-H.----QV-YKT...-AI--DI-K-Y-E-NVT--YGV---V-G--K-H-.........-S..-V--.P----L--TM-I-R-------TSE---N-CIK..N-......S......... 391
01_AE.TH.11.40436v09_04 --H--K-------TS-V.T.-.--QV-YKT...-DI--DI-K-Y-E-NGT---KV---VTE--K-H-.........-N..-V--.P----L-T---H-R-------T-K---GG..NK....NE--......... 402
01_AE.TH.90.CM240     --H--K-------TS-V.T.-.--QV-YKT...-DI--DI-K-Y-E-NGT---KV---VTE--K-H-.........-N..-V--.P----L--T---N-CLG..NE....TMA......... 402
02_AG.CM.10.DE00210CM013 -L---KEP-K-------.--GV.HL.----QV-YKT...-DIT-DI-----T-K--QGV-TQ--.-OM--.........P-AT..-A..NH----L-T---------STSA-----LF..NG....TYS....TWH- 425
02_AG.DE.09.701114    -----KEP--------.--GV.HL.-QTLYAT...-DI--DI-----T-KET--M--QKVSEQ--.-K-T-.........-AT..S..-R-T--------I-------TSI-VDN-FN..TS....NES......SSG 417
02_AG.KR.12.12MHI11_10746 -----TNP-G--------.--SA.H.-QT-YAT...-DI--Y--K------TD-T-A-QEV-RQ--QTY-.........-T--T..-NA-----L-T--------TSI-VDN-FN..TS....NES......SSG 410
02_AG.KR.12.12MHR9_   -----KP-Q--------.--SV.----QT-YAT...-EI--DI-----TV-ETE-T--RKV-IQ--.--H-.........K-N..-A..NPL----L-T--------TS-----D..NS....TWDN.NSTW.-NN 422
02_AG.LR.x.POC44951   -----DKP-R--------.--SV.----QT-YAT...-DI--DI-Y--V--TE--T--QKV-TK--.-K--.........G-NH..-T-N..-T-A--L-T--------TSI----D..VN......T......I 394
02_AG.NG.12.12NG060304 -----SKP-R------.--TSW.H.-Q--Y-T...-EV---S-EVN-TE--R-V-LQ--.-KY-.........Q-E..-I..-K-Q.PPPP---L--MM-----A-------TSG---A-RI..NG....TVNH.......TE- 404
02_AG.NG.x.IBNG       -----ANP-K-------.--GV.H-.---QV-YAT...-DI--V-KTE--K--H-KV-TQ--KTY-.........K..-A..NPL---V--T--------TSK-----NS..NG....ST.........A 397
02_AG.PK.15.PK032     -----TEPIK-I------.--S-.H.-QT-YAT...-EI---I---Y--V--NN-TG--SKV-TQ--.-KH-.........QTNT..-N-A..SH----L-VT------T-K-----NS..NG....TWNS...TG.KDN 415
02_AG.SE.14.098GN     ------EP-Q-T--S----.-TSV.--QT-YA-...-EI--DI------VDATQ--K--QKVVGQ--.-NH-.........-VS..N..-T-A..S-----L-T--------TSA----N-......-N........N 395
```

Env

HIV-1/SIVcpz Proteins

glycosylation NNT
glycosylation NCT
V3 tip
V3
glycosylation NTS
glycosylation NNT
glycosylation NKT
CD4 binding
glycosylation NST
glycosylation NST
glycosylation NST
V4

B.FR.83.HXB2	IIVQLNTSVEINCTRPNNN.TRKRI.RIQRGPGRAFVTI.....GK.IGNMRQAHCNISRAKWNNTLKQIASKL..REQFG.........NNK.....T..IIFK..QSSGGDPEIVTHSFNCGGE.FFYCNSTQLFNSTWFN.....ST.......WST.......EGS	405
02_AG.SN.13.9580	--I--DKP-R----G----SV.H-----TWYGT---DI---I-----V-KKE--A--QKVVE--.GQY------------K-------G----V-N-----------------TLG----NWEWHL---------------DN	402
03_AB.RU.97.KAL153_2	-----KEP-----------G-.H-----YAT-----DIT---I------IT-------VI---K--------G----V-N-----------M--------T-K----NG----TE----------E	389
04_cpx.CY.94.94CY032_3	-----AKA-K------G----SV.H-----LTWYAT---EI--DI----GND--D---V-SEE--KRL---P------K-A.PPV--L--I---V-KE--QSH---------T-P----HMQ---NG----TNI----T	406

[Large multiple sequence alignment continues — approximately 100 HIV-1 and SIVcpz Env protein sequences with residue positions numbered on the right margin]

O.CM.96.LA51YBF35	401
O.CM.96.LA52YBF39	393
O.CM.99.99CMU4122	406
O.CM.x.pCMO2_3	407
O.ES.01.Read25_HIV_GroupO	382
O.FR.06.RBF206	402
O.GA.10.10Gab1190	408
O.GA.11.11Gab6352	415
O.SN.99.99SE_MP1300	404
O.US.10.LTNP	400
N.CM.15.S4858	391
N.FR.11.N1_FR_2011	390
P.CM.06.U14788	423
P.FR.09.RBF168	413
CPZ.TZ.06.TAN5	402
CPZ.US.85.US_Marilyn	404
GOR.CM.12.SIVgor_BQID2	418
GOR.CM.13.SIVgor_BPID15	423

HIV-1 Proteins

Env

```
                                    V4                           V5
                                                glycosylation NIT  glycosylation NES                      fusion peptide
                 glycosylation NNT    CD4 binding  glycosylation NIT  CD4 binding   CD4 binding           gp120 end gp41 start
B.FR.83.HXB2            NNTEGSDTITLPCRIKQIINMWQKVGKAMYAPPISGQIRCSSNITGLLLTRDGGNS......NNESEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRRVV.....QREKRA.VG..IGALFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIE 560
A1.CD.02.LA01AlPr      SLVNSTGN------------------R--Q------P-I-K-E--------I---IN.....N-SMN-T---------------------------Q------------E------.F--V--------------I--A-----F-----S------ 563
A1.CM.08.886_24        ...TENA------------------L--R--Q------P-I-N-T-----------A......SE-T-----N-K------------------R------.E------.--V-I-------I--A-------------S------ 547
A1.ES.15.100_117       QESNST-N---------------L-RT-Q-I----K-V-Q-I-------KD..NST-NT---------I--------------S-----R-R-----.G------.---V-----X--------I-----------S---Q-- 556
A1.KE.11.DEMA111KE002  PD-VTNG------K---------Q--R--Q-------MI------A.....D-RT-T----------------V-----R-R-----.E------.------------------I------------S------ 560
A1.PK.15.PK034         ...-GSN------------------E-L--N-T---------N...N.TD-T--T------------------------------.L--V-I-------------S------ 552
A1.RW.11.DEMA111RW002  SS-GLNG----Q-K--------RT-R--------P-T-M-N-S---------G--SN-T---------------------SW-R-----.E-G----------V------S---K-- 568
A1.SE.13.078SE         .SST-GQ---------K-----RT-Q------E-V---V----I-----PN.....STN---N-K-----N----Q-N--I------E.....E--V-.-L--I-------I------------S------ 548
A1.UG.11.DEMA110UG009  ...DSTEN-------------R--Q------E-I--D-----------V-----ST-ETQ---I----K---N------I------E-G----------V------S------ 551
A2.CM.01.01CM_1445MV   SER-SN-N---Q-K----Q--R--R------A-V-K-I---MI----KR.....NDSS-T-------------------I-------E.....-L--A----------S------ 559
A3.SN.01.DDI579        -VQNSTYN-------R--VH---R--Q------Q-V-N-T-------DN.....SSTN-T---N-K----------------------E.....-L--A----------S------ 558
A4.CD.97.97CD_KCC2     IGSREN--M------------R-EQ------K-N-T-V----------TN.....ETR-N-T---I--E---------------E..G------.-M--F-------I-------A-----S------ 559
A6.BY.13.PV85          SS---NKN--IS-------------R--Q------R-V-K-E---------GN.....STQGN-T-------------I-I---T-R-----.E------I.-L--A----------S---Q-- 561
A6.IT.02.60000         STNSAEEN----------------R--Q-I----A-X-E-X----------GG.....NSSN-T---------------I---I---T-----------E.-L--A----------S------ 564
A6.RU.11.11RU6950      .TNNTE-NV----------------R--Q------R-S-K-V--------I----GG.....N-SNN-T---I--N---------Q--I-I---T-----.E------I.-L--A----------S---Q-- 560
A6.UA.12.DEMA112UA024  DS-QVDV------------------R--Q------G-V--E-------------AN.....STN-T---I---P--------S----I--E.....IQV--I---P----------S------ 567
B.AR.14.DEMB14AR003    TENSTNFN--I--K----V-R--E----------R-I-Q-T-------------E.....---I-------------------------M---S--------V------S------ 557
B.BR.10.10BR_MG035     I-NNDTTN------------RR-E--T------K-V--V--------II---KVN..AS--DT-----A-------------X-X------------.T--L--V-------V------S------ 559
B.CU.14.14CU005        KEIG--RE----Q------Q--E--R----------------V--------I----V-------T----------------R--------.I.F--V-------V--------S---N--Q 558
B.DE.13.366396         TG-GN-TN------------F----------K----------PN.....KT-NDT---T---N------------------D.L-.-L----I--------V------S------ 538
B.DE.13.947915         T-NTDNTN----Q-------------G--R------R-L-K---------NDT-NT---T-----N-----Q-K---I----------.T.-F--M-------L---------- 564
B.DO.11.DEMB11DR001    TAN-NDT---Q-------------G------R--K-V-D----------TG...DTL-----N-K----R---------------R----.TL------VA---------------- 562
B.ES.15.100600         A-SSEAK---------------G------Q------R---K-V-D--------I---DLN-------TG-----------------------N.L.TL--V---------------K--D 562
B.FR.11.DEMB11FR001    -TGGNDTI--Q-------------G------R------------I----DLN.....TQN-T--------------------------E----------S------ 544
B.JP.12.DEMB12JP001    DTMPNNN------------V---------N--A-L----------V----T.....T-TTT-T------------------------------LM------------ 567
B.PH.15.DEMB15PH003    TEESN-TV----------------E------R-R-N-T----------G.....STN-T---Q--NI-----------.M--W----A-------V---Q---- 553
B.RU.11.11RU21n        -ELNNTEI----------------A----------A.....--DT-T----T--------------------------I----------------------- 546
B.SE.15.102SE          TEGPNNA--I-------V--L--E----------R-N-S----------D----T---N-K------R------I--I--T-R----.T.-L--M-------------------- 550
B.TH.10.DEMB10TH002    PEKNNTKN----------------------A-K-V------------GG.....NNT-GN-T-------------------------.A.-L--M---I-----V------S------ 570
B.US.16.2609           L-GINNTN---------------------N--K-V-Q-T---II---S-N-GI.....G---N-T---A---------------R----I-.T.-L--M-I------V---------S---K-- 557
C.BR.11.DEMC11BR035    ...NNTEN--I-----------E--R---S--K-P-T-R---------TDGG..GEKD-NT-T----------------E-K---I-----.E------A----------S------ 565
C.CN.10.YNFL19         .....NA--I---------------E--R------N--A-N-T-K-------V---EG......T-ATN-T----------E-K--L---E-------A----------S------ 565
C.DE.10.622166         ...NDTTS----A--FV-----E--R------A-N-T-K-----I--V--V.......G--T---------N-K---------E-K--L-------I-------V-------S------ 552
C.ES.15.100_118        ...TEGG---I--K---F-----E--R------E-M-T-N----------SE.....G-NT----------E-K--L---E--------A-------------S------ 548
C.ET.08.ET104          S-IPDN-------K---YL----Q--R------A-R-S-N----------KG.....N-KQN-T---I-----N---------E-K--L--G-------A--M------I--------S------ 547
C.IN.15.NIRT008        .....NS---I--S---V----G--R------A-N-T-K--------V---TN.....STT-V----------------E-K--L--G-E--------I-------------S------ 539
C.MW.09.703010256_CH256.w96  ....KT-GN-----K------------G--R------A-N-T-T----------V.....T-DT-----------------E-K--I-------IA------------------- 550
C.NG.10.10NG020523     TVINNT-P--IQ--------X----Q-------E-N-T-E-S---------M---QN.....EEV-KT-T---------N------E-K---I---S--M----.A.V--M-----H--------S------ 568
C.NP.11.11NP016        .....NA---I---------V----EI-R------A-N-T-R---------V---KK......ANNT-DT-T-----------------E-K---I-----EK----------------------S---M-- 560
C.PK.14.DEMC14PK009    STKNS-SV--I-----------V--E----------A-N-T-N-T---II-V--E.....NNT-NT-T-----------------E-K-V-I--T---------EK------F----F--------------S------ 558
C.SE.15.100ET          .VNEAN-----N-FL-----G--R------E-E-T-K----------V.....RDKGN--T--A---------------ELK-----I-----------F---F--------------S------ 542
C.TZ.08.707010457_CH457.w8  ..DN-SN--I--K---------G--R------A-N-T-T----I------V-----SDSDP-T--T---------------EV----I--G-----Q.-M--H---------------S------ 557
C.US.14.M0118v1d14_5M04_C5  ..NSTSN---Q-K---------G--R------A-N-T-R------V---TDNS.TNKSDPE-T-------------E-K--I--E-R------------I------A----V------S------ 566
C.ZA.13.DEMC13ZA152    GTEGSNA-----------------G--Q------A-N-T-K-----------GNGN.ETES-RT---A-------------EVQ--------G---------------------V------------ 563
C.ZM.11.DEMC11ZM006    TSNNTNS---IQ-----------R--Q------A-N-T-K--------------AK.....KNESDPP-T-----------------EVK---L----Q---Q------------------S------ 560
D.BR.10.10BR_RJ095     STWNQTNVVKI--------V-L--T--T-------RE-L-----------I-----KN.....SST-----------------Q--I-I--G---------R--I.-L-VM---------- 554
D.CD.03.LA17MuBo       SIESNNTI--I------------------A-L-N-T--------K---KN.....-SNN-T-------------T-----------------I.-L--M---------- 572
D.CM.10.DEMD10CM009    TR-N-TGDTII-------------G------K--K--TV---I-V-EV-N.....SSPEAT-------------------------I--------.-L--M---------- 548
D.CY.06.CY163          ----LNG---I--T------I-C-G--V--T--PA-I-Q-K----ITG.....S-CN-T------------------I---R----.-L--M------IA---A----V---------- 561
D.KE.11.DEMD11KE003    ...SEGN-------------G--R------A-N-T-R-----------QEG.....---IN-T-----N-K---------R-I-I---------------I.-L--M----------------------- 554
D.SE.12.077UG          TTKNRGGN-PI-W--------GKGIRR-----SSAS--C*--TVP----AG---KD.....A-DS-T---E--N-----SN-----T----I-I--------L---------V---------- 563
D.TZ.04.CO6405V4       .PNNTDPI---Q-------------G--R-I---A-L-K---------AD.....NSSHN-T----N-S----------------R--------.E------.-L--M--------------V---------S----Q 559
D.UG.10.DEMD10UG004    -Y--ENK----------------RI----------E-L-K----------E---SVE.....-SQ-TV--V--E--N--------R-K-I-L--R-------EKQ---I----------A--------V---------- 551
D.UG.11.DEMD11UG003    STEKNNG----------------G------E-E-K-I------------N.....-ATN-T---------T-----I-----G---------.E------I.-L--M------------------- 561
D.YE.02.02YE516        TGNNTN----I----------G--R------A-I-K----------SN.....NT-ST-T-------------------R--------.E------I.-L--M---------------------- 550
F1.AO.06.AO_06_ANG32   ...TSNR----------------V---G--Q------A-VA-K-T-N----------SHT-T-----N-K--------EL-------R---.ARG------.M---I--A---G----F------ 544
F1.AR.02.ARE933        ...DNG--------------V----G--R-NNH--HA-N-T-G--------------TGNCTDGGTG-CT-T----N-K------E-------R-Q----.M--M---I--A---G----F------ 549
F1.BR.11.DEMF111BR037  ...NDTI-LP--Q------------R----TS--A-N-T-S----------TD--NT-------A--------------E-I-------Q---.K--E.----V--T-------------------- 559
F1.CY.08.CY222         ...SNNG----------K------E--Q------A-K-T-N-S---------Q-NK..SKD-G------R-NIK---------E-I-------Q------K----V-T------------- 548
F1.ES.11.VA0053_nfl    ...TQN---I-------V--L--E--R------A-N-T-H------------QR.....NET-TT----------------------E---------K------------S---------- 550
F1.FR.04.LA22LeRe      ...TNE------------Q-----G--R------A--A-N-T-H-----------Q-.....T-GSI-T-------------------------C---------E-Q--I--A----Q----------------AT-------- 546
F1.RO.03.LA20DuCl      .....NN----------VK--G--Q---A---A-T-K-N-S-----------QDNK.NGT-RT-T---A--N-----------N-K---------R------E-Q---A.-L--M--------------------K-- 567
F1.RU.08.D88_845       -GSSLNG------------------V---G--Q------A-V--V-N-S--------XN.....DTRTN-T--A--N------------------------H.-E-Q-.L-.M------------------I----- 551
F2.CM.10.DEMF210CM007  .GNNGS-F-----R--------V---G--R------A-V---N-S----------ED.....-SSTDTL---------A----------------------N--------V-----T------------I--A--------- 552
F2.CM.11.DEURF11CM026  XMSDNKTE-VI--------------R--R------A-T-Q-I--------------T.....TN-TL----K---------------R--------------.M--M-I------V-TA---T---------K-- 559
G.CD.03.LA23LiEd       ...N-TTN------K----VR--R--Q------A--T-I---------------GG-SSN-T---------------------------K------R--R-----.G------.L--V------------V-------- 535
G.CM.08.789_10         ...NTAEN---G----------VR--R--QG------A-N-T-I------------N.....ITE-------R----------------K------R--R-----------------V---T----------I-------V--------- 541
G.CM.10.DEMG10CM008    ...NSTGN--------------VR--R--Q------A-N-T-R-------------N.....TGN-T-------------I--K--I---R--R------------.T.M------------V---------S------ 551
G.CN.08.GX_2084_08     -TNNT--------------VR--R--Q------A-N-T-N------------A--N.....SSTN-T----------I--K--I---R--R------------E------.-L--V------------V-------- 556
G.CN.16.224GX          DT-LENTNL-----K----VR--R--Q------A-N-T-N--------K-SEGN.....SSTN-T-----------I--K--I---R--R------------.G------.L--V------------V-------- 560
G.ES.14.EUR_0033       .NN-EV--------------VR--R--Q------A-N-T-N------------N.....ST-GT-----A-------------------K--------------.G------.L--VL------------------- 557
G.GW.08.LA57LmNe       ...DNET------------VR--R--Q------A-N-T-K-----I-I----QG.....NDT-RT-T---A--------------------K--I--I--R--R-----.G------.L--VL------------------- 553
G.KE.09.DEMG09KE001    ...S-NE------------VR--R--Q------E-K-T-N------------EN.....-VS-GT--------------IR-K--------I--I---R--R----.G------.L--VL------------------- 556
G.NG.12.12NG060409     S--SS---N-I--------VR--R--Q------A-N-S-T------------GN.....DTSGN-T--------------------------K----------E------.V--VLF------------------- 568
H.CD.04.LA19KoSa       TG-GLNG----------------R--Q------R-D-T-E------I--S-YV-T.....TDT-G-VTL--T----------------I--SR-R-----.E------.-L--------A----------S---K-- 563
H.CF.02.LA25LeMi       ...TNNTM---------K---R--Q------D-T-I----I--N-E--M.....SA-NFT---V--E----------------R--------.M--F------L----------S---K-- 554
J.CD.03.LA26DiAn       --NATN---I------VR--R--Q-I-----A-N-T-V------------I----QGGN.ETNKT-A-T---T-------------EV---------A-------------------I------------S---N-- 569
J.CM.04.04CMU11421     ...STANYTII--K-R--VR--------Q-I-----E-N-T---------DKN.....STS--ST-----T---K----------------L----V----T------------I------------S---K-- 564
K.CD.97.97ZR_EQTB11    ....-EDT---I-----------------Q-I----TA-N-T-R------MI-------D.....NNTRTE-T----------Q------I--R-R-------------------V--F---------- 555
K.CM.96.96CM_MP535     ....ENG------------------------N--A-S-N-------ML--------N.....THN-T---------------------Q------I----R--R-------------V--F------------------ 546
01_AE.CM.11.1156_26    ..STN-SE-I--------V-----RA-Q------R-V-N-T------I-----FV-A.....TTDVT---T--N--N---------Q----I----TSR-----.E-G------MIF-----------I---------S------ 546
01_AE.CN.12.DE00112CN011  .GDH-NG--------------RA-Q------V-N-V------I-------T.....SSSN-T-------NLK---KN------------Q------------.D------.-MIF----- 563
01_AE.IR.10.10IR.THR48F   ...-NG--I---N---------ET-Q------V--I-N-L------I--S-T-GKNN.GTGE-GT-T---A--N-K------------E-I------------.-R-R----.-MIF----- 564
01_AE.JP.11.DE00111JP003  .QYN-------------------R--Q------T-N-N-V------I------GT.....STN--Q--NIK--K---------Q---Q--I---------.E-I------.-R-R----.-MIF----- 557
01_AE.MM.14.fKSDU26    .RHNG--I---------VR--G--QT------E-T-N-T-----I---------N.....-G-X-T------XXK------------Q------------.E-I------.-R-R----.-MIF----- 574
01_AE.PH.15.DE00115PH012  ..GCNG---I-----------G--Q------VR-R-N-V------I-------TN.....SSM-ETN-T---E--NIK-----N---------Q------.-R-R----.-MIF----- 554
01_AE.SE.12.070SE      ..TCNN---I--K---------G--Q------G--N-T-L------I------T.....TGN-T--------NIK-----------Q------------.E-I------.-L-MIF-----------I---------S---K--Q 554
01_AE.TH.10.DE00110TH001  ..RDNG---I--K---------G--R------G-K-V-K-L------I------D.....STENN-T-------NIK--N------------Q------.EKQ-----.-L-MIF-----------I---------S------ 561
01_AE.TH.11.40436v09_04  .MCNSNV---K---------G--R------O-K-N-V-----------VDSN.ETS--T---T---NIK---N----------Q------------.E------.-MIF----- 547
01_AE.TH.90.CM240      ...GCN-------I----------G-Q------R-N-V-----------VN.....-TDN-T---------NIK--------------Q------------.E------.-MIF-----------I---------S------ 554
02_AG.CM.10.DE00210CM013  SWNSSAEN---Q---------FV----Q------R-M--E----------K.....NTSAN-T------------------I-----------Q-I.-L-VMF----- 581
02_AG.DE.09.701114     A-NDT-KN---Q--------R--R--Q------K-E--R----------K.....SSGN-T--A-----------------------H-R-Q------.G------.L--V-------------I-------- 572
02_AG.KR.12.12MHII_10746  HTGWSNESL----N--V--------Q-I-----N-V----------F-------RNG.GDSGGDN-T--------------V--R-------------------V-------------- 563
02_AG.KR.12.12MHR9     S--NS-G--FQ-K-------V----Q------L-Q-V--I-------------N.....N-SNN-T-------------------------Q---------.-L--VL------------------- 566
02_AG.LR.x.POC44951    -I--SN-------------------Q------K-I------------D-I.....-I--N-T------------------------EK---I----S.-L-.M--F--------------I-------- 549
02_AG.NG.12.12NG060304  -TNHTNN--I-----------Q--R--R------R-I---I-----------E.....A-N-T---------------------------------.M--LF------ 559
02_AG.NG.x.IBNG        -H-GSN---Q-----------Q------Q------Q-I--D-------------N.....STN-T---------------------------------.E------.-R-I-----------V------S------ 551
02_AG.PK.15.PK032      STGT-N--------K---------R--Q-I----P-V-Q-E------------VN.....STTN-T--------------------V--NR-R----.E------I.-M--F---------I-----H-------S---K-- 570
02_AG.SE.14.098GN      EIN-TNK----Q---------R--Q------P-A-K-E---------------TT-T-------------------SR-----.E------.V--MA-V------------I---------S---K-- 550
```

V4 · glycosylation NNT · CD4 binding · glycosylation NIT · V5 · glycosylation NES · CD4 binding · CD4 binding · fusion peptide · gp120 end | gp41 start

```
                              V4                                      glycosylation NIT        V5
                  glycosylation NNT        CD4 binding          glycosylation NIT  glycosylation NES                    fusion peptide
                                                              CD4 binding        CD4 binding          gp120 end gp41 start
B.FR.83.HXB2          NNTEGSDTITLPCRIKQIINMWQKVGKAMYAPPISGQIRCSSNITGLLLTRDGGNS......NNESEIFRPGGGDMRDNWRSELYKYKVVKIEPLGVAPTKAKRRVV.....QREKRA.VG..IGALFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIE 560
02_AG.SN.13.9580     -SSISYARV--Q-K----R--Q-----P-EL--D--------KI......ILMN-T--------------------------F------H------------------------M----------------I------------S--K---- 557
03_AB.RU.97.KAL153_2 L-NTEG-IV-----------E---R-----A-----------Q......-S-VT--T----T------------R-----------------------------------------I--------------S--K---- 544
04_cpx.CY.94.CY032_3 STDSTNS---Q--L--FVR--E--Q---S--A-S-N--I---------.-T-NT-----------------I---N--R-R----------*--M----------------------------------------- 558
05_DF.BE.x.VI310     F-NDSDKN-I----K---VR--E--Q-----A--A-N-T--I--A----GN.DSS-DT----------------E-Q---I----R--Q---K----M---A----------IA-----------------Q---- 553
06_cpx.AU.96.BFP90   RGNDTNT------K----VR--E--Q---A-N-T-V----II-----N---.E--V--T--------------------G-----------L--M-----T----------I----V---------------S---- 563
07_BC.CN.98.98CN009  ....NS--SI--I-----------E--R---E-N-T-K-----V---TE---.D-T-T-----------------E-K----A----------G-----------------I----V----------------S---- 556
08_BC.CN.97.97CNGX_6F ....NS-SI--I-----------E--R---E-N-T-K-----V---RT---.EP-NT-----------N--N--E-K----E------------L--V----------------------------------S---- 561
```

(Full sequence alignment matrix continues for all listed isolates; names listed below with their accession identifiers and C-terminal residue counts)

```
09_cpx.GH.96.96GH2911 546
10_CD.TZ.96.96TZ_BF061 544
11_cpx.CM.95.95CM_1816 569
12_BF.AR.99.ARMA159 546
13_cpx.CM.96.96CM_1849 564
14_BG.ES.05.X1870 562
15_01B.TH.99.99TH_MU2079 551
16_A2D.KR.97.97KR004 559
17_BF.AR.99.ARMA038 552
18_cpx.CU.99.CU76 561
19_cpx.CU.99.CU7 506
20_BG.CU.99.Cu103 546
21_A2D.KE.99.KER2003 556
22_01A1.CM.01.01CM_0001BBY 545
23_BG.CU.03.CB118 552
24_BG.ES.08.X2456_2 557
25_cpx.CM.02.1918LE 554
26_A5U.CD.02.02CD_MBTB047 555
27_cpx.FR.04.04CD_FR_KZS 560
28_BF.BR.99.BREPM12609 572
29_BF.BR.01.BREPM16704 574
31_BC.BR.04.04BR142 546
32_06A6.EE.01.EE0369 554
33_01B.ID.07.JKT189_C 556
34_01B.TH.99.OUR1969P 553
35_AD.AF.07.169H 551
36_cpx.CM.00.00CMNYU830 553
37_cpx.CM.00.00CMNYU926 552
38_BF1.UY.03.UY03_3389 548
39_BF.BR.04.04BRRJ179 548
40_BF.BR.05.05BRRJ055 555
41_CD.TZ.05.CO6577V5 554
42_BF.LU.03.luBF_01_03 547
43_02G.SA.03.J11223 543
44_BF.CL.00.CH80 548
45_cpx.FR.04.04FR_AUK 549
46_BF.BR.07.07BR_FPS625 539
47_BF.ES.08.P1942 552
48_01B.MY.07.07MYKT021 564
49_cpx.GM.03.N26677 559
50_A1D.GB.10.12792 560
51_01B.SG.11.11SG_HM021 555
52_01B.MY.03.03MYKL018_1 564
53_01B.MY.11.11FIR164 552
54_01B.MY.09.09MYSB023 561
55_01B.CN.10.HNCS102056 563
56_cpx.FR.10.URF5_patient_A 549
57_BC.CN.09.09YNLX19sg 567
58_01B.MY.09.09MYPR37 561
59_01B.CN.09.09LNA423 556
60_BC.IT.11.BAV499 556
61_BC.CN.10.JL100010 545
62_BC.CN.10.YNFL13 552
63_02A.RU.10.10RU6637 552
64_BC.CN.09.YNFL31 556
65_cpx.CN.10.YNFL01 566
67_01B.CN.11.ANHUI_HF115 544
68_01B.CN.11.ANHUI_WH73 545
69_01B.JP.05.05JPMYC113SP420 570
70_BF1.BR.10.10BR_PE004 554
71_BF1.BR.10.10BR_PE008 560
72_BF1.BR.10.10BR_MG002 551
73_BG.DE.01.9196_01 553
74_01B.MY.10.10MYPR268 556
77_cpx.MY.14.14MYNBB090 557
78_cpx.CN.13.YNTC19 571
79_0107.CN.15.SX15DT013 561
82_cpx.MM.13.mSSDU12 549
83_cpx.MM.13.mSSDU94 559
85_BC.CN.14.14CN_SCYB2 575
86_BC.CN.13.15YNHS18 541
87_cpx.CN.12.DH32 554
88_BC.CN.05.05YNRL25sg 590
90_BF1.BR.11.BRG06043 546
O.CM.96.LA51YBF35 559
O.CM.96.LA52YBF39 555
O.CM.99.99CMU4122 567
O.CM.x.pCMO2_3 563
O.ES.01.Read25_HIV_Group0 536
O.FR.06.RBF206 565
O.GA.10.10Gab1190 565
O.GA.11.11Gab6352 575
O.SN.99.99SE_MP1300 562
O.US.10.LTNP 561
N.CM.15.S4858 543
N.FR.11.N1_FR_2011 540
P.CM.06.U14788 581
P.FR.09.RBF168 573
CPZ.TZ.06.TAN5 554
CPZ.US.85.US_Marilyn 532
GOR.CM.12.SIVgor_BQID2 578
GOR.CM.13.SIVgor_BPID15 583
```

HIV-1 Proteins

Env

glycosylation NAS glycosylation NHT
immunodominant region   glycosylation NKS
glycosylation NYT
gp41 cytoplasmic tail start
transmembrane domain

```
B.FR.83.HXB2                    AQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICT.TAVPWNAS.WS.NK............SLEQIWNHTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIKLFIMIVGGLVGLRIVFAVLSIVNRVRQGYSP  714
A1.CD.02.LA01AlPr                                                                                                                                                                          717
A1.CM.08.886_24                                                                                                                                                                            701
A1.ES.15.100_117                                                                                                                                                                           711
A1.KE.11.DEMA111KE002                                                                                                                                                                      714
A1.PK.15.PK034                                                                                                                                                                             706
A1.RW.11.DEMA111RW002                                                                                                                                                                      722
A1.SE.13.078SE                                                                                                                                                                             702
A1.UG.11.DEMA110UG009                                                                                                                                                                      705
A2.CM.01.01CM_1445MV                                                                                                                                                                       713
A3.SN.01.DDI579                                                                                                                                                                            712
A4.CD.97.97CD_KCC2                                                                                                                                                                         713
A6.BY.13.PV85                                                                                                                                                                              715
A6.IT.02.60000                                                                                                                                                                             718
A6.RU.11.11RU6950                                                                                                                                                                          714
A6.UA.12.DEMA112UA024                                                                                                                                                                      721
```

[HIV-1 Env protein sequence alignment — B.FR.83.HXB2 reference across top, with subtype representatives A1, A2, A3, A4, A6, B, C, D, F1, F2, G, H, J, K and CRF 01_AE, 02_AG sequences shown as differences from reference]

Env

HIV-1/SIVcpz Proteins

glycosylation NAS glycosylation NHT gp41 cytoplasmic tail start

immunodominant region glycosylation NKS glycosylation NYT transmembrane domain

```
                             AQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGIWGCSGKLICT.TAVPWNAS.WS.NK............SLEQIWNHTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELLELDKWASLWNWFNITNWLWYIKLFIMIVGGLVGLRIVFAVLSIVNRVRQGYSP
B.FR.83.HXB2                                                                                                                                                                                             714
```

HIV-1 Proteins

Env

```
                                                                                                                              Env end
                                                                                                                              gp41 end
                  glycosylation NGS                                          glycosylation NAT                  cytoplasmic tail end
                                                                                          *                                              *
B.FR.83.HXB2         .LSFQTHL.PT.PR..GPDRPEGIEEEGGERDRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELLGRR.......GWEALKYWWNLLQYW.SQELKNSAVSLLNATAIAV.....AEGTDRVIEVVQGACRAIRH.IPRRIRQG.LERILL*   856
A1.CD.02.LA01AlPr    .......I..N..E...L...R-------QGKI.......-S-F---A----------------------FV--AA-T----HSSLKGLRL--G---LG---L---GR---I--IN--DTI-V----GW-----IG-R----LLN...........F-A---   867
A1.CM.08.886_24      -----ILT..N..E...-V--G--G----P------------A--F---A---------------------FI--AA-T----HSSLKGLRL--G---HL---L---L---GR-------IN--DTI----GW--W---LG-RI------N............-A---   851
A1.ES.15.100_117     -----LI..S..G..-L--R--G----QGSS----D-F---A---------------------------FVS-A--T----HSSLKGLRL--G---LG---I---IR---I----FDTL--VI----GW-----IG-RTG-----N...........-A---   864
A1.KE.11.DEMA111KE002 -----T..S...DL---R-G--Q------------A--F-TVV-------------------------FI-AA-TL---HSSLKGLRL--G---LG---L-----L--I---IN--DNI----GW-----LG-RF-----L............-A---   864
A1.PK.15.PK034       ------A..S......-R-D--D-G----------A--F-TVV---------------------------FI-AA-TL---SSLKGLRL--G-----L--GR--T---I--FDTI--VI____GW-----EG-LN---------.........-A-AF-   856
A1.RW.11.DEMA111RW002 ----I-T..N.QG..E..-R--G--G--QG------AS--F----A----------------------SSLKGLRL--G---LG--T---I-----I--T--FDTI--T---GW-----I--RI----LN...T------F-A---   872
A1.SE.13.078SE       -----LT..N..E...-L--GR-------Q-T------S--F--A---------------------V--AA-T----HNSLKGLR#......#----L------IN-FDTI----GW-----IG--I---LN...........F----   844
A1.UG.11.DEMA110UG009 -----LT..S...N...-R-----QG------S--F--V-E---------------------FG--AA-T---SSLKGLRLV--G---L-----L----I-ID-F-T---GW-----IG-RIG--LN.V-T-------A---   855
A2.CM.01.01CM_1445MV  ----L-IPT..R..E...L---R-K-G--Q--T-V---S-F---A-------------------CIS-AA-T-T---HSSLKGLRL--G--GI----L---L---GR----N-VD-I-VT----GW-----IG-RI---LN...........F-A---   863
A3.SN.01.DDI579      -----LT..N....---R-----Q------S-F---A------------------------FT--AA-T------------QG--L--V-L---GR---S-I--------V--W-----IG-R-G--LL....--V----A---   851
A4.CD.97.97CD_KCC2   ------PI..S..E...-GR--G--QG-T-------S-F-E-A----------------V--S--FI-AA-T----HSSLKGLRL--G--L----V---G---X--IN--DTI----G----I--RI---LN...........-A---   863
A6.BY.13.PV85        .V----LT..H..E...-GR-N-------QG-GS-T--S-F---A----------------------IS-AA-A---T-SSLKGLRL--G--L----V---G---X--IN--DTI----G---I--I--RX---CN...........A-A-Q-   865
A6.IT.02.60000       -----LT..N..E...-GR-----QG-------S-F---A----------------------K-FVS-AA-T------------------IDTI----GW-----IG-RV-----L............F-A---   868
A6.RU.11.11RU6950    --L--LT..H..E...-GR-S-----QG-----S-F---A---------------------FI-AA-T---SSLKGLRLV--G---LG--G--G---IN-VDTI----GW-----IG-RV-------CN...........A-A-Q-   864
A6.UA.12.DEMA112UA024 .V-L--LT..H..E...-GRTK-----QG---T-S-F---A---------------------FI-AA-T---SSLKGLRL--G---LG---V--G---IN-IDTI-V---GW-----IG--I-------CN...........A-A-Q-   871
B.AR.14.DEMB14AR003  ...........A....--------T---------G-K---T-F------------------H------------------G-I------L-----I-----I----T----I--G-RRIL--FL...........F-A---   854
B.BR.10.DEMB_MG035   ...-L--L-..-........-------M----D--G--RP----FFV-V----W---S---------------A----I-----XS---------K-I---------------------I--R---L-.........-A---   856
B.CU.14.14CU005      ------RH..A.Q....-----G---------GG-E--D-F-I--V------------------S-L----------------S-----C------------------------V--------I--I--R---FL........F-L---   855
B.DE.13.366396       ------R-..A.Q....-------G---------GI--D-F----V----------I--T------------I-T--------------L--A---L---G-----------------G-RIL-G-L..........-A-Q-   835
B.DE.13.947915       ------H..A.Q....-------G---------RP--H-F--V---------F--I--T---S------------------------L--A---L--------------------------I--R----FL......-A---   861
B.DO.11.DEMB11DR001  ------R-..E.....-G-D---------S-AH-F-V-----F-----------------------------DTI---------D--V-V---------------I-A-RRL--FLN...........A-A-V-   859
B.ES.15.100600       ------R..A.Q....-------G----KS--T-GC--D-F-TI--V---N--R------------S-AA--------------------------------I--I---I------I......-V----A---   841
B.FR.11.DEMB11FR001  ------R-..A.Q....-------G----KS--T-GC--D-F-TI--V---N--R------------S-AA------------------------K-I-------------I--RI----I..-V-------A---   841
B.JP.12.DEMB12JP001  ------RF..A.....-------G-----S--S-F-P--------------------------------------I-----K-I----------------------I--I-RL---L-.......-A---   865
B.PH.15.DEMB15PH003  ------IRP..A.....-------G---------EP-A--F-T-F-T----F-----------------S--L-V--F--------------------RT--I-----------I--I--RIG--LN.......--S---   850
B.RU.11.11RU21n      ------R-..A.....-----------T-V----LF------------L-----AA-----L------------------I----------I-------T----G--------RTF--VL..........F-A---   843
B.SE.15.102SE        ------N-..........-----T----Q---SP--A--IF-V---F-FS--L---------A--T-----------------------G---------------I-F--I----G------------I-AI-RTG--L...T-------A-Q-   847
B.TH.10.DEMB10TH002  ------RF..A.....-------G---------GG-E--AS-F-T--V---D-------S------------I--A-V---------------T-----------I-A--V---------#-.A----A-A---   854
B.US.16.2609         ---L-IRP..A.....-----G---D----GG-E--AS-F-T--V---D--------R--T--S---S------------I----------------------T----------G----------I--L-RFG-G-L..........-A---   854
C.BR.11.DEMC11BR035  ------LI..S........-----G-------------G-F-----A--------------------AV-A------SSLKGLQR---I--LGS-V----GL--K-I-DTV--V----------------I---IW---CN..........F-AA-Q-   869
C.CN.10.YNFL19       ------LI..N.QG...-L-LGR------Q-KN--V---S-F---A----------------FI-VTA-A----NSLRGLQT--R---LGG-V----GL--K-I-DT-----------------I-AIX-R-W-.....T--G-D-P-VA-Q-   865
C.DE.10.622166       ------LI..N.Q...E--KLG-------Q-----S-F-P-A----------------------FI-A-L---SSLKGLQR---I--LGS-V----GL--R--IN-FDTI----------I-A-I-TI------N.........F-TA---   856
C.ES.15.100_118      ------LI..S........-LG------Q--N-----S-F---A----------------------S-AA-T---SSLKGLQR---I--LGS-V----GL----IN-FDTI----------I--IW--LCNI...........F-TA-Q-   853
C.ET.08.ET104        ------LI....AH-L--------AS-R-P-C--K-FF.VLLG----------H--I----A---TA---------GT--F.Y-PF---GK-E-K-VL---CYS-TP---------V-R-TGIM--L--K-W-DFQ.......-F--SFP-   837
C.IN.15.NIRT008      ------LT..-.....-E--SLGR------Q-----S-F---A----------------------FISVAA-A----HSSLRGLQK----I----I--DSI--T-------IL-AI-RIWT--N.........-A-IA-Q-   843
C.MW.09.703010256_CH256.w96 ------LI..N.Q...E--LGR------P--N--------T---T--------------------FI-VTA-V----SSFRGLQR-------LGS-V----GL--K-I-DTI--V----------I--FL-RI----Y...........F-AA-Q-   854
C.NG.10.10NG020523   ------LT..A.....-E--R--R--GD--Q--E--V---T-F---A---------------------FI-VTA-V----H---------R-----LG--A----GL--K-I-DTI--T----------IL--ILRF-----N...T-------F-AA---   865
C.NP.11.11NP016      ------LT..A.Q....-------G----Q-----S-F--A-E---N--------------------ISVTA-G------SS---------LGS-V----GL--K-I-DT---------------I------RI------N.........F-AA-Q-   859
C.PK.14.DEMC14PK009  .-P--IPT..N.-G..-A--G------DKNNSS---S-F---A-------------L----------LTA-A---SSLKGLQK-----LGS-I----GL-I-K--IN-FDT-----------I------NIY----N.V-------AS---   846
C.SE.15.100E1        ------LT..N....-LGR--G----Q-G-----S-F---T-----------------LA-A----SSIRGLQR---V--LGS-V----GL-I-K-IN-FDT-------------I--I-RIG--LN.........--VC----F-AA-Q-   862
C.TZ.08.707010457_CH457.w8 ------LT..N....-LGR--G--G--Q-N--T--S-F---A-------F----------FT--A-V----SSLKGLQR---T--LGS-V----GL--K-I--DTI--T----------I--LI--I--HN..V-T-------A-AA-Q-   846
C.US.14.M0118v1d14_5M04_C5 ------LT..N....-LGR-------Q---------S-F--F-E--N----------CI--AA-V----NSLKGLQR-----LKS-V----GL--K--IG--DT-------------I--LI--I--HN............F-AA---   870
C.ZA.13.DEMC13ZA152  ------LTPT..N....-LGR-------Q-G--V---S-F--V-I----------Q---FI-AV-T----SSLKGLQR---I--LGG-A-----GL--K-I-DTI--V---------I--FI-RF------CN.V-T-------F-AA-Q-   868
C.ZM.11.DEMC11ZM006  ------LT..S.S..-E--LGR------Q-G--KG----S-F---A--------H------FI--AA-AA---SSLRGLQR---I--LGS-I----GL--K-I-DTI--RT---------I--I-LNI-----LS...........F-AA-Q-   864
D.BR.10.DEMB_RJ095   ------L-..A.S........-----D-T-------LG-------T-FSS-----------N--------------------A-----L---------------R--T--L-----------------I---A-R-G---LN.V--------A-F---   851
D.CD.03.LA17MuBo     ------L-..A.....-------------DN--G--T--------N----------------N------R---V--A-T--------R--T--L---------------------------I---R---VL..........-A---   869
D.CM.10.DEMD10CM009  ------L-..A.....-------G--G-----------FS-------------N---------------I-AA-T-N-------R------------FD----------A-LR-F--L-........-A---   845
D.CY.06.CY163        .P--H-L-..A.S....-P-A--P--G--Q-------FS-----------------------N----------I-AA-T-G------------R---------R--IN--D-----------T---RT---VLR.........--A---   859
D.KE.11.DEMD11KE003  ------L-..A.....-------G-----------QG-------LS--------------RH--------------R---------I---------IN--DTI--VI---L--IGLRIG---IN............-A---   851
D.SE.12.077UG        ------L-..A.....-------G--------Q-----LS----------------------N--------R-------------V---------I-V-------------AI--F--L..........-A---   860
D.TZ.04.CO6405V4     ------L-..A.....-------------Q-------FS-----------------N----R--------------R---------I-----------------------I--R----VIN.V-V-------L---   857
D.UG.10.DEMD10UG004  ------L-..A.....-------QG----LG-LST-------N--I---------I-AL--L----------SR--L--------I--V-----I--FGT--V------------I--LRIG--LN...T-------A-AF---   848
D.UG.11.DEMD11UG003  ------L-..A.....-------Q----FFP--------N---Y--Q----I--A-T-D----------R-------------I--------I-F-T--V-----------I--I-R-G--LN...........--A---   858
D.YE.02.02YE516      ------L-L-..A.....-PGKGT-------FS--F-E---N--Y--Q--I--A-T-D---------R----------------I------I--F------------I--I-R-G--LN...........--A---   847
F1.AO.06.AO_06_ANG32 ------L--LI..S........-----GD--Q--K-V---S-F-P-V----------------N------RH----FI--AA-T-NRGL-S--------------R-----IN-DT--------------L-R-G--VL.........--L---   841
F1.AR.02.ARE933      ------L--LI..S........-----G----QGK----T-F--A----------------N------RH----FI--AA-T--DRGL--------------LLG--TL----RI--I-F-T--V-------------A-R-V--VCN.........--AF---   846
F1.BR.11.DEMF111BR037 ------L--LI..S........-----G----QG----V---S-F--A------------N------RH----FI--AA-A-NRGLS--------V---LG--T----C--R--I---T--------------A--ALRR-VT--LN.........--L---   856
F1.CY.08.CY222       ------L--LI..S........-----G-------FFS--V---------N------RH----FI--AA--L----------------I-L--I---I--DTI--V------LL--L--G--VLN.V--K-------A---   845
F1.ES.11.VA0053_nfl  ------I--LI..S........-----E----QG--K-V---S-F-S-V----------N------RH----FI--AA-A-NRGL---------------I-T--------G----I--DTI--V--------------A-R-V--VLN...........F-A---   847
F1.FR.04.LA22LeRe    ------L--LI..S........-----E----QGK----V-F-T-V-----------N------RH----FI--AA---DRGL--L-------------I--I-T--G-----------I--DT--VL.....-R-----IS-RIG--VLN.V-----------F-A---   843
F1.RO.03.LA20DuCl    ------L--LI..S........-----G----QG----V---S-F--A-----------N------RH----FI--AA-T-NRGLKG--------V-------------L-S-T----R---------------I--AL-R-G--VL.A--------A---   864
F1.RU.08.D88_845     ------L--L-..-...Q....-------------G----Q----S-F-TIV------------N------N-RLW--FI--AA-V--DRGLK--------------------G-----------------W----L-AL-R-G--VL........--L---   848
F2.CM.10.DEMF210CM007 ------L--L-..-...Q....-------------G----Q-----Q----G----S-F---A--------N------RH----FI-AA-T-NRGLE-----------K----L--A-H---------G------I--DT---V----------IL-GL-RVG--VL...........F-A---   848
F2.CM.11.DEURF11CM026 .V-L--LI..N.S..-EP----GE--G---QGK-----V---S-F--A------------N------RH----FI-LA--T-DRGL-E------------K----L--T-----G-------I--F--I-------------IL-GL-RVG--VL...........F-A---   857
G.CD.03.LA23LiEd     ------LT..HH.R...-E-----G----K-----S-F---A----------------------FT--AA-T----SSLKGLRL--G--L-----L--G-----IN-IDTV----NW------G--RI---FLN...........F----V-   839
G.CM.08.789_10       ------LT..HH.Q...-E-----R--G--QG-G----S-F---A---------------------H--FV--AA-T----SSLKGLRL--S--HL----L---G-----IN--DTI----DW-----A-RVG--FLN............-A---   845
G.CM.10.DEMG10CM008  ------LT..HQ.Q...-E-----R--G--Q---V---S-F---A---------N---WC--------FI--AA-T--I---SSLQGL---V------L-----G-----IN--DTI----NW-----AHR----CN............-A-V-   848
G.CN.08.GX_2084_08   ------T.HH.Q...-E-----R--G--Q-K--V---S-F---A-----------------FI--AA-TI---K-SSLQGLRLV-G--L----L---G-----V-----IN--DTI----GNW-----GTYR------CN..........----I-   864
G.ES.14.EUR_0033     ------PS..HH.Q...-E-----R--------S--FFT-A----------------------FI--AA-T--I------------L---S---L-----G-----IN--T----NW--SA----RRIG--FLN............-A---   861
G.GW.08.LA57LmNe     ------LT..HH.Q...-E-G--GR-----Q-K--T--S-F---A----------C--------FI-LAA-T---GL--L----------G-----IN--DTI----N-----A----LRVG--FLN...........-A---   857
G.KE.09.DEMG09KE001  ------LT..HH.Q...-E-----R--G--Q-K--T--S-F---A----------------FI--AA-T---GL--L----------S---G-----IN--DTV----NW-----I--R----VL...........F-SA---   853
G.NG.12.12NG060409   ------RP..HH.Q...-ET--G--P-K--V---S-F---A--------------------IS-AA-TA---SSLQGLRL--G--L--I-L--G-----IN-YDTI----NW-----I-R-G--LN...........-A---   872
H.CD.04.LA19KoSa     ------LI..N.....-------QG-G----QG-F-P-F-I--N---------N----------------A-T--I---------V-----L----GK---I--T---------I-IG-RI---LN...........-A---   860
H.CF.02.LA25LeMi     ------LI..S........-G--R--G--Q-----S-F---A----------------------RL---A--T---Q------------L-S-----G----K--IN--T--------I-IG-R-G--I...........-A---   851
J.CD.03.LA26DiAn     ------LT..N..T..EA--G--G--Q--T-----S-F-A------------------FV--AA-T--T--H---------I--LGA-AC---GR--R--I--F---------------I----A-R-V--VL...........F-IA---   866
J.CM.04.04CMU11421   ------LT..H.LA..EA---T--G--QG-TX-----FF--A-E---N----------------FI--A-T--T---------I--LGA-AW----G---I-----DT---VX---------I-IXRR-F--VL............-A---   861
K.CD.97.97ZR_EQTB11  ------LT..S........-------G----Q-K---V---S-F---A-------------N------RH----V--A--LDRGLKG-------S--------L---IL---G--I----IN--T----------I--I-YR-F--LL.........F-L---   852
K.CM.96.96CM_MP535   ------LT..A.....-------G------Q-KN----S-F---A-------------N------RQ--N-I----L-RGL-G------------L---------G--I----IN--T-------------G---I-IG-R-F--LL............-A---   843
01_AE.CM.11.1156_26  ------LT..HH.Q...-E-----GR-----------Q--S-FF-IA--------------------T---NSLKGLRR--G---LG---S---G----T--TT--D--------GW-----F--I-R----VL.........--A---   850
01_AE.CN.12.DE00112CN011 -----L--PS.HH.Q...-EL----R--G--Q---V---S-F---A----------------FI--AA-T----SSLKGLRR---V---LG---L---G----I--FD-V---------I--AA-R-G--L-.........--A-Q-   867
01_AE.IR.10.10IR.THR48F -----L-IPT.HQ.Q...-E-----R--G-----V---T-F-IA-----------------K--A--AA-T----HSSLKGLRR--G---LG---A-----I---S---D-----GW------A-R-W--LL..........--A---   868
01_AE.JP.11.DE00111JP003 ------LT..HH.Q...-E-----R--G-----Q-----S-F---A----------------FI--------T----HSILKGLRR--G--LG----I---G----I---I---DT-------GW-----I--A-RVG--LN..........--A---   861
01_AE.MM.14.fKSDU26  ------L-IPS.HH.Q...-E-----R--G--Q----V---S-F---V---N--------------FI--A-T-I---HSSLKGLRR--G---LG---L---G----I---I---DT-----NW-----I-XA--XW---L..-------A---   878
01_AE.PH.15.DE00115PH012 ------L-IPS.HH.Q...-E-----R--G--QG---V---VT--S-F---V---N--------N------FI---A-T-I--HSSLKGLRR--G---LG---L-----GR--RI--I--D----T---GW-----I-XA--XW---L...........-A---   858
01_AE.SE.12.070SE    ------I-Y.S.....-E-----R--G-----QG----S-F---A-------------------V--A--T----SSLKGLRL--G---LG---L---------------D-S----GW-----A--RVW---L...........-A-Q-   854
01_AE.TH.10.DE00110TH001 .V---IPT.HH.Q...-ET---------QGK-----V--S-F---V---N----------N------FI--AA-T-I---HSSLQGLRL--G---LG---L---------------N---A----A----GW-----A-R-W--L-.T.........--A---   865
01_AE.TH.11.40436v09_04 .V-L--PT.HH.QG...-E--LGR------HS-GG--T---S-F---A---------------N------FI--AA-T-I---HSSLKGLRL--G---LG---L---I--I--DTA---I---A--GW-----A-R-G-G-L-.........--S---   851
01_AE.TH.90.CM240    ------LT..HH.QK..-E-----------QG---V---S-F---A------------------FI--AA-T----HSSLKGLRL--G---LG---L--------------D---A----CA----GW-----A-R-G--LN...........-A---   858
02_AG.CM.10.DE00210CM013 ------PI..HH.QW..-E---GRT--G--Q--N-V---S-F---N--I-----------FV----T---HSSLKGLRL--G---LG---L-----S--------TN--DTI----NW-----I--R----YN...........-A---   885
02_AG.DE.09.701114   ------PI..HH.QW..-E---------Q-------FA--F---V---N--I---------L----FV----T---HSSLKGLRL--G-----L-----S--------I--FDTAT----GNW--I--IG-ATG----N..........--A---   876
02_AG.KR.12.12MHI11_10746 -----LA.HH.Q...-E--K----G-------N----------RQ---L--AA-T-------------N-----------D---L-----S-----G----------TN--DTI----NW-S---KG-R-----N.T.........-A-Q-   867
02_AG.KR.12.12MHR9_  ------LT..HH.Q...-E--------Q-G-K-V----S-F---A-------------------V------FI--A-T----HGSLKGLRL--------HL--S--A----G-----IN--FDTI-----NW-----I--R----N-.T.........-A---   870
02_AG.LR.x.POC44951  ------LT..HH.Q...-E-----R--G--Q-K--V--LS-F---A-------------------FA--AA-T----HSSLKGLRL--G---LG--I-S--G---I--IN--DTI----NW-----IG-R-G--LN...T-------KT---   853
02_AG.NG.12.12NG060304 ------LT..HH.Q...-E---L--R--DG--Q----S-F---A----------------R----FV----L---HNSLKGLRL--G---LG-I-S--G--I----IN--TI----NW-----IG-R-G--LL...T-------A---   863
02_AG.NG.x.IBNG      ------LT..HH.Q...-E-----R--G--Q-K--V---S-F---A----------------I--AA-T----HNCLKGLQK--G----L--IS---V------IN--TI--VI----NW-----A-IG-RVG-----N............F-A---   855
02_AG.PK.15.PK032    -----ILT..HQ.Q...-L-----R--G--Q-K--V---T-F---A------------------CA--A-T----YSSLKGLRL--G----L-S--S--G-----IN--DT-----NW-----LG-RIG--FLN.-V------A---   874
02_AG.SE.14.098GN    ------LT..HH.Q...-------R--G--Q------M--SS-F-E-T------------------C-----A--AA-T----HSSLKGLRL--GV----L--F-S---G-------IN--TI----NW--SI--IG-R-G--LN...........--S---   854
```

Env

HIV-1/SIVcpz Proteins

glycosylation NGS glycosylation NAT

Env end
gp41 end
cytoplasmic tail end

```
B.FR.83.HXB2                  .LSFQTHL.PT.PR..GPDRPEGIEEEGGERDRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELLGRR.......GWEALKYWWNLLQYW.SQELKNSAVSLLNATAIAV.....AEGTDRVIEVVQGACRAIRH.IPRRIRQG.LERILL*    856
02_AG.SN.13.9580              ---P-LT.HH.Q-..E----GR---G--Q--E--V---S-F---A-N------I------NFA--AA--T-DI---SQSQGTET-G----L---S---G--G--IN---T--V----GNW------IG-R-G---N.V-V--------LCYK       861
03_AB.RU.97.KAL153_2          ------R--.A.Q-..------D-----T------RF-----F-I--H---AA--------------------------------L---S---G------IN-IGTI------GW------IG-RF---M-N.--------------A-KA-Q-    840
04_cpx.CY.94.94CY032_3        ---L-LIPT-.Q..-L--G-T---Q--S-------Q-G-------N------RH-N----A--T--I----------L--F-L--G------IN-F-T-------------I--A-R----CN.----------------A----           856
05_DF.BE.x.VI1310             ------PF..A..-------ET--G--Q-G-----FS---------N-------------------------S---R--I--T--V.VI----------------I--AL-R-G--LN.-------------------         862
06_cpx.AU.96.BFP90            ------L-LI.N--T..--A---GE---G--Q--V---S-F---A-N-----FG--AA-T--I--------I--LG--IC--G--Q---I-FD-A----NW-----A----RIF-FLN.V-------F-A----              860
07_BC.CN.98.98CN009           ------LT.-N.-G--.....-------G----Q-K-------F---A-----FI-VAA-V----NSLRGLQR------LGS-V---G---K-TI--VDTI-------------I--L---L---YN.-------F-AA-Q-              860
08_BC.CN.97.97CNGX_6F         ----ILT.-N.-G--G-LGR------Q-KTS----F---A-----N------FI-LTA-G----NSLRGLQR------LGS-V---GL--K-TI--VD-I-------------I-NR---I--HN.V-------F-AA-Q-            865
09_cpx.GH.96.96GH2911         ----ILT.-N.-G----------Q-KV----S-F---A-N------FI--AA-T--I-H----V-QI--LG--A---G------I---T---V-----G------I-R-G--L-.--------A---             843
10_CD.TZ.96.96TZ_BF061        ---LH-L-.A...--------G--G-----D-F-T------N----I--A-------------M---I---I-----------I--DTI--E--GW-----I--R-V--LN.--T--------A---        841
11_cpx.CM.95.95CM_1816        .V-L--LI.-S.HK..EA-S-G----Q---S-F-----N---Y--X---I--LGT---------------T--LG--T-------I----T--V----G------I--IHRXL--VL.-------A---              866
12_BF.AR.99.ARMA159           ------L-.-I.-S...E-------Q-QGKV--V--L-T-------------V-----------------------R---I---T--V----G-----IL-AL-RIG--VLN.A---V----A---           843
13_cpx.CM.96.96CM_1849        ------LT.HP.K-..E-------Q----S-F---A------FI-A--T--HSSLKGLRL--G--L--G----GR---I-IT--DS------1849-----I---TG-RIG-GL-.--------F---      868
14_BG.ES.05.X1870             ------PT.HH.Q-..E-------Q-K------S-F---A------X--FI-VAA-T----NSLKGLRL--G--L--G----L---GR----I---DTI-------NW----A-X----VG--LN.V-V-------F---         866
```

(alignment table continues — remaining rows omitted)

HIV-1 Proteins

Nef

myristoylation
Nef start
phosphorylation
acidic cluster poly-P helix
HXB2 premature Nef end

```
B.FR.83.HXB2               MGGKWSKSS.VIGWPTVRERMRR................AEPAADRV......GAASRDLEKHGAITSSNTAATNAACAWLEAQ..EE..EEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLWIYHTQGYFPD*QNYTPGPGVRYPLTFGWC 141
A1.CD.02.LA01AlPr          --------R-.MP---EI------TQTPPAA.........QTP----VG---PV-Q--D-Y--V-IN.....NQ-S--Q-DE-....-..-D-----R----------G-F--GY-------D-----KK--Q-----V-N----W-------I------ 146
A1.CM.08.886_24            --------.IV--QI--L-QAS...............----EG----KV-Q--D---V-I-I..NHPSNV.....K--.-D-----R-----------L-----------TY-RK-----V-N---F---W-------I 140
A1.ES.15.100_117           --------R-REAE-SE------QQTP...........A-XSG---V-Q--D---V--L..NHPS-V--.....-D.----.-----R----------F-G-F-X-F------D---W-RK-------V---------T 141
A1.KE.11.DEMA111KE002      -------IV--A-------TP............S--PG---V-Q--D---V--L..NHPS---.....-.---G-----R----------F-G-F----------D-Y-K------V-N----W-----I 140
A1.PK.15.PK034             --N-------.IV---EI---L--ATP.........TT--TG---V-Q--AR--V---I.NHPSYT.....-D.----.----R----------F-G-F-----------D-Y-RK------V------I-F- 142
A1.RW.11.DEMA111RW002      --S-----.IV--A-----QARRAT.........PP---AG---V--------V..NHPS-.....-D.----.-------R----------F-------------D-Y-RK-----V-N---F---W-------I 145
A1.SE.13.078SE             --N-----.RR--TE---I-QTP...........T--AG---V-----R------..NHPSVT--A--.-DN.----R----------F--I--E------R------D-Y-RK--E---V-N----W 141
A1.UG.11.DEMA110UG009      --S------.IV--A----I-QTP...........A--TG---V-Q--AR--V---..NHPS-V----....-D.---R----------G-L------------D-Y-RK-----V-N----W---I- 140
A2.CM.01.01CM_1445MV       --S-L-.--.---A-------TP...............--EG----V-Q--GTR--V-IR-----SD----....SD-------R-----------G---------------------------------F- 139
A3.SN.01.DDI579            --------R.QE--HK--D---QTPP...........T---EG---V-Q--S---V--T-Q..AS-S-T.......-------R----------G-F---------D--Y--Q------V---------T 142
A4.CD.97.97CD_KCC2         --------.LV--A--D----TP.............A-Q--D--Y--V--N-----PD----....-------R----------------D---W--K----V---------I-F- 142
A6.BY.13.PV85              --------.IV--QI----APAPA.........P----TG----PV-Q--D---V-----XX--D-----H..-D.....----R----------G--------------D-Y-K------V----------I 147
A6.IT.02.60000             --------.IV--E---I--APA...........PA---RG---PV-Q--D-Y--V--R-ST---I..-D.....---R----------G-F----------D-Y-RK--E------V------W--K---I 145
A6.RU.11.11RU6950          --------.IV--A------APA...........PA---SG---PV-Q--DR--V---.-D.....---R----------G-F----------D-Y-K-------V-N---F---W-------I 145
A6.UA.12.DEMA112UA024      -------NS----AQI----APA...........P----PG---PV-Q--D-Y--V--R-----D--....---K----------G-F---------D-Y-RK--E---V-N----W-------T-F-I 143
B.AR.14.DEMB14AR003        --N-----R.-P--S-I----AE.............P---VG-------R--L----PTN--T---Q-..-.E-------.---------------Q----------Q 145
B.BR.10.10BR_MG035         --S-L--AF.-R--D----I-ARPAV.........E----VG------------SD----K.Q--ED----------R------------S-L---#----X.......Y--D------V 148
B.CU.14.14CU005            ----L--RI.G---A------AE.............--EG-------Q--L---D----Q-.....-E-D--R-----------S-L---#----X. 99
B.DE.13.366396             --------NW.RA--S------II-TDPRA....GRR..GQ----G-------V-Q--AA------N----------V----.---G----------G-L------------V-N--------T 143
B.DE.13.947915             --------NW.RA--S------II-TDPRA....GRR..GQ----G-------V-Q--AA------N----------V----.---G----------G-L--------I-Y--Q------V-N----W-------T-F- 153
B.DO.11.DEMB11DR001        --R-----.AG--V-----A--G............--EG-------V-Q--------L-T--PGN-PD--.-S--.-DS.-----R----------F-----------VY--Q-R-------V-N----W-------T---C-- 145
B.ES.15.100600             --------.IV--A--D----TR.............--EG-------V-----R------L-T--NPSN---.-.---R----------F-----------VY--Q---------V-N----W 145
B.FR.11.DEMB11FR001        ----IF--R.GG---A----TGTEPKAERLGKERVRRT......--EG----------R---T---DN--DS----.D......--R----------G-L------------D-Y--Q------V-N----W-------Q 159
B.JP.12.DEMB12JP001        --------.IV--A------I-AEPR.........--EG-------V----------I..N--D-V-----.-----R----------F-E-F---------D-Y-K--E------V---------I---W-E-----T-F 144
B.PH.15.DEMB15PH003        --N-----.LV--KI---K-............T---EG-------Q--Y--L-----D--.---D--R----------G-L------------D-Y-K--E------V------W-------P-F- 142
B.RU.11.11RU21n            --------.IV--A---Q-TE.............--EG-------V---------Q-----D--R----------F-----------Y-K-------V-N----W-------T 142
B.SE.15.102SE              --------RRRGDE-QA-K-K-KQTDPT.........A--EG----------Y-F----V-N-G..----R----------F-G-L-------VY-----------V---------W 145
B.TH.10.DEMB10TH002        --S-----.LG------AQP..........Q---G-------------NPDS----.EQ------R----------G-L------------D-Y-K--E------V-N----W-----T-W 144
B.US.16.2609               -------.LG--SN----------.--T-GEGAVSQDR-V----------N------V----.--R----------G-L------------Y-K------M-V---------WH-----F--C-- 144
```

[alignment continues for additional subtypes: C, D, F1, F2, G, H, J, K, and CRF 01_AE, 02_AG groups]

myristoylation
Nef start
phosphorylation
acidic cluster   poly-P helix
HXB2 premature Nef end

HIV-1/SIVcpz Proteins

B.FR.83.HXB2   MGGKWSKSS.VIGWPTVRERMRR................AEPAADRV......GAASRDLEKHGAITSSNTAATNAACAWLEAQ..EE..EEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQRRQDILDLWIYHTQGYFPD*QNYTPGPGVRYPLTFGWC   141

```
                                                                   normal Nef end
B.FR.83.HXB2                YKLVPVEPDKIEE.ANKGENTSLLHPVSLHGMDD.PEREVLEWRFDSRLAFHHVARELHPEYF.KN......C*      205
A1.CD.02.LA01AlPr          F-----D-AGV--..TE---NC----ICQ----E-.D-K---M-K---S---K---------FY-...D...-- 211
A1.CM.08.886_24            F-------EEV-R-.TE---N-----MCQ------.D---T-I-K---H---K-I---M---FY-..D....-- 205
A1.ES.15.100_117           F-E----EEV-K-.TE---NC----ICQ------.D-K-T-R-K---K--LK-L---K---FY-..DLRL.--  209
A1.KE.11.DEMA11KE002       F-----D-GEV-Q.-.-E--DN-----ICQ------E---T-R-E---RR-L----Y-..D...--         205
A1.PK.15.PK034             F-----D-GEV-Q.-.-E--DN-----ICQ------E---T-R-E---RR-L----Y-..D...--         207
A1.RW.11.DEMA111RW002      F-----N-EEV-R-.-E--N-----ICQ------E-G-T-M------LR-I----E---FY-..D....--    210
A1.SE.13.078SE             F-----D-EEV-K-.TE---NT----ICQ------E-K---M------LK-R---I---FY-..D...--     206
A1.UG.11.DEMA110UG009      F-----D-EEV-K-.TE---N-----ICQ------E---T-K-K------LK-R---K---FY-..D...--   205
A2.CM.01.01CM_1445MV       F-----N-SEV-..TE--N-----ICQ--I----------K-----H--LR-R----Y-DFY-..D...--    204
A3.SN.01.DDI579            F-----DQASV--..TE---..-----CQ------K-K---M--------LK-T---M---FY-..D-       204
A4.CD.97.97CD_KCC2         F-------EAV--.TG--N-----ICQ--------------M---K--LK-R-K-----FY-..D...--     207
A6.BY.13.PV85              ------D-AEV--.TE---N-----ICQ------E-.E-K---R-K------LR-L---Q---FY-..D--    212
A6.IT.02.60000             ------D---V---.TG--N-----ICQ------D-K---M-K---S--RR-Q-M---FY-..D-         210
A6.RU.11.11RU6950          ------D-ADV--.TE---N-----ICQ------E-K---K-Q------LT-R---M---Y-..D---      210
A6.UA.12.DEMA112UA024      F-----KEV---.TE--N-----ICQ------E---M-K------LK-R-Y-M---FY-..D...--       208
B.AR.14.DEMB14AR003        F-------E-----QR--A----MH-------.-G--Q-K-----------M------Y-..D..--        210
B.BR.10.10BR_MG035         F-----D--QV-K-..E---NC----MHQ--I---.-K---M-K---K-------K-I--FY-...--       213
B.CU.14.14CU005            ..................................................................        99
B.DE.13.366396             F-------E-V--------N----A-M-Q-------------------------M---K---Y-...--      208
B.DE.13.947915             F-----EDV-K-..E---NC----M-Q--I--.-K---M-K------M-----------..D-           218
B.DO.11.DEMB11DR001        F---M---V---..-----NC--N------E-.Q-K--V-R-K----------R--K----------..D-   207
B.ES.15.100600             F----V-VHP-Q.VDN.-TNC----G-Q-I--.-K---M-K---H---Q-K----------..D-         209
B.FR.11.DEMB11FR001        F-----DKE-----..-Q--NC----M-Q------------V-K--L--VRYM----K---FY-...D-     224
B.JP.12.DEMB12JP001        F-----D--QV-K.D-E--IN----M-Q---------*M-K--H--LQ----K-----Y-..D--         208
B.PH.15.DEMB15PH003        F-----D--QV-K.V-E---AC----M-Q---------M-K------------K-K----Y-..--         207
B.RU.11.11RU21n            F-----EE--K.-TT---NC----ANM-------------V-K------Y----K---FY-..D-         207
B.SE.15.102SE              F-E-MK-EQ---.ET----T----A-Q-I-------------V-K--L--R-M-K------Y-..--       210
B.TH.10.DEMB10TH002        F-------E....NSE--TN-----A----E-.-G--V-K--------------M------Y-..D-        205
B.US.16.2609               F-----A-EDV-K.VTE--RNC----MN-------------M-K--K----------Q---Y-..--        208
C.BR.11.DEMC11BR035        F-----D-REV--.-TA--DNR----M----E-.ADG---Q-K--L--RR-----I----Y-..D--       208
C.CN.10.YNFL19             F-----D-KEV--.-E---N-----CQ----E-.EH----K-Q--IQ--HR-M------FY-..D--       206
C.DE.10.622166             F-----D-REV--.-Q---NC----M-Q------DD---K-K---S--RR-L-------Y-..D--        213
C.ES.15.100_118            F-----D-RRV--.-E---N-----M-Q------E-.EHG----------L--RR-M------FY-..D--    211
C.ET.08.ET104              F-----D-EEV-K.N-E---NC----I----E-.EHK----V-------S--RR--------FX.          209
C.IN.15.NIRT008            F-----D-REV--.TSQ--NC----C----E-.EY----Q-K--Q--HR-----R---Y-..D--         208
C.MW.09.703010256_CH256.w96 F-----D-REV--.-E---NC----M-----O---DH----K-K--Q--HH-M-------FY-..D--     208
C.NG.10.10NG020523         ------D-KEV--.-G---NC----L-QY-A-.-DH----M-K---S--RR-M-------Y-..D--        209
C.NP.11.11NP016            F-----D-REV--.-E--DNC----CQ----E-.EH----K-K--Q--RR-M-------FY-..D--        208
C.PK.14.DEMC14PK009        F-----D-REV--.-E---SC----M-Q------E-.-M----R-H--RR-M-------Y-..D--         209
C.SE.15.100ET              F-----D-SEV--.-E---N-----MCQ--E-.EN----K-Q--Q--RR-L-------WY-..--          207
C.TZ.08.707010457_CH457.w8 ------D-KEV--.ETE--N-----M----E-.EH----K---M--RR-M---K---Y-..D--         208
C.US.14.M0118v1d14_5M04_C5 F-----D-SEV--.-E-D-NC----M-Q--E-.EHK----Q-K---S--RR-M--------Y-..D--      212
C.ZA.13.DEMC13ZA152        F-----D-REV--------DNC----M-Q--WE-.QYG---Q-K--L--RR-I-----Y---Y-..D--      214
C.ZM.11.DEMC11ZM006        F-----D-QEA--.H--RDDSC-M--M-Q--IE-.EH--I-M-K--TQ--RR-L------WY-..--        206
D.BR.10.10BR_RJ095         F-----D-QEV-K.LTA--DNC----LNQ------S-----Q-K--------E-K--QM---FY-..D--     211
D.CD.03.LA17MuBo           FE----D-QVVG--.TE--NC----ICQ--AE-.S------V--N-------E-K---KY-.-..D--       207
D.CM.10.DEMD10CM009        FE----D-RVV-G.ETE---C-----Q--E-.T------V-K-N------E-K---QK--F-.-..D--      211
D.CY.06.CY163              ..................................................................        45
D.KE.11.DEMD11KE003        FE---L-EEV-K.-TE---NC---------.......Q-----I----N-----E-K--Q----Y------    207
D.SE.12.077UG              FE----D-KEV--.NI-.--SC----G-Q-T--.-----V---N------E-K-LK---Y.Q-..-        214
D.TZ.04.CO6405V4           FE----D-QEV---.-E--TNC----M-Q--IE-.Q--Q--K---N-K---E-K--QM---FY.R..D--     212
D.UG.10.DEMD10UG004        FE----D-KEV--.DTE---NC----MCQ------E-.Q------K---N------E-K-KIQ----Y-..D-- 207
D.UG.11.DEMD11UG003        F-----IDAKEA--.DTE---NC----ACQ-----E-.T------M---N-----E-K--V---FY-..D--   208
D.YE.02.02YE516            FE----D-KMV--..E-E--DKC----MHQ---G-.-K---V--N-K---E-K--KT---F-.-..D--      205
F1.AO.06.AO_06_ANG32       ..................................................................        96
F1.AR.02.ARE933            ..................................................................        90
F1.BR.11.DEMF111BR037      F-----D-EDV-K.-E---NC----M----E-.EDK----V-K------LR-I----R-D-Y-...D--      217
F1.CY.08.CY222             ..................................................................        45
F1.ES.11.VA0053_nfl        F-----D-EEV-K.-E---NC----M-Q------.EDG----K-Q------LK-I----R----Y.Q-.D--   216
F1.FR.04.LA22LeRe          F-----D-EEV-K.-E---NC----M-Q------.E-K----Q-K------LR-I----K---FY.Q..D--   206
F1.RO.03.LA20DuCl          F-----D-EEV-K.-E-D-NC----M-Q--------Q------K---------K-K----FY.Q-          207
F1.RU.08.D88_845           F-----D-EEV-K.-E---C-----M-Q------.EDG----R-K------------LR-L----Y-..D--   215
F2.CM.10.DEMF210CM007      F-----D-EEV-K.-E--TNC----I-Q--IE-.ED----M-K---T--MR-I-K-K---FY-.-.D--      206
F2.CM.11.DEURF11CM026      F-----S-EEV-K..-E---NC----M-----E-.EHG---Q-K---S--RR-L-K-----FY.T..E#...-- 211
G.CD.03.LA23LiEd           F-----L--EE---.-T----N-----ICQ--IE-.AD----V----S--RR-L-----DFY-...D--      211
G.CM.08.789_10             F-E---D-AE---.-E---TN----ICQ--IE-.EH------V-------S--RR-L-----FY-...--     206
G.CM.10.DEMG10CM008        F-----MD-AE---.-----N-----ICQ--E-.ED----V-----S--RR-L-----FY-...D--        213
G.CN.08.GX_2084_08         F-----MD-AEV---.---E--N-----ICQ--Q-.ED-----V------S--RR-R------FY-..D--    208
G.CN.16.224GX              F-----MDQAA--.K.-E---N-----ICQ--QE-.ED-----V------S--RR-L------FY-...D--   208
G.ES.14.EUR_0033           F-----ID-ADV--.E----TN----ICQ--XE-.ADK---V-----S--RR-L---M----Y-..D--      210
G.GW.08.LA57LmNe           F-E-MD-AEV---.------N-----ICQ--AE-.ADK---V-----S--RR-R------DFY-...D--      207
G.KE.09.DEMG09KE001        F-----D-AE---.TE--RNT----ICQ--E-.SD----K-K---S--RR-I----Y---Y-..--         219
G.NG.12.12NG060409         F----LD-AE---.-TR---N-----ICQ--V--.DDK---V-----S--RR-I------DFY-...--      211
H.CD.04.LA19KoSa           F-----D-QDV-K.-E---NC----ICQ------E-------M-K---H----T---QFY-..D--         205
H.CF.02.LA25LeMi           F-----D-QEV-K.-E--TNC----ICQ--E-.E------I-K------T-T--R---FY-..D--         208
J.CD.03.LA26DiAn           F-----D-REV-K.-E--DNC----ICQ--E-.E------M-K------S--RR-I-----Y-..D--       209
J.CM.04.04CMU11421         ------D-REV--.-E---NC----X-Q--IX-.X-X---M-K--XS--RR-X--------FY-..D--      210
K.CD.97.97ZR_EQTB11        ------D-REV--.-TE---NC----NQ---E-.EH----K-K---S--RK-----M-----Y-..D--      207
K.CM.96.96CM_MP535         F-----D-AEV--.TTE--DNC----INQ----E-.EH--I-M-K---S--RR-I-----Y-..D--        216
01_AE.CM.11.1156_26        ------D-REV--.D----NC------Q-I--.ED----M-K---S--RR-I---RY---Y-..D--        204
01_AE.CN.12.DE00112CN011   F-----D-REV--.D-----SC----INQ--IE-.E-----M-K---A--RK-I----R---Y-..D--      207
01_AE.IR.10.10IR.THR48F    F-----DRQEV---.------NC----M-Q--QE-.E------R-K---X--RK---------Y-..D--      205
01_AE.JP.11.DE00111JP003   F-----D-REV--.DT-E--NC----M-Q--V--.-DK---M-K---A--RK-I---K---FY-..--       206
01_AE.MM.14.fKSDU26        ..................................................................        110
01_AE.PH.15.DE00115PH012   F-----D-KEV--.D---NC----M-Q--QE-.ED----M-K---S--RR-R---K---FY-..D--        209
01_AE.SE.12.070SE          F-----D-REV---.--E--NC----M-Q--E-.E-----M-K---S--RK-I--------Y-..D--       205
01_AE.TH.10.DE00110TH001   F-I-ID-REV--.N----NC----M-Q--IE-.E------M-K---S--RK-I---I----Y-..D--       206
01_AE.TH.11.40436v09_04    F-----D-REV--.S----NC----M-Q--I--.E------M-K---S--RK-I---Q-DFY-..D--       207
01_AE.TH.90.CM240          F-----DQREV--.D-----NC----M-Q--IE-.E------M-K---A--RK-----Q----Y-..D--     205
02_AG.CM.10.DE00210CM013   F-----LD-VE--R.-E---N-----ICQ------ED----M-K---S--RR-I-------Y-..D--       216
02_AG.DE.09.701114         F-----MD-ADV--.H-----N-----ICQ------DHG---V----RS--SK-I---K---FY-..D--     211
02_AG.KR.12.12MHI11_10746  F-----MD-EE---.-------N-----ICQ----E-.ED----V---S--RR-I-----Y-..D--        207
02_AG.KR.12.12MHR9         F-----LD-AEV--.---E--N-----CQ----E-.ED----V---G------T-K-K------FY-..D--    221
02_AG.LR.x.POC44951        F-----D-REV--.-E--N-----ICQ--IE-.TD-Q--V-----S--K-R-Q-----FY-..D--         210
02_AG.NG.12.12NG060304     F-----D-QEV--.-Q--DNN----ICQ------DH----I-K---T---R-M---M---FY-..D--       206
02_AG.NG.x.IBNG            F-----MD-AEV--.-------N-----ICQ------ED----I---------M---FY-..D--          207
02_AG.PK.15.PK032          F-----MD-TEV-K.-E---NC----IAQ---G-.ED-----V------S--LK-R------FY-..D--     209
02_AG.SE.14.098GN          F-----MD-AEV--.-------N-----ICQ---E-.EN----V---RS--SK-I---M---FY-..D--     207
```

```
                                                                    normal Nef end
B.FR.83.HXB2                   YKLVPVEPDKIEE.ANKGENTSLLHPVSLHGMDD.PEREVLEWRFDSRLAFHHVARELHPEYF.KN......C*         205
02_AG.SN.13.9580               ------MD-SEV--..-E---N------CQ------.E-D------V---------LK-R-Q--L--F-.-.......-          209
03_AB.RU.97.KAL153_2           ------D-AEV--..-TE---N------ICQ------.E-K---M-K------LT-R-----FY-..D....-               203
04_cpx.CY.94.94CY032_3         F----D-QEV--..-TE----C---I-Q--E.------YK---------Y--..D....-                           220
05_DF.BE.x.VI1310              F---N-EEV-K..-E--DNC---M----E.DD----Q--S--LR-I---R--FY.Q..D....-                       207
06_cpx.AU.96.BFP90             ------D-EEV--.LT---NC------ICQ--AE.E-----K-K---S--RR-I---K--FY-..D....-                 206
07_BC.CN.98.98CN009            F----D-REV--..-E---N------CQ---E.EH----K-K---Q--HR-R------FY-..D....-                   207
08_BC.CN.97.97CNGX_6F          F----D-REV--..-E-DNC---CQ---E.EH----K-K---Q--HR-R------FY-..D....-                     207
09_cpx.GH.96.96GH2911          F----D-REV--.-TG---N------MCQ--V.-.KHG---M-----T--LK-I---K--FY-..D....-                 207
10_CD.TZ.96.96TZ_BF061         F----D-REV--......-NC---M----E.-----V-K---T--HK-I--------Y-..D....-                     208
11_cpx.CM.95.95CM_1816         F----D-REV--.--A---NC---L-Q-T--.E----M--S--RR-R-I-----DFY-..D....-                     209
12_BF.AR.99.ARMA159            F----D-VEV-K..-E---NC---M----E.ED----I-K-----LR-L---K---WY-..D....-                     206
13_cpx.CM.96.96CM_1849         F----ID-KEV--..-E---NC---M-Q-IE-.A-------M-K---S--RR-I-----FY-..D....-                   200
14_BG.ES.05.X1870              F-E-D-AEV--.-TG---N------ICQ--IE-.ADK---R-K---G--RR-I----R-DFY-..D....-                 208
15_01B.TH.99.99TH_MU2079       F----D-SDV--.N------NC---M-Q---E.E-----K-Q----A--RK-I-----------Y-..D....-             201
16_A2D.KR.97.97KR004           ------S-AEV--..-TE--TN------ICQ--IE-.------K-V---H--LV-K-------FY-..D....-               217
17_BF.AR.99.ARMA038            -------EE--K..-E---NC---M-Q-E---.ED----K-K---S----R---R--FY.Q-.-                        204
18_cpx.CU.99.CU76              F-----DSNEV--..-E---NC---ICQ-----.EH----K-K---A--RR-I---Q---FY-..D....-                 207
19_cpx.CU.99.CU7               FE---D-KEA-K.DTE---NC---ACQ-------.-E-K---V---N-----R-K--IK----FY-..D....-             207
20_BG.CU.99.Cu103              F-----MD-AEV--.-N---N------LC---E.-N------I-----S--RR-L------.Y-..D....-                204
21_A2D.KE.99.KER2003           F-----D-SEV--.DT---N------MCQQ-I-..------R-K--PH--LK-R-Q-K---FY-..D....-                208
22_01A1.CM.01.01CM_0001BBY                                                                                          130
23_BG.CU.03.CB118              F-----D-ADV--.-SN---N------ICQ----E.DDK---M-----S--RR-I---R--DFY-..D....-               209
24_BG.ES.08.X2456_2            F-----D-TEV--..-S--TN------ICQ----E.SDK---I-K---S--RR-I-----DFY-..D....-                224
25_cpx.CM.02.1918LE                                                                                                 137
26_A5U.CD.02.02CD_MBTB047      F-----D-AEV--.-N----N----AM-----E.S-G---M-K---Q--RK-T-------..D....-                   208
27_cpx.FR.04.04CD_FR_KZS       F-----D-KEV--..-E---N------E.E-----M-K----LR----K---FY-..D....-                        208
28_BF.BR.99.BREPM12609         F-----D-EQV-A..-E---N------M-Q-------.-K---M-K---------M-------Y-..D....-               210
29_BF.BR.01.BREPM16704         ------D--QV--.E----N------M-Q------.-----M-K----------F---.-..D....-                    212
31_BC.BR.04.04BR142            F-----D-REV--.--A---N------M-Q-----E.-THG---M-K---L--RR-M-------Y-..D....-              223
32_06A6.EE.01.EE0369           ------D-REV--.-T----NC---IC---AE.E-G---M-----S--RR-I---K--FY-..D....-                   207
33_01B.ID.07.JKT189_C          F-----D-TEV--..-E---NC---M-Q---E.ED----K-Q---S--RR-I---Q---Y-..D....-                  205
34_01B.TH.99.OUR1969P          F-----D-GEV--......-SC---M-Q-IE-.------M-K---A--RK-I-------FY-..D....-                  203
35_AD.AF.07.169H               -------D-AEV-K.D-E---N------MCQ-------.E---T-M-K-------T-K------FY-..D....-             206
36_cpx.CM.00.00CMNYU830                                                                                             131
37_cpx.CM.00.00CMNYU926                                                                                             132
38_BF1.UY.03.UY03_3389         F-----D--EV-K..-E---NC---M-Q-A--.EDK---V-------LR-I---R---FY.Q..D....-                 206
39_BF.BR.04.04BRRJ179          F-----S-EEV-R..-E--DNC---M-Q-----.ED----V-K---K-----------Y-..D....-                   216
40_BF.BR.05.05BRRJ055          F-----D-EEV-K..-E---NC---M-----E-.E-K---Q----AT----------FH.-..-                       213
41_CD.TZ.05.CO6577V5           F-----D-REV--..-E--TNC---M-Q-----E.ED----I-----Q--HR-----I----Y-..D....-               208
42_BF.LU.03.luBF_01_03         F-----Q-QV-K.-TE..----------M-----E-.T-----V-K---K--------H-X-----Y-..D....-           207
43_02G.SA.03.J11223            F----MD-ADV--.-E--N------ICQ----E.ED----V---S--RR-------Y-..D....-                     215
44_BF.CL.00.CH80               F-----D-EEV-K..-E--NC---M-Q-----.ADK---M-K--S--LR-I--RF--FY.Q.DD.....-                 207
45_cpx.FR.04.04FR_AUK          F-----D-KEV--..-TE----C---I-Q--IE-.T----I-K--H--QR-L---K--FY-..D....-                  208
46_BF.BR.07.07BR_FPS625        F-----N-EEV-R..-E---NC---I-Q---E.E-Q---K--S-----R------DFY-..CCIELN.-                  213
47_BF.ES.08.P1942              F-----D-EEV-K..-E---NC---M-Q---E.ED----M-K--RH--SK------Y-..D....-                     207
48_01B.MY.07.07MYKT021         F-----D-QEV--.D-----SC---M-Q---E.-E----M-K---S--RR-I---R---Y-..D....-                  206
49_cpx.GM.03.N26677                                                                                                 99
50_A1D.GB.10.12792             ........QEEV-K..-TE-G-NC----ICQ-------.E-K---Q-K---S--RK-I-------FY-..D....-            205
51_01B.SG.11.11SG_HM021        FX----D-QV-K..-EK--------M-Q--N-----.--K------M-K----LR-M-K-K--D-Y-..-                 211
52_01B.MY.03.03MYKL018_1       F-----D-KEV--.D----SC---I-Q--QE-.E-G---M-K---A--RT-T-------FY-..D....-                 205
53_01B.MY.11.11FIR164          F-----D-AEV--..-E---NC---M-Q---E.-------I-K---S--RR-I-------Y-..D....-                  206
54_01B.MY.09.09MYSB023         F-----D-REV--......-NC---M-Q--AE-.E-K---M-K---A--RT-----M------.-..D....-              213
55_01B.CN.10.HNCS102056        F-----D-REV--.TT-E--NC---M-Q------.E-K---M-K---A--RK-R-------FY-..D....-               206
56_cpx.FR.10.URF5_patient_A                                                                                         84
57_BC.CN.09.09YNLX19sg         F-----D-RE----.-E---NC---CQ----E.EH----R-K---Q--HR-R--------FY-R..D....-               208
58_01B.MY.09.09MYPR37          F-----D-REV--......-NC---M-Q-IE-.N-----M-K---A--RK-I---K---Y-..D....-                  206
59_01B.CN.09.09LNA423                                                                                               129
60_BC.IT.11.BAV499             F------TKD.....IEE-SSC----MQQ-RVE-.-------M-K---S-------M--K-------..D....-             202
61_BC.CN.10.JL100010           F------KEV--..-TE---NC---CQ---E.EH----M-K---H---S------K----WY-..D....-                207
62_BC.CN.10.YNFL13             F-----D-AE---.-E---NC---M----E.EHG---R-K---Q--HK-R-------FY-..D....-                   208
63_02A.RU.10.10RU6637          F----MD-AE---.-TE---N------ICQ-LE-.ED----V---K--Q--D-----R--FY-..D....-                208
64_BC.CN.09.YNFL31             F-----D-REV--.--E---NC---CQ---E.EH----K-K---Q--HR-R--------FY-..D....-                 208
65_cpx.CN.10.YNFL01            F-----D-REV--.-S-E---NC---M-Q-----.EHG---V-K---M--HR-M----Q---FY-..D....-              212
67_01B.CN.11.ANHUI_HF115       F-----D-REV--.D-----NC---M-Q-IE---.E-----M---A--RK-I--------X-Y-..D....-               208
68_01B.CN.11.ANHUI_WH73        F-----D-RDV--.D--R-DSC----I-Q--I--.E-----M-K---A--RK-I------D-Y-..D....-               208
69_01B.JP.05.05JPMYC113SP420   F-----D-REV--.ET----NC----MN----E.E-G---R-E---S--RK-I-------Y-..D....-                208
70_BF1.BR.10.10BR_PE004                                                                                             97
71_BF1.BR.10.10BR_PE008        F-----D-E-V--.---K--S------M-----E-.A-K---V-K---S-G------------Y-..-                   205
72_BF1.BR.10.10BR_MG002        F----L--EEV-K.-SE#--NC----AN------PG-G---V----------M-------.Y-..-                     221
73_BG.DE.01.9196_01            F-----D-AE---.-T----N------ICQ---E.XDN---R-T---S--RR-------DFY-..D....-                208
74_01B.MY.10.10MYPR268         F-------EQV--..--A---NC---M-Q--V-.--------M-K----M--K-------..D....-                   209
77_cpx.MY.14.14MYNBB090        F-----D-REV--.L-E---NC---M-Q------.EH----M-K--RQ--RR-L------WY-..D....-                212
78_cpx.CN.13.YNTC19            F-----D-REV--.DT-E--NC---M-Q--I--.G------M-K----T--RR-M--------Y-..D....-              200
79_0107.CN.15.SX15DT013        F-----D-G-V--.--A--DNC---MCQ------.EH----M-K---Q--RR-M--------FY-..D....-              206
82_cpx.MM.13.mSSDU12                                                                                                109
83_cpx.MM.13.mSSDU94                                                                                                109
85_BC.CN.14.14CN_SCYB2         F-----D-REV--..-E---NC----MCQ--T-.DH----R-K---A--RR-M-------FY-..D....-                207
86_BC.CN.13.15YNHS18           F-----D-REV--..-E---NC---M-Q-----.EH----K-K---H--YR-R------FY-..D....-                 210
87_cpx.CN.12.DH32              F-----S-EE----.-E---NC---MWQ----E.-------M-K---Q--RR-M------FY-..-                     208
88_BC.CN.05.05YNRL25sg         F-E--D-SEV--..-SE---NC----AWQ--E.EH----K-K---H---K-W------FY-..D....-                  217
90_BF1.BR.11.BRG06043                                                                                               213
O.CM.96.LA51YBF35              F-----TEEEA-GLGTTS-KA----AHN--FE-.DQG-I-K-Q--RS-GLK---L------L-P--..-                  213
O.CM.96.LA52YBF39              F-----SEEEA-RLG--C-RAK----ACA--YG-.QHG-M-K-Q--RS-GST---LIT---L-H--..-                  216
O.CM.99.99CMU4122              F-----SEAEA--LG--CVTAK----ACN--SE-.-HG-I-K-Q--RS-GNI---KVT---L-L--..-                  212
O.CM.x.pCMO2_3                 F-----SEAEA--LG--C-RAK----ACN--FG-.-HG-I-K-Q--S-GRT---MIT--DL-H--..-                   212
O.ES.01.Read25_HIV_Group0      F-----SEEEA-RLG-TCVRA----ACA--WG-.-HG-I-M-K--RS-GNT---MITN-L-Q--K.....-                216
O.FR.06.RBF206                 F-----SEEEA-RLG--C-RA----ACD--HE-.-HG-M-K-Q--RA-GIN---LQ----L-P--..-                   213
O.GA.10.10Gab1190              F-----SEAEA-RLG-TC-DAG----CA--FE-.EHK-I-M-K--RS--ST---LIT---L-Q--..-                   204
O.GA.11.11Gab6352              F-----TE-EAKRLG-ECVRAH----ACA--FE-.-HG-I-K-K--S--GLT--SKVI---L-T--..D....-             214
O.SN.99.99SE_MP1300            F-----SEAEA--LG--CDRAH----CN--FE-.-HK-M-K-Q--RS-GST---LIT---L-L--..D....-              212
O.US.10.LTNP                                                                                                        103
N.CM.15.S4858                  F-----MSTEEV--..------NA----ICQ--V--.AHK---V-Q----S--RR--------FY-..D....-             209
N.FR.11.N1_FR_2011             F-----LSAEAV--.-TE-----N------ICQ--V--.DHK---V-----S--RR---K----L-..-                  212
P.CM.06.U14788                 F-----SEVEA--MGD-Q-KAK----ACTY-FS-.-HK-I-V-K---S-GRE---LQK---L-I--..D....-             218
P.FR.09.RBF168                 F-----SEVEV--MGDEQ-KAK----ACTY-DS-.HH--I-M-K--RS-GER---LQK---L-T--..D....-             210
CPZ.TZ.06.TAN5                 F-----D-PE......ED-RNL----ACS--K--..-G-L-I-----S--RR-I---R-------.-..-                 196
CPZ.US.85.US_Marilyn           F-----LTEEEV-Q.------TNI----MCQ----E.LHQ--I-Q--TE--RR-R-K-------.R-.......-             210
GOR.CM.12.SIVgor_BQID2         F-----SEAEA--MG--N-RAK----ACN--YG-.EHGQI-K-Q--RS-GNN---LA------..-                     211
GOR.CM.13.SIVgor_BPID15        F-----SEAEA-YMG--D-RAK----ACTY-FS-.-HK-I-V-KY-RS-GTQ---LMK---L-I--..D....-             211
```