Sequences in the *Tat* Alignment

COMMON NAME	LOCUS	ACC #	FIRST AUTHOR	REFERENCE
SUBTYPE A:				
U455	HIVU455	M62320	Oram,J.D.	ARHR 6, 1073 (1990)
UG273A	HIVUG273A	L22957	Louwagie,J.J.	JVI 69, 263 (1995)
UG275A	HIVUG275A	L22951	Louwagie,J.J.	JVI 69, 263 (1995)
DJ263A	HIVDJ263A	L22941	Louwagie,J.J.	JVI 69, 263 (1995)
DJ264A	HIVDJ264A	L23064	Louwagie,J.J.	JVI 69, 263 (1995)
DJ258A	HIVDJ258A	L22939	Louwagie,J.J.	JVI 69, 263 (1995)
SF170	HIVSF170	M66535	Evans,L.	PNAS 85, 2815 (1988)
SUBTYPE B:				
SF2	HIVSF2	K02007	Sanchez-Pescador,R.	Science 227, 484 (1985)
TH475A	HIVTH475A	L31963	Neumann,M.	JVI 69, 2159 (1995)
LAI	HIVLAI	K02013	Wain-Hobson,S.	Cell 40, 9 (1985)
HXB2R	HIVHXB2R	K03455	Starcich,B.	Science 227, 538 (1985)
NL43	HIVNL43	M19921	Buckler,C.E.	JVI 59: 284 (1986)
NY5CG	HIVNY5CG	M38431	Theodore,T.	PNAS 83: 5038 (1986)
MN	HIVMN	M17449	Gurgo,C.	Virol. 164, 531 (1988)
SC	HIVSC	M17450	Gurgo,C.	Virol. 164, 531 (1988)
BAL1	HIVBAL1	M68893	Gartner,S.	Science 233: 215 (1986)
JRCSF	HIVJRCSF	M38429	Koyanagi,S.	Science 236: 819 (1987)
JRFL	HIVJRFL	M74978	O'Brien,W.A.	Nature 348, 69 (1990)
NH5	HIVNH5	L07424	Weidt,G.	Unpublished (1992)
OYI	HIVOYI	M26727	Wain-Hobson,S.	AIDS 3: 707 (1989)
SF162	HIVSF162	M38428	Cheng-Mayer,C.	JVI 64, 4390 (1990)
CAM1	HIVCAM1	D10112	McIntosh,A.	Unpublished (1991)
CDC42	HIVCDC42	M13137	Desai,S.M.	PNAS 83, 8380 (1986)
SF33	HIVSF33	M38427	York-Higgins,D.	JVI 64, 4016 (1990)
RJS	HIVRJS	M37491	Daniels,R.S.	JVI 65, 5574 (1991)
HAN	HIVHAN	U43141	Sauermann,U.	ARHR 6, 813 (1990)
D31	HIVD31	U43096	Dietrich,U.	Unpublished (1992)
RF	HIVRF	M17451	Starcich,B.R.	Cell 45, 637 (1986)
YU2	HIVYU2	M89974	Li,Y.	JVI 65, 3973 (1991)
BCSG3C	HIVBCSG3C	L02317	Ghosh,S.K.	Virol. 194, 858 (1993)
T810144	HIVT810144	U01471	Sabino,E.	AIDS 8, 901 (1994)
3202A11	HIV3202A11	M64489	Groenink,M.	JVI 65, 1968 (1991)
P896	HIVP896	M96155	Collman,R.	JVI 66, 7517 (1992)
TATAB	HIVTATAB	M64490	Groenink,M.	JVI 65, 1968 (1991)
3202A12	HIV3202A12	U34603	Guillon,C.	ARHR 11, 1537 (1995)
3202A21	HIV3202A21	U34604	Guillon,C.	ARHR 11, 1537 (1995)
MANC	HIVMANC	U23487	Zhu,T.	Nature 3745, 503 (1995)
WEAU160	HIVWEAU160	U21135	Ghosh,S.	Unpublished (1995)
SUBTYPE C:				
SM145	HIVSM145	L11803	Louwagie,J.J.	AIDS 7, 769 (1993)
DJ259	HIVDJ259	L11764	Louwagie,J.J.	AIDS 7, 769 (1993)
DJ373A	HIVDJ373A	L23065	Louwagie,J.J.	JVI 69, 263 (1995)
UG268A	HIVUG268A	L11799	Louwagie,J.J.	AIDS 7, 769 (1993)
SUBTYPE D:				
ELI	HIVELI	K03454	Alizon,M.	Cell 46, 63 (1986)
Z2Z6	HIVZ2Z6	M22639	Theodore,T.	Unpublished (1988)
Z6	HIVZ6	K03458	Srinivasan,A.	Gene 52, 71 (1987)
UG269A	HIVUG269A	L22949	Louwagie,J.J.	JVI 69, 263 (1995)
NDK	HIVNDK	M27323	Spire,B.	Gene 81, 275 (1989)

HIV1 TAT

SE365	HIVSE365	L11797	Louwagie,J.J.	AIDS 7, 769 (1993)
UG274A2	HIVUG274A2	L22950	Louwagie,J.J.	JVI 69, 263 (1995)
MAL	HIVMAL	K03456	Alizon,M.	Cell 46, 63 (1986)
SUBTYPE F:				
BZ126A	HIVBZ126A	L22082	Louwagie,J.J.	ARHR 10, 561 (1994)
BZ163A	HIVBZ163A	L22085	Louwagie,J.J.	ARHR 10, 561 (1994)
SUBTYPE O:				
ANT70C	HIVANT70C	L20587	Vanden Haesevelde,M.	JVI 68, 1586 (1994)
MVP5180	HIVMVP5180	L20571	Gurtler,L.G.	JVI 68, 1581 (1994)
SUBTYPE U:				
ZAM184	HIVZAM184	L22955	Louwagie,J.J.	JVI 69, 263 (1995)
CPZ:				
CPZGAB	SIVCPZGAB	X52154	Huet,T.	Nature 345, 356 (1990)
CPZANT	SIVCPZANT	U42720	Vanden Haesevelde,M.	Virol. In Press (1996)

Subtypes E, G, and H are not currently represented in TAT. The sequence ZAM184 could not easily be
assigned to any of the three TAT subtypes, A, B, C, D and F.

```
                                   intramolecular    3'sj   3'sj
                                  disulfide bonding   \/     \/
                                  |              | rev cds ->/<- nls ->/
CONSENSUS-A   M?PVDPnLEPWnHPGSqPtTaCskCYCK?CCwHCqlCFLnKGLGISYGrKKR..r?RRgtPQs?kDhQnp     64
U455          -E---------K-----------N----V----------K--------K---..KP---P--GS----TL     68
UG273A        -D--------------K---P-------V--Y--PV---K------------..SQ-------N------     68
UG275A        -D------------------P-N-----K--Y--TV----------------..KQ--R---CN------     68
DJ263A        -E----S---------------------I-----------------------..-R-------RQ-----     68
DJ264A        -E----S----------------N----I-----------------------..-R-------RQ-R---     68
DJ258A                         x---------K-----------------------..-R-------RQ-----     50
SF170         -D------------------P-------V--Y--PA---R------------..-Q-------N------     68

CONSENSUS-B   MePVdprLEPWkHPGSQPktactnCYCKkCCfHCqvCFttKgLgISYGRKKR..RQRrrapqdSqthQvs     68
SF2           ------N-----------R---N-----------YA---R------------..--------------A-     68
TH475A        -----------------------T--------------M--A----------..------H-N-----A-     68
LAI           -----------------------T----------------A----------..-----P--G-------     68
HXB2R         ----------------------------------------I--A----------..------H-N-----A-     68
NL43          ----------------------------------------M--A----------..------H-N-----A-     68
NY5CG         ----------------------------R------------A----------..----------E-----     68
MN            ---------------------T---------------K-A----------..-------E--------     68
SC            -D-----------------A---S----------------------------..----------------     68
BAL1          ------K----------------T--------------I-------------..----------E-----     68
JRCSF         ------S------------------------L--------------------..-----P----------     68
JRFL          ------S----------------T-------L------I-------------..--------------A-     68
NH5           --------------------P--T-------L--------------------..----------------     68
OYI           ---------------------SN-----R--L-------K------------..----------K-----     68
SF162         ----------------------------------------I-------------..-------P--EV----     68
CAM1          ------------------R---------Q---------I-------------..-----T--S-K---A-     68
CDC42         ----------------------------------------K-A----------..------H----N--A-     68
SF33          ------N-----------R-------------------I-------------..---Q---DS--N--D-     68
RJS           --------------------E------------------------------..------H-N--N--IP     68
HAN           -------------------R-P----------L-------R---S--------..----------------     68
D31           -----------E--------------------------K------------..-------E--------     68
RF            ----------------------N--------Y------L-------------..----GP--G-------     68
YU2           ------N-----------R---N----------------K------------..-----P--------S-     68
BCSG3C        -----------------------K-------L------M-------------..--------------A-     68
T810144       ----------------------------------------K-A----------..-----P--N--A----     68
3202A11       ------K---------------N----------------K------------..--------------A-     68
P896          ----N-S----------------------------A--I-------------..-----P----------     68
TATAB         ------K---------------N----------------K------------..--------------A-     68
3202A12       ------K---------------N----------------K------------..--------------A-     68
3202A21       ------K---------------N----------------K------------..-----S----E-----     68
MANC          -----------N-----------K-------L------I-------------..-------P----R---     68
WEAU160       -----H----------------------R---------M-------------..-----S--N-----D-     68

CONSENSUS-C   m?pvDP?LEPWnHpGsQPKTaCtkCycKkCsYHClVCFqtKGLGISYGRKKR..RqRRsaP?sSEDHQnp     65
SM145A                   x---C------NN----L-----------------------..-------Q--------     57
DJ259A        SV$I--R---------------N----------------------------..-------P------T-     67
DJ373A        -E----N-----------------------------------------------..-------S-------L     68
UG268A                     x------P----FG-------Q---LN------------..----N--P-------L     55

CONSENSUS-D   mdpvDPnlEpWNHPGSQP?TpCNkChCKKCCYHCqvCFitKgLGISYGRKKR..RQRRrppqggQaHQ?p     66
ELI           ------------------R---------------P---LN------------..----G---------V-     68
Z2Z6          -------I----------K-A--R----------------------------..------S----T--D-     68
Z6            ------------------K-A--R----------------------------..------S----T--D-     68
UG269A                                                 x----------..-----S-K------V-     27
NDK           ---------S--------R-A------------------------------..----K----D----V-     68
SE365A        ------------------K----N-Y---------N----------------..-----T---N-T--VS     68
UG274A        SV$I--R-----------S------Y---------L----------------..--------SS----D-     67
MAL           ------------------R------Y---------M----------------..---------N----D-     68

CONSENSUS-F   MELVDPNLDPWNHPGSQPTTPCTKCYCKRCCFHCQWCFTTKGLGISYGRKKR..KQRHRTPQSSQIHQDL     68
BZ163A        ----------------------------------------------------..----------------     68
BZ126A        ----------------------------------------------------..----------------     68

CONSENSUS-O   MDPVDPE?PPWHHPGS?PQ?PCNNCYCKRCCYHCYVCF??KGLGIS?GRKKR...????AAAS?PD?KDP     55
ANT70         -------V--------Q--I------------------VR------Y-----....GRP----H--H---     66
MVP5180       -------M--------K--T------------------TK------H-----...RRPA----Y--N---     67

CONSENSUS-U   MDPVDPKLEPWNHPGSQPKTACTKCYCKKCCYHCPVCFLNKGLGISYGRKKR..RPRRRSPQNSEDHQNP     68
ZAM184        ----------------------------------------------------..----------------     68

CONSENSUS-CPZ MDP?D????PW?HP???P?T?CNNCYCK?CCYHC??CFTKKGLGISYGRK?R????RRT????S?NNQD?     45
CPZGAB        ---I-PDLE--K--GSQ-R-V-------A-----IY--------------K-T.TR---APAG-K----S     69
CPZANT        ---V-AETP--L--PAT-A-P-------C-----PL--------------R-ARRN---TAES-E----P     70
```

HIV1 TAT

```
            exon \/ exon
CONSENSUS-A   ipKQplPqtqg??ptgpkESkKkVeSKteTDrf?$                          95
U455          --------S-RV.SA-QE---------AK----A                          101
UG273A        -----I-----I.S------Q------------A-                         101
UG275A        -----I-----V.S---E---------------T-                         101
DJ263A        V------T-R-D.--D------E-A------QGD-                         101
DJ264A        VS-----T-R-N.---------E-----K--PCD-                         101
DJ258A        V------T-G-NH--D------E-A---K--QYA-                          84
SF170         ----SI-----V.--------------A-----D                          101

CONSENSUS-B   LsKqP?sqprGD.PtgPkesKkkVErEtEtdP?d?                          99
SF2           -----A--S---.----T--------------F-                          101
TH475A        -----T--S---.------$------------F-$                         100
LAI           -----T------.------$------------F-$                         100
HXB2R         -----T------.------$------------F-$                         100
NL43          -----T--S---.------$------------F-$                         100
NY5CG         -----A------.-------------------DH                          101
MN            -P---AP-F---.-----------------H-V-                          101
SC            -P---A--A---.-------------------V-                          101
BAL1          -----T-H----.----E--------------VH                          101
JRCSF         -P---S--Q---.-------------------DN                          101
JRFL          -----S------.-------------------D-                          101
NH5           -----S------.-DS----------------VH                          101
OYI           -----A------.-------------------E-                          101
SF162         -P---A---Q--.-------------------VH                          101
CAM1          -----A--FQ--.------------G----H-G-                          101
CDC42         -----S--T---.------P--E----A----L-                          101
SF33          -----S------.---------E---------L-                          101
RJS           -P---A------.-------------------VH                          101
HAN           -P---S--Q---.-DS--K----------A--F-                          101
D31           -----A------.------------T------I-                          101
RF            -----T------.--------E----------AVQ                         102
YU2           -----T--L---.----T--------------VH                          101
BCSG3C        ---R-T------.-------------------V-                          101
T810144       ---$                                                         71
3202A11       -----A------.----T-P------------Y-$                         101
P896          -----S------.------Q------------VH$                         101
TATAB         -----A------.----T-P------------Y-$                         101
3202A12       -----A------.----T-P------------Y-$                         101
3202A21       -----A------.-------------------V-$                         101
MANC          -----T------.----E--------------VA$                         101
WEAU160       -----T------.-------------------E-$                         101

CONSENSUS-C   IsKQpLpqtrgd.pTGpEeSKKkVESKtetDpfD$                          98
SM145A        ------------.---S-----------K------                          90
DJ259A        ------------.----------------A-R---                         100
DJ373A        ------SR-Q--.---------------KA--L--                         101
UG268A        --------P---.-----K-------------$--                          87

CONSENSUS-D   iPkQPSSQpRGd?PTGPKe?KKKVESkAEtDpfDw$                          99
ELI           ------------.------Q------E-----$-C-                         101
Z2Z6          ------------.------$------E---------                         101
Z6            -----------N.------$                                         86
UG269A        ------------.------$---------A----$                          59
NDK           --E-----S---.-----K$------E---------                         101
SE365A        V-------H---.------$-----------Q---                         101
UG274A        -----------N.-----K$---------A-----                         100
MAL           L-E-----H---H------$---------A-Q----                         102

CONSENSUS-F   VPKQPISQARGN.PTGPK?SKKEVESKA??DP??$$                          96
BZ163A        ------------.-----K---------ET--$D--                         100
BZ126A        ------------.-----E---------TR--CA--                         101

CONSENSUS-O   VP?QS???T?RK.Q?RQEEQE??VEK??GP?G?P????SC??CTR?S?Q$            83
ANT70         --K--PTI-K--.-E------EE---KA--G-Y-RRKG--HC---T-E-            114
MVP5180       --E--LSH-G--.-K------KK---ET--S-Q-CHQD--NS---I-G--           115

CONSENSUS-U   IPKQSLPHTQRV.STGPEESKKEVESKAETDRFD$                          101
ZAM184        ------------.----------------------                         101

CONSENSUS-CPZ ??K??L??????.?????K??K?V??KT???????                            52
CPZGAB        IP-QP-SQSRGN.KEGSE-ST-E-AS--EADQ$D$                           101
CPZANT        VS-.S-PKTSRI.QSSQK-$E-K-EE--GSGGGPC                           102
```

```
                                   intramolecular    3'sj   3'sj
                                  disulfide bonding   \/     \/
                                  |              | rev cds ->/<- nls ->/
CONSENSUS-A   M?PVDPnLEPWnHPGSqPtTaCskCYCK?CCwHCqlCFLnKGLGISYGrKKR..r?RRgtPQs?kDhQnp     64
CONSENSUS-B   -e----r----k------k---tn----k--f---v--tt------------..-Q--ra--dSqt--vs     68
CONSENSUS-D   -d----------------?-p-N--h--K--Y---v--it------------..-Q--rp--ggQa--?-     66
CONSENSUS-F   -EL-----D-----------P-T-----R--F---W--TT------------..KQ-HR----SQI--DL     68
CONSENSUS-O   -D----E?P--H----?-Q?P-NN----R--Y--YV--??------?-----...-???AAA--P-?KD-     55
CONSENSUS-U   -D----K-----------K---T-----K--Y--PV----------------..-P--RS--NSE-----     68
CONSENSUS-CPZ -D-?-????--?--???-?-?-NN-------Y--??--TK----------?-???---T????S?NN-D?     45

            exon \/ exon
CONSENSUS-A   ipKQplPqtqg??ptgpkESkKkVeSKteTDrf?$                                       95
CONSENSUS-B   Ls---?s-pr-D.------------rE----P?d?                                       99
CONSENSUS-D   -----SS-pR-d-------?-------A---p-Dw$                                       99
CONSENSUS-F   V----IS-AR-N.-----?---E----A??-P?--$                                       96
CONSENSUS-O   V-?-S???-?RK.Q?RQE-QE??--K??GP?G?P????SC??CTR?S?Q$                         83
CONSENSUS-U   ----S--H--RV.S---E----E----A-----D-                                       101
CONSENSUS-CPZ ??-??-?????-.?????K??-?-??--?????-?                                        52
```