

# V

## HIV-1/SIVcpz Proteins

### Contents

---

|       |                              |            |
|-------|------------------------------|------------|
| V-1   | Introduction . . . . .       | <b>317</b> |
| V-2   | Annotated features . . . . . | <b>318</b> |
| V-3   | Sequences . . . . .          | <b>320</b> |
| V-4   | Alignments . . . . .         | <b>326</b> |
| V-4.1 | Gag . . . . .                | 326        |
| V-4.2 | Pol . . . . .                | 334        |
| V-4.3 | Vif . . . . .                | 348        |
| V-4.4 | Vpr . . . . .                | 352        |
| V-4.5 | Tat . . . . .                | 354        |
| V-4.6 | Rev . . . . .                | 356        |
| V-4.7 | Vpu . . . . .                | 358        |
| V-4.8 | Env . . . . .                | 360        |
| V-4.9 | Nef . . . . .                | 374        |

---

### 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 strange 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.

## V-2 Annotated features

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

| Feature                     | Protein | Location | Page |
|-----------------------------|---------|----------|------|
| Gag start, p17 start        | Gag     | 1        | 326  |
| membrane binding            | Gag     | 1-30     | 326  |
| phosphorylation site        | Gag     | 112      | 326  |
| p17 end                     | Gag     | 132      | 326  |
| p24 start                   | Gag     | 133      | 326  |
| CyPA binding                | Gag     | 205-241  | 328  |
| major homology region       | Gag     | 285-304  | 328  |
| p24 end                     | Gag     | 364      | 330  |
| p2 start                    | Gag     | 365      | 330  |
| p2 end                      | Gag     | 377      | 330  |
| p7 start                    | Gag     | 378      | 330  |
| Zn motif                    | Gag     | 392-404  | 330  |
| Zn motif                    | Gag     | 413-425  | 330  |
| p7 end                      | Gag     | 432      | 330  |
| p1 start                    | Gag     | 433      | 330  |
| p1 end                      | Gag     | 448      | 332  |
| p6 start                    | Gag     | 449      | 332  |
| Vpr binding                 | Gag     | 455-460  | 332  |
| Vpr binding                 | Gag     | 489-494  | 332  |
| p6 end, Gag end             | Gag     | 501      | 332  |
| Pol start                   | Pol     | 1        | 334  |
| Gag-Pol TF start            | Pol     | 1        | 334  |
| Gag-Pol TF end              | Pol     | 56       | 334  |
| protease start              | Pol     | 57       | 334  |
| protease end                | Pol     | 155      | 336  |
| p66, p51 RT start           | Pol     | 156      | 336  |
| M41L                        | Pol     | 196      | 336  |
| D67N                        | Pol     | 222      | 336  |
| K70R                        | Pol     | 225      | 336  |
| D110 catalytic site         | Pol     | 265      | 336  |
| polymerase motif            | Pol     | 337-342  | 338  |
| T215Y                       | Pol     | 370      | 338  |
| K219Q                       | Pol     | 374      | 338  |
| p51 RT end                  | Pol     | 595      | 340  |
| p15 RNase H start           | Pol     | 596      | 340  |
| p66 RT, p15 Rnase H end     | Pol     | 715      | 342  |
| p31 Integrase start         | Pol     | 716      | 342  |
| p31 Integrase end           | Pol     | 1004     | 346  |
| Pol end                     | Pol     | 1004     | 346  |
| Vif start                   | Vif     | 1        | 348  |
| Vif end                     | Vif     | 193      | 350  |
| Vpr start                   | Vpr     | 1        | 352  |
| oligomerization             | Vpr     | 1-41     | 352  |
| amphipathic $\alpha$ -helix | Vpr     | 17-33    | 352  |
| H(S/N)RIG motifs            | Vpr     | 71-83    | 352  |
| frameshift in HXB2          | Vpr     | 72       | 352  |
| Vpr end in HXB2             | Vpr     | 79       | 352  |
| Vpr end                     | Vpr     | 98       | 352  |

| Feature                  | Protein | Location | Page |
|--------------------------|---------|----------|------|
| Tat start                | Tat     | 1        | 354  |
| C-rich region            | Tat     | 22-37    | 354  |
| nuclear localization     | Tat     | 49-57    | 354  |
| exon 1 end               | Tat     | 72       | 354  |
| exon 2 start             | Tat     | 73       | 354  |
| Tat end                  | Tat     | 102      | 354  |
| Rev start                | Rev     | 1        | 356  |
| exon 1 end               | Rev     | 25       | 356  |
| exon 2 start             | Rev     | 26       | 356  |
| NLS                      | Rev     | 34-49    | 356  |
| Leu-rich effector domain | Rev     | 75-83    | 356  |
| Rev end                  | Rev     | 117      | 356  |
| Vpu start                | Vpu     | 1        | 358  |
| transmembrane domain     | Vpu     | 1-27     | 358  |
| cytoplasmic domain       | Vpu     | 28-82    | 358  |
| $\alpha$ -helix          | Vpu     | 43-51    | 358  |
| phos                     | Vpu     | 53       | 358  |
| phos                     | Vpu     | 57       | 358  |
| $\alpha$ -helix          | Vpu     | 58-70    | 358  |
| Vpu end                  | Vpu     | 83       | 358  |
| Env start                | Env     | 1        | 360  |
| signal peptide end       | Env     | 30       | 360  |
| gp120 start              | Env     | 31       | 360  |
| glycosylation NVT        | Env     | 88-90    | 360  |
| CD4 binding              | Env     | 124      | 360  |
| V1                       | Env     | 131-156  | 360  |
| glycosylation NDT        | Env     | 136-138  | 360  |
| glycosylation NSS        | Env     | 141-143  | 360  |
| glycosylation NCS        | Env     | 156-158  | 362  |
| V2                       | Env     | 158-196  | 362  |
| glycosylation NIS        | Env     | 160-162  | 362  |
| glycosylation NDT        | Env     | 186-188  | 362  |
| glycosylation NTS        | Env     | 197-199  | 362  |
| CD4 binding              | Env     | 196      | 362  |
| glycosylation NKT        | Env     | 230-232  | 362  |
| glycosylation NGT        | Env     | 234-236  | 362  |
| glycosylation NVS        | Env     | 241-243  | 362  |
| glycosylation NGS        | Env     | 262-264  | 362  |
| glycosylation NFT        | Env     | 276-278  | 362  |
| CD4 binding              | Env     | 279      | 362  |
| glycosylation NTS        | Env     | 289-291  | 364  |
| glycosylation NCT        | Env     | 295-297  | 364  |
| V3                       | Env     | 296-331  | 364  |
| glycosylation NNT        | Env     | 301-303  | 364  |
| V3 tip                   | Env     | 312-315  | 364  |
| glycosylation NNT        | Env     | 339-341  | 364  |
| glycosylation NKT        | Env     | 356-358  | 364  |
| CD4 binding              | Env     | 365      | 364  |
| V4                       | Env     | 385-418  | 364  |
| glycosylation NST        | Env     | 386-388  | 364  |
| glycosylation NST        | Env     | 392-394  | 364  |
| glycosylation NST        | Env     | 397-399  | 364  |
| glycosylation NNT        | Env     | 406-408  | 364  |

| Feature                     | Protein | Location | Page |
|-----------------------------|---------|----------|------|
| CD4 binding                 | Env     | 425      | 366  |
| glycosylation NIT           | Env     | 448-450  | 366  |
| CD4 binding                 | Env     | 455      | 366  |
| V5                          | Env     | 460-471  | 366  |
| glycosylation NES           | Env     | 463-465  | 366  |
| CD4 binding                 | Env     | 469      | 366  |
| fusion peptide              | Env     | 512-527  | 366  |
| gp120 end                   | Env     | 511      | 366  |
| gp41 start                  | Env     | 512      | 366  |
| immunodominant region       | Env     | 588-607  | 368  |
| glycosylation NAS           | Env     | 611-613  | 368  |
| glycosylation NKS           | Env     | 616-618  | 368  |
| glycosylation NHT           | Env     | 624-626  | 368  |
| glycosylation NYT           | Env     | 637-639  | 368  |
| transmembrane domain        | Env     | 685-704  | 368  |
| gp41 cytoplasmic tail start | Env     | 705      | 368  |
| glycosylation NGS           | Env     | 750-752  | 370  |
| glycosylation NAT           | Env     | 816-818  | 370  |
| cytoplasmic tail end        | Env     | 857      | 372  |
| gp41 end                    | Env     | 857      | 372  |
| Env end                     | Env     | 857      | 372  |
| Nef start                   | Nef     | 1        | 374  |
| myristoylation              | Nef     | 2-7      | 374  |
| acidic cluster              | Nef     | 62-65    | 374  |
| poly-P helix                | Nef     | 69-78    | 374  |
| phosphorylation             | Nef     | 77-81    | 374  |
| HXB2 premature Nef end      | Nef     | 124      | 374  |
| normal Nef end              | Nef     | 207      | 376  |

### 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. | <i>Nature</i> <b>313</b> (6000):277-284 (1985)   |
| A1.AU.03.PS1044_Day0      | DQ676872  | All      | Li, B.         | <i>J Virol</i> <b>81</b> (1):193-201 (2007)  |
| A1.CH.03.HIV_CH_BID_V3538 | JQ403028  | All      | Henn, M.R.     | <i>PLoS Pathog</i> <b>8</b> (3):E1002529 (2012)  |
| A1.ES.06.X2110            | FJ670523  | All      | Cuevas, M.T.   | <i>ARHR</i> <b>26</b> (9); 1019-25 (2010)  |
| A1.IT.02.60000            | EU861977  | All      | Riva, C.       | <i>ARHR</i> <b>24</b> (10); 1319-25 (2008)   |
| A1.KE.11.DEMA111KE002     | KF716474  | All      | Sanchez, A.M.  | <i>J Immunol Methods</i> <b>2014</b> Jan 19 pii: S0022-1759(14)00014-3 doi: 101016/jjim201401004 |
| A1.RU.11.11RU6950         | JX500694  | All      | Baryshev, P.B. | <i>ARHR</i> <b>30</b> (6); 592-7 (2014)  |
| A1.RW.11.DEMA111RW002     | KF716472  | All      | Sanchez, A.M.  | <i>J Immunol Methods</i> <b>2014</b> Jan 19 pii: S0022-1759(14)00014-3 doi: 101016/jjim201401004 |
| A1.SE.95.SE8538           | AF069669  | All      | Carr, J.K.     | <i>AIDS</i> <b>13</b> (14); 1819-26 (1999)   |
| A1.TZ.01.A341             | AY253314  | All      | Arroyo, M.A.   | <i>ARHR</i> <b>20</b> (8):895-901 (2004)   |
| A1.UA.01.01UADN139        | DQ823357  | All      | Saad, M.D.     | <i>ARHR</i> <b>22</b> (8):709-714 (2006)   |

| Name                     | Accession | Proteins | Author         | Reference  |
|--------------------------|-----------|----------|----------------|--|
| A1.UG.11.DEMA110UG009    | KF716486  | All      | Sanchez, A.M.  | <i>J Immunol Methods</i> <b>2014</b> Jan 19<br>pii: S0022-1759(14)00014-3 doi:<br>101016/jjim201401004 |
| A1.ZA.04.04ZASK162B1     | DQ396400  | All      | Rousseau, C.M. | <i>J Virol Methods</i> <b>136</b> (1-2):118-125<br>(2006)  |
| A2.CD.97.97CDKTB48       | AF286238  | All      | Gao, F.        | <i>ARHR</i> <b>17</b> (8):675-688 (2001)   |
| A2.CM.01.01CM_1445MV     | GU201516  | All      | Carr, J.K.     | <i>Retrovirology</i> <b>2010</b> Apr 28;7:39   |
| A2.CY.94.94CY017_41      | AF286237  | All      | Gao, F.        | <i>ARHR</i> <b>17</b> (8):675-688 (2001)   |
| B.AR.04.04AR143170       | DQ383750  | All      | Pando, M.A.    | <i>Retrovirology</i> <b>3</b> , 59 (2006)  |
| B.AU.04.PS1038_Day174    | DQ676871  | All      | Li, B.         | <i>J Virol</i> <b>81</b> (1):193-201 (2007)  |
| B.BR.06.06BR1119         | JN692480  | All      | Sanabani, S.S. | <i>PLoS ONE</i> <b>6</b> (10):E25869 (2011)  |
| B.CA.07.502_1191_03      | JF320424  | All      | Rolland, M.    | <i>Nat Med</i> <b>17</b> (3); 366-71 (2011)  |
| B.CH.08.M2_0803101_NFLG8 | KC797225  | All      | Castro, E.     | <i>AIDS</i> <b>2014</b> Jun 6  |
| B.CN.10.DEMB10CN002      | JX140658  | All      | Sanchez, A.M.  | <i>J Immunol Methods</i> <b>2014</b> Jan 19<br>pii: S0022-1759(14)00014-3 doi:<br>101016/jjim201401004 |
| B.CO.01.PCM001           | AY561236  | All      | Sanchez, G.I.  | <i>Am J Trop Med Hyg</i> <b>74</b> (4); 674-7<br>(2006)  |
| B.CU.99.Cu19             | AY586542  | All      | Sierra, M.     | <i>JAIDS</i> <b>45</b> (2):151-160 (2007)  |
| B.CY.09.CY266            | JF683807  | All      | Kousiappa, I.  | <i>ARHR</i> <b>27</b> (11); 1183-99 (2011)   |
| B.DE.04.HIV_DE_BID_V4131 | JQ403037  | All      | Henn, M.R.     | <i>PLoS Pathog</i> <b>8</b> (3):E1002529<br>(2012)   |
| B.DK.07.PMVL_011         | FJ694790  | All      | Vinner, L.     | <i>APMIS</i> <b>119</b> (8); 487-97 (2011)   |
| B.DO.05.05DO_160884      | EU839597  | All      | Nadai, Y.      | <i>PLoS ONE</i> <b>4</b> (3):E4814 (2009)  |
| B.ES.10.DEMB10ES002      | KC473842  | All      | Sanchez, A.M.  | <i>J Immunol Methods</i> <b>2014</b> Jan 19<br>pii: S0022-1759(14)00014-3 doi:<br>101016/jjim201401004 |
| B.FR.11.DEMB11FR001      | KF716496  | All      | Sanchez, A.M.  | <i>J Immunol Methods</i> <b>2014</b> Jan 19<br>pii: S0022-1759(14)00014-3 doi:<br>101016/jjim201401004 |
| B.GE.03.03GEMZ004        | DQ207940  | All      | Zarandia, M.   | <i>ARHR</i> <b>22</b> (5):470-476 (2006)   |
| B.HT.05.05HT_129389      | EU839602  | All      | Nadai, Y.      | <i>PLoS ONE</i> <b>4</b> (3):E4814 (2009)  |
| B.JM.05.05JM_KJ108       | EU839605  | All      | Nadai, Y.      | <i>PLoS ONE</i> <b>4</b> (3):E4814 (2009)  |
| B.JP.12.DEMB12JP001      | KF716498  | All      | Sanchez, A.M.  | <i>J Immunol Methods</i> <b>2014</b> Jan 19<br>pii: S0022-1759(14)00014-3 doi:<br>101016/jjim201401004 |
| B.KR.12.12KYY10_10742    | KF561441  | All      | Cho            | Unpublished  |
| B.NL.00.671_00T36        | AY423387  | All      | Geels, M.J.    | <i>J Virol</i> <b>77</b> (23):12430-12440 (2003)   |
| B.PE.07.502_2649_wg8     | JF320019  | All      | Rolland, M.    | <i>Nat Med</i> <b>17</b> (3); 366-71 (2011)  |
| B.PY.03.03PY_PSP0115     | JN251906  | All      | Eyzaguirre     | Unpublished  |
| B.RU.11.11RU21n          | JX500708  | All      | Baryshev       | Unpublished  |
| B.TH.07.AA040a_WG11      | JX447156  | All      | Rolland, M.    | <i>Nature</i> <b>490</b> (7420); 417-20 (2012)   |
| B.TT.01.01TT_CRC50069    | EU839608  | All      | Nadai, Y.      | <i>PLoS ONE</i> <b>4</b> (3):E4814 (2009)  |
| B.TW.94.TWCYS_LM49       | AF086817  | All      | Huang          | Unpublished  |
| B.UA.01.01UAKV167        | DQ823362  | All      | Saad, M.D.     | <i>ARHR</i> <b>22</b> (8):709-714 (2006)   |
| B.US.11.ES39_42          | KF384810  | All      | Salgado, M.    | <i>J Virol</i> <b>88</b> (6); 3340-52 (2014)   |
| B.UY.02.02UY_TSU1290     | JN235958  | All      | Eyzaguirre     | Unpublished  |
| B.YE.02.02YE507          | AY795904  | All      | Saad, M.D.     | <i>ARHR</i> <b>21</b> (7):644-648 (2005)   |
| C.AR.01.ARG4006          | AY563170  | All      | Carrion, G.    | <i>ARHR</i> <b>20</b> (9):1022-1025 (2004)   |
| C.BR.07.DEMC07BR003      | JX140663  | All      | Sanchez, A.M.  | <i>J Immunol Methods</i> <b>2014</b> Jan 19<br>pii: S0022-1759(14)00014-3 doi:<br>101016/jjim201401004 |
| C.BW.00.00BW07621        | AF443088  | All      | Novitsky, V.   | <i>J Virol</i> <b>76</b> (11):5435-5451 (2002)   |

| Name                        | Accession | Proteins | Author               | Reference  |
|-----------------------------|-----------|----------|----------------------|--|
| C.CN.09.09YNLX133sg         | KC898996  | All      | Han, X.              | <i>PLoS ONE</i> <b>8</b> (5):E65337 (2013)   |
| C.CY.09.CY260               | JF683803  | All      | Kousiappa, I.        | <i>ARHR</i> <b>27</b> (11); 1183-99 (2011)   |
| C.ES.07.X2118_2             | EU884500  | All      | Fernandez-Garcia, A. | <i>ARHR</i> <b>25</b> (1); 93-102 (2009)   |
| C.ET.02.02ET_288            | AY713417  | All      | Brown, B.K.          | <i>J Virol</i> <b>79</b> (10):6089-6101 (2005)   |
| C.GE.03.03GEMZ033           | DQ207941  | All      | Zarandia, M.         | <i>ARHR</i> <b>22</b> (5):470-476 (2006)   |
| C.IL.98.98IS002             | AF286233  | All      | Rodenburg, C.M.      | <i>ARHR</i> <b>17</b> (2):161-168 (2001)   |
| C.IN.09.T125_2139           | KC156210  | All      | Hora                 | Unpublished  |
| C.KE.00.KER2010             | AF457054  | All      | Dowling, W.E.        | <i>AIDS</i> <b>16</b> (13):1809-1820 (2002)  |
| C.MM.99.mIDU101_3           | AB097871  | All      | Takebe, Y.           | <i>AIDS</i> <b>17</b> (14):2077-87 (2003)  |
| C.MW.09.703010256_CH256.w96 | KC156214  | All      | Hora                 | Unpublished  |
| C.SN.90.90SE_364            | AY713416  | All      | Brown, B.K.          | <i>J Virol</i> <b>79</b> (10):6089-6101 (2005)   |
| C.SO.89.89SM_145            | AY713415  | All      | Brown, B.K.          | <i>J Virol</i> <b>79</b> (10):6089-6101 (2005)   |
| C.TZ.08.707010457_CH457.w8  | KC156220  | All      | Hora                 | Unpublished  |
| C.US.11.17TB4_4G8           | KF526226  | All      | Ho, Y.-C.            | <i>Cell</i> <b>155</b> (3); 540-51 (2013)  |
| C.UY.01.TRA3011             | AY563169  | All      | Carrion, G.          | <i>ARHR</i> <b>20</b> (9):1022-1025 (2004)   |
| C.YE.02.02YE511             | AY795906  | All      | Saad, M.D.           | <i>ARHR</i> <b>21</b> (7):644-648 (2005)   |
| C.ZA.07.704010042_CH042_TF  | KC156123  | All      | Hora                 | Unpublished  |
| C.ZM.11.DEMC11ZM006         | KF716467  | All      | Sanchez, A.M.        | <i>J Immunol Methods</i> <b>2014</b> Jan 19 pii: S0022-1759(14)00014-3 doi: 101016/jjim201401004 |
| D.CD.83.ELI                 | K03454    | All      | Alizon, M.           | <i>Cell</i> <b>46</b> (1):63-74 (1986)   |
| D.CM.10.DEMD10CM009         | JX140670  | All      | Sanchez, A.M.        | <i>J Immunol Methods</i> <b>2014</b> Jan 19 pii: S0022-1759(14)00014-3 doi: 101016/jjim201401004 |
| D.CY.06.CY163               | FJ388945  | All      | Kousiappa, I.        | <i>ARHR</i> <b>25</b> (8); 727-40 (2009)   |
| D.KE.11.DEMD11KE003         | KF716476  | All      | Sanchez, A.M.        | <i>J Immunol Methods</i> <b>2014</b> Jan 19 pii: S0022-1759(14)00014-3 doi: 101016/jjim201401004 |
| D.KR.04.04KBH8              | DQ054367  | All      | Cho, Y.-K.           | <i>ARHR</i> <b>29</b> (4); 738-43 (2013)   |
| D.SN.90.SE365               | AB485648  | All      | Takekawa             | Unpublished  |
| D.TD.99.MN011               | AJ488926  | All      | Vidal, N.            | <i>JAIDS</i> <b>33</b> (2):239-246 (2003)  |
| D.TZ.01.A280                | AY253311  | All      | Arroyo, M.A.         | <i>ARHR</i> <b>20</b> (8):895-901 (2004)   |
| D.UG.11.DEMD11UG003         | KF716480  | All      | Sanchez, A.M.        | <i>J Immunol Methods</i> <b>2014</b> Jan 19 pii: S0022-1759(14)00014-3 doi: 101016/jjim201401004 |
| D.YE.02.02YE516             | AY795907  | All      | Saad, M.D.           | <i>ARHR</i> <b>21</b> (7):644-648 (2005)   |
| D.ZA.90.R1                  | EF633445  | All      | Jacobs, G.B.         | <i>ARHR</i> <b>23</b> (12):1575-8 (2007)   |
| F1.AO.06.AO_06_ANG125       | FJ900269  | All      | Guimaraes, M.L.      | <i>Retrovirology</i> <b>6</b> , 39 (2009)  |
| F1.AR.02.ARE933             | DQ189088  | All      | Aulicino, P.C.       | <i>ARHR</i> <b>21</b> (2):158-164 (2005)   |
| F1.BE.93.VI850              | AF077336  | All      | Laukkanen, T.        | <i>Virology</i> <b>269</b> (1):95-104 (2000)   |
| F1.BR.07.07BR844            | FJ771010  | All      | Sanabani, S.S.       | <i>Virol J</i> <b>2009</b> Jun 16;6:78   |
| F1.CY.08.CY222              | JF683771  | All      | Kousiappa, I.        | <i>ARHR</i> <b>27</b> (11); 1183-99 (2011)   |
| F1.ES.11.DEMF110ES001       | JX140671  | All      | Sanchez, A.M.        | <i>J Immunol Methods</i> <b>2014</b> Jan 19 pii: S0022-1759(14)00014-3 doi: 101016/jjim201401004 |
| F1.FI.93.FIN9363            | AF075703  | All      | Laukkanen, T.        | <i>Virology</i> <b>269</b> (1):95-104 (2000)   |
| F1.FR.96.96FR_MP411         | AJ249238  | All      | Triques, K.          | <i>ARHR</i> <b>16</b> (2):139-151 (2000)   |
| F1.RO.96.BCI_R07            | AB485658  | All      | Takekawa             | Unpublished  |
| F1.RU.08.D88_845            | GQ290462  | All      | Fernandez-Garcia, A. | <i>ARHR</i> <b>25</b> (11):1187-1191 (2009)  |

| Name                      | Accession | Proteins | Author               | Reference  |
|---------------------------|-----------|----------|----------------------|--|
| F2.CM.10.DEMF210CM007     | JX140673  | All      | Sanchez, A.M.        | <i>J Immunol Methods</i> <b>2014</b> Jan 19 pii: S0022-1759(14)00014-3 doi: 101016/jjim201401004 |
| F2.CM.97.CM53657          | AF377956  | All      | Carr, J.K.           | <i>Virology</i> <b>286</b> (1):168-181 (2001)  |
| G.BE.96.DRCBL             | AF084936  | All      | Oelrichs, R.B.       | <i>ARHR</i> <b>15</b> (6):585-589 (1999)   |
| G.CM.10.DEMG10CM008       | JX140676  | All      | Sanchez, A.M.        | <i>J Immunol Methods</i> <b>2014</b> Jan 19 pii: S0022-1759(14)00014-3 doi: 101016/jjim201401004 |
| G.CN.08.GX_2084_08        | JN106043  | All      | Liu, W.              | <i>Zhonghua Liu Xing Bing Xue Za Zhi</i> <b>34</b> (1); 53-6 (2013)                              |
| G.CU.99.Cu74              | AY586547  | All      | Sierra, M.           | <i>JAIDS</i> <b>45</b> (2):151-160 (2007)  |
| G.ES.09.X2634_2           | GU362882  | All      | Cuevas, M.T.         | <i>ARHR</i> <b>26</b> (9); 1019-25 (2010)  |
| G.GH.03.03GH175G          | AB287004  | All      | Takekawa             | Unpublished  |
| G.KE.09.DEMG09KE001       | KF716477  | All      | Sanchez, A.M.        | <i>J Immunol Methods</i> <b>2014</b> Jan 19 pii: S0022-1759(14)00014-3 doi: 101016/jjim201401004 |
| G.NG.09.09NG_SC62         | JN248593  | All      | Charurat, M.         | <i>J Infect Dis</i> <b>2012</b> Feb 21   |
| G.PT.x.PT3306             | FR846409  | All      | Freitas, F.B.        | <i>ARHR</i> <b>2012</b> Sep 25   |
| G.SE.93.SE6165_G6165      | AF061642  | All      | Carr, J.K.           | <i>Virology</i> <b>247</b> (1):22-31 (1998)  |
| H.BE.93.VI991             | AF190127  | All      | Janssens, W.         | <i>AIDS</i> <b>14</b> (11):1533-1543 (2000)  |
| H.BE.93.VI997             | AF190128  | All      | Janssens, W.         | <i>AIDS</i> <b>14</b> (11):1533-1543 (2000)  |
| H.CF.90.056               | AF005496  | All      | Gao, F.              | <i>J Virol</i> <b>72</b> (7):5680-5698 (1998)  |
| H.GB.00.00GBAC4001        | FJ711703  | All      | Holzmayr, V.         | <i>ARHR</i> <b>25</b> (7):721-726 (2009)   |
| J.CD.97.J_97DC_KTB147     | EF614151  | All      | Abecasis, A.B.       | <i>J Virol</i> <b>81</b> (16):8543-8551 (2007)   |
| J.CM.04.04CMU11421        | GU237072  | All      | Yamaguchi, J.        | <i>ARHR</i> <b>26</b> (6); 693-7 (2010)  |
| J.SE.93.SE9280_7887       | AF082394  | All      | Laukkanen, T.        | <i>ARHR</i> <b>15</b> (3):293-297 (1999)   |
| J.SE.94.SE9173_7022       | AF082395  | All      | Laukkanen, T.        | <i>ARHR</i> <b>15</b> (3):293-297 (1999)   |
| K.CD.97.97ZR_EQTB11       | AJ249235  | All      | Triques, K.          | <i>ARHR</i> <b>16</b> (2):139-151 (2000)   |
| K.CM.96.96CM_MP535        | AJ249239  | All      | Triques, K.          | <i>ARHR</i> <b>16</b> (2):139-151 (2000)   |
| 01_AE.AF.07.569M          | GQ477441  | All      | Sanders-Buell, E.    | <i>ARHR</i> <b>26</b> (5):605-608 (2010)   |
| 01_AE.CN.10.CYM059        | JX112796  | All      | Feng, Y.             | <i>AIDS</i> <b>27</b> (11):1793-1802 (2013)  |
| 01_AE.HK.04.HK001         | DQ234790  | All      | Tsui                 | Unpublished  |
| 01_AE.JP.x.DR0492         | AB253423  | All      | Sakamoto             | Unpublished  |
| 01_AE.TH.04.BKM           | DQ314732  | All      | Thitithyanont        | Unpublished  |
| 01_AE.TH.09.AA111a_WG11   | JX448059  | All      | Rolland, M.          | <i>Nature</i> <b>490</b> (7420); 417-20 (2012)   |
| 01_AE.TH.90.CM240         | U54771    | All      | Carr, J.K.           | <i>J Virol</i> <b>70</b> (9):5935-5943 (1996)  |
| 01_AE.VN.98.98VNND15      | FJ185235  | All      | Liao, H.             | <i>Virology</i> <b>391</b> (1):51-56 (2009)  |
| 02_AG.CM.10.DE00210CM013  | KF859739  | All      | Sanchez, A.M.        | <i>J Immunol Methods</i> <b>2014</b> Jan 19 pii: S0022-1759(14)00014-3 doi: 101016/jjim201401004 |
| 02_AG.ES.06.P1261         | EU786671  | All      | Fernandez-Garcia, A. | <i>ARHR</i> <b>25</b> (1); 93-102 (2009)   |
| 02_AG.FR.91.DJ263         | AF063223  | All      | Carr, J.K.           | <i>Virology</i> <b>247</b> (1):22-31 (1998)  |
| 02_AG.GH.03.03GH181AG     | AB286855  | All      | Sakamoto             | Unpublished  |
| 02_AG.KR.12.12MHI11_10746 | KF561437  | All      | Cho                  | Unpublished  |
| 02_AG.LR.x.POC44951       | AB485636  | All      | Takekawa             | Unpublished  |
| 02_AG.NG.x.IBNG           | L39106    | All      | Howard, T.M.         | <i>ARHR</i> <b>10</b> (12):1755-1757 (1994)  |
| 02_AG.SN.98.98SE_MP1211   | AJ251056  | All      | Toure-Kane, C.       | <i>ARHR</i> <b>16</b> (6):603-609 (2000)   |
| 02_AG.US.06.502_2696_FL01 | JF320297  | All      | Rolland, M.          | <i>Nat Med</i> <b>17</b> (3); 366-71 (2011)  |
| 03_AB.RU.97.KAL153_2      | AF193276  | All      | Liitsola, K.         | <i>ARHR</i> <b>16</b> (11):1047-1053 (2000)  |
| 04_cpx.CY.94.94CY032_3    | AF049337  | All      | Gao, F.              | <i>J Virol</i> <b>72</b> (12):10234-10241 (1998)   |
| 05_DF.BE.x.VI1310         | AF193253  | All      | Laukkanen, T.        | <i>Virology</i> <b>269</b> (1):95-104 (2000)   |
| 06_cpx.AU.96.BFP90        | AF064699  | All      | Oelrichs, R.B.       | <i>ARHR</i> <b>14</b> (16):1495-1500 (1998)  |

| Name                       | Accession | Proteins | Author               | Reference  |
|----------------------------|-----------|----------|----------------------|--|
| 07_BC.CN.98.98CN009        | AF286230  | All      | Rodenburg, C.M.      | <i>ARHR</i> <b>17</b> (2):161-168 (2001)   |
| 08_BC.CN.06.nx2            | HM067748  | All      | Miao                 | Unpublished  |
| 09_cpx.GH.96.96GH2911      | AY093605  | All      | McCutchan, F.E.      | <i>ARHR</i> <b>20</b> (8):819-826 (2004)   |
| 10_CD.TZ.96.96TZ_BF061     | AF289548  | All      | Koulinska, I.N.      | <i>ARHR</i> <b>17</b> (5):423-431 (2001)   |
| 11_cpx.CM.95.95CM_1816     | AF492624  | All      | Wilbe, K.            | <i>ARHR</i> <b>18</b> (12):849-56 (2002)   |
| 12_BF.AR.99.ARMA159        | AF385936  | All      | Carr, J.K.           | <i>AIDS</i> <b>15</b> (15):F41-F47 (2001)  |
| 13_cpx.CM.96.96CM_1849     | AF460972  | All      | Wilbe, K.            | <i>ARHR</i> <b>18</b> (12):849-56 (2002)   |
| 14_BG.ES.05.X1870          | FJ670522  | All      | Cuevas, M.T.         | <i>ARHR</i> <b>26</b> (9); 1019-25 (2010)  |
| 15_01B.TH.99.99TH_MU2079   | AF516184  | All      | Viputtijul, K.       | <i>ARHR</i> <b>18</b> (16):1235-1237 (2002)  |
| 16_A2D.KR.97.97KR004       | AF286239  | All      | Gao, F.              | <i>ARHR</i> <b>17</b> (8):675-688 (2001)   |
| 17_BF.AR.99.ARMA038        | AY037281  | All      | Carr, J.K.           | <i>AIDS</i> <b>15</b> (15):F41-F47 (2001)  |
| 18_cpx.CU.99.CU76          | AY586540  | All      | Thomson, M.M.        | <i>AIDS</i> <b>19</b> (11):1155-63 (2005)  |
| 19_cpx.CU.99.CU7           | AY894994  | All      | Casado, G.           | <i>JAIDS</i> <b>40</b> (5):532-537 (2005)  |
| 20_BG.CU.99.Cu103          | AY586545  | All      | Sierra, M.           | <i>JAIDS</i> <b>45</b> (2):151-160 (2007)  |
| 21_A2D.KE.99.KER2003       | AF457051  | All      | Dowling, W.E.        | <i>AIDS</i> <b>16</b> (13):1809-1820 (2002)  |
| 22_01A1.CM.01.01CM_0001BBY | AY371159  | All      | Kijak, G.H.          | <i>ARHR</i> <b>20</b> (5):521-530 (2004)   |
| 23_BG.CU.03.CB118          | AY900571  | All      | Sierra, M.           | <i>JAIDS</i> <b>45</b> (2):151-160 (2007)  |
| 24_BG.ES.08.X2456_2        | FJ670526  | All      | Cuevas, M.T.         | <i>ARHR</i> <b>26</b> (9); 1019-25 (2010)  |
| 25_cpx.CM.02.1918LE        | AY371169  | All      | Kijak, G.H.          | <i>ARHR</i> <b>20</b> (5):521-530 (2004)   |
| 26_AU.CD.02.02CD_MBTB047   | FM877782  | All      | Vidal, N.            | <i>ARHR</i> <b>25</b> (8):823-832 (2009)   |
| 27_cpx.FR.04.04CD_FR_KZS   | AM851091  | All      | Vidal, N.            | <i>ARHR</i> <b>24</b> (2):315-321 (2008)   |
| 28_BF.BR.99.BREPM12609     | DQ085873  | All      | Sa Filho, D.J.       | <i>ARHR</i> <b>22</b> (1):1-13 (2006)  |
| 29_BF.BR.01.BREPM16704     | DQ085876  | All      | Sa Filho, D.J.       | <i>ARHR</i> <b>22</b> (1):1-13 (2006)  |
| 31_BC.BR.04.04BR142        | AY727527  | All      | Sanabani, S.         | <i>ARHR</i> <b>22</b> (2):171-176 (2006)   |
| 32_06A1.EE.01.EE0369       | AY535660  | All      | Adojaan, M.          | <i>JAIDS</i> <b>39</b> (5):598-605 (2005)  |
| 33_01B.ID.07.JKT189_C      | AB547463  | All      | SahBandar, I.N.      | <i>ARHR</i> <b>2010</b> Oct 19   |
| 34_01B.TH.99.OUR2478P      | EF165541  | All      | Tovanabutra, S.      | <i>ARHR</i> <b>23</b> (6):829-833 (2007)   |
| 35_AD.AF.07.169H           | GQ477446  | All      | Sanders-Buell, E.    | <i>ARHR</i> <b>26</b> (5):605-608 (2010)   |
| 36_cpx.CM.00.00CMNYU830    | EF087994  | All      | Powell, R.L.         | <i>ARHR</i> <b>23</b> (8):1008-1019 (2007)   |
| 37_cpx.CM.00.00CMNYU926    | EF116594  | All      | Powell, R.L.         | <i>ARHR</i> <b>23</b> (7):923-933 (2007)   |
| 38_BF1.UY.03.UY03_3389     | FJ213783  | All      | Ruchansky, D.        | <i>ARHR</i> <b>25</b> (3); 351-6 (2009)  |
| 39_BF.BR.04.04BRRJ179      | EU735535  | All      | Guimaraes, M.L.      | <i>AIDS</i> <b>22</b> (3):433-435 (2008)   |
| 40_BF.BR.05.05BRRJ055      | EU735537  | All      | Guimaraes, M.L.      | <i>AIDS</i> <b>22</b> (3):433-435 (2008)   |
| 42_BF.LU.03.luBF_05_03     | EU170155  | All      | Struck               | Unpublished  |
| 43_02G.SA.03.J11223        | EU697904  | All      | Badreddine, S.       | <i>ARHR</i> <b>23</b> (5):667-674 (2007)   |
| 44_BF.CL.00.CH80           | FJ358521  | All      | Delgado, E.          | <i>ARHR</i> <b>26</b> (7); 821-6 (2010)  |
| 45_cpx.FR.04.04FR_AUK      | EU448295  | All      | Frange, P.           | <i>Retrovirology</i> <b>2008</b> Aug 1;5:69  |
| 46_BF.BR.07.07BR_FPS625    | HM026456  | All      | Sanabani, S.S.       | <i>Virol J</i> <b>2010</b> Apr 16;7:74   |
| 47_BF.ES.08.P1942          | GQ372987  | All      | Fernandez-Garcia, A. | <i>ARHR</i> <b>26</b> (7); 827-32 (2010)   |
| 48_01B.MY.07.07MYKT021     | GQ175883  | All      | Li, Y.               | <i>JAIDS</i> <b>54</b> (2):129-136 (2010)  |
| 49_cpx.GM.03.N26677        | HQ385479  | All      | de Silva, T.I.       | <i>Retrovirology</i> <b>7</b> (1):82 (2010)  |
| 51_01B.SG.11.11SG_HM021    | JN029801  | All      | Ng, O.T.             | <i>ARHR</i> <b>2011</b> Sep 23   |
| 52_01B.MY.03.03MYKL018_1   | DQ366664  | All      | Tee, K.K.            | <i>JAIDS</i> <b>43</b> (5):523-529 (2006)  |
| 53_01B.MY.11.11FIR164      | JX390610  | All      | Chow, W.Z.           | <i>J Virol</i> <b>86</b> (20):11398-11399 (2012)   |
| 54_01B.MY.09.09MYSB023     | JX390976  | All      | Ng, K.T.             | <i>J Virol</i> <b>86</b> (20):11405-11406 (2012)   |
| 55_01B.CN.10.HNCS102056    | JX574661  | All      | Han, X.              | <i>Genome Announc</i> <b>1</b> (1):E00050-12 (2013)                                      |
| 57_BC.CN.09.09YNLX19sg     | KC899008  | All      | Han, X.              | <i>PLoS ONE</i> <b>8</b> (5):E65337 (2013)   |
| 61_BC.CN.10.JL100010       | KC990124  | All      | Li, X.               | <i>Genome Announc</i> <b>2013</b> Jun 27;1(3) pii: e00326-13 doi: 101128/genomeA00326-13 |



| Name                  | Accession | Proteins | Author                   | Reference  |
|-----------------------|-----------|----------|--------------------------|--|
| O.BE.87.ANT70         | L20587    | All      | Vanden<br>Haesevelde, M. | <i>J Virol</i> <b>68</b> (3):1586-1596 (1994)      |
| O.CM.91.MVP5180       | L20571    | All      | Gurtler, L.G.            | <i>J Virol</i> <b>68</b> (3):1581-1585 (1994)      |
| O.CM.98.98CMA104      | AY169802  | All      | Yamaguchi, J.            | <i>ARHR</i> <b>19</b> (11):979-988 (2003)          |
| O.FR.92.VAU           | AF407418  | All      | Vartanian, J.P.          | <i>J Gen Virol</i> <b>83</b> (Pt 4):801-805 (2002) |
| O.GA.11.11Gab6352     | JX245015  | All      | Liegeois, F.             | <i>ARHR</i> <b>2013</b> Mar 8                      |
| O.SN.99.99SE_MP1299   | AJ302646  | All      | Toure-Kane, C.           | <i>ARHR</i> <b>17</b> (12):1211-1216 (2001)        |
| O.US.10.LTNP          | JN571034  | All      | Buckheit, R.W.3.         | <i>ARHR</i> <b>30</b> (6); 511-513 (2014)          |
| N.CM.02.DJO0131       | AY532635  | All      | Bodelle, P.              | <i>ARHR</i> <b>20</b> (8):902-908 (2004)           |
| N.CM.04.04CM_1015_04  | DQ017382  | All      | Yamaguchi, J.            | <i>ARHR</i> <b>22</b> (1):83-92 (2006)             |
| N.CM.06.U14842        | GQ324958  | All      | Vallari, A.              | <i>ARHR</i> <b>26</b> (1):109-115 (2010)           |
| N.CM.95.YBF30         | AJ006022  | All      | Simon, F.                | <i>Nat Med</i> <b>4</b> (9):1032-1037 (1998)       |
| N.CM.97.YBF106        | AJ271370  | All      | Roques, P.               | <i>AIDS</i> <b>18</b> (10):1371-1381 (2004)        |
| N.FR.11.N1_FR_2011    | JN572926  | All      | Delaugerre, C.           | <i>Lancet</i> <b>378</b> (9806); 1894 (2011)       |
| P.CM.06.U14788        | HQ179987  | All      | Vallari, A.              | <i>J Virol</i> <b>85</b> (3); 1403-7 (2011)        |
| P.FR.09.RBF168        | GU111555  | All      | Plantier, J.-C.          | <i>Nat Med</i> <b>15</b> (8); 871-2 (2009)         |
| CPZ.CD.06.BF1167      | JQ866001  | All      | Li, Y.                   | <i>J Virol</i> <b>86</b> (19):10776-10791 (2012)   |
| CPZ.CM.05.SIVcpzMT145 | DQ373066  | All      | Keele, B.F.              | <i>Science</i> <b>313</b> (5786):523-526 (2006)    |
| CPZ.GA.88.GAB1        | X52154    | All      | Huet, T.                 | <i>Nature</i> <b>345</b> (6273):356-359 (1990)     |
| CPZ.TZ.06.SIVcpzTAN13 | JQ768416  | All      | Takehisa                 | Unpublished  |
| CPZ.US.85.US_Marilyn  | AF103818  | All      | Gao, F.                  | <i>Nature</i> <b>397</b> (6718):436-441 (1999)     |



B. FR. 83. HXB2
G. KE. 09. DEM609KE001
G. NG. 09. 09NG SC62
G. PT. x. PT3306
G. SE. 93. SE6165 G6165
H. BE. 93. V1991
H. BE. 93. V1997
H. CF. 98. 056
H. GB. 00. 00GBA4001
J. CD. 97. J. 97DC KTB147
J. CM. 04. 04CMU1421
J. SE. 93. SE9280 7887
J. SE. 94. SE9173 7022
K. CD. 97. 97ZR E0TB11
K. CM. 96. 96CM MP535
01 AE. AF. 07. 569M
01 AE. CM. 10. CYN050
01 AE. HK. 04. HK001
01 AE. JP. x. DR0492
01 AE. TH. 04. BKM
01 AE. TH. 09. AA111a WG11
01 AE. TH. 98. C8240
01 AE. VN. 98. 08VND15
02 AG. CM. 10. DE00210CM013
02 AG. ES. 06. P1261
02 AG. FR. 91. D3263
02 AG. GH. 03. 03GH181A6
02 AG. KR. 12. 12MH11 10746
02 AG. LR. x. POC44951
02 AG. NG. x. IBNG
02 AG. SN. 98. 98SE MP1211
02 AG. US. 06. 502 2696 FL01
03 AG. RU. 97. 97R0003
04 CPX. CY. 94. 94CY032 3
05 DF. BE. x. V11310
06 CPX. AU. 96. BFP90
07 BC. CN. 98. 98CN009
09 BC. CN. 06. mx2
09 CPX. GH. 96. 96GH2911
10 CD. TZ. 96. 96TZ BF061
11 CPX. CM. 95. 95CM 1816
12 BF. AR. 99. ARMA159
13 CPX. CM. 96. 96CM 1849
14 BG. ES. 04. 01870
15 01B. TH. 99. 99TH MU2079
16 A2D. KR. 97. 97KR004
17 BF. AR. 99. ARMA038
18 CPX. CU. 99. CU76
19 CPX. CU. 99. CU7
20 BG. CU. 99. CU103
21 A2D. KE. 99. KER2003
22 01A1. CM. 01. 01CM 0001BBY
23 BG. CU. 03. CB118
24 BG. ES. 08. X2456 2
25 CPX. CM. 02. 0219LE
26 AU. CD. 02. 02CD MBT047
27 CPX. FR. 04. 04CD FR KZ5
28 BF. BR. 99. BREPM12609
29 BF. BR. 01. BREPM16704
31 BC. BR. 04. 04BR142
32 06A1. EE. 01. EE0369
33 01B. ID. 07. JKT189 C
34 01B. TH. 99. OUR2478P
35 AD. AF. 07. 169H
36 CPX. CM. 00. 00CMNYU030
37 CPX. CM. 00. 00CMNYU026
38 BF1. UY. 03. UY03 3389
39 BF. BR. 04. 04BRRJ179
40 BF. BR. 05. 05BRRJ055
42 BF. LU. 03. LU06 95 03
43 02G. SA. 03. J1123
44 BF. CL. 00. CH60
45 CPX. FR. 04. 04FR AUK
46 BF. BR. 07. 07BR FP5625
47 BF. ES. 08. P1942
48 01B. MY. 07. 07MYKT021
49 CPX. GM. 03. N26677
51 01B. SG. 11. 11SG HM021
52 01B. MY. 03. 03MYKL018 1
53 01B. MY. 11. 11MYR164
54 01B. MY. 09. 09MYR023
55 01B. CN. 10. 10HNC102056
57 BC. CN. 09. 09NXL193g
61 BC. CN. 10. JL100010
0. BE. 87. ANT70
0. CM. 91. MWP5180
0. CM. 98. 98CMA104
0. FR. 92. VAU
0. GA. 11. 11GAb6352
0. SN. 99. 99SE MP1299
0. US. 10. LTNP
N. CM. 02. DJ00131
N. CM. 04. 04CM 1015 04
N. CM. 06. U1404
N. CM. 95. YBF30
N. CM. 97. YBF106
N. FR. 11. N1 FR 2011
P. CM. 06. U14788
P. FR. 09. RBF168
CPZ. CD. 06. BF1167
CPZ. CM. 05. SIVcpzMT145
CPZ. GA. 98. GAB1
CPZ. TZ. 06. SIVcpzTAN13
CPZ. US. 85. US\_Marilyn

Table with columns: membrane binding, phosphorylation site, and p17 end\_p24 start. It contains a large list of protein sequences with their corresponding amino acid residues and domain annotations.



B. FR. 83. HXB2
G. KE. 09. DEMG09KE001
G. NG. 09. 09NG SC62
G. PT. x. PT3306
G. SE. 93. SE6165 G6165
H. BE. 93. VI991
H. BE. 93. VI997
H. CF. 98. 056
H. GB. 00. 00GBAC4001
J. CD. 97. J 97DC KTB147
J. CM. 04. 04CMU11421
J. SE. 93. SE9280 7887
J. SE. 94. SE9173 7022
K. CD. 97. 97ZR E0TB11
K. CM. 96. 96CM MP535
01 AE. AF. 07. 569M
01 AE. CN. 10. CYM059
01 AE. HK. 04. HK001
01 AE. JP. x. DR0492
01 AE. TH. 04. BKM
01 AE. TH. 09. AA111a WG11
01 AE. TH. 98. C8VND15
01 AE. VN. 98. PL
02 AG. CM. 10. DE00210CM013
02 AG. ES. 06. P1261
02 AG. FR. 91. DJ263
02 AG. GH. 03. 03GH181AG
02 AG. KR. 12. 12MK11 10746
02 AG. LR. x. POC44951
02 AG. NG. x. IBNG
02 AG. SN. 98. 98SE MP1211
02 AG. US. 06. 502 2696 FL01
02 AB. RU. 97. K153
04 cpx. CY. 94. 94CY032 3
05 DF. BE. x. VI1310
06 cpx. AU. 96. BFP90
07 BC. CN. 98. 98CN009
09 BC. CN. 06. mx2
09 cpx. GH. 96. 96GH2911
10 CD. TZ. 96. 96TZ BF061
11 cpx. CM. 95. 95CM 1816
12 BF. AR. 99. ARMA159
13 cpx. CM. 96. 96CM 1849
14 BG. ES. 04. X1870
15 01B. TH. 99. 99TH MU2079
16 A2D. KR. 97. 97KR004
17 BF. AR. 99. ARMA038
18 cpx. CU. 99. CU76
19 cpx. CU. 99. CU7
20 BG. CU. 99. CU103
21 A2D. KE. 99. KER2003
22 01A1. CM. 01. 01CM 0001B8Y
23 BG. CU. 03. CB118
24 BG. ES. 08. X2456 2
25 cpx. CM. 02. 0219LE
26 AU. CD. 02. 02CD MBT047
27 cpx. FR. 04. 04CD FR KZ5
28 BF. BR. 99. BREPM12609
29 BF. BR. 01. BREPM16704
31 BC. BR. 04. 04BR142
32 06A1. EE. 01. E0369
33 01B. ID. 07. JKT189 C
34 01B. TH. 99. OUR2478P
35 AD. AF. 07. 169H
36 cpx. CM. 00. 00CMNYU030
37 cpx. CM. 08. 08CMNYU026
38 BF1. UY. 03. UY03 3389
39 BF. BR. 04. 04BRRJ179
40 BF. BR. 05. 05BRRJ055
42 BF. LU. 03. LU03 05 03
43 02G. SA. 03. J11223
44 BF. CL. 00. CH60
45 cpx. FR. 04. 04FR AUK
46 BF. BR. 07. 07BR FP5625
47 BF. ES. 08. P1942
48 01B. MY. 07. 07MYKT021
49 cpx. GM. 03. N26677
51 01B. SG. 11. 11SG HM021
52 01B. MY. 03. 03MYKL018 1
53 01B. MY. 11. 11F1R164
54 01B. MY. 09. 09MYSB023
55 01B. CN. 18. 18CNCS102056
57 BC. CN. 09. 09YNLX195g
61 BC. CN. 10. J1100010
O. BE. 87. ANT70
O. CM. 91. MWP5180
O. CM. 98. 98CMA104
O. FR. 92. VAU
O. GA. 11. 11GAb6352
O. SN. 99. 99SE MP1299
O. US. 10. LTNP
N. CM. 02. DJ00131
N. CM. 04. 04CM 1015 04
N. CM. 06. U14012
N. CM. 95. YBF30
N. CM. 97. YBF106
N. FR. 11. N1 FR 2011
P. CM. 06. U14788
P. FR. 09. RBF168
CPZ. CD. 06. BF1167
CPZ. CM. 05. SIVcpzMT145
CPZ. GA. 08. GAB1
CPZ. TZ. 06. SIVcpzTAN13
CPZ. US. 85. US Marilyn

Table with columns for sequence alignment and major homology region. Headers include 'CyPA binding' and 'major homology region'. The table contains a dense grid of amino acid residues and gaps for various HIV-1/SIVcpz proteins.



Table with columns: major homology region, p24 end p2 start, p2 end p7 start, Zn motif, Zn motif, p7 end p1 start. The table contains a list of HIV-1 sequences and their corresponding amino acid residues.

|                             | p1 end_p6 start | Vpr binding | Vpr binding | p6 end_Gag end |
|-----------------------------|-----------------|-------------|-------------|----------------|
| B.FR.83.HXB2                | GRPGNFLQ        | SRPEPT      | APPEE.SFRS  | GVE            |
| A1.AU.03.PS1044.Day0        | P               | N-L         | A-LGM       | E              |
| A1.CH.03.HIV.CH.BID.V3538   | P               | N-L         | A-LGM       | E              |
| A1.ES.06.X2110              | P               | N-L         | A-LGM       | E              |
| A1.IT.02.60000              | P               | N-L         | A-LGM       | E              |
| A1.KE.11.DEMA111KE002       | P               | N-L         | A-LGM       | E              |
| A1.RU.11.RU6950             | P               | N-L         | A-LGM       | E              |
| A1.RW.11.DEMA111RW002       | P               | N-L         | A-LGM       | E              |
| A1.SE.95.SE8538             | P               | N-L         | A-LGM       | E              |
| A1.TZ.01.A341               | P               | N-L         | A-LGM       | E              |
| A1.UA.01.01UADN139          | P               | N-L         | A-LGM       | E              |
| A1.UG.11.DEMA110UG009       | P               | N-L         | A-LGM       | E              |
| A1.ZA.04.04ZASK162B1        | P               | N-L         | A-LGM       | E              |
| A2.CD.97.97CDKTB48          | P               | N-L         | A-LGM       | E              |
| A2.CM.01.01CM.1445MV        | P               | N-L         | A-LGM       | E              |
| A2.CY.94.94CY017.41         | P               | N-L         | A-LGM       | E              |
| B.AR.04.04AR143170          | P               | N-L         | A-LGM       | E              |
| B.AU.04.PS1038.Day174       | P               | N-L         | A-LGM       | E              |
| B.BR.06.06BR1119            | P               | N-L         | A-LGM       | E              |
| B.CA.07.502.1191.03         | P               | N-L         | A-LGM       | E              |
| B.CH.08.M2.0803101.NFLG8    | P               | N-L         | A-LGM       | E              |
| B.CN.10.DEMB10CN002         | P               | N-L         | A-LGM       | E              |
| B.CO.01.PCM001              | P               | N-L         | A-LGM       | E              |
| B.CU.99.CU19                | P               | N-L         | A-LGM       | E              |
| B.CY.09.CY266               | P               | N-L         | A-LGM       | E              |
| B.DE.04.HIV.DE.BID.V4131    | P               | N-L         | A-LGM       | E              |
| B.DK.07.PMVL.011            | P               | N-L         | A-LGM       | E              |
| B.DO.05.05DO.160884         | P               | N-L         | A-LGM       | E              |
| B.ES.10.DEMB10ES002         | P               | N-L         | A-LGM       | E              |
| B.FR.11.DEMB11FR001         | P               | N-L         | A-LGM       | E              |
| B.GE.03.03GEMZ004           | P               | N-L         | A-LGM       | E              |
| B.HT.05.05HT.129389         | P               | N-L         | A-LGM       | E              |
| B.JM.05.05JM.KJ108          | P               | N-L         | A-LGM       | E              |
| B.JP.12.DEMB12JP001         | P               | N-L         | A-LGM       | E              |
| B.KR.12.12KYY18.18742       | P               | N-L         | A-LGM       | E              |
| B.NL.00.671.00T36           | P               | N-L         | A-LGM       | E              |
| B.PE.07.502.2649.wq8        | P               | N-L         | A-LGM       | E              |
| B.PY.03.03PY.PSP0115        | P               | N-L         | A-LGM       | E              |
| B.RU.11.11RU21n             | P               | N-L         | A-LGM       | E              |
| B.TH.07.AA0409.WG11         | P               | N-L         | A-LGM       | E              |
| B.TT.01.01TT.CRC50069       | P               | N-L         | A-LGM       | E              |
| B.TW.94.TWCYS.LM49          | P               | N-L         | A-LGM       | E              |
| B.UA.01.01UAKV167           | P               | N-L         | A-LGM       | E              |
| B.US.11.E539.42             | P               | N-L         | A-LGM       | E              |
| B.UY.02.02UY.TSUI290        | P               | N-L         | A-LGM       | E              |
| B.YE.02.02YE507             | P               | N-L         | A-LGM       | E              |
| C.AR.01.ARG4006             | P               | N-L         | A-LGM       | E              |
| C.BR.07.DEMC07BR003         | P               | N-L         | A-LGM       | E              |
| C.BW.00.00BW07621           | P               | N-L         | A-LGM       | E              |
| C.CN.09.09YNLX1335g         | P               | N-L         | A-LGM       | E              |
| C.CY.09.CY266               | P               | N-L         | A-LGM       | E              |
| C.ES.07.X2118.2             | P               | N-L         | A-LGM       | E              |
| C.ET.02.02ET.268            | P               | N-L         | A-LGM       | E              |
| C.GE.03.03GEMZ033           | P               | N-L         | A-LGM       | E              |
| C.IL.98.98IS002             | P               | N-L         | A-LGM       | E              |
| C.IN.09.T125.2139           | P               | N-L         | A-LGM       | E              |
| C.KE.00.KER2016             | P               | N-L         | A-LGM       | E              |
| C.MM.99.mIDU101.3           | P               | N-L         | A-LGM       | E              |
| C.MW.09.703010256.CH256.w96 | P               | N-L         | A-LGM       | E              |
| C.SN.90.90SE.364            | P               | N-L         | A-LGM       | E              |
| C.SO.89.89SM.145            | P               | N-L         | A-LGM       | E              |
| C.TZ.08.707010457.CH457.w8  | P               | N-L         | A-LGM       | E              |
| C.UY.11.17B4.468            | P               | N-L         | A-LGM       | E              |
| C.UY.01.TRA3011             | P               | N-L         | A-LGM       | E              |
| C.YE.02.02YE511             | P               | N-L         | A-LGM       | E              |
| C.ZA.07.704010042.CH042.TF  | P               | N-L         | A-LGM       | E              |
| C.ZM.11.DEMC11ZM006         | P               | N-L         | A-LGM       | E              |
| D.CD.83.ELI                 | P               | N-L         | A-LGM       | E              |
| D.CM.10.DEMD10CM009         | P               | N-L         | A-LGM       | E              |
| D.CY.06.CY163               | P               | N-L         | A-LGM       | E              |
| D.KE.11.DEMD11KE003         | P               | N-L         | A-LGM       | E              |
| D.KR.04.04KBH8              | P               | N-L         | A-LGM       | E              |
| D.SN.90.SE365               | P               | N-L         | A-LGM       | E              |
| D.TD.99.MN011               | P               | N-L         | A-LGM       | E              |
| D.TZ.01.A280                | P               | N-L         | A-LGM       | E              |
| D.UG.11.DEMD11UG003         | P               | N-L         | A-LGM       | E              |
| D.YE.02.02YE516             | P               | N-L         | A-LGM       | E              |
| D.ZA.90.R1                  | P               | N-L         | A-LGM       | E              |
| F1.A0.06.A0.06.ANG125       | P               | N-L         | A-LGM       | E              |
| F1.AR.02.ARE933             | P               | N-L         | A-LGM       | E              |
| F1.BE.93.VT850              | P               | N-L         | A-LGM       | E              |
| F1.BR.07.07BR844            | P               | N-L         | A-LGM       | E              |
| F1.CY.08.CY222              | P               | N-L         | A-LGM       | E              |
| F1.ES.11.DEMF110ES001       | P               | N-L         | A-LGM       | E              |
| F1.FI.93.FIN9363            | P               | N-L         | A-LGM       | E              |
| F1.FR.96.96FR.NP411         | P               | N-L         | A-LGM       | E              |
| F1.R0.96.BC1.R07            | P               | N-L         | A-LGM       | E              |
| F1.RU.08.D88.845            | P               | N-L         | A-LGM       | E              |
| F2.CM.10.DEMF210CM007       | P               | N-L         | A-LGM       | E              |
| F2.CM.97.CM53657            | P               | N-L         | A-LGM       | E              |
| G.BE.96.DRCBL               | P               | N-L         | A-LGM       | E              |
| G.CM.10.DEMG10CM008         | P               | N-L         | A-LGM       | E              |
| G.CN.08.GX.2084.08          | P               | N-L         | A-LGM       | E              |
| G.CU.99.CU74                | P               | N-L         | A-LGM       | E              |
| G.ES.09.X2634.2             | P               | N-L         | A-LGM       | E              |
| G.GH.03.03GH175G            | P               | N-L         | A-LGM       | E              |

500  
500  
499  
499  
499  
497  
495  
500  
497  
498  
497  
501  
504  
498  
497  
497  
497  
501  
501  
501  
503  
504  
503  
499  
502  
503  
510  
503  
501  
501  
499  
496  
500  
503  
500  
513  
504  
500  
510  
501  
504  
501  
506  
504  
498  
503  
492  
491  
493  
492  
498  
491  
497  
493  
491  
491  
494  
486  
504  
492  
508  
488  
489  
489  
499  
501  
504  
500  
501  
500  
507  
497  
501  
501  
501  
464  
495  
494  
507  
493  
495  
496  
497  
495  
501  
494  
493  
501  
501  
499  
501  
501  
501  
493  
499  
501  
497  
502  
499



|                            | p1 end   | p6 start | Vpr binding |      | Vpr binding |        | p6 end, Gag end |     |    |                   |
|----------------------------|----------|----------|-------------|------|-------------|--------|-----------------|-----|----|-------------------|
|                            | GRPGNFLQ | SRPEPT   | APPEE       | SFRS | GVE         | TTTTPP | QKQPEI          | DKE | LY | PLTSLRSLFGNDPSSQ* |
| B.FR.83.HXB2               |          |          |             |      |             |        |                 |     |    |                   |
| G.KE.09.DEMG09KE001        |          |          |             |      |             |        |                 |     |    |                   |
| G.NG.09.09NG SC62          |          |          |             |      |             |        |                 |     |    |                   |
| G.PT.x.PT3306              |          |          |             |      |             |        |                 |     |    |                   |
| G.SE.93.SE6165 G6165       |          |          |             |      |             |        |                 |     |    |                   |
| H.BE.93.VI991              |          |          |             |      |             |        |                 |     |    |                   |
| H.BE.93.VI997              |          |          |             |      |             |        |                 |     |    |                   |
| H.CF.98.056                |          |          |             |      |             |        |                 |     |    |                   |
| H.GB.00.00GBAC4001         |          |          |             |      |             |        |                 |     |    |                   |
| J.CD.97.J.97DC KTB147      |          |          |             |      |             |        |                 |     |    |                   |
| J.CM.04.04CMU11421         |          |          |             |      |             |        |                 |     |    |                   |
| J.SE.93.SE9280 7887        |          |          |             |      |             |        |                 |     |    |                   |
| J.SE.94.SE9173 7022        |          |          |             |      |             |        |                 |     |    |                   |
| K.CD.97.97ZR E0TB11        |          |          |             |      |             |        |                 |     |    |                   |
| K.CM.96.96CM MP535         |          |          |             |      |             |        |                 |     |    |                   |
| 01 AE.AF.07.569M           |          |          |             |      |             |        |                 |     |    |                   |
| 01 AE.CM.10.CYM059         |          |          |             |      |             |        |                 |     |    |                   |
| 01 AE.HK.04.HK001          |          |          |             |      |             |        |                 |     |    |                   |
| 01 AE.JP.x.DR0492          |          |          |             |      |             |        |                 |     |    |                   |
| 01 AE.TH.04.BKM            |          |          |             |      |             |        |                 |     |    |                   |
| 01 AE.TH.09.AA111a WG11    |          |          |             |      |             |        |                 |     |    |                   |
| 01 AE.TH.98.CH240          |          |          |             |      |             |        |                 |     |    |                   |
| 01 AE.VN.98.98VND15        |          |          |             |      |             |        |                 |     |    |                   |
| 02 AG.CM.10.DE00210CM013   |          |          |             |      |             |        |                 |     |    |                   |
| 02 AG.ES.06.P1261          |          |          |             |      |             |        |                 |     |    |                   |
| 02 AG.FR.91.DJ263          |          |          |             |      |             |        |                 |     |    |                   |
| 02 AG.GH.03.03GH181AG      |          |          |             |      |             |        |                 |     |    |                   |
| 02 AG.KR.12.12MH11 10746   |          |          |             |      |             |        |                 |     |    |                   |
| 02 AG.LR.x.P0C44951        |          |          |             |      |             |        |                 |     |    |                   |
| 02 AG.NG.x.IBNG            |          |          |             |      |             |        |                 |     |    |                   |
| 02 AG.SN.98.98SE MP1211    |          |          |             |      |             |        |                 |     |    |                   |
| 02 AG.US.06.502 2696 FL01  |          |          |             |      |             |        |                 |     |    |                   |
| 03 AB.RU.97.KAL15 3        |          |          |             |      |             |        |                 |     |    |                   |
| 04 cpx.CY.94.94CY032 3     |          |          |             |      |             |        |                 |     |    |                   |
| 05 DF.BE.x.VI1310          |          |          |             |      |             |        |                 |     |    |                   |
| 06 cpx.AU.96.BFP90         |          |          |             |      |             |        |                 |     |    |                   |
| 07 BC.CN.98.98CN009        |          |          |             |      |             |        |                 |     |    |                   |
| 08 BC.CN.06.mx2            |          |          |             |      |             |        |                 |     |    |                   |
| 08 cpx.GH.96.96GH2911      |          |          |             |      |             |        |                 |     |    |                   |
| 10 CD.TZ.96.96TZ BF061     |          |          |             |      |             |        |                 |     |    |                   |
| 11 cpx.CM.95.95CM 1816     |          |          |             |      |             |        |                 |     |    |                   |
| 12 BF.AR.99.ARMA159        |          |          |             |      |             |        |                 |     |    |                   |
| 13 cpx.CM.96.96CM 1849     |          |          |             |      |             |        |                 |     |    |                   |
| 14 BG.ES.05.X1870          |          |          |             |      |             |        |                 |     |    |                   |
| 15 01B.TH.99.99TH MU2079   |          |          |             |      |             |        |                 |     |    |                   |
| 16 A2D.KR.97.97KR004       |          |          |             |      |             |        |                 |     |    |                   |
| 17 BF.AR.99.ARMA038        |          |          |             |      |             |        |                 |     |    |                   |
| 18 cpx.CU.99.CU76          |          |          |             |      |             |        |                 |     |    |                   |
| 19 cpx.CU.99.CU7           |          |          |             |      |             |        |                 |     |    |                   |
| 20 BG.CU.99.CU103          |          |          |             |      |             |        |                 |     |    |                   |
| 21 A2D.KE.99.KER2003       |          |          |             |      |             |        |                 |     |    |                   |
| 22 01A1.CM.01.01CM 0001BBY |          |          |             |      |             |        |                 |     |    |                   |
| 23 BG.CU.03.CB118          |          |          |             |      |             |        |                 |     |    |                   |
| 24 BG.ES.08.X2456 2        |          |          |             |      |             |        |                 |     |    |                   |
| 25 cpx.CM.02.1919LE        |          |          |             |      |             |        |                 |     |    |                   |
| 26 AU.CD.02.02CD MBT047    |          |          |             |      |             |        |                 |     |    |                   |
| 27 cpx.FR.04.04CD FR KZ5   |          |          |             |      |             |        |                 |     |    |                   |
| 28 BF.BR.99.BREPM12609     |          |          |             |      |             |        |                 |     |    |                   |
| 29 BF.BR.01.BREPM16704     |          |          |             |      |             |        |                 |     |    |                   |
| 31 BC.BR.04.04BR142        |          |          |             |      |             |        |                 |     |    |                   |
| 32 06A1.EE.01.EE0369       |          |          |             |      |             |        |                 |     |    |                   |
| 33 01B.ID.07.JKT189 C      |          |          |             |      |             |        |                 |     |    |                   |
| 34 01B.TH.99.OUR2478P      |          |          |             |      |             |        |                 |     |    |                   |
| 35 AD.AF.07.169H           |          |          |             |      |             |        |                 |     |    |                   |
| 36 cpx.CM.00.00CMNYU030    |          |          |             |      |             |        |                 |     |    |                   |
| 37 cpx.CM.00.00CMNYU026    |          |          |             |      |             |        |                 |     |    |                   |
| 38 BF1.UY.03.UY03 3389     |          |          |             |      |             |        |                 |     |    |                   |
| 39 BF.BR.04.04BRR179       |          |          |             |      |             |        |                 |     |    |                   |
| 40 BF.BR.05.05BRR1055      |          |          |             |      |             |        |                 |     |    |                   |
| 42 BF.LU.03.LuBF 05 03     |          |          |             |      |             |        |                 |     |    |                   |
| 43 02G.SA.03.J11223        |          |          |             |      |             |        |                 |     |    |                   |
| 44 BF.CL.00.CH60           |          |          |             |      |             |        |                 |     |    |                   |
| 45 cpx.FR.04.04FR AUK      |          |          |             |      |             |        |                 |     |    |                   |
| 46 BF.BR.07.07BR FPS625    |          |          |             |      |             |        |                 |     |    |                   |
| 47 BF.ES.08.P1942          |          |          |             |      |             |        |                 |     |    |                   |
| 48 01B.MY.07.07MYKT021     |          |          |             |      |             |        |                 |     |    |                   |
| 49 cpx.GM.03.NZ6677        |          |          |             |      |             |        |                 |     |    |                   |
| 51 01B.SG.11.11SG HM021    |          |          |             |      |             |        |                 |     |    |                   |
| 52 01B.MY.03.03MYKL018 1   |          |          |             |      |             |        |                 |     |    |                   |
| 53 01B.MY.11.11F1R164      |          |          |             |      |             |        |                 |     |    |                   |
| 54 01B.MY.09.09MYSB023     |          |          |             |      |             |        |                 |     |    |                   |
| 55 01B.CN.18.HNC102056     |          |          |             |      |             |        |                 |     |    |                   |
| 57 BC.CN.09.09YNLX19sg     |          |          |             |      |             |        |                 |     |    |                   |
| 61 BC.CN.10.JL100010       |          |          |             |      |             |        |                 |     |    |                   |
| 0.BE.87.ANT70              |          |          |             |      |             |        |                 |     |    |                   |
| 0.CM.91.MVP5180            |          |          |             |      |             |        |                 |     |    |                   |
| 0.CM.98.98CMA104           |          |          |             |      |             |        |                 |     |    |                   |
| 0.FR.92.VAU                |          |          |             |      |             |        |                 |     |    |                   |
| 0.GA.11.11GAb6352          |          |          |             |      |             |        |                 |     |    |                   |
| 0.SN.99.99SE MP1299        |          |          |             |      |             |        |                 |     |    |                   |
| 0.US.10.LTNP               |          |          |             |      |             |        |                 |     |    |                   |
| N.CM.02.DJ00131            |          |          |             |      |             |        |                 |     |    |                   |
| N.CM.04.04CM 1015 04       |          |          |             |      |             |        |                 |     |    |                   |
| N.CM.06.U14842             |          |          |             |      |             |        |                 |     |    |                   |
| N.CM.95.YBF30              |          |          |             |      |             |        |                 |     |    |                   |
| N.CM.97.YBF106             |          |          |             |      |             |        |                 |     |    |                   |
| N.FR.11.N1 FR 2011         |          |          |             |      |             |        |                 |     |    |                   |
| P.CM.06.U14788             |          |          |             |      |             |        |                 |     |    |                   |
| P.FR.09.RBF168             |          |          |             |      |             |        |                 |     |    |                   |
| CPZ.CD.06.BF1167           |          |          |             |      |             |        |                 |     |    |                   |
| CPZ.CM.05.SIVcpzMT145      |          |          |             |      |             |        |                 |     |    |                   |
| CPZ.GA.08.GAB1             |          |          |             |      |             |        |                 |     |    |                   |
| CPZ.TZ.06.SIVcpzTAN13      |          |          |             |      |             |        |                 |     |    |                   |
| CPZ.US.85.US_Marilyn       |          |          |             |      |             |        |                 |     |    |                   |

500  
497  
493  
498  
498  
501  
500  
488  
497  
498  
498  
496  
496  
499  
496  
499  
498  
496  
496  
499  
498  
495  
495  
496  
496  
499  
495  
495  
495  
495  
495  
499  
501  
503  
500  
496  
487  
497  
497  
491  
497  
494  
508  
497  
498  
499  
499  
501  
497  
499  
500  
498  
497  
499  
499  
510  
506  
499  
497  
498  
497  
499  
501  
499  
498  
493  
499  
502  
500  
503  
498  
501  
499  
496  
499  
501  
499  
501  
499  
499  
500  
501  
501  
500  
500  
501  
513  
513  
513  
513  
513  
512  
513  
495  
497  
501  
499  
509  
501  
507  
526

HIV-1 Proteins

Table with 3 main columns: Pol start / Gag-Pol TF start, Gag-Pol TF end, protease start. Rows list various HIV-1 protein sequences and their amino acid compositions across different positions.

Pol start  
Gag-Pol TF start

Gag-Pol TF end, protease start

Table of HIV-1/SIV sequences with columns for various proteins and domains (e.g., Pol start, Gag-Pol TF start, Gag-Pol TF end, protease start) and corresponding amino acid sequences.

HIV-1/SIV gpz Proteins

|                                | protease end_p66, p51 RT start   | M41L     | K70R<br>D67N <sup>TM</sup> | D110 catalytic site |        |     |
|--------------------------------|--|----------|----------------------------|---------------------|--------|-----|
| B. FR. 83. HXB2                | IKVQRQDQLLEICGHKAIGTLVVGPTPVNIIGRNLLTIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTVPVFAIKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKSVTLVDVGDVYFVSVPLDEDF | MXX      | X                          | A                   | T D I  | 279 |
| A1. AU. 03. PS1044 Day0        | ---K---S---KR  | ---      | ---                        | ---                 | ---N   | 279 |
| A1. CH. 03. HIV_CH_BID_V3538   | ---K---K---K   | M---     | T---                       | ---                 | ---H-S | 279 |
| A1. ES. 06. X2110              | ---K---A---K   | M L---   | ---                        | R---                | ---I   | 279 |
| A1. IT. 02. 60000              | ---K---K---K   | M---     | X---                       | X---                | ---X   | 279 |
| A1. KE. 11. DEM111KE002        | ---K---K---K   | M---     | ---                        | ---                 | ---X   | 279 |
| A1. RU. 11. DERU6950           | ---K---K---K   | M L---   | T---                       | ---                 | ---S   | 279 |
| A1. RW. 11. DEM111RW002        | ---K---K---K   | M---     | ---                        | ---                 | ---G   | 279 |
| A1. SE. 95. SE8538             | ---K---S---K   | M---     | ---                        | ---                 | ---I   | 279 |
| A1. TZ. 01. A341               | ---K---K---K   | M---     | I---                       | ---                 | ---NR  | 279 |
| A1. UA. 01. 01UA0N139          | ---K---K---K   | M---     | ---                        | ---                 | ---S   | 279 |
| A1. UG. 11. DEM1110UG009       | ---K E---  | M---     | Q---                       | ---                 | ---V   | 279 |
| A1. ZA. 04. 04ZASK162B1        | ---K---K---K   | M---     | K R---                     | ---                 | ---I   | 278 |
| A2. CD. 97. 97CDKT848          | ---V---KR  | M V L--- | ---                        | ---                 | ---D   | 272 |
| A2. CM. 01. 01CM_1445MV        | ---V---EKR   | I V L--- | ---                        | ---                 | ---Y   | 279 |
| A2. CY. 94. 94CY017_41         | ---A---KR  | M V L--- | ---                        | ---                 | ---A   | 279 |
| B. AR. 04. 04AR143170          | ---K E VT---   | VT I---  | ---                        | ---                 | ---    | 279 |
| B. AU. 04. PS1038 Day174       | ---S X---VT X I---   | LX---    | ---                        | ---                 | ---    | 279 |
| B. BR. 06. 06BR1119            | ---M---  | ---      | ---                        | IP---               | X K I  | 279 |
| B. CA. 07. 502_1191_03         | ---K P---T---  | ---      | ---                        | ---                 | ---    | 279 |
| B. CH. 08. M2_0803101_NFLG8    | R K P---V I---   | VM L---  | ---                        | ---                 | Q N    | 283 |
| B. CN. 10. DEMB10CN002         | ---K H---V T I---  | M L---   | ---                        | ---                 | ---    | 279 |
| B. CO. 01. PCM001              | ---H V---  | ---      | ---                        | ---                 | ---    | 279 |
| B. CU. 99. CU19                | ---P---  | ---      | ---                        | ---                 | ---    | 282 |
| B. CY. 09. CY266               | ---P---  | L D---   | ---                        | ---                 | ---    | 279 |
| B. DE. 04. HIV_DE_BID_V4131    | ---P---  | ---      | ---                        | ---                 | ---    | 279 |
| B. DK. 07. PMVL_011            | ---E VP---   | V I---   | LX---                      | ---                 | ---    | 279 |
| B. DO. 05. 05DO_160884         | ---V S V---  | I M---   | ---                        | ---                 | ---    | 281 |
| B. ES. 10. DEM10E5002          | ---V H---  | ---      | ---                        | ---                 | ---    | 278 |
| B. FR. 11. DEMB11FR001         | ---V H T---  | I---     | ---                        | ---                 | ---    | 279 |
| B. GE. 03. 03GEMZ004           | ---V H T---  | I---     | ---                        | ---                 | ---    | 279 |
| B. HT. 05. 05HT_129389         | ---E N P---  | L X---   | ---                        | ---                 | ---    | 278 |
| B. JM. 05. 05JM_KJ108          | ---K V V---  | L X---   | ---                        | ---                 | ---    | 278 |
| B. JP. 12. DEMB12JP001         | ---K V V---  | L X---   | ---                        | ---                 | ---    | 278 |
| B. KR. 12. 12KYY18_18742       | ---P S---T---  | L---     | ---                        | ---                 | ---    | 279 |
| B. NL. 00. 671_00T36           | ---P---  | L D---   | ---                        | ---                 | ---    | 289 |
| B. PE. 07. 502_2649_wq8        | ---E P M---  | ---      | ---                        | ---                 | ---    | 283 |
| B. PY. 03. 03PY_PSP0115        | ---V P---  | ---      | ---                        | ---                 | ---    | 280 |
| B. RU. 11. 11RU21n             | ---P---  | L---     | ---                        | ---                 | ---    | 283 |
| B. TH. 07. A0A0489_WG11        | ---E P---  | V I---   | ---                        | ---                 | ---    | 279 |
| B. TT. 01. 01TT_CRC50069       | ---P---  | D---     | ---                        | ---                 | ---    | 278 |
| B. TW. 94. TWCYS_LM49          | ---V---  | I---     | ---                        | ---                 | ---    | 279 |
| B. UA. 01. 01UAKV167           | ---T V---  | I L---   | ---                        | ---                 | ---    | 286 |
| B. US. 11. E539_42             | ---E P S---  | S R---   | ---                        | ---                 | ---    | 279 |
| B. UY. 02. 02UY_T5U1290        | ---E P S---  | S R---   | ---                        | ---                 | ---    | 280 |
| B. YE. 02. 02YE507             | ---V M K---  | M L---   | ---                        | ---                 | ---    | 275 |
| C. AR. 01. ARG4006             | ---P K---  | M L---   | ---                        | ---                 | ---    | 275 |
| C. BR. 07. DEMC07BR003         | ---P K---  | M L---   | ---                        | ---                 | ---    | 275 |
| C. BW. 00. 00BW07621           | ---P K R---  | M L---   | ---                        | ---                 | ---    | 275 |
| C. CN. 09. 09YNLX1335g         | ---C N---  | K---     | ---                        | ---                 | ---    | 275 |
| C. CY. 09. CY266               | X---X K---   | X I---   | LX S---                    | ---                 | ---    | 275 |
| C. ES. 07. X2118_2             | ---H K---  | M L---   | ---                        | ---                 | ---    | 275 |
| C. ET. 02. 02ET_288            | ---V K---  | S I---   | M L---                     | ---                 | ---    | 281 |
| C. GE. 03. 03GEMZ033           | ---C K---  | K I---   | M L---                     | ---                 | ---    | 275 |
| C. IL. 98. 98IS002             | ---P K---  | T I---   | M L---                     | ---                 | ---    | 279 |
| C. IN. 09. T125_2139           | ---C K---  | T I---   | M L---                     | ---                 | ---    | 276 |
| C. KE. 00. KERA203             | ---S M K---  | T I---   | M L---                     | ---                 | ---    | 275 |
| C. MM. 99. MIDU101_3           | ---S K---  | M L---   | ---                        | ---                 | ---    | 275 |
| C. MW. 09. 703010256_CH256_w96 | ---P K---  | M L---   | ---                        | ---                 | ---    | 277 |
| C. SN. 90. 90SE_364            | ---P K---  | M L---   | ---                        | ---                 | ---    | 271 |
| C. SO. 89. 89SM_145            | ---V K---  | S I---   | M L---                     | ---                 | ---    | 286 |
| C. TZ. 08. 707010457_CH457_w8  | ---V K---  | S I---   | M L---                     | ---                 | ---    | 289 |
| C. UY. 01. TRA3011             | ---V K---  | S I---   | M L---                     | ---                 | ---    | 275 |
| C. YE. 02. 02YE511             | ---K R---  | M L---   | ---                        | ---                 | ---    | 273 |
| C. ZA. 07. 704010042_CH042_TF  | ---E P K---  | M L---   | ---                        | ---                 | ---    | 270 |
| C. ZM. 11. DEMC11ZM006         | ---P O---  | ---      | ---                        | ---                 | ---    | 280 |
| D. CD. 83. ELT                 | ---P O---  | ---      | ---                        | ---                 | ---    | 278 |
| D. CM. 10. DEMD10CM009         | ---G Y---  | I V R--- | ---                        | ---                 | ---    | 284 |
| D. CY. 06. CY163               | ---K N---  | P V T--- | ---                        | ---                 | ---    | 278 |
| D. KE. 11. DEMD11KE003         | ---K N---  | Q---     | A---                       | ---                 | ---    | 278 |
| D. KR. 04. 04KBH8              | ---K H---  | S V D--- | T D---                     | ---                 | ---    | 278 |
| D. SN. 90. SE365               | ---E V V---  | I---     | L---                       | ---                 | ---    | 278 |
| D. TD. 99. MM011               | ---P L---  | ---      | ---                        | ---                 | ---    | 284 |
| D. UG. 11. DEMD11UG003         | ---K T D---  | I---     | ---                        | ---                 | ---    | 280 |
| D. YE. 02. 02YE516             | ---K T D---  | I---     | ---                        | ---                 | ---    | 280 |
| D. ZA. 90. R1                  | ---K T D---  | I---     | ---                        | ---                 | ---    | 278 |
| F1. A0. 06. A0_06_ANG125       | ---K N---  | T C---   | F I M---                   | H V K---            | I S--- | 281 |
| F1. AR. 02. ARE933             | ---K N---  | ---      | ---                        | ---                 | ---    | 279 |
| F1. BE. 93. VT850              | ---K N---  | ---      | ---                        | ---                 | ---    | 279 |
| F1. BR. 07. 07BR844            | ---K N D---  | ---      | ---                        | ---                 | ---    | 285 |
| F1. CY. 08. CY222              | ---K T D---  | T---     | ---                        | ---                 | ---    | 279 |
| F1. ES. 11. DEMF110ES001       | ---K E N---  | Y T---   | ---                        | ---                 | ---    | 279 |
| F1. FI. 93. FIN9363            | ---K H---  | ---      | V---                       | ---                 | ---    | 278 |
| F1. FR. 96. 96FR_MP411         | ---K T D---  | T---     | ---                        | ---                 | ---    | 281 |
| F1. RO. 96. BC1_R07            | ---K V---  | T---     | ---                        | ---                 | ---    | 279 |
| F1. RU. 08. D88_845            | ---K V---  | ---      | ---                        | ---                 | ---    | 279 |
| F2. CM. 10. DEMF210CM007       | ---P O---  | ---      | ---                        | ---                 | ---    | 278 |
| F2. CM. 97. CM53657            | ---P O---  | ---      | S---                       | ---                 | ---    | 279 |
| G. BE. 96. DRCLB               | ---K E P---  | S K R--- | I M---                     | ---                 | ---    | 278 |
| G. CM. 10. DEMG10CM008         | ---K V P---  | E K T--- | I M---                     | ---                 | ---    | 278 |
| G. CN. 08. GX_2084_08          | ---K V P---  | E K T--- | I M---                     | ---                 | ---    | 283 |
| G. CU. 99. CU74                | ---K V P---  | E K T--- | I M---                     | ---                 | ---    | 279 |
| G. ES. 09. X2634_2             | ---K V P---  | E K T--- | I M---                     | ---                 | ---    | 283 |
| G. GH. 03. 03GH175G            | ---K V P---  | E K T--- | I M---                     | ---                 | ---    | 279 |

Table with columns: Accession ID, protease end\_p66, p51 RT start, M41L, K70R D67N, D110 catalytic site, and a final column of numbers (279-281). Rows list various HIV-1/SIVcpz protein sequences and their corresponding amino acid residues.

|                             |   | polymerase motif                  | T215Y                  | K219Q    |       |
|-----------------------------|---|-----------------------------------|------------------------|----------|-------|
| B.FR.83.HXB2                |   | RYKYAFTIPSNINNETPGIRYQYNVLPQGWKGS | PAIFQSSMTKILEPFRKONPDI | VIYQYMDL | YVGS  |
| A1.AU.03.P51044 Day0        |   | -----                             | -----                  | -----    | ----- |
| A1.CH.03.HIV_CH_BID_V358    |   | -----                             | -----                  | -----    | ----- |
| A1.ES.06.X2110              |   | -----                             | -----                  | -----    | ----- |
| A1.IT.02.60000              |   | -----                             | -----                  | -----    | ----- |
| A1.KE.11.DEMA111KE002       | X | X                                 | X                      | X        | X     |
| A1.RU.11.DEMR06950          |   | -----                             | -----                  | -----    | ----- |
| A1.RW.11.DEMA111RW002       |   | -----                             | -----                  | -----    | ----- |
| A1.SE.95.SE8538             |   | -----                             | -----                  | -----    | ----- |
| A1.TZ.01.A341               |   | -----                             | -----                  | -----    | ----- |
| A1.UA.01.01UADN139          |   | -----                             | -----                  | -----    | ----- |
| A1.UG.11.DEMA110UG009       |   | -----                             | -----                  | -----    | ----- |
| A1.ZA.04.04ZASK162B1        |   | -----                             | -----                  | -----    | ----- |
| A2.CD.97.97CDKT848          |   | -----                             | -----                  | -----    | ----- |
| A2.CM.01.01CM.1445MV        |   | -----                             | -----                  | -----    | ----- |
| A2.CY.94.94CY017 41         |   | -----                             | -----                  | -----    | ----- |
| B.AR.04.04AR143170          |   | -----                             | -----                  | -----    | ----- |
| B.AU.04.P51038 Day174       |   | -----                             | -----                  | -----    | ----- |
| B.BR.06.06BR1119            |   | -----                             | -----                  | -----    | ----- |
| B.CA.07.502.1191.03         |   | -----                             | -----                  | -----    | ----- |
| B.CH.08.M2.0803101.NFLG8    |   | -----                             | -----                  | -----    | ----- |
| B.CN.10.DEMB10CN002         |   | -----                             | -----                  | -----    | ----- |
| B.CO.01.PCM001              |   | -----                             | -----                  | -----    | ----- |
| B.CU.99.CU19                |   | -----                             | -----                  | -----    | ----- |
| B.CY.09.CY266               |   | -----                             | -----                  | -----    | ----- |
| B.DE.04.HIV.DE.BID.V4131    |   | -----                             | -----                  | -----    | ----- |
| B.DK.07.PMVL.011            |   | -----                             | -----                  | -----    | ----- |
| B.DO.05.05DO.160884         |   | -----                             | -----                  | -----    | ----- |
| B.ES.10.DEMB10ES002         |   | -----                             | -----                  | -----    | ----- |
| B.FR.11.DEMB11FR001         |   | -----                             | -----                  | -----    | ----- |
| B.GE.03.03GEMZ004           |   | -----                             | -----                  | -----    | ----- |
| B.HT.05.05HT.129389         | X | X                                 | X                      | X        | X     |
| B.JM.05.05JM.KJ108          |   | -----                             | -----                  | -----    | ----- |
| B.JP.12.DEMB12JP001         |   | -----                             | -----                  | -----    | ----- |
| B.KR.12.12KYY18.18742       |   | -----                             | -----                  | -----    | ----- |
| B.NL.00.671.00T36           |   | -----                             | -----                  | -----    | ----- |
| B.PE.07.502.2649.wg8        |   | -----                             | -----                  | -----    | ----- |
| B.PY.03.03PY.P5P0115        |   | -----                             | -----                  | -----    | ----- |
| B.RU.11.11RU21T             |   | -----                             | -----                  | -----    | ----- |
| B.TH.07.AA0409.WG11         |   | -----                             | -----                  | -----    | ----- |
| B.TT.01.01TT.CRC50069       |   | -----                             | -----                  | -----    | ----- |
| B.TW.94.TWCYS.LM49          |   | -----                             | -----                  | -----    | ----- |
| B.UA.01.01UAKV167           |   | -----                             | -----                  | -----    | ----- |
| B.US.11.E539.42             |   | -----                             | -----                  | -----    | ----- |
| B.UY.02.02UY.T5U1290        |   | -----                             | -----                  | -----    | ----- |
| B.YE.02.02YE507             |   | -----                             | -----                  | -----    | ----- |
| C.AR.01.ARG4006             |   | -----                             | -----                  | -----    | ----- |
| C.BR.07.DEMC07BR003         |   | -----                             | -----                  | -----    | ----- |
| C.BW.00.00BW07621           |   | -----                             | -----                  | -----    | ----- |
| C.CN.09.09YNLX1335g         |   | -----                             | -----                  | -----    | ----- |
| C.CY.09.CY266               |   | -----                             | -----                  | -----    | ----- |
| C.ES.07.X2118.2             |   | -----                             | -----                  | -----    | ----- |
| C.ET.02.02ET.268            |   | -----                             | -----                  | -----    | ----- |
| C.GE.03.03GEMZ033           |   | -----                             | -----                  | -----    | ----- |
| C.IL.98.98IS002             |   | -----                             | -----                  | -----    | ----- |
| C.IN.09.T125.2139           |   | -----                             | -----                  | -----    | ----- |
| C.KE.00.KER209.WG11         |   | -----                             | -----                  | -----    | ----- |
| C.MM.99.MIDU101.3           |   | -----                             | -----                  | -----    | ----- |
| C.MW.09.703010256.CH256.w96 |   | -----                             | -----                  | -----    | ----- |
| C.SN.90.90SE.364            |   | -----                             | -----                  | -----    | ----- |
| C.SO.89.89SM.145            |   | -----                             | -----                  | -----    | ----- |
| C.TZ.08.707010457.CH457.w8  |   | -----                             | -----                  | -----    | ----- |
| C.US.11.17T84.468           |   | -----                             | -----                  | -----    | ----- |
| C.UY.01.TRA3011             |   | -----                             | -----                  | -----    | ----- |
| C.YE.02.02YE511             |   | -----                             | -----                  | -----    | ----- |
| C.ZA.07.704010042.CH042.TF  |   | -----                             | -----                  | -----    | ----- |
| C.ZM.11.DEMC11ZM006         |   | -----                             | -----                  | -----    | ----- |
| D.CD.83.ELI                 |   | -----                             | -----                  | -----    | ----- |
| D.CM.10.DEMD10CM009         |   | -----                             | -----                  | -----    | ----- |
| D.CY.06.CY163               |   | -----                             | -----                  | -----    | ----- |
| D.KE.11.DEMD11KE003         |   | -----                             | -----                  | -----    | ----- |
| D.KR.04.04KBH8              |   | -----                             | -----                  | -----    | ----- |
| D.SN.90.SE365               |   | -----                             | -----                  | -----    | ----- |
| D.TD.99.MM011               |   | -----                             | -----                  | -----    | ----- |
| D.TZ.07.X2180               |   | -----                             | -----                  | -----    | ----- |
| D.UG.11.DEMD11UG003         |   | -----                             | -----                  | -----    | ----- |
| D.YE.02.02YE516             |   | -----                             | -----                  | -----    | ----- |
| D.ZA.90.R1                  |   | -----                             | -----                  | -----    | ----- |
| F1.A0.06.A0.06.ANG125       |   | -----                             | -----                  | -----    | ----- |
| F1.AR.02.ARE933             |   | -----                             | -----                  | -----    | ----- |
| F1.BE.93.VT850              |   | -----                             | -----                  | -----    | ----- |
| F1.BR.07.07BR844            |   | -----                             | -----                  | -----    | ----- |
| F1.CY.08.CY222              |   | -----                             | -----                  | -----    | ----- |
| F1.ES.11.DEMF110ES001       |   | -----                             | -----                  | -----    | ----- |
| F1.FI.93.FIN9363            |   | -----                             | -----                  | -----    | ----- |
| F1.FR.96.96FR.MP411         |   | -----                             | -----                  | -----    | ----- |
| F1.R0.96.BC1.R07            |   | -----                             | -----                  | -----    | ----- |
| F1.RU.08.D88.845            |   | -----                             | -----                  | -----    | ----- |
| F2.CM.10.DEMF210CM007       |   | -----                             | -----                  | -----    | ----- |
| F2.CM.97.CM53657            |   | -----                             | -----                  | -----    | ----- |
| G.BE.96.DRCBL               |   | -----                             | -----                  | -----    | ----- |
| G.CM.10.DEMG10CM008         |   | -----                             | -----                  | -----    | ----- |
| G.CN.08.GX.2084.08          |   | -----                             | -----                  | -----    | ----- |
| G.CU.99.CU7.R07             |   | -----                             | -----                  | -----    | ----- |
| G.ES.09.X2634.2             |   | -----                             | -----                  | -----    | ----- |
| G.GH.03.03GH175G            |   | -----                             | -----                  | -----    | ----- |

Table of HIV-1/SIVcpz proteins with sequence motifs. Includes columns for protein name, sequence, and motif (polymerase motif, T215Y, K219Q). Rows list various strains and their specific amino acid variations.

Table of HIV-1/SIVcpz proteins with sequence motifs. Includes columns for protein name, sequence, and motif (polymerase motif, T215Y, K219Q). Rows list various strains and their specific amino acid variations.









B. FR. 83. HXB2
G. KE. 09. DEMG09KE001
G. NG. 09. 09NG SC62
G. PT. x. PT3306
G. SE. 93. SE6165 G6165
H. BE. 93. VI991
H. BE. 93. VI997
H. CF. 90. 056
H. GB. 00. 00GBAC4001
J. CD. 97. J. 97DC KTB147
J. CM. 04. 04CMU11421
J. SE. 93. SE9280 7887
J. SE. 94. SE9173 7022
K. CD. 97. 97ZR E0TB11
K. CM. 96. 96CM MP535
01 AE. AF. 07. 509M
01 AE. CN. 10. CYN059
01 AE. HK. 04. HK001
01 AE. JP. x. DR0492
01 AE. TH. 04. BKM
01 AE. TH. 09. AA111a WG11
01 AE. TH. 98. C0240
01 AE. VN. 98. 03VND15
02 AG. CM. 10. DE00210CM013
02 AG. ES. 06. P1261
02 AG. FR. 91. DJ263
02 AG. GH. 03. 03GH181AG
02 AG. KR. 12. 12MH11 10746
02 AG. LR. x. POC44951
02 AG. NG. x. IBNG
02 AG. SN. 98. 98SE MP1211
02 AG. US. 06. 502 2696 FL01
03 AB. RU. 97. KA153 7
04 cpx. CY. 94. 94CY032 3
05 DF. BE. x. VI1310
06 cpx. AU. 96. BFP90
07 BC. CN. 98. 98CN009
08 BC. CN. 06. mx2
09 cpx. GH. 96. 96GH2911
10 CD. TZ. 96. 96TZ BF061
11 cpx. CM. 95. 95CM 1816
12 BF. AR. 99. ARMA159
13 cpx. CM. 96. 96CM 1849
14 BG. ES. 05. X1870
15 01B. TH. 99. 99TH MU2079
16 A2D. KR. 97. 97KR004
17 BF. AR. 99. ARMA038
18 cpx. CU. 99. CU76
19 cpx. CU. 99. CU17
20 BG. CU. 99. CU103
21 A2D. KE. 99. KER2003
22 01A1. CM. 01. 01CM 0001B8Y
23 BG. CU. 03. CB118
24 BG. ES. 08. X2456 2
25 cpx. CM. 02. 02CM 1919LE
26 AU. CD. 02. 02CD MBT0047
27 cpx. FR. 04. 04CD FR KZ5
28 BF. BR. 99. BREPMI2609
29 BF. BR. 01. BREPMI6704
31 BC. BR. 04. 04BR142
32 06A1. EE. 01. EE0369
33 01B. ID. 07. JKT189 C
34 01B. TH. 99. OUR2478P
35 AD. AF. 07. 169H
36 cpx. CM. 00. 00CMNYU030
37 cpx. CM. 00. 00CMNYU026
38 BF1. UY. 03. UY03 3389
39 BF. BR. 04. 04BRRJ179
40 BF. BR. 05. 05BRRJ055
42 BF. LU. 03. LU0BF 05 03
43 02G. SA. 03. J11223
44 BF. CL. 00. CH80
45 cpx. FR. 04. 04FR AUK
46 BF. BR. 07. 07BR FPS625
47 BF. ES. 08. P1942
48 01B. MY. 07. 07MYKT021
49 cpx. GM. 03. N26677
51 01B. SG. 11. 11SG HM021
52 01B. MY. 03. 03MYKL018 1
53 01B. MY. 11. 11FJR164
54 01B. MY. 09. 09MYSB023
55 01B. CN. 10. HNC102056
57 BC. CN. 09. 09YNLX195g
61 BC. CN. 10. JL100010
O. BE. 87. ANT70
O. CM. 91. MWP5180
O. CM. 98. 98CMA104
O. FR. 92. VAU
O. GA. 11. 11Gabb352
O. SN. 99. 99SE MP1299
O. US. 10. LTNP
N. CM. 02. DJ00131
N. CM. 04. 04CM 1015 04
N. CM. 06. U14A-2
N. CM. 95. YBF30
N. CM. 97. YBF106
N. FR. 11. N1 FR 2011
P. CM. 06. U14788
P. FR. 09. RBF168
CPZ. CD. 06. BF1167
CPZ. CM. 05. SIVcpzMT145
CPZ. GA. 08. GAB1
CPZ. TZ. 06. SIVcpzTAN13
CPZ. US. 85. US Marilynn

Sequence alignment table for HIV-1/SIVcpz proteins. Header: p66 RT, p15 RNase H end\_p31 Integrase start. Columns: Accession ID, Protein Name, and Amino Acid Sequence. The table shows a highly conserved region across various HIV-1 and SIVcpz strains.

|                                |   |     |
|--------------------------------|---|-----|
| B. FR. 83. HXB2                | GKVLIVAVHVASGYIEAEVIPAETGQETAYLLKLAGRWPKVIHTDNGSNFTGATVRAACWAGIKQEFGIPYNPQSGQVSMNKLKIIQGVDRQDAEHLKTAVQMAVFIHNFKRK. GGIGGYSAGERIVDIIATD. . . . . IQTKELQKQITKIQN | 937 |
| A1. AU. 03. PS1044 Day0        | -----VV-----S-A-K-----N-Q-----E-----I-----H-----  | 937 |
| A1. CH. 03. HIV_CH_BID_V358    | -----I-----RV-----S-A-K-----N-Q-----E-----I-----H-----  | 937 |
| A1. ES. 06. X2110              | -----I-----RV-----P-S-AFK-----SVQ-----E-----I-----R-----  | 937 |
| A1. IT. 02. 60000              | X-----X-----I-----VV-----P-S-A-K-----N-Q-X-X-----X-X-----V-X-----R-----   | 937 |
| A1. KE. 11. DEMB111KE002       | -----I-----R-----V-----P-S-A-K-----N-Q-----E-----I-----R-----N-----   | 937 |
| A1. RU. 11. IRU6950            | -----I-----V-----P-SNA-K-----N-Q-----E-----I-----I-----   | 937 |
| A1. RW. 11. DEMA111RW002       | -----I-----V-----V-----S-A-K-----D-----E-----I-----I-----   | 937 |
| A1. SE. 95. SE8538             | -----I-----V-----V-----S-AFK-----NVO-Y-----E-----I-----I-----   | 937 |
| A1. TZ. 01. A341               | -----M-----I-----A-----V-----S-A-K-----N-Q-----E-----I-----R-----   | 937 |
| A1. UA. 01. 01UADN139          | -----I-----V-----P-SNA-K-----N-Q-----E-----I-----I-----   | 937 |
| A1. UG. 11. DEMB110UG009       | -----I-----A-----V-----G-IS-AFK-----DVO-----E-----I-----I-----N-K-----  | 937 |
| A1. ZA. 04. 04ZASK162B1        | -----I-----A-----RV-----S-A-K-----NVO-----E-----I-----I-----  | 936 |
| A2. CD. 97. 97CDKT848          | -----I-V-----V-----P-S-A-K-----DV-----E-----I-----I-----  | 930 |
| A2. CM. 01. 01CM_1445MV        | -----I-----I-----V-----P-S-A-K-----VQ-----D-----I-----I-----R-----  | 937 |
| A2. CY. 94. 94CY017_41         | -----T-----D-----I-----V-----P-IS-K-----Q-----E-----I-----I-----R-----  | 937 |
| B. AR. 04. 04AR143170          | -----I-----X-----I-----X-----G-IST-K-----I-----L-----E-----I-----R-----   | 937 |
| B. AU. 04. PS1038 Day174       | -----I-----I-----X-----SXA-K-----V-----T-----E-----I-----I-----V-----   | 937 |
| B. BR. 06. 06BR1119            | -----I-----I-----R-----P-IS-K-----T-----E-----I-----I-----  | 937 |
| B. CA. 07. 502_1191_03         | -----I-----I-----R-----P-IS-K-----E-----I-----I-----  | 937 |
| B. CH. 08. M2_0803101_NFLG8    | -----I-----I-----I-----P-SN-K-----E-----I-----I-----  | 941 |
| B. CN. 10. DEMB10CN002         | -----I-----I-----Q-----STA-K-----V-----N-----Q-----R-----I-----S-----L-----A-----N-----   | 937 |
| B. CO. 01. PCM001              | -----I-----P-----I-----V-----N-----E-----R-----I-----S-----   | 937 |
| B. CU. 99. CU19                | -----I-----I-----V-----P-ISP-A-K-----V-----E-----I-----S-----   | 940 |
| B. CY. 09. CY266               | -----I-----R-V-----P-SNA-K-----N-Q-Q-E-----E-----I-----S-----S-----   | 937 |
| B. DE. 04. HIV_DE_BID_V4131    | -----I-----T-----V-----P-SNA-K-----N-----K-----I-----S-----R-----   | 937 |
| B. DK. 07. PMVL_011            | -----I-----T-----P-SN-K-----E-----I-----I-----H-----  | 937 |
| B. DO. 05. 05DO_160884         | -----I-----V-----P-SN-K-----R-----E-----I-----I-----  | 939 |
| B. ES. 10. DEMB10E5002         | -----I-----A-----V-----P-SNA-K-----D-----E-----I-----I-----   | 937 |
| B. FR. 11. DEMB11FR001         | -----I-----I-----V-----P-SNA-K-----D-----R-----I-----I-----   | 936 |
| B. GE. 03. 03GEMZ004           | -----I-V-----I-----I-----S-V-K-----T-----E-----I-----I-----H-----   | 937 |
| B. HT. 05. 05HT_129389         | -----I-----I-----R-----S-V-K-----E-----E-----E-----I-----I-----H-----   | 936 |
| B. JM. 05. 05JM_KJ108          | -----I-----I-----V-----S-A-K-----N-----E-----M-----T-----R-----   | 936 |
| B. JP. 12. DEMB12JP001         | -----I-----I-----I-----SN-K-----N-----E-----I-----I-----N-----  | 937 |
| B. KR. 12. 12KYY18_18742       | -----I-----I-----I-----SN-K-----E-----I-----I-----  | 947 |
| B. NL. 00. 671_00T36           | -----I-----I-----P-ISTA-K-----E-----E-----I-----I-----  | 955 |
| B. PE. 07. 502_2649_wq8        | -----I-----I-----I-----ST-K-----N-----E-----LQ*CRGKNSRHNSNR#L-----V-----L-----  | 938 |
| B. PY. 03. 03PY_P5P0115        | -----I-----I-----I-----P-IS-A-K-----Q-----I-----I-----  | 941 |
| B. RU. 11. 11RU21n             | E-I-----A-----I-----V-----P-STA-K-----Q-Q-----I-----Q-----  | 937 |
| B. TH. 07. A0A089_WG11         | -----I-----I-----A-----I-----V-----S-G-SNV-K-----K-----I-----Q-E-----E-X-----K-----R-----T-----   | 936 |
| B. TT. 01. 01TT_CRC50069       | -----I-----I-----I-----VV-S-G-SNV-K-----K-----I-----Q-E-----E-X-----#-----I-----M-----  | 937 |
| B. TW. 94. TWCYS_LM49          | -----I-----I-----IV-----SNA-K-----K-----I-----E-----I-----  | 944 |
| B. UA. 01. 01UAKV167           | -----I-----I-----V-----P-IST-K-----I-----E-----I-----M-----   | 937 |
| B. US. 11. E539_42             | -----I-----I-----V-----T-SNA-K-----I-----E-----I-----   | 937 |
| B. UY. 02. 02UY_T5U1290        | -----I-----I-----V-----S-A-K-----E-----I-----I-----H-----   | 938 |
| B. YE. 02. 02YE507             | -----I-----I-----V-----S-A-K-----Q-----I-----I-----   | 933 |
| C. AR. 01. ARG4006             | -----I-----I-----I-----V-----SNA-K-----Q-----I-----I-----   | 933 |
| C. BR. 07. DEMC07BR003         | -----I-----I-----P-ISN-K-----Q-----I-----I-----N-----I-----   | 933 |
| C. BW. 00. 00BW07621           | -----T-----M-----YI-----I-----S-A-K-----Q-----R-----I-----R-----R-----IQ-----   | 933 |
| C. CN. 09. 09YNLX1335g         | -----I-----I-----V-----IS-K-----Q-----R-----E-----I-----R-----V-----  | 940 |
| C. CY. 09. CY260               | -----I-----I-----YI-----X-----V-----S-A-K-----X-----X-----X-----R-----I-----  | 933 |
| C. ES. 07. X2118_2             | -----I-----I-----YI-----X-----V-----S-A-K-----X-----X-----X-----R-----D-C-----I-----N-----L-----  | 939 |
| C. ET. 02. 02ET_288            | -----I-----XX-----YI-----V-----P-SNA-K-----Q-----E-----I-----I-----   | 933 |
| C. GE. 03. 03GEMZ033           | -----I-----YI-----V-----P-SNA-K-----Q-----E-----I-----I-----  | 937 |
| C. IL. 98. 98IS002             | -----I-----M-----YI-----V-----P-S-A-K-----Q-----E-----I-----N-----T-----  | 934 |
| C. IN. 09. T125_2139           | -----I-----M-----YI-----V-----S-A-K-----Q-----R-----I-----R-----IQ-----V-----   | 933 |
| C. KE. 00. KERA208             | -----I-----I-----YI-----V-----S-A-K-----Q-----R-----I-----R-----IQ-----   | 933 |
| C. MM. 99. mIDU101_3           | -----I-----M-----YI-----V-----P-AAK-----H-----R-----I-----I-----  | 935 |
| C. MW. 09. 703010256_CH256_w96 | -----I-----M-----T-----YI-----V-----P-S-A-K-----Q-----T-----I-----Q-----  | 929 |
| C. SN. 90. 90SE_364            | -----I-----I-----V-----S-A-K-----Q-----E-----R-----I-----I-----   | 944 |
| C. SO. 89. 89SM_145            | -----I-----E-----YI-----V-----SNA-K-----H-----R-----D-T-----I-----S-----N-----L-----  | 947 |
| C. TZ. 08. 707010457_CH457_w8  | -----I-----E-----YI-----V-----R-SN-K-----E-----R-----D-T-----I-----S-----#-----N-----I-----   | 931 |
| C. US. 11. 11T84_468           | -----I-----I-----YI-----V-----P-SNA-K-----Q-----E-----I-----I-----  | 928 |
| C. UY. 01. TRA3011             | -----I-----I-----QV-----P-SN-K-----Q-----E-----I-----I-----   | 938 |
| C. YE. 02. 02YE511             | -----I-----YI-----V-----S-A-K-----Q-----E-----I-----I-----  | 936 |
| C. ZA. 07. 704010042_CH042_TF  | -----I-----YI-----V-----P-S-A-K-----Q-----E-----I-----I-----  | 936 |
| C. ZM. 11. DEMC11ZM006         | -----I-----YI-----V-----S-A-K-----Q-----R-R-----I-----I-----  | 942 |
| D. CD. 83. ELI                 | -----I-----I-----VV-----S-A-K-----R-----R-----I-----I-----  | 936 |
| D. CM. 10. DEMD10CM009         | -----I-----K-----I-----VV-----S-A-K-----V-----R-----L-----I-----R-----  | 936 |
| D. CY. 06. CY163               | -----I-----I-----VV-----S-A-K-----Q-----R-----L-----I-----R-----  | 936 |
| D. KE. 11. DEMD11KE003         | -----I-----I-----VV-----S-K-----Q-----E-----I-----R-----I-----  | 936 |
| D. KR. 04. 04KBH8              | -----I-----I-----V-----S-V-K-----D-----E-----E-----I-----R-----I-----   | 942 |
| D. SN. 90. SE365               | -----I-----I-----VV-----S-K-----D-----E-----E-----I-----R-----I-----  | 938 |
| D. TD. 99. MN011               | -QI-----I-----V-----R-----S-K-----R-----E-----I-----S-----L-----I-----  | 938 |
| D. UG. 11. DEMD11UG003         | -----I-----I-----I-----VV-----IS-A-K-----E-----E-----I-----I-----   | 938 |
| D. YE. 02. 02YE516             | -----I-----I-----VV-----SGA-K-----E-----I-----I-----  | 938 |
| D. ZA. 90. R1                  | -----I-----I-----VV-----S-A-K-----L-----I-----I-----  | 936 |
| F1. AO. 06. AO_06_ANG125       | -----I-----L-----A-----I-----I-----S-A-K-----Q-----I-----M-----R-----   | 939 |
| F1. AR. 02. ARE933             | -----I-----E-----I-----I-----A-----P-S-A-K-----Q-----I-----R-----I-----   | 937 |
| F1. BE. 93. VT850              | -----I-----I-----I-----I-----S-A-K-----Q-----I-----R-----I-----   | 938 |
| F1. BR. 07. 07BR844            | -----I-----I-----I-----T-----S-A-K-----Q-----I-----M-----I-----   | 943 |
| F1. CY. 08. CY222              | E-I-----K-----I-----I-----T-----SGA-K-----Q-----E-----K-I-K-----I-----K-----I-----  | 936 |
| F1. ES. 11. DEMF110ES001       | -----I-----I-----I-----P-IS-A-K-----VQ-----I-----I-----I-----I-----K-----   | 937 |
| F1. FI. 93. FIN9363            | -----I-----I-----D-----I-----M-----S-A-K-----Q-----I-----I-----   | 936 |
| F1. FR. 96. 96FR_MP411         | -----I-----I-----I-----R-----SNA-K-----Q-----I-----R-----V-----I-----   | 937 |
| F1. GU. 96. BC1_R07            | -----I-----I-----I-----R-----SNA-K-----Q-----I-----D-----I-----   | 937 |
| F1. RU. 08. D88_845            | -----I-----I-----I-----I-----ST-K-----R-----I-----S-----  | 937 |
| F2. CM. 10. DEMF210CM007       | -----I-----I-----I-----I-----S-A-K-----Q-----I-----S-----   | 936 |
| F2. CM. 97. CM53657            | -----I-----I-----I-----I-----STV-K-----Q-----I-----S-----   | 937 |
| G. BE. 96. DRCBL               | -----I-----I-----I-----I-----S-A-K-----S-T-----I-----S-----   | 933 |
| G. CM. 10. DEMG10CM008         | -QI-----A-----I-----V-----R-SNA-K-----N-T-----E-----E-----I-----S-----  | 936 |
| G. CN. 08. GX_2084_08          | -----I-----I-----I-----I-----SN-K-----N-T-----E-----I-----SR-----   | 941 |
| G. CU. 99. CU7                 | -----I-----I-----I-----I-----S-A-K-----N-T-----E-----I-----   | 941 |
| G. ES. 09. X2634_2             | -----I-----I-----I-----I-----S-A-K-----N-T-----M-----R-----L-----   | 941 |
| G. GH. 03. 03GH175G            | -----I-----I-----I-----I-----VL-----S-A-K-----N-T-----E-----I-----S-----  | 937 |

B. FR. 83. HXB2  
G. KE. 09. DEMG09KE001  
G. NG. 09. 09NG SC62  
G. PT. x. PT3306  
G. SE. 93. SE6165 G6165  
H. BE. 93. V1991  
H. BE. 93. V1997  
H. F. 90. 056  
H. GB. 00. 00GBAC4001  
J. CD. 97. J. 97DC KTB147  
J. CM. 04. 04CMU11421  
J. SE. 93. SE9280 7887  
J. SE. 94. SE9173 7022  
K. CD. 97. 97ZR E0TB11  
K. CM. 96. 96CM MP535  
01 AE. AF. 07. 509M  
01 AE. CM. 10. CYN059  
01 AE. HK. 04. HK001  
01 AE. JP. x. DR0492  
01 AE. TH. 04. BKM  
01 AE. TH. 09. AA111a WG11  
01 AE. TH. 98. CM240  
01 AE. VN. 98. 98VND15  
02 AG. CM. 10. DE00210CM013  
02 AG. ES. 06. P1261  
02 AG. FR. 91. DJ263  
02 AG. GH. 03. 03GH181A6  
02 AG. KR. 12. 12MHI11 10746  
02 AG. LR. x. POC44951  
02 AG. NG. x. IBNG  
02 AG. SN. 98. 98SE MP1211  
02 AG. US. 06. 502 2696 FL01  
03 AB. RU. 97. KAL153 2  
04 cpx. CY. 94. 94CY032 3  
05 DF. BE. x. V11310  
06 cpx. AU. 96. BFP90  
07 BC. CN. 98. 98CN009  
08 BC. CN. 06. mx2  
09 cpx. GH. 96. 96GH2911  
10 CD. TZ. 96. 96TZ BF061  
11 cpx. CM. 95. 95CM 1816  
12 BF. AR. 99. ARMA159  
13 cpx. CM. 96. 96CM 1849  
14 BG. ES. 05. X1870  
15 01B. TH. 99. 99TH MU2079  
16 A2D. KR. 97. 97KR004  
17 BF. AR. 99. ARMA038  
18 cpx. CU. 99. CU76  
19 cpx. CU. 99. CU7  
20 BG. CU. 99. CU103  
21 A2D. KE. 99. KER2003  
22 01A1. CM. 01. 01CM 0001BBY  
23 BG. CU. 03. CB118  
24 BG. ES. 08. X2456 2  
25 cpx. CM. 02. 0219LE  
26 AU. CD. 02. 02CD MBT047  
27 cpx. FR. 04. 04CD FR KZ5  
28 BF. BR. 99. BREPM12609  
29 BF. BR. 01. BREPM16704  
31 CB. BR. 04. 04BR142  
32 06A1. EE. 01. EE0369  
33 01B. ID. 07. JKT189 C  
34 01B. TH. 99. OUR2478P  
35 AD. AF. 07. 169H  
36 cpx. CM. 00. 00CMNYU830  
37 cpx. CM. 00. 00CMNYU926  
38 BF1. UY. 03. UY03 3389  
39 BF. BR. 04. 04BRRJ179  
40 BF. BR. 05. 05BRRJ055  
42 BF. LU. 03. LuBF 05 03  
43 02G. SA. 03. J11223  
44 BF. CL. 00. CH60  
45 cpx. FR. 04. 04FR AUK  
46 BF. BR. 07. 07BR FPS625  
47 BF. ES. 08. P1942  
48 01B. MY. 07. 07MYKT021  
49 cpx. GM. 03. N26677  
51 01B. SG. 11. 11SG HM021  
52 01B. MY. 03. 03MYKL018 1  
53 01B. MY. 11. 11FIR164  
54 01B. MY. 09. 09MYSB023  
55 01B. CN. 10. HNC3102056  
57 BC. CN. 09. 09YNLX195g  
61 BC. CN. 10. JL100010  
O. BE. 87. ANT70  
O. CM. 91. MP5180  
O. CN. 98. 98CMA104  
O. FR. 92. VAU  
O. GA. 11. 11Gabb352  
O. SN. 99. 99SE MP1299  
O. US. 10. LTNP  
N. CM. 02. DJ00131  
N. CM. 04. 04CM 1015 04  
N. CM. 06. U1484  
N. CM. 95. YBF30  
N. CM. 97. YBF106  
N. FR. 11. N1 FR 2011  
P. CM. 06. U14788  
P. FR. 09. RBF168  
CPZ. CD. 06. BF1167  
CPZ. CM. 05. SIVcpzMT145  
CPZ. GA. 08. GAB1  
CPZ. TZ. 06. SIVcpzTAN13  
CPZ. US. 85. US\_Marilyn

Sequence alignment table for HIV-1 proteins. Header: GKVILVAVHVASGYIEAEVIPAQETAYFLLLKLAGRWPKTIHTDNGSNFTGATVRAACVAGWJKQEFIPYNPQSQGVVSMNKLKKIKIQVDRQDAEHLKTAVQMAVFIHNFRRK. GGIGGYSAGERIVDIIATD.....IQTKEKQKQITKIQN 937. Rows: Amino acid positions 1-937. Columns: 1000+ different HIV-1 strains (see left column for IDs). Characters: Standard amino acid codes (S, A, G, T, C, A, P, Q, R, K, M, L, I, V, N, D, E, F, Y, W, H, G, D, K, E, Q, P, S, L, M, T, N, Q, R, G, A, V, I, P, S, A, K, V, H, C, Y, F, W, L, M, I, D, E, F, G, H, K, N, Q, R, S, T, V, Y, W, X, Z, -, \*, ~, |, ^, &, &, &).

p31 Integrase end  
Pol end\*

Table with 3 columns: Identifier, Sequence, and Reference. Contains multiple rows of protein sequences from various sources like B.FR.83.HXB2, A1.AU.03.PS1044, etc.

Table with 2 columns: Reference and Value. Contains numerical values corresponding to the rows of the main table.



|                                |  |           |  |  |     |
|--------------------------------|--|-----------|--|--|-----|
|                                |  | Vif start |  | MENRWQVMIVWQVDMRRTWKS LVKHHMYS GKGARWGFRHYHSEPHF . RISSEVHIPLG . D . AR . LVITTTYV . GLHTGERDWHLGQGVSI EWRKRYSTQVDP ELDLQI LHYLYFDFCS DAIKALKG LHVISPRCEYQAGH . NKGVLSQYLALALAIL . . . PKKIK . | 160 |
| B. FR. 83. HXB2                |  |           |  |  | 160 |
| A1. AU. 03. PS1044 Day0        |  |           |  |  | 160 |
| A1. CH. 03. HIV CH BID V3538   |  |           |  |  | 160 |
| A1. ES. 06. X2110              |  |           |  |  | 160 |
| A1. IT. 02. 60000              |  |           |  |  | 160 |
| A1. KE. 11. DEM111KE002        |  |           |  |  | 160 |
| A1. RU. 11. DEMR06950          |  |           |  |  | 160 |
| A1. RW. 11. DEM111RW002        |  |           |  |  | 160 |
| A1. SE. 95. SE8538             |  |           |  |  | 160 |
| A1. TZ. 01. A341               |  |           |  |  | 160 |
| A1. UA. 01. 01UA0N139          |  |           |  |  | 160 |
| A1. UG. 11. DEM111UG009        |  |           |  |  | 160 |
| A1. ZA. 04. 04ZASK16ZB1        |  |           |  |  | 160 |
| A2. CD. 97. 97CDKTB48          |  |           |  |  | 160 |
| A2. CM. 01. 01CM 1445MV        |  |           |  |  | 160 |
| A2. CY. 94. 94CY017 41         |  |           |  |  | 160 |
| B. AR. 04. 04AR143170          |  |           |  |  | 160 |
| B. AU. 04. PS1038 Day174       |  |           |  |  | 160 |
| B. BR. 06. 06BR119             |  |           |  |  | 160 |
| B. CA. 07. 502 1191 03         |  |           |  |  | 160 |
| B. CH. 08. M2 0803101 NFLG8    |  |           |  |  | 160 |
| B. CN. 10. DEMB10CN002         |  |           |  |  | 160 |
| B. CO. 01. PCM001              |  |           |  |  | 160 |
| B. CU. 09. Cui19               |  |           |  |  | 160 |
| B. CY. 09. CY266               |  |           |  |  | 160 |
| B. DE. 04. HIV DE BID V4131    |  |           |  |  | 160 |
| B. DK. 07. PMVL 011            |  |           |  |  | 160 |
| B. DO. 05. 05DO 160884         |  |           |  |  | 160 |
| B. FS. 10. DEM10FS002          |  |           |  |  | 160 |
| B. FR. 11. DEMB11FR001         |  |           |  |  | 160 |
| B. GE. 03. 03GEMZ004           |  |           |  |  | 160 |
| B. HT. 05. 05HT 129389         |  |           |  |  | 160 |
| B. JM. 05. 05JM KJ108          |  |           |  |  | 160 |
| B. JP. 12. DEMB12JP001         |  |           |  |  | 160 |
| B. KR. 12. 12KYY18 18742       |  |           |  |  | 160 |
| B. NL. 00. 671 00T36           |  |           |  |  | 160 |
| B. PE. 07. 502 2649 wq8        |  |           |  |  | 160 |
| B. PY. 03. 03PY P5P0115        |  |           |  |  | 160 |
| B. RU. 11. 11RU21n             |  |           |  |  | 160 |
| B. TH. 07. A0A089 W611         |  |           |  |  | 160 |
| B. TT. 01. 01TT CRC50069       |  |           |  |  | 160 |
| B. TW. 94. TWCYS L449          |  |           |  |  | 160 |
| B. UA. 01. 01UAUK167           |  |           |  |  | 160 |
| B. US. 11. E539 42             |  |           |  |  | 160 |
| B. UY. 02. 02UY T5U1290        |  |           |  |  | 160 |
| B. YE. 02. 02YE507             |  |           |  |  | 160 |
| C. AR. 01. ARG4006             |  |           |  |  | 160 |
| C. BR. 07. DEMC07BR003         |  |           |  |  | 160 |
| C. BW. 00. 00BW07621           |  |           |  |  | 160 |
| C. CN. 09. 09YNLX13359         |  |           |  |  | 160 |
| C. CY. 09. CY266               |  |           |  |  | 160 |
| C. ES. 07. X2118 2             |  |           |  |  | 160 |
| C. ET. 02. 02ET 268            |  |           |  |  | 160 |
| C. GE. 03. 03GEMZ033           |  |           |  |  | 160 |
| C. IL. 98. 98IS002             |  |           |  |  | 160 |
| C. IN. 09. T125 2139           |  |           |  |  | 160 |
| C. KE. 00. KERA2016            |  |           |  |  | 160 |
| C. MM. 99. M1DU101 3           |  |           |  |  | 160 |
| C. MW. 09. 703010256 CH256.w96 |  |           |  |  | 160 |
| C. SN. 90. 90SE 364            |  |           |  |  | 160 |
| C. SO. 89. 89SM 145            |  |           |  |  | 160 |
| C. TZ. 08. 707010457 CH457.w8  |  |           |  |  | 160 |
| C. US. 11. 17T84 468           |  |           |  |  | 160 |
| C. UY. 01. TRA3011             |  |           |  |  | 160 |
| C. YE. 02. 02YE511             |  |           |  |  | 160 |
| C. ZA. 07. 704010042 CH042_TF  |  |           |  |  | 160 |
| C. ZM. 11. DEMC11ZM006         |  |           |  |  | 160 |
| D. CD. 83. ELT                 |  |           |  |  | 160 |
| D. CM. 10. DEM10CM009          |  |           |  |  | 160 |
| D. CY. 00. CY163               |  |           |  |  | 160 |
| D. KE. 11. DEMD11KE003         |  |           |  |  | 160 |
| D. KR. 04. 04KRBH8             |  |           |  |  | 160 |
| D. SN. 90. SE365               |  |           |  |  | 160 |
| D. TD. 99. MN011               |  |           |  |  | 160 |
| D. TZ. 01. A289                |  |           |  |  | 160 |
| D. UG. 11. DEMD11UG003         |  |           |  |  | 160 |
| D. YE. 02. 02YE516             |  |           |  |  | 160 |
| D. ZA. 90. R1                  |  |           |  |  | 161 |
| F1. AO. 06. AO 06 ANG125       |  |           |  |  | 160 |
| F1. AR. 02. ARE933             |  |           |  |  | 160 |
| F1. BE. 93. V7850              |  |           |  |  | 160 |
| F1. BR. 07. 07BR0844           |  |           |  |  | 160 |
| F1. CY. 08. CY222              |  |           |  |  | 160 |
| F1. ES. 11. DEMF110E5001       |  |           |  |  | 160 |
| F1. FI. 93. FIN9363            |  |           |  |  | 160 |
| F1. FR. 96. 96FR NP411         |  |           |  |  | 160 |
| F1. GO. 96. BG 007             |  |           |  |  | 160 |
| F1. RU. 08. D88 845            |  |           |  |  | 160 |
| F2. CM. 10. DEMF210CM007       |  |           |  |  | 160 |
| F2. CM. 97. CM53657            |  |           |  |  | 160 |
| G. BE. 96. DRCBL               |  |           |  |  | 160 |
| G. CM. 10. DEMG10CM008         |  |           |  |  | 160 |
| G. CN. 08. GX 2084 08          |  |           |  |  | 160 |
| G. GU. 99. GUT 007             |  |           |  |  | 160 |
| G. ES. 09. X2634 2             |  |           |  |  | 160 |
| G. GH. 03. 03GH175G            |  |           |  |  | 160 |



B. FR. 83. HXB2
G. KE. 09. DEMG09K001
G. NG. 09. 09NG 56C2
G. PT. X. PT3306
G. SE. 93. SE6165 G6165
H. BE. 93. V1991
H. BF. 93. V1997
H. CF. 90. 056
H. GB. 00. 00GBCA4001
J. CD. 97. J. 97DC KTB147
J. CM. 04. 04CMU1421
J. SE. 93. SE9280 7887
J. SE. 94. SE9173 7822
K. CD. 97. 97ZR E0TB11
K. CM. 96. 96CM MP535
01 AE. AF. 07. 509M
01 AE. CN. 10. CYN059
01 AE. HK. 04. HK001
01 AE. JP. X. DR0492
01 AE. TH. 04. BKM
01 AE. TH. 09. AA111a WG11
01 AE. TH. 98. C818
01 AE. VN. 98. 98VND15
02 AG. CM. 10. DE00210CM013
02 AG. ES. 06. P1261
02 AG. FR. 91. DJ263
02 AG. GH. 03. 03GH181AG
02 AG. KR. 12. 12MH11 10746
02 AG. LR. X. POC44951
02 AG. NG. X. IBNG
02 AG. SN. 98. 98SE MP1211
02 AG. US. 06. 502 2696 FL01
03 AB. RU. 97. 97AB12
04 cpx. CY. 94. 94CY032 3
05 DF. BE. X. V11310
06 cpx. AU. 96. BFP90
07 BC. CN. 98. 98CN009
09 BC. CN. 06. mx2
09 cpx. GH. 92. 96GH2911
10 CD. TZ. 96. 96TZ BF061
11 cpx. CM. 95. 95CM 1816
12 BF. AR. 99. ARMA159
13 cpx. CM. 96. 96CM 1849
14 BG. ES. 04. 01870
15 01B. TH. 99. 99TH MU2079
16 AZD. KR. 97. 97KR004
17 BF. AR. 99. ARMA038
18 cpx. CU. 99. CU76
19 cpx. CU. 99. CU103
20 BG. CU. 99. CU103
21 AZD. KE. 99. KER2003
22 01A1. CM. 01. 01CM 0001BBY
23 BG. CU. 03. C8118
24 BG. ES. 08. X2456 2
25 cpx. CM. 02. 0219LE
26 AU. CD. 02. 02CD MBT0047
27 cpx. FR. 04. 04CD FR KZ5
28 BF. BR. 99. BREPM12609
29 BF. BR. 01. BREPM15704
31 BC. BR. 04. 04BR142
32 096A1. EE. 01. 01EE0369
33 01B. ID. 07. JKT189 C
34 01B. TH. 99. 09UR2478P
35 AD. AF. 07. 169H
36 cpx. CM. 00. 00CMYU030
37 01B. MY. 07. 07MYKT021
49 cpx. GM. 03. N26677
51 01B. SG. 11. 11SG HM021
52 01B. MY. 03. 03MYKL018 1
53 01B. MY. 11. 11FJR164
54 01B. MY. 01. 01MYNS023
55 01B. CN. 10. 00CNS102956
57 BC. CN. 09. 09YNLX195g
61 BC. CN. 10. JL100010
0. BE. 87. ANT70
0. CM. 91. MP5180
0. CM. 99. 98CMA104
0. FR. 92. VAU
0. GA. 11. 11GAb6352
0. SN. 99. 99SE MP1299
0. US. 10. LTNP
N. CM. 02. DJ00131
N. CM. 04. 04CM 1015 04
N. CM. 06. U1402
N. CM. 95. YBF30
N. CM. 97. YBF106
N. FR. 11. N1 FR 2011
P. CM. 06. U14788
P. FR. 09. RBF168
CPZ. CD. 06. BF1167
CPZ. CM. 05. SIVcpzMT145
CPZ. GA. 08. GAB1
CPZ. TZ. 06. SIVcpzTAN13
CPZ. US. 85. US\_Mar15N

Vif start
MENRWQVMIWVQVDMRRTWKSVLKHHMYVSGKARGVFRHYHSEPH. RISSEVHIPLG. D. AR. LVITTYV. GLHTGERDWHGQGVSIERWKKRYSTQVPELADQLIHYDFCFSDSAIRKALLGHVSPRCEYQAGH. NKGVLSQYLALAAALIT. . . PKKIK. 160

Table of amino acid sequences for HIV-1/SIVcpz proteins. Each row represents a different protein or variant, with columns showing the amino acid sequence. The table is organized into columns of 10 amino acids each, with a column number (160) at the end of each row.











B.FR.83.HXB2
G.KE.09.DEMG09KE001
G.NG.09.09NG SC62
G.PT.X.PT3306
G.SE.93.SE6165 G6165
H.BE.93.VI991
H.BE.93.VI997
H.CF.98.056
H.GB.00.00GBCA4001
J.CD.97.J.97DC KTB147
J.CM.04.04CMU1421
J.SE.93.SE9280 7887
J.SE.94.SE9173 7822
K.CD.97.97ZR E0TB11
K.CM.96.96CM MP535
01 AE.AF.07.509M
01 AE.CM.10.CYM059
01 AE.HK.04.HK001
01 AE.JP.X.DR0492
01 AE.TH.04.BKM
01 AE.TH.09.AA111a WG11
01 AE.TH.98.CM240
01 AE.VN.98.98VND15
02 AG.CM.10.DE00210CM013
02 AG.ES.06.P1261
02 AG.FR.91.DJ263
02 AG.GH.03.03GH181AG
02 AG.KR.12.12MH11 10746
02 AG.LR.X.P0C44951
02 AG.NG.X.IBNG
02 AG.SN.98.98SE MP1211
02 AG.US.06.502 2696 FL01
03 AB.RU.97.KAL153 2
04 cpx.CY.94.94CY032 3
05 DF.BE.X.VI1310
06 cpx.AU.96.BFP90
07 BC.CN.98.98CN009
08 BC.CN.06.mx2
09 cpx.GH.96.96GH2911
10 CD.TZ.96.96TZ BF061
11 cpx.CM.95.95CM 1816
12 BF.AR.99.ARMA159
13 cpx.CM.96.96CM 1849
14 BG.ES.05.X1870
15 01B.TH.99.99TH MU2079
16 A2D.KR.97.97KR004
17 BF.AR.99.ARMA038
18 cpx.CU.99.CU76
19 cpx.CU.99.CU7
20 BG.CU.99.CU103
21 A2D.KE.99.KER2003
22 01A1.CM.01.01CM 0001B8Y
23 BG.CU.03.CB118
24 BG.ES.08.X2456 2
25 cpx.CM.02.1918LE
26 AU.CD.02.02CD MBT047
27 cpx.FR.04.04CD FR KZ5
28 BF.BR.99.BREPM12609
29 BF.BR.01.BREPM16704
31 BC.BR.04.04BR142
32 06A1.EE.01.EE0369
33 01B.ID.07.JKT189 C
34 01B.TH.99.OUR2478P
35 AD.AF.07.169H
36 cpx.CM.00.00CMNYU030
37 cpx.CM.00.00CMNYU026
38 BF1.UY.03.UY03 3389
39 BF.BR.04.04BRRJ179
40 BF.BR.05.05BRRJ055
42 BF.LU.03.LuBF 05 03
43 02G.SA.03.J11223
44 BF.CL.00.CH60
45 cpx.FR.04.04FR AUK
46 BF.BR.07.07BR FPS625
47 BF.ES.08.P1942
48 01B.MY.07.07MYKT021
49 01B.GM.03.N26677
51 01B.SG.11.11SG HM021
52 01B.MY.03.03MYKL018 1
53 01B.MY.11.11FJR164
54 01B.MY.09.09MYSB023
55 01B.CN.10.HNC3102056
57 BC.CN.09.09NXL195g
61 BC.CN.10.JL100010
O.BE.87.ANT70
O.CM.91.MWP5180
O.CM.98.98CM104
O.FR.92.VAU
O.GA.11.11Gabb352
O.SN.99.99SE MP1299
O.US.10.LTNP
N.CM.02.DJ00131
N.CM.04.04CM 1015 04
N.CM.06.U14842
N.CM.95.YBF30
N.CM.97.YBF106
N.FR.11.N1 FR 2011
P.CM.06.U14788
P.FR.09.RBF168
CPZ.CD.06.BF1167
CPZ.CM.05.SIVcpzMT145
CPZ.GA.98.GAB1
CPZ.TZ.06.SIVcpzTAN13
CPZ.US.85.US\_Marilyn

Table with columns: Tat start, C-rich region, nuclear localization, exon 1 end, exon 2 start, Tat end. Contains a large list of HIV-1 protein sequences and their corresponding amino acid alignments.





Table with columns: Rev start, exon 1 end, exon 2 start, NLS, Leu-rich effector domain, Rev end. Contains a dense list of protein sequences and their corresponding amino acid positions.









| Accession    | Sequence   | Accession    | Sequence   | Accession    | Sequence   | Accession    | Sequence   | Accession    | Sequence   | Accession    | Sequence   | Accession    | Sequence   | Accession    | Sequence   | Accession    | Sequence   | Accession    | Sequence   |              |  |              |  |
|--------------|--|--------------|--|--------------|--|--------------|--|--------------|--|--------------|--|--------------|--|--------------|--|--------------|--|--------------|--|--------------|--|--------------|--|
| B.FR.83.HXB2 | ..SSGS..RIMMEKGEIKNCSFNISTSIRKGVQ..KEYAFFYKLDIIPIDND..TT..SYKLTSCNTSVITQACPKVSEFPIPIHYCAAGPAFAILKNMKTFTNGTGPCTNVSTVQCHTGIRPVSTQLLLNGSLAEE..E | B.FR.83.HXB2 | ..SSGS..RIMMEKGEIKNCSFNISTSIRKGVQ..KEYAFFYKLDIIPIDND..TT..SYKLTSCNTSVITQACPKVSEFPIPIHYCAAGPAFAILKNMKTFTNGTGPCTNVSTVQCHTGIRPVSTQLLLNGSLAEE..E | B.FR.83.HXB2 | ..SSGS..RIMMEKGEIKNCSFNISTSIRKGVQ..KEYAFFYKLDIIPIDND..TT..SYKLTSCNTSVITQACPKVSEFPIPIHYCAAGPAFAILKNMKTFTNGTGPCTNVSTVQCHTGIRPVSTQLLLNGSLAEE..E | B.FR.83.HXB2 | ..SSGS..RIMMEKGEIKNCSFNISTSIRKGVQ..KEYAFFYKLDIIPIDND..TT..SYKLTSCNTSVITQACPKVSEFPIPIHYCAAGPAFAILKNMKTFTNGTGPCTNVSTVQCHTGIRPVSTQLLLNGSLAEE..E | B.FR.83.HXB2 | ..SSGS..RIMMEKGEIKNCSFNISTSIRKGVQ..KEYAFFYKLDIIPIDND..TT..SYKLTSCNTSVITQACPKVSEFPIPIHYCAAGPAFAILKNMKTFTNGTGPCTNVSTVQCHTGIRPVSTQLLLNGSLAEE..E | B.FR.83.HXB2 | ..SSGS..RIMMEKGEIKNCSFNISTSIRKGVQ..KEYAFFYKLDIIPIDND..TT..SYKLTSCNTSVITQACPKVSEFPIPIHYCAAGPAFAILKNMKTFTNGTGPCTNVSTVQCHTGIRPVSTQLLLNGSLAEE..E | B.FR.83.HXB2 | ..SSGS..RIMMEKGEIKNCSFNISTSIRKGVQ..KEYAFFYKLDIIPIDND..TT..SYKLTSCNTSVITQACPKVSEFPIPIHYCAAGPAFAILKNMKTFTNGTGPCTNVSTVQCHTGIRPVSTQLLLNGSLAEE..E | B.FR.83.HXB2 | ..SSGS..RIMMEKGEIKNCSFNISTSIRKGVQ..KEYAFFYKLDIIPIDND..TT..SYKLTSCNTSVITQACPKVSEFPIPIHYCAAGPAFAILKNMKTFTNGTGPCTNVSTVQCHTGIRPVSTQLLLNGSLAEE..E | B.FR.83.HXB2 | ..SSGS..RIMMEKGEIKNCSFNISTSIRKGVQ..KEYAFFYKLDIIPIDND..TT..SYKLTSCNTSVITQACPKVSEFPIPIHYCAAGPAFAILKNMKTFTNGTGPCTNVSTVQCHTGIRPVSTQLLLNGSLAEE..E | B.FR.83.HXB2 | ..SSGS..RIMMEKGEIKNCSFNISTSIRKGVQ..KEYAFFYKLDIIPIDND..TT..SYKLTSCNTSVITQACPKVSEFPIPIHYCAAGPAFAILKNMKTFTNGTGPCTNVSTVQCHTGIRPVSTQLLLNGSLAEE..E | B.FR.83.HXB2 | ..SSGS..RIMMEKGEIKNCSFNISTSIRKGVQ..KEYAFFYKLDIIPIDND..TT..SYKLTSCNTSVITQACPKVSEFPIPIHYCAAGPAFAILKNMKTFTNGTGPCTNVSTVQCHTGIRPVSTQLLLNGSLAEE..E | B.FR.83.HXB2 | ..SSGS..RIMMEKGEIKNCSFNISTSIRKGVQ..KEYAFFYKLDIIPIDND..TT..SYKLTSCNTSVITQACPKVSEFPIPIHYCAAGPAFAILKNMKTFTNGTGPCTNVSTVQCHTGIRPVSTQLLLNGSLAEE..E |



HIV-1 Proteins

Table of HIV-1 glycosylation NNTs and NTSs. Includes headers for glycosylation NNT, glycosylation NTS, glycosylation NCT, and V3 tip V3. Lists protein IDs and their corresponding amino acid sequences.

Table of HIV-1 glycosylation NNTs and NTSs (continued). Includes headers for glycosylation NNT, glycosylation NTS, glycosylation NCT, CD4 binding, glycosylation NNT, and glycosylation NTS. Lists protein IDs and their corresponding amino acid sequences.





Table listing HIV-1 protein sequences for various proteins including V4, CD4 binding, glycosylation NIT, CD4 binding, glycosylation NES, CD4 binding, gp120 end, gp41 start, and fusion peptide. Each row represents a protein entry with its identifier, sequence, and a numerical index on the right.

Table with columns for protein regions (V4, CD4 binding, glycosylation NIT, glycosylation NES, CD4 binding, gp120 end, gp41 start, fusion peptide) and a corresponding list of amino acid sequences for various HIV-1 strains. The table is organized into sections based on these regions and lists specific amino acid positions and substitutions for each strain.







B.FR.83.HXB2
G.KE.09.DEM609KE001
G.NC.09.09NG.5C62
G.PT.x.PT3306
G.SE.93.SE6165.66165
H.BE.93.VI991
H.BE.93.VI997
H.CF.98.856
H.GB.00.00GBCA4001
J.CD.97.J.97DC.KTB147
J.CM.04.04CMU11421
J.SE.93.SE9280.7887
J.SE.94.SE9173.7822
K.CD.97.97ZR.E0TB11
K.CM.96.96CM.MP535
01 AE.AF.07.509M
01 AE.CM.10.CYM059
01 AE.HK.04.HK001
01 AE.JP.x.DR0492
01 AE.TH.04.BKM
01 AE.TH.09.AA111a.WG11
01 AE.TH.98.CM240
01 AE.VH.98.03VWMD15
02 AG.CM.10.DE00210CM013
02 AG.ES.06.P1261
02 AG.FR.91.DJ263
02 AG.GH.03.03GH181AG
02 AG.KR.12.12M111.10746
02 AG.LR.x.P0C44951
02 AG.NG.x.IBNG
02 AG.SN.98.98SE.MP1211
02 AG.US.06.502.2696.FL01
02 AG.UY.97.97UY032.3
05 DF.BE.x.VI1310
06 cpx.AU.96.BFP90
07 BC.CN.98.98CN009
07 BC.CN.06.mx2
09 cpx.GH.96.96GH2911
10 CD.TZ.96.96TZ.BF061
11 cpx.CM.95.95CM.1816
12 BF.AR.99.ARMA159
13 cpx.CM.96.96CM.1849
14 BG.ES.06.01876
15 01B.TH.99.99TH.MU2079
16 AZD.KR.97.97KR004
17 BF.AR.99.ARMA038
18 cpx.CU.99.CU76
19 cpx.CU.99.CU73
20 BG.CU.99.CU103
21 AZD.KE.99.KER2003
22 01A1.CM.01.01CM.0001BBY
23 BG.CU.03.CB118
24 BG.ES.08.X2456.2
25 cpx.CM.02.02CM19LE
26 AU.CD.02.02CD.MBT0047
27 cpx.FR.04.04CD.FR.KZ5
28 BF.BR.99.BREP12609
29 BF.BR.01.BREP16704
31 BG.BR.04.04BR142
32 06A1.EE.01.E0369
33 01B.ID.07.JKT189.C
34 01B.TH.99.09UR2478P
35 AD.AF.07.07AF.169H
36 cpx.CM.00.00CM.VNUY830
37 CM.CM.06.06CM.VNUY926
38 BF1.UY.03.UY03.3389
39 BF.BR.04.04BRR1719
40 BF.BR.05.05BRR1055
42 BF.LU.03.LuBF.05.03
43 02G.SA.03.31123
44 BF.CL.00.CH80
45 cpx.FR.04.04FR.AUK
46 BF.BR.07.07BR.FP5625
47 BF.ES.08.P194
48 01B.MY.07.07MY.KT021
49 01B.MY.03.N26677
51 01B.SG.11.11SG.HM021
52 01B.MY.03.03MY.KL018.1
53 01B.MY.11.11FR1364
54 01B.MY.00.00MY.SB023
55 01B.CM.18.08CM.N2056
57 BC.CN.09.09YNLX195g
61 BC.CN.10.JL100010
0.BE.87.ANT70
0.CM.91.MP5180
0.CM.99.09CM104
0.FR.92.VAU
0.GA.11.11GAB6352
0.SN.99.99SE.ASB1299
0.US.10.LTNP
N.CM.02.DJ00131
N.CM.04.04CM.1015.04
N.CM.06.U14012
N.CM.95.YBF30
N.CM.97.YBF106
N.FR.11.N1.FR.2011
P.CM.06.U14788
P.FR.09.RBF168
CPZ.CD.06.BF1167
CPZ.CM.05.SIVcpzMT145
CPZ.GA.98.GAB1
CPZ.TZ.06.SIVcpzTAN13
CPZ.US.85.US.Marilyn

Table with columns for glycosylation NGS and glycosylation NAT. Rows list protein sequences with amino acid positions and corresponding glycosylation sites. Includes sequences for GYSPLSFQTHL, PTP, R, GPDRPEGIEEEGERRDRSIRLVNGLSALITWDDLRSLCLFSYHRLRDLILLVTRIVE.LLGR, and GWEALKYWNMLQYW.

|                             | Env end              |     |
|-----------------------------|----------------------|-----|
|                             | gp41 end             |     |
|                             | cytoplasmic tail end |     |
| B.FR.83.HXB2                | *                    | 856 |
| A1.AU.03.PS1044 Day0        | -                    | 871 |
| A1.CH.03.HIV_CH_BID_V3538   | -                    | 855 |
| A1.ES.06.X2110              | -                    | 862 |
| A1.IT.02.60000              | -                    | 868 |
| A1.KE.11.DEMA111KE002       | -                    | 864 |
| A1.RU.11.1RU6950            | -                    | 864 |
| A1.RW.11.DEMA111RW002       | -                    | 872 |
| A1.SE.95.SE8538             | -                    | 855 |
| A1.TZ.01.A341               | -                    | 854 |
| A1.UA.01.01UADN139          | -                    | 869 |
| A1.UG.11.DEMA111UG009       | -                    | 855 |
| A1.ZA.04.04ZASK162B1        | -                    | 860 |
| A2.CD.97.97CDKT848          | -                    | 851 |
| A2.CM.01.01CM_1445MV        | -                    | 863 |
| A2.CY.94.94CY017_41         | -                    | 866 |
| B.AR.04.04AR143170          | -                    | 849 |
| B.AU.04.PS1038 Day174       | -                    | 849 |
| B.BR.06.06BR1119            | -                    | 864 |
| B.CA.07.502_1191_03         | -                    | 863 |
| B.CH.08.M2_0803101_NFLG8    | -                    | 886 |
| B.CN.10.DEMB10CN002         | -                    | 859 |
| B.CO.01.PCM001              | -                    | 851 |
| B.CU.99.Cu19                | -                    | 864 |
| B.CY.09.CY266               | -                    | 861 |
| B.DE.04.HIV_DE_BID_V4131    | -                    | 875 |
| B.DK.07.PMVL_011            | -                    | 880 |
| B.DO.05.05DO_160884         | -                    | 843 |
| B.ES.10.DEMB10ES002         | -                    | 864 |
| B.FR.11.DEMB11FR001         | -                    | 841 |
| B.GE.03.03GEMZ004           | -                    | 858 |
| B.HT.05.05HT_129389         | -                    | 863 |
| B.JM.05.05JM_KJ108          | -                    | 855 |
| B.JP.12.DEMB12JP001         | -                    | 865 |
| B.KR.12.12KYY10_10742       | -                    | 863 |
| B.NL.00.671_00T36           | -                    | 856 |
| B.PE.07.502_2649_wq8        | -                    | 849 |
| B.PY.03.03PY_PSP0115        | -                    | 853 |
| B.RU.11.11RU21n             | -                    | 845 |
| B.TH.07.AA0409_WG11         | -                    | 851 |
| B.TT.01.01TT_CRC50069       | -                    | 861 |
| B.TW.94.TWCYS_LM49          | -                    | 847 |
| B.UA.01.01UAKV167           | -                    | 866 |
| B.US.11.E539_42             | -                    | 864 |
| B.UY.02.02UY_T5U1290        | -                    | 868 |
| B.YE.02.02YE507             | -                    | 862 |
| C.AR.01.ARG4006             | -                    | 857 |
| C.BR.07.DEMC07BR003         | -                    | 887 |
| C.BW.00.00BW07621           | -                    | 851 |
| C.CN.09.09YNLX1335g         | -                    | 870 |
| C.CY.09.CY266               | -                    | 861 |
| C.ES.07.X2118_2             | -                    | 850 |
| C.ET.02.02ET_268            | -                    | 858 |
| C.GE.03.03GEMZ033           | -                    | 866 |
| C.IL.98.98IS002             | -                    | 852 |
| C.IN.09.T125_2139           | -                    | 884 |
| C.KE.00.KER2010             | -                    | 863 |
| C.MM.99.mIDU101_3           | -                    | 866 |
| C.MW.09.703010256_CH256.w96 | -                    | 854 |
| C.SN.90.90SE_364            | -                    | 842 |
| C.SO.89.89SM_145            | -                    | 862 |
| C.TZ.08.707010457_CH457.w8  | -                    | 846 |
| C.US.11.17T84_468           | -                    | 855 |
| C.UY.01.TRA3011             | -                    | 859 |
| C.YE.02.02YE511             | -                    | 863 |
| C.ZA.07.704010042_CH042_TF  | -                    | 865 |
| C.ZM.11.DEMC11ZM006         | -                    | 864 |
| D.CD.83.ELI                 | -                    | 854 |
| D.CM.10.DEMD10CM009         | -                    | 845 |
| D.CY.06.CY163               | -                    | 859 |
| D.KE.11.DEMD11KE003         | -                    | 851 |
| D.KR.04.04KBH8              | -                    | 843 |
| D.SN.90.SE365               | -                    | 851 |
| D.TD.99.MN011               | -                    | 846 |
| D.TZ.01.A280                | -                    | 853 |
| D.UG.11.DEMD11UG003         | -                    | 858 |
| D.YE.02.02YE516             | -                    | 847 |
| D.ZA.90.R1                  | -                    | 859 |
| F1.AO.06.AO_06_ANG125       | -                    | 846 |
| F1.AR.02.ARE933             | -                    | 846 |
| F1.BE.93.VI850              | -                    | 852 |
| F1.BR.07.07BR844            | -                    | 850 |
| F1.CY.08.CY222              | -                    | 845 |
| F1.ES.11.DEMF110ES001       | -                    | 845 |
| F1.FI.93.FIN9363            | -                    | 843 |
| F1.FR.96.96FR_MP411         | -                    | 844 |
| F1.RO.96.BCI_R07            | -                    | 864 |
| F1.RU.08.D88_845            | -                    | 863 |
| F2.CM.10.DEMF210CM007       | -                    | 848 |
| F2.CM.97.CM53657            | -                    | 850 |
| G.BE.96.DRCBL               | -                    | 855 |
| G.CM.10.DEMG10CM008         | -                    | 848 |
| G.CN.08.GX_2084_08          | -                    | 860 |
| G.CU.99.CU74                | -                    | 866 |
| G.ES.09.X2634_2             | -                    | 859 |
| G.GH.03.03GH175G            | -                    | 863 |



|                            |                      |  |
|----------------------------|----------------------|--|
|                            | Env end              |  |
|                            | gp41 end             |  |
|                            | cytoplasmic tail end |  |
| B.FR.83.HXB2               | *                    |  |
| G.KE.09.DEMG09KE001        | .                    |  |
| G.NG.09.09NG_SC62          | .                    |  |
| G.PT.x.PT3306              | .                    |  |
| G.SE.93.SE6165_G6165       | .                    |  |
| H.BE.93.VI991              | .                    |  |
| H.BE.93.VI997              | .                    |  |
| H.CF.90.056                | .                    |  |
| H.GB.00.00GBAC4001         | .                    |  |
| J.CD.97.J_97DC_KTB147      | .                    |  |
| J.CM.04.04CMU11421         | .                    |  |
| J.SE.93.SE9280_7887        | .                    |  |
| J.SE.94.SE9173_7022        | .                    |  |
| K.CD.97.97ZR_E0TB11        | .                    |  |
| K.CM.96.96CM_MP535         | .                    |  |
| 01.AE.AF.07.569M           | .                    |  |
| 01.AE.CM.10.CYM059         | .                    |  |
| 01.AE.HK.04.HK001          | .                    |  |
| 01.AE.JP.x.DR0492          | .                    |  |
| 01.AE.TH.04.BKM            | .                    |  |
| 01.AE.TH.09.AA111a_WG11    | .                    |  |
| 01.AE.TH.98.CH240          | .                    |  |
| 01.AE.VN.98.98VND15        | .                    |  |
| 02.AG.CM.10.DE00210CM013   | .                    |  |
| 02.AG.ES.06.P1261          | .                    |  |
| 02.AG.FR.91.DJ263          | .                    |  |
| 02.AG.GH.03.03GH181AG      | .                    |  |
| 02.AG.KR.12.12MH11_10746   | .                    |  |
| 02.AG.LR.x.P0C44951        | .                    |  |
| 02.AG.NG.x.IBNG            | .                    |  |
| 02.AG.SN.98.98SE_MP1211    | .                    |  |
| 02.AG.US.06.502_2696_FL01  | .                    |  |
| 03.AB.RU.97.KAL153_2       | .                    |  |
| 04.cpx.CY.94.94CY032_3     | .                    |  |
| 05.DF.BE.x.VI1310          | .                    |  |
| 06.cpx.AU.96.BFP90         | .                    |  |
| 07.BC.CN.98.98CN009        | .                    |  |
| 08.BC.CN.06.mx2            | .                    |  |
| 09.cpx.GH.96.96GH2911      | .                    |  |
| 10.CD.TZ.96.96TZ_BF061     | .                    |  |
| 11.cpx.CM.95.95CM_1816     | .                    |  |
| 12.BF.AR.99.ARMA159        | .                    |  |
| 13.cpx.CM.96.96CM_1849     | .                    |  |
| 14.BG.ES.05.X1870          | .                    |  |
| 15.01B.TH.99.99TH_MU2079   | .                    |  |
| 16.A2D.KR.97.97KR004       | .                    |  |
| 17.BF.AR.99.ARMA038        | .                    |  |
| 18.cpx.CU.99.CU76          | .                    |  |
| 19.cpx.CU.99.CU7           | X                    |  |
| 20.BG.CU.99.CU103          | .                    |  |
| 21.A2D.KE.99.KER2003       | .                    |  |
| 22.01A1.CM.01.01CM_0001BBY | .                    |  |
| 23.BG.CU.03.CB118          | .                    |  |
| 24.BG.ES.08.X2456_2        | .                    |  |
| 25.cpx.CM.02.1918LE        | .                    |  |
| 26.AU.CD.02.02CD_MBT047    | .                    |  |
| 27.cpx.FR.04.04CD_FR_KZS   | .                    |  |
| 28.BF.BR.99.BREPMI2609     | .                    |  |
| 29.BF.BR.01.BREPMI6704     | .                    |  |
| 31.BC.BR.04.04BR142        | .                    |  |
| 32.06A1.EE.01.EE0369       | .                    |  |
| 33.01B.ID.07.JKT189_C      | .                    |  |
| 34.01B.TH.99.OUR2478P      | .                    |  |
| 35.AD.AF.07.169H           | .                    |  |
| 36.cpx.CM.00.00CMNYU830    | .                    |  |
| 37.cpx.CM.00.00CMNYU926    | .                    |  |
| 38.BF1.UY.03.UY03_3389     | .                    |  |
| 39.BF.BR.04.04BRRJ179      | .                    |  |
| 40.BF.BR.05.05BRRJ055      | .                    |  |
| 42.BF.LU.03.LuBF_05_03     | .                    |  |
| 43.02G.SA.03.J11223        | .                    |  |
| 44.BF.CL.00.CH60           | .                    |  |
| 45.cpx.FR.04.04FR_AUK      | .                    |  |
| 46.BF.BR.07.07BR_FP5625    | .                    |  |
| 47.BF.ES.08.P1942          | .                    |  |
| 48.01B.MY.07.07MYKT021     | .                    |  |
| 49.cpx.GM.03.N26677        | .                    |  |
| 51.01B.SG.11.11SG_HM021    | .                    |  |
| 52.01B.MY.03.03MYKL018_1   | .                    |  |
| 53.01B.MY.11.11FIR164      | .                    |  |
| 54.01B.MY.09.09MY58023     | .                    |  |
| 55.01B.CN.10.HNCS102056    | .                    |  |
| 57.BC.CN.09.09YNLX19sg     | .                    |  |
| 61.BC.CN.10.JL100010       | .                    |  |
| 0.BE.87.ANT70              | .                    |  |
| 0.CM.91.MVP5180            | .                    |  |
| 0.CM.98.98CMA104           | .                    |  |
| 0.FR.92.VAU                | .                    |  |
| 0.GA.11.11Gabb352          | .                    |  |
| 0.SN.99.99SE_MP1299        | .                    |  |
| 0.US.10.LTNP               | .                    |  |
| N.CM.02.DJ00131            | .                    |  |
| N.CM.04.04CM_1015_04       | .                    |  |
| N.CM.06.U14842             | .                    |  |
| N.CM.95.YBF30              | .                    |  |
| N.CM.97.YBF106             | .                    |  |
| N.FR.11.N1_FR_2011         | .                    |  |
| P.CM.06.U14788             | .                    |  |
| P.FR.09.RBF168             | X                    |  |
| CPZ.CD.06.BF1167           | .                    |  |
| CPZ.CM.05.SIVcpzMT145      | .                    |  |
| CPZ.GA.88.GAB1             | .                    |  |
| CPZ.TZ.06.SIVcpzTAN13      | .                    |  |
| CPZ.US.85.US_Marilyn       | X                    |  |

|     |
|-----|
| 856 |
| 853 |
| 857 |
| 873 |
| 867 |
| 863 |
| 856 |
| 849 |
| 856 |
| 850 |
| 861 |
| 852 |
| 854 |
| 852 |
| 843 |
| 859 |
| 849 |
| 860 |
| 866 |
| 872 |
| 868 |
| 858 |
| 868 |
| 885 |
| 861 |
| 870 |
| 867 |
| 867 |
| 853 |
| 855 |
| 862 |
| 857 |
| 840 |
| 856 |
| 862 |
| 860 |
| 860 |
| 870 |
| 843 |
| 841 |
| 866 |
| 843 |
| 868 |
| 866 |
| 855 |
| 863 |
| 849 |
| 864 |
| 802 |
| 850 |
| 853 |
| 849 |
| 856 |
| 850 |
| 858 |
| 859 |
| 864 |
| 869 |
| 871 |
| 850 |
| 851 |
| 860 |
| 844 |
| 855 |
| 858 |
| 856 |
| 845 |
| 845 |
| 852 |
| 832 |
| 847 |
| 845 |
| 846 |
| 836 |
| 849 |
| 865 |
| 856 |
| 852 |
| 861 |
| 857 |
| 865 |
| 863 |
| 871 |
| 849 |
| 860 |
| 873 |
| 881 |
| 874 |
| 882 |
| 879 |
| 863 |
| 843 |
| 834 |
| 854 |
| 853 |
| 840 |
| 845 |
| 886 |
| 880 |
| 889 |
| 856 |
| 855 |
| 876 |
| 837 |







