HMMER Sequences in the Tat Alignment

A_U455	HIVU455	M62320	Oram,J.D.	ARHR 6, 1073 (1990)
B_HXB2R	HIVHXB2R	K03455	Starcich,B.	Science 227, 484 (1985)
C_UG268A2	HIVUG268A2	L22948	Louwagie,J.J.	JVI69, 263 (1995)
D_ELI	HIVELI	K03454	Alizon,M.	Cell 46, 63 (1986)
F_BZ163A	HIV1BZ163A	L22085	Louwagie,J.J.	ARHR 10, 561 (1994)
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)
CPZANT	SIVCPZANT	U42720	Vanden Haesevelde,M.	Virology 221, 346 (1996)
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)

HIV1 TAT

The following alignment was generated using the HMMER program as described in the introduction to this Part and in Part III. For simplicity, only representative types and subtypes are shown. An ordinary consensus sequence (lowercase signifies majority, uppercase signifies 50% or better) was created from these representative sequences using MASE; this is not a "most likely sequence" generated from a HMMER model (Part II). Annotation is based on HIV-1s, therefore the user should be cautious about its applicability to other PIV sequences. Putative splice sites are designated by '\/'.

```
                                 <- vpR end
TAT_CONSENSUS ATGga?cca?taga?cc?gag????????????tagagccctggaacca?cc?????????????????     35
A_U455        -----G---G----T--TA-C............C-------------A--C--GGGAAGT..........     48
B_HXB2R       -----G---G----T--TAGA............C-------------G--T--AGGAAGT..........     48
C_UG268A2                                                     %-AGGAAGT..........      8
D_ELI         -----T---G----T--TA-C............C----------------T--AGGAAGT..........     48
F_BZ163A      -----G-T-G----T--TA-C............T----T-----------T--AGGAAGC..........     48
O_ANT70C      -----T---G----T--T---............G-GCCC--T---C----C--TGGAAGT..........     48
O_MVP5180     -----T---G----T--T---............A-GCCC--T---C-T--C--TGGGAGC..........     48
CPZGAB        -----T---A----T--T--C............C-G-----------A--C--AGGCAGT..........     48
CPZANT        -----C--CG----CG-T--A............AC-CCT--A---TTA--C--TCCTGCT..........     48
A_ROD         -----GA--CCCTTGAAG-C-CCAGAGAGCTCAT--A--T----C---G-G--CTTTTCACGCACTTCAG     70
B_EHOA        -----AAT-CCCTTGAAG---CAGGAGAGCTCAT--A-CT---CC-G-GGG-ACTCTTCCTCCACTTCAG     70
SD_MM251      -----GA--CCCTTGAGG---CAGGAGAACTCAT----AT---CC---G-G-GCTCTTCATGCATTTTAG     70
STM_STM       -----GA--CCCTTGAAG---CAGGAGAGCTCAT--AGAT---CC-G-G-G--CTCTTCATGCACTTCAG     70
VER_AGM3      ---..........--CAAG-G-............GAG--CGAGCA-GG-GCGTACCACCAG..........     39
GRI_AGM677    ---..........--CAAG---............GAG-------------..--C--CCTCCTC..........     33
SAB_SAB1C     ---..........--C-AG---CAGGAGGCCCGCCCCC--GT----G-GG-G-TTCAGGAG..........     51
SYK_SYK       ---TCCT--ACG--C-AGATA............TGCC--A-ACA--GGGTA--CCCATCCTTCTTAGAAG     58

TAT_CONSENSUS ??????????????????????????????????????????????????????????????????CC     37
A_U455        ...................................................................CAG...--     53
B_HXB2R       ...................................................................CAG...--     53
C_UG268A2     ...................................................................CAG...--     13
D_ELI         ...................................................................CAG...--     53
F_BZ163A      ...................................................................CAG...--     53
O_ANT70C      ...................................................................CAG...--     53
O_MVP5180     ...................................................................AAG...--     53
CPZGAB        ...................................................................CAG...--     53
CPZANT        ...................................................................ACA...--     53
A_ROD         AGCAGGATGTGGCCACTCAAGAATTGGCCAGACAAGGGGAGGAAATCCTCTCTCAGCTATAC...CGA--    137
B_EHOA        AGGGGGTTGCCAACACTCAAGGATTGGACAACCGGGGGGAGGAAATCCTCTCTCAGCTATAC...CGC--    137
SD_MM251      AGGCGGATGCAACCACTCCAGAATCGGCCAACCTGGGGGAGGAAATCCTCTCTCAACTATAC...CGC--    137
STM_STM       AGGCGGTTGCCGCCACTCCAGGATTGGCCAACCAGGAGGAGGAAATCCTTTGGCAACTATAC...CGC--    137
VER_AGM3      ........................................GACTTGATTGAACAACTCAAA...GCA--     65
GRI_AGM677    ...................................................CAGGACTTGCAT...AGG--     50
SAB_SAB1C     ....................................................GAGCTGCAT...CGG--     65
SYK_SYK       GAACATTCCTAGAGAAA..........................................GGA...--     80

TAT_CONSENSUS ?c?aac?cc?TGtaataa?a??TGctattGtAAaa??TGttgcTatCAtTGcca??ttTGcTTc?taaaa     95
A_U455        TAC---TG-T----GC--C...--T---------GTG-------GG--------AC-------TC-G---    120
B_HXB2R       TAA---TG-T----CC--T..............AG-------T---------AG----T---A---C-    120
C_UG268A2     TAA---T--T----C-...-AG--T-T-G------AA---A-T--------T--AG-------TC-G--C     80
D_ELI         TAGG--T--T-----C...-AG--TC---------AG----------------CAG--------T----C    120
F_BZ163A      TAC---T--T----CA...-AA------------CGA-------T------T--ATGG-----TAC--CG    120
O_ANT70C      C-AG-TC--T--------T...---------C----GA--C------------T-TG----T---G---G-    120
O_MVP5180     C-A---C--T--------T...---------C----GA--C------------T-TG----T---AC---G    120
CPZGAB        AAG---AGTT--------T...---------C---GCT-----T-----C--TATATA-------AC----    120
CPZANT        AGC---G--A--C-----C...--T--C------TGC-----T--C-------CTC-C------AC----    120
A_ROD         C-T-GAAA-A--C-----CTCA-----------GCGA--C-----C-----T--GA-G--T--TC----C    207
B_EHOA        T-TT-AGG-A--C-G---T-CA-----C-------AA----C---C--------GC-------TC-T---    207
SD_MM251      T-TTGAGG-G--CT----C-CA-------------AG--------C--------GT----T--TC-T---    207
STM_STM       A-T-GAGGAG---TG---C-AG----T---C----AA-----------C-----GC-G--T--TG-T-C-    207
VER_AGM3      C-TG-AG-GG----CA--C-AG------------TGT--C--T-----C--T--GC-------TT--C--    135
GRI_AGM677    ATT-CAA--A----C---T-AA--------C----AA--C--T-----C--TG-GC--------C-GC-G    120
SAB_SAB1C     G-TCCAGG-C---G----T-CT----TC--C---GTT--C----T------TATTC-C------CAT---    135
SYK_SYK       A-C---A--A------...-AA----T---C----AC-----------------GC----T---C-TC--    147
```

```
                 \/
                \/     \/    Rev ->    \/
TAT_CONSENSUS AAgGgctTaGG?aT?tcctATggcaggaaga???g???????aga????g?cgaaga????ctcct?aaa    143
A_U455        --A--------C--T----------A-----AGC-G......-A-CCCC-A---G--...C-----C--G    181
B_HXB2R       --A-C------C--C----------------AGC-G.........---CAGC-A------...G---A-C-G-    181
C_UG268A2     --A--------C--C----------------AGC-G.........---CAGC-A----AC...G-----CC--    141
D_ELI         --A--------C--C----------------AGC-G.........---CAGC-A---G--...C-----C--G    181
F_BZ163A      -----------C--C----------------AGC-G......-A-CAGC-G-ACC--...A-----C-G-    181
O_ANT70C      -----T--G--A--C----------------AGC-A......G--..........---CC-...G--G--GC--    175
O_MVP5180     -----T--G--A--C---C------------AGC-A.....---...A-A-C-GC-...G--G--GC--    178
CPZGAB        --A--------A--C----------------AGC-A......-CCACAA-A---C-CACGG-----GC-G    184
CPZANT        --A-----G--G--T----------------GAA-AGCTCGA--GAACA-A----C-...A--G--G---    187
A_ROD         -----GC-C--G--A-GT----AAC-A---GGCA-A......C--.......A----G...A----AA-G-    262
B_EHOA        ------C-G--G--A-GT----AAC-CTC-............--.......AAG---...T--T-AA---    256
SD_MM251      -----A--G--G--A-GT----AGCA-TCA............C--AAGA-AA-----...A----GA---    262
STM_STM       --A--TC-T--A--AA------AGC-ATCA............--.......A-----...GT-AAGA---    256
VER_AGM3      -----A-----TG-TA-----CATGCC...CCTA-G......-TC......A-----...AAGAAGATTG    187
GRI_AGM677    -----------AG-GCGT---CATGTCTCT............--G......AA----...AGAAAAAC--    169
SAB_SAB1C     ----C-C----C--TCG----TATGTC...CCTA-G......CC-......A-----...G--T--A-G-    187
SYK_SYK       -----A--G--A--CA-T----C---A...CCAA-G......-AG.........---GC-...G---GCAGT-    199

                                                 exon \/ exon
TAT_CONSENSUS gca?t?aggat?atcaag?tcctgtAcca???aagcaaccc?tatccaaa????g?ggg?ac...a??ca    194
A_U455        ---G-A-----C-----AC--T-A-----...---------T-GC--C--TCCCAGA--GT-...TCGGC    245
B_HXB2R       A--G-C--AC-C------C-T--C--T--...---------ACC---C--CCCC-A---G--...CCGAC    245
C_UG268A2     ---G-G-----C-----AA--T-A--T--...---------T--C--C--CCCC-A---G--...CCGAC    205
D_ELI         --GG-C---C-C------T----A-----...---------TCC---C-GCCCC-A---G--...CCGAC    245
F_BZ163A      ---GCC--ATAC------A--T-------...---------A-----C--GCCC-A---A--...CCGAC    245
O_ANT70C      --CA-CCA---C--A---A----------...------T--CC-A---TC...ACCAA-AGG...-AG--    236
O_MVP5180     --TA-CCA---A--A---A----------...G-----T--C-G---C-C...ACC--CAGG...-AG--    239
CPZGAB        ---G-A-AA--A------A-T--A-----...---------T-----C-GTCCA-A---A--...-AGG-    248
CPZANT        ---G-G-AA--A------A-------T--...--ATCC-TACCCAAA-CCAGCA-GATCC-G...-GCAG    251
A_ROD         AA-C-A--AC-C---CGTC----AC----...G-CA--T--A------C-AGGACC---G--...-GC--    326
B_EHOA        -GGC-A--AC-AC-AC-TC-T---C----AAC--CAG-T--C-----GC-AGGACC---G--...-GC--    323
SD_MM251      AGGC-A---C-A--AC-TC-T---C-T--...--CA---T-A--C----C...A-GACCCGG...CACTG    323
STM_STM       AGGC-A--AC-T---C-AT-T---C-T--AAC--CAG-T--C------C-AGGGCCA--A--...-GC--    323
VER_AGM3      CTCCGCTT---...-GCTT-----A--A-...--A---T--A------CCCGTG-AA--G--...-GC--    248
GRI_AGM677    ---C-C-A---A------AC---A---GG...C-----T--A------C-GTTCAGC--A--...GGC--    233
SAB_SAB1C     AG-T-TC-C--A------T-T--C---AC...--C                                     219
SYK_SYK       TTTC-G-----......-A-AGC-C---T...-CAGG-A--C--C---GG...GCC--AAGG...-CC--    254

TAT_CONSENSUS ?g?accg?ag?Aa?agaaGaagaaG???GTGGaga??aaggcaga?aCAgatcc?????????????????    237
A_U455        A-GC-A-G-AG--TC----------...-------GC------A-G-------GA...............    297
B_HXB2R       A-GC---A--G--T-----------...-------GAG--A----G--------A...............    297
C_UG268A2     A-GC---G--A--TC----------...-------GC---A----G--------G...............    257
D_ELI         A-GC---A--G--C-----------...-------GCG-------G--------G...............    297
F_BZ163A      A-GC---A--A--TC-------G--...-------GC--------G--------G...............    297
O_ANT70C      G-A--GCC--G--G-AC--G--G--GAG-------AG-------GCC---G-GGATACCCTCGCCGCAAG    306
O_MVP5180     GAA--GCC--G--G-AC--G-----AAG-------AGG--A---GCC--AG-GGACAGCCTTGCCACCAG    309
CPZGAB        A-GCT--G--A--TC--C----G--...----C--GC---A----AG------AG...............    300
CPZANT        CCAGAA-A-AT--G-----...---...-----AGAA---A---GAT---G-GGA...............    300
A_ROD         GCC-A--A--A--C---------C-...-----AGCA-C--TG--G------A-T...............    378
B_EHOA        GCC-A--A--A--C--------G--...-----A-CA-C-AGG-CT-------TTGGCCCTGGCAGATCG    390
SD_MM251      CCA---AG--A-GGCA------G--ACG-------AGGC--TG-CA----C---T...............    378
STM_STM       ACC-AA-A--G--C--------G--...--------CAG---TG--AT--AC---T...............    375
VER_AGM3      GAC-A--C--A-GGGCC--G-----...-------CA-GC---AGA----C---GAGCCTTGGCAGAAAG    315
GRI_AGM677    GAC-A--G--G-GGGC----CAG--...------                                      264
SYK_SYK       G-C-AACCC-C-GAC----------GCA-------CAG--AAG--CT--ACGT-TCTACCT.........    315

TAT_CONSENSUS ?????????????????????????????tg????ga?t?a                                243
A_U455        ..............................-TC......GCT                              303
B_HXB2R       ..............................-TC...--T-AG                              306
C_UG268A2     ..............................--A...--T-A-                              266
D_ELI         ..............................--AGATTGC-G-                              309
F_BZ163A      ..............................--AGATT-G-AG                              309
O_ANT70C      GGTTCT...TGCCACTGTTGTACACGGACC-CA...--ACA-                              342
O_MVP5180     GATTCTTGCAACAGTTGTACACGGATCTCAG-A...C-A-A-                              348
CPZGAB        ..............................--A...--T-AG                              309
CPZANT        ..............................G-G...CCT-GC                              309
A_ROD         ..............................G-CCCT-GCCG-                              390
B_EHOA        AAT............ACATCCACTTCCCGA-TCGCCA-C-G-                              420
SD_MM251      ..............................G-CCTT-GCAG-                              390
STM_STM       ..............................G-CCTT-GCAA-                              387
VER_AGM3      AATCTGGCACAGCAGAGTGGAAGAGCAACTG-T...-CAAGC                              354
SYK_SYK       .......................TCAGCAGAG...A-TCTT                                330
```

HIV1 TAT CONSENSUS

```
                                 <- vpR
CONSENSUS-A   ATGGA?CC?GTAGATCCTAaCCTAGAGCCCTGGAAcCA?CCGGGAAGTcAGCCTACAACTgCTTGTAgCA     67
CONSENSUS-B   -----g--A----------ga--------------g--T--A-------------ag----------c--     70
CONSENSUS-C   ???-?-??-?---------??-----------------T--A-------------A-----------?T-     58
CONSENSUS-D   -----t--a-----------------------------T--A-------------gg---c------A--     70
CONSENSUS-F   -----G-TA------------T----T--------?--T--A-----C------------C------CA-     69
CONSENSUS-O   -----T--A---------G-G?-GCCC--T---C-?--C--T--?--??----CCA?-?CC------AT-     63
CONSENSUS-U   -----T--A-----C-----A-----------------T----------------A-----------C--     70
CONSENSUS-CPZ ---------?----??--G-???????--?---??A--C--??????-???--A??---????--?-AT-     39

                                                                 3'sj  3'sj     rev cds
                                                                  \/    \/        ->
CONSENSUS-A   AgTGTTAcTGTAAAatgTGtTGCTggCATTGccaA?ttTGCTTTCTgaacAAaGGCTTAGGCATcTCCTA    136
CONSENSUS-B   -t--C--T-------a--------tt---------G----t--caca-cA------------------    140
CONSENSUS-C   -------T-------aa---A---At-----t-t-G---------a--ca--------------t-----    128
CONSENSUS-D   -----c-T-------A--------AT---------g-------Ca-A-cg--------------------    140
CONSENSUS-F   -A--C--T------CGA-------TT-----T---TGG------ACA-C?--G-----------------    138
CONSENSUS-O   -T--C--T--C----GA--C----AT------T-TG----T--C??A-??--G--T--G--A------?-    128
CONSENSUS-U   -------T-------A--------AT-------C-G-G--------A----------------------    140
CONSENSUS-CPZ -?--?--?--?---???-----T-A?--?--????-??-----CACA--A--------?--?--?-----     90

                   \/ 3'sj
CONSENSUS-A   TGGCAgGAAGAAGCGG......AgaC?cCGACGAgGAaCTCCTCAaaGcagTaAGGAtCaTCAAAaTCcT    199
CONSENSUS-B   -----------------......----AG------a--g--------ga----c--ac-------gt-t--    204
CONSENSUS-C   -----------------......----AG------A-CG------c-------G---------------?-    191
CONSENSUS-D   -----------------......----AG------a--c--------g--g--C---C-------G?----    203
CONSENSUS-F   -----------------......-A--AG--?-ACC----------G-----?C--ATA------G---T-    200
CONSENSUS-O   ---------------A.........?-??---?-?C-G--G--GC----?A-CCA---?--A--G-----    181
CONSENSUS-U   -----------------......----C-------A--T---------A---CG-----------------    204
CONSENSUS-CPZ -----------???-A?C????-??A-??----?AC??--?--G?-?-----?-AA--A-----G--?--    139

                   exon \/ exon
CONSENSUS-A   aTAcCAAAGCAAcCCtTaCCCcaaaCccaagGGgtC???cCGaCAGgCCcGaAAGAATCGaAGAAGaAGG    266
CONSENSUS-B   c--t-----------?Cct----gc---g-----A-...--------------G----------------    270
CONSENSUS-C   ---T------------------------g-----A-...------------G-g----------------    258
CONSENSUS-D   ----------------CCT----gc---G-----A------------------G----A-----------    270
CONSENSUS-F   G--------------A--T-----G---G----AA-...--------------G?-----------G---    266
CONSENSUS-O   G-----?-----T--C???--??C---???A--AAG...-A-?A-C---A-G-----CA-G--?--?---    236
CONSENSUS-U   ------------T--C-------C-----GA-----...T-----------G--------------G---    271
CONSENSUS-CPZ ?--?----?????--?--?--?-??--AG??--A?-...?A-????--??-?-GA---????----?---    178

CONSENSUS-A   TGGagAGCAAGaCAgAGACAGATCgattCgcTTAG                                       301
CONSENSUS-B   -------aG---------------Cgg---a----                                       305
CONSENSUS-C   --------------?--?------c-----A----                                       291
CONSENSUS-D   -----------G------------c-----A--g-TGA                                    308
CONSENSUS-F   -----------G--??--?-----CG-G?-?----T??                                    297
CONSENSUS-O   ------AG?--?---GCC--?G-G--?A?C--?GCC?C?AGG?TTCTTGC?AC?GTTGTACACGGA?CTC    294
CONSENSUS-U   -----------G------------------A----                                       306
CONSENSUS-CPZ ---?????-------?A?---?-????G???--??                                       196
```