HMMER Sequences in the Nef Alignment

A_U455	HIVU455	M62320	Oram,J.D.	ARHR 6, 1073 (1990)
B_SF2	HIVSF2	K02007	Sanchez-Pescador,R.	Science 227, 484 (1985)
D_ELI	HIVELI	K03454	Alizon,M.	Cell 46, 63 (1986)
O_ANT70C	HIVANT70C	L20587	Vanden Haesevelde,M.	JVI 68, 1586 (1994)
O_MVP5180	HIVMVP5180	L20571	Gurtler,L.G.	JVI 68, 1581 (1994)
CPZGAB	SIVCPZGAB	X52154	Huet,T.	Nature 345, 356 (1990)
A_ROD	HIV2ROD	M15390	Clavel,F.	Nature 324, 691 (1986)
B_EHOA	HIV2EHOA	U27200	Rey-Cuille,M.A.	Virology 202, 471 (1994)
SD_MM251	SIVMM251	M19499	Franchini,G.	Nature 328, 539 (1994)
STM_STM	SIVSTM	M83293	Novembre,F.J.	Virology 186, 783 (1992)
VER_AGM3	SIVAGM3	M30931	Baier,M.	Virology 176, 216 (1990)
GRI_AGM677	SIVAGM677	M66437	Fomsgaard,A.	Virology 182, 397 (1991)
SAB_SAB1C	SIVSAB1C	U04005	Jin,M.J.	EMBO J. 13, 2935 (1994)
SYK_SYK	SIVSYK	L06042	Hirsch,V.M.	JVI 67, 1517 (1993)

The following alignment and most-likely sequence were generated using the HMMER program as described in Part III. For simplicity, only representative types and subtypes are shown. The annotation is based on HIV1s, therefore the user should be cautious about its applicability to other PIVs. Cysteines are indicated by '*'. Two conserved gapless blocks were found using the BLOCKMAKER/Motif program; these are shown with shading. The Motif "Cobbler" sequence follows this alignment. For further information about Nef, see the section by S. Wain-Hobson in Part II of this compendium.

```
most-likely MGGKWSK....RSK.SGWPAVRERMRRAEPET.....YGPAADGVG.AVSR.DLEKHGA.........ITSS..     51
A_U455      -------....K-R.VE--E--K---ETPA..........--K---.---Q.--D-Y--.........V---..     46
B_SF2       -------....--M.G--S-I---------.R.....AE-------.----.------------------..     50
D_ELI       -------....S-I.V----I---I--TN-..........------.----.------------------..     46
O_ANT70     --NALR-.....G-FE--A--------TRTFPE....SE-C-P---.QI--.E-AAR-G.........-P--..     52
O_MVP5180   --NA---.......--FA--SE--D----SSSDP.....QQ-C-P---.-----.E-ATR-G.........-S--..     51
CPZGAB      --T----....S-L.V---E--R-I-E-PT..........--E---.E--K.---R---.........---R..     46
A_ROD       --ASG--...KH-RPPRG..LQ--LL--RAGA.....C-GYWNES-GEY--.FQ-GSDR.........EQK-PS     54
B_EHOA      --SAG--...KQ--QQPG..L---LL--RRGPR....GESSGERQE.RSLQ.YPGGSDK.........GLN-PS     54
SD_MM251    ---AI-M...R---PA-D..L-QKLL--RG--......--RLLGE-EDGS-Q.S-GGL-K.........GL--RS     54
STM_STM     --ASG--...KQR-QH-E..L---LL--RG--......--KLLE-L-EGSGP.SQGASDK.........GLN-HS     54
VER_AGM3    --LGN--PQHKKQL.-L-H-LH....KTRATR.....--LL--PLIGQS-T.LQ-ECDK.........ALKESL     54
GRI_AGM677  --SSN--.....-QQ.Q-LLKLW.-GL-GK-GAD....WVLLS-PLIGQS-T.VQ-EC-K.........ALKK..     52
SAB_SAB1C   ----S--...QQQRH-L-..LWSKL-Q-PVIQ.....-DML--PLLGQS-H.IQ-ECAKSLRDGLIRQGD--..     61
SYK_SYK     --STS--.........-QQ..L-SEGKY-IGWRLFGKQ-T-LP-ELS.RPLQPCRGGFDK.........AWR-TL     54

                                 *
most-likely ............NTA.ANNADCAWL..EAQ..........E..EEEVGFPVRPQVPLRPMTYKAAVDLSHFLKE     98
A_U455      ............--S.ST--S----..---..........-..-GD-------------------F---F----     93
B_SF2       ............---.-T-------..---..........-..----------------------L-I------     97
D_ELI       ............---.ST-------..---..........-.ESD------------------E-L--------     94
O_ANT70     ............H-P.Q---AL-F-..-SH..........Q..--------A-----------G-F---F----     99
O_MVP5180   ............H-P.Q---AL-F-..DSH.........K..D-D---------------F---F---F----     98
CPZGAB      ............--P.ET-QTL---..-EM..........D..N------------T--------F--------     93
A_ROD       CEGRQYQQGDFM--PWKDP-AEREK..NLYRQQNMDDVDSD.DDDQ-RVS-T-K-------HRL-I-M--LI-T    125
B_EHOA      CDGQ........K-L......G-EGG.GK-....DSDEDD-..DN---VR---G-------F-L---M------    107
D_FO784                                                                            X-      2
SD_MM251    CEGQKYNQGQYM--PWR-P-EEKEK..L-YRKQNMDDIDE-.DDDL--VS---K----A----L-I-M---I--    125
STM_STM     CEPQRYNEGQFM--PWK-P-AESAK..LEYRQQNMDDVDE-.DDNL--VA-H-R----E----L-I-----I-S    125
VER_AGM3    IRK.....RNGKM-P.....EGRK-Q.-GD....KWDEWSD.E-D--------R----Q----L---F------    112
GRI_AGM677  ......SWGKGKM-P.....-GRR-Q.-GD....TFDEWDD.D--------Q-R----Q----L---F-----S    109
SAB_SAB1C   ............R-E.....EGVKM..KH-GRQP..SWYD-.D----------CL---A----L-I-FG-----    113
SYK_SYK     TEP......IDPHGP.....-RD-GHSGG-KFSPGDIVQD-GDTGL-----C--T---TL---L-I-----I-N    117

                                                          *                       *
most-likely KGGLEGLIYSQRRQDILDLWVYHTQGYFPDWQNYTPGPGIRYPLTFGWCFKLVPVEPEKVEEA.NEG.ENNCLL    170
A_U455      ----D---H--K--E-----------F----------------------Y-----D-AE----.TG-.---S--    165
B_SF2       --------W-----E-----I-------------------------------------------.---.---S--    169
D_ELI       --------W-KK--E-------N---I----------------------YE----D-QE---D.T--.-T-S--    166
O_ANT70     ----------HK-AE-------N---F------------T-F------L------SE-EA-RLG-TC.-RAN--    172
O_MVP5180   ----D-----HK-AE-----I-----F-----C------P-F------L------SA-EA-RLG-TN.-DAS--    171
CPZGAB      -------V--R---E-----------F--------T---T-F--C---------LTE-Q--Q-.---.D-----    165
A_ROD       R-----MF--E--HK--NIYLEKEE-IIA------H---V---MF---LW-----DVP..Q-G.EDT.-TH--V    195
B_EHOA      --E---IF--E--HK---TYLENEE-IVSG-----H---V---KF---LW----INMI..A-P.EDE.-TH--V    177
D_FO784     ------IY-NH--HR---IYLQNEE-II-------S---E-L-MMY--LW-----DVP..D--.QGD.-RH--V     72
SD_MM251    ------IY--A--HR---MYLEKEE-II----D--S-------K----LW-----NVS..D--.Q-D.-RHY-M    195
STM_STM     ------IY--E--HR---MYLEKEE-IV-------A-------KQ---LW-----DMS..N--.Q-DDGTHY-V    196
VER_AGM3    ----D-IY--D--NQ--N-YALNEW-IID--NAWSE-------RC--F-------DLH..----.ETC.-RH--V    182
GRI_AGM677  ----D-IY--E--EK--N-YALNEW-IID---A-S--------RV--F-------DLH..----.RNC.-RH--M    179
SAB_SAB1C   ------IY--E--KK----YALNEW-IVDG-----D---T---KC----------DLS..----.KNS.--H---    183
SYK_SYK     ----Q-MN-CEK-DE--H-YLQNEH-II.-RI---S---T----I---LWE---N-I-..GCL.EYE.-HTL--    186
```

HIV1 NEF

```
                                                        *
most-likely HPMSQHG.MDDP....EGEVLVWKFDSRLAFHHVARELHPEYYKDC                                211
A_U455      --IC---.V--E....-K---M-----T--LK-R-Y-----F---                                 205
B_SF2       ----L--.-E-A....-K-----R---K-----M------------                                210
D_ELI       --IC---.-E--....-RQ--K-R-N-----E-K---M---F--N                                 206
O_ANT70     --ACA--.FE-T....HK-I-M----RS-GNT---MIT---LFQKD                                213
O_MVP5180   --ACN--.AE-A....H--I-K-Q--RS-GLT-I-LQK---LFPK                                 211
CPZGAB      --IC---.-E-E....DK-----R------LR-I---Q-------                                  205
A_ROD       --.A-TSKF--....H--T---E--PL--YSYE-FIRY--EFGHKSGLPEEEWK.......ARLKARGIPFS     256
B_EHOA      --.A-TSAW--....HE-T---Q---L--YDY--FSRF--EFGYQSGMPEKEWK.......AKLRARGIPTE     238
D_FO784     --.A-TYQX--....XE----X---P---YNYE-FIKY--EFGSKSGLPEDEVRRRLTARGLYKTADKKETG     140
SD_MM251    Q-.A-TSKW--....W----A----PT--YTYE-YARY--EXGSKSGLSEEEVRRR                     247
STM_STM     --.A-THQW--....W---------PL--HTYE-FVR---EFGSKSGLPKEEVERRLTARGLLKMADKKETS     264
VER_AGM3    --.A-VR..E--DGINH---------PM--VQYDPN...RK-LT-MHDLGKR                         228
GRI_AGM677  --.A-M-..E--DGIDH---------PK--VEYRPD.....MF--....MHEHAKR                      223
SAB_SAB1C   --.A-VAYE--A....WK-T------PL--VDY--WR----QVPSAQG                              226
SYK_SYK     --A-XDR.GVRA....W.-N-M-N-NPH--I-Q-G-........WPACSXRDKQ...ENHKSCSQHLVRTSP     241
```

```
      **COBBLER sequence from MOTIF**
>NEF   STM_STM, with embedded consensus blocks
mgasgskkqrkqhgelrerllrargetygkllleglgegsgpsqgasdkglnshscepqry
negqfmntpwknpaaesakleyrqqnmddvdeeddnlvgvavhprvplremtyklaidls
hfikskgglegiyYSERRHEILDLWVYNEEGIIPDWQNYTPGPGVRYPLTFGWCWKLVPV
Dmsneaqeddgthy LLHPACTHGxDDPHGEVLVWKFDPRLAYEYVARxKHPEefgsksgl
pkeeverrltargllkmadkkets
```

```
CONSENSUS-A MGGKWSKsSiVgWPeVrkRmRqT..............?PtAAkGVGAvSQD.....LDkhGAiTSSNt??     48
CONSENSUS-B -------?-??---?--e---ra??????????????-Ep--d------r-.....-e----------aa     46
CONSENSUS-C --------------A--E-I-R-...............QP--E----A---.....---Y--L-----PT     50
CONSENSUS-D --------------AI-E-I-r-?????..........dP--D------R-.....-E----------as     50
CONSENSUS-E --------------Q--E-ik--...............-p-te--------.....------v----m..     48
CONSENSUS-F --?-----------AI-E-L-?-...............-P--E--------.....-ERR----?---R-     46
CONSENSUS-G --------------A--E-I-R-PPAAERK........EA--E--------.....-AR---V-----AA     57
CONSENSUS-O --NA??-?KF?--??--?---R?.........???P?-?PC-P---??-RE.....-A?R-G-?--H-PQ     38
CONSENSUS-U --?----????---??-E-I-?-???............-P???----?---?????-?-?--??--??A-     31

                *       SH3-binding |  |  |  | SH3-binding
CONSENSUS-A tnpsCaWLE?Aqe?.d..e?.VGFPVRPQVPLRPMTYKgAvDLShFLKEKGGLDGLIyS?kRQEILDLWV    110
CONSENSUS-B --ad-----.----.e??-e?-----------------a-?------------e---?-q---d------    108
CONSENSUS-C S-AD----Q.---..EE.GE.-------------------F---F-                             91
CONSENSUS-D --ad-----.---..ES.-E.-----------------e--------------E---W-K----------    115
CONSENSUS-E n-aD-v--r.---..e??-g.-------------------f---F--------------Kr---------    112
CONSENSUS-F N--DL----.---..E..?E.--------------?-----?------------E-----K?---------    106
CONSENSUS-G N--D-----.---..EE.SE.---------------F-S-F---F-                             98
CONSENSUS-O N-AAL-F-?.SH?..?..--.-----?---------?-?-F---F--------?-----H--A------?     93
CONSENSUS-U N-??-???-.??-..E?.-E.---------------?---F---?-----------??------------     83

                                        SH3-binding
                                      *    |  |
CONSENSUS-A YnTQGfFPDWQNYTPGPGtRf.PLTFGWCfKLVPvDPaEVE.eat?GEnNSLLHPICQHGmdDe?revLm    176
CONSENSUS-B -h---y------------?-y?-------------e-ek--.--ne---------msl-----pE----?    174
CONSENSUS-D -----I------------I-Y.--------e------q---.---E--t-c----?-----E-pE-q--k    182
CONSENSUS-E ------------------i-y.--C------------r---.-dnk----C----ms---ie--e-----    180
CONSENSUS-F -H                                                                         108
CONSENSUS-O -?---------?------?--.------L------S?E-A-RLGNT?-?A?----A-?--?E-?H?-I-?    150
CONSENSUS-U -H---?----?-------?-?.--?---------??-?---.--N-----C----?S----?-?E----?    138

CONSENSUS-A WkFDSrLAlkHrA?ElHPEfY.KDC$                                                199
CONSENSUS-B -r------fh-m-r-----y-.---?$HRCVYKGLSAGAFPGRRGWAGLGSGEPSDAA                229
CONSENSUS-D -R-N----fE-K-R-m-----.---                                                  206
CONSENSUS-E -----a--R--i-R-----y-.---                                                  204
CONSENSUS-O -?--RS-G?T-?--??---LF?-?                                                   166
CONSENSUS-U -----S--??-?-R-?---?-.---                                                  157
```