
HMMER Sequences in the Gag Alignment

A_U455	HIVU455	M62320	Oram,J.D.	ARRH 6, 1073 (1990)
B_HXB2R	HIVHXB2R	K03455	Starck,B.	Science 227, 538 (1985)
C_UG268	HIVUG268	L11799	Louwagie,J.J.	AIDS 7, 769 (1993)
D_ELI	HIVELI	K03454	Alizon,M.	Cell 46, 63 (1986)
F_BZ163B	HIV1BZ163B	S0585	Louwagie,J.J.	ARRH 10, 561 (1994)
G_LBV217	HIVLBV217	L11778	Louwagie,J.J.	AIDS 7, 769 (1993)
H_VI557	HIVVI557	L11793	Louwagie,J.J.	AIDS 7, 769 (1993)
O_ANT70C	HIVANT70C	L20587	Vanden Haesevelde,M.	JVI 68, 1586 (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)
TAN_AGM17	SIVAGM17	L19250	Hirsch	Virology 197, 426 (1993)
SAB_SAB1C	SIVSAB1C	U04005	Jin,M.J.	EMBO J. 13, 2935 (1994)
SYK_SYK	SIVSYK	L06042	Hirsch,V.M.	JVI 67, 1517 (1993)

HIV1 GAG

The following alignment and most-likely sequence were generated using the HMMER program as described in Part III. For simplicity, only representative type and subtype sequences are shown. The annotation is based on HIV1s, therefore the user should be cautious about its applicability to other PIV sequences. Nuclear localization signal is indicated by nls. MHR is the major homology region. Cleavage sites are indicated by '\/'. Four conserved, gapless blocks were generated using the BLOCKMAKER and Motif programs; these are shown by shading (The Gibbs program identified five blocks that overlap the four Motif blocks). The Motif "Cobbler" sequence follows this alignment.

	<- nls ->	
	<- membrane binding ->	
p17	->	
most-likely	MGARAS.VLSGGKLDWEKIRLRPGGKKKYRLKHLVWASRELERFALNPGLLLETKEGCGQILGQLQPALQTGS	72
A_U455	-----K--S-----N-----K-T-----A-----T	72
B_HXB2R	-----E--R-----K--I-----V-----S--R-----S-----	72
C_UG268	-----R-----T---K-----C-MM-----G---S---K--MK-----T	72
D_ELI	-----K-----I-----Y-----S--K--I---I---T	72
F_BZ163B	-----S-----I-----S---K-I---SS-----	72
G_LBV217	-----E-----M-----T-----Q---S-----T	72
H_VI557	-----D--D-A---L-LIE---K--T	72
O_ANT70	---S---T-S---Q---K-S-----C--E---A-NEKL-Q--E---K---	72
O_MVP51	---T-S---R---S-A-----Y-C---A-TE-L-Q--E---K---	72
CPZGAB	---T---R--V---R-R-MM-----CD---M-S---TKL-Q--E---K---	72
CPZANT	---G---R-E---T-S-----MI---RS--Q---SSS---S---EKAIH--S-SIEIR-	72
A_ROD	---N---R-K-A-EL-R-----I---ANK-D--G-AES---S---K--TV-D-MVP---	72
B_EHOA	---G---K-T-EL--V-----R-M--I---VN---G-AESR-GS---RK-RKV-G-LVP---	72
C_2238	AN--D---AES-----K--TV-E-LVP---	35
D_FO784	AN--D--G-AES---N---K--SV-A-LVP---	35
SD_MM251	---N-.-K-A-EL-----M--V---AN--D--G-AES---N---K--SV-A-LVP---	72
STM_STM	---S-.-K-A-EL--V-----M--V---AN--D--G-AES---S---K-ITV-E-LVP---	72
VER_AGM3	---AT-.A-NRRQ--KF-H---T---QI---I--GK-M---G-HER---SE---KK-IEV-Y-LEP---	72
GRI_AGM677	---GGH-.A--RS--TF-----N---QI---I--GK-M---G-HEK-----K-IEV-T-LEP---	72
TAN_AGM17	---GH-.A--RRN--TF-----N---Q---IX-GK-M---G-HEK-----K-IEV-A-LEP---	72
SAB_SAB1C	---SN-.---R---F-SV---N---K-R-----K--D--S-SAN-----VVK--SV-L-LVP---	72
SYK_SYK	---AG-AI-T-RE--RY-----K--R-LVR---KK--D--G-SDQ-M-S---EK--TV-L-LEAN--	73
most-likely	EELKSLFNTVAVLYCVHQRIEVKDTKEALDK.....IE...E.....EQNK	110
A_U455	---R--Y-----D-----N-----M---	110
B_HXB2R	---R--Y---T-----I-----	110
C_UG268	D--R--Y---T---KG--R-----	110
D_ELI	---R--Y---T---KG-D-----E.....M-----	110
F_BZ163B	---R--Y-----F---V-----L-----	110
G_LBV217	-----A---W-I---G-----EE.....V....K.....R-KN	110
H_VI557	---Q---LL-T-----D-----E.....LK.....A---	110
O_ANT70	DS-Q--W-AIV--W---N-YKIG--QQ-IQ-.....LK.....VMGS	110
O_MVP51	-D---W-AI--W---N-FDIR--QQ-IQ-.....LK.....VMAS	110
CPZGAB	-G-R-----L---W-I-SD-T-E--QK--EQLK..RHHG.....Q-S-	114
CPZANT	P-II-----IC--W---KGEKI---EQ-VKT...VKMK...V.....M-TQ	112
A_ROD	-N-----C-IW-I-AEEK---EG-KQIVR..RHLV...A.....TGT	114
B_EHOA	-N---Y---C-IF-L-AEEK---E-K.....AQR.H.....LAAD	111
C_2238	-N---Y--TC-IW-L-AEQK--H-E-RNEVVE..RHLA...A.....TKN	77
D_FO784	-N---Y---C-IW-L-AEEK---EG-KQIVQ..RHLV...V.....TGT	77
SD_MM251	-N---Y---C-IW-I-AEEK--H-E--KQIVQ..RHLV...V.....TGT	114
STM_STM	-N-----C-IW-I-AEEK--H-E--KQVVK..RHLV...V.....TGT	114
VER_AGM3	-G-----L-C--F---KDK-----E--VAIVRQCCHLV...-KER.....NA-R-T	121
GRI_AGM677	-G--A---LCC-IW-I-AEQK---E--VVT.....VKQHYH.....LVD-	113
TAN_AGM17	-G-----LCC-IW-I-AEQK---E--VVIVKQHCHLV.....KEKT	116
SAB_SAB1C	-N-IA---LCC--A-I-AE-K---E--KA-VK..EEVP...A.....MTE	114
SYK_SYK	-N-----GIIS-VWA--AKK--E--EQ-KQ-.....VK...-ACNWKDDPPATSGGQSENSSQNMA--TSS	136

	p17 \ / p24	
most-likely	SKQK.....AQA..AADTGN...S.....SQVSNYP I V Q N L Q G Q M V H Q P L S P R T L N A	154
A_U455	N--R.....T-----N-----A---P---A-----	150
B_HXB2R	--K-----H-----N-----I-----AI-----	154
C_UG268	IQ-----T..E..T--K-----K-----A-----	149
D_ELI	--K-----N-----AI-----	154
F_BZ163B	-Q-----TK-----K-----SI-----	150
G_LBV217	-Q-----I-----K-D..N.....A-----I-----	152
H_VI557	-QNR.....T-----TG-K---G.....NKI-----A---P---AI-----	154
O_ANT70	R-.S.....-DA..KE--S...A.....R-AG-----S-A-----AI-----	152
O_MVP51	R-.S.....-EA..KEE-S...P.....R-T-----T-A-----AI-----	152
CPZGAB	TE.SNSGSR..EGGAS-G...S-SA-----I-G---L---A-----AI-----	162
CPZANT	AE.TG.....SS-T-.SRGMLL...RLLLLNKQWCQRHL-GEGR---IVDAG-IAR---T-----	170
A_ROD	AE.-.....MPST..SRP-AP...-.....EKGG---HVG-NYT-I-----	156
B_EHOA	TE.-.....MPAM..SKPSKP...T.....RLA---QIA-NYS-L-----	151
C_2238	AE.-.....MPAT..SRP-AP...-.....GGRG---QVA-NY--H-----	119
D_FO784	AD.-.....MPST..SRP-AP...P.....DRGR---QVG-NY--L-----	119
SD_MM251	AE.T.....MPKT..SRP-AP...-.....GRGG---QIG-NY--L-----	156
STM_STM	AN.-.....MPAT..SRP-AP...P.....GRGG---QVG-NY--L-----	156
VER_AGM3	TETS.....SG-K..KN-K-.VTVP.....PGG---F-.A-QQGNAWI-V-----	166
GRI_AGM677	NEKA.....-KKK..NET-AP...P.....GGE-R---V-QNNAW-----	156
TAN_AGM17	AA.TPSGG...QKQNYNT..-PP-.....GNHG---VQQNN-W--T-----	162
SAB_SAB1C	-A.TATSSGQTKELQ-KKK..NEP-VT...P.....GG-R---SVNN-W-----	167
SYK_SYK	GQKV.....V--EKQK-A-PP...P.....R...G---LLR-P-N-WI-TGVPV---KT	179
most-likely	WVKVVEEKAFSPEVIPMFSALESEGATPQDLNLTMLNTVGGHQAMQMLKETINEEAAEWDRLHPVHAGP.IPPG	226
A_U455	-----D-----M---V-----D-----	222
B_HXB2R	-----V-----A--	226
C_UG268	---I-----T-----D-----VA--	221
D_ELI	---I-----A--	226
F_BZ163B	---I-----D-----Q-----	222
G_LBV217	-----D-----I--QQ-----	224
H_VI557	-----A---I-----D-----V-----	226
O_ANT70	---A-----N---I---M-----ISY-I---AI---G-L-V--V---V---T--PPV--L---	224
O_MVP51	---A-----N---I---M-----V-Y-I---AI---G-L-V--V-----T--PAM--L---	224
CPZGAB	-----L--V-----A---G---V--V-----T-----A--	234
CPZANT	---C---N-N-----H-----A-D-G---V--V-----T-----VQA-	242
A_ROD	---L---K-GA--V-G-Q---C--Y-I-Q---C-D-----IIR-I-----VQ--.IP--.L-A-	227
B_EHOA	---L---K-GA--V-G-Q---C--Y-I-Q---C-E-----IIR-I-----D--QQ--.SP--.M-A-	222
C_2238	---L---K-GA--V-G-Q--T-C--Y-I-QL--C-D-----IIR-I-----D--EH--.P--.L-A-	190
D_FO784	---L--D-K-GA--V-G-Q---C--Y-I-Q---C-E-----IIR-I--D--D--QQ--.QP--.L-A-	190
SD_MM251	---LI---K-GA--V-G-Q---C--Y-I-Q---C-D-----IIRDI-----D--LQ--.QPA-Q..Q-	226
STM_STM	---L---K-GA--V-G-Q---C--Y-I-QL--C-E-----II--I-----D--MQ--QPP--.L-A-	228
VER_AGM3	---A---K-GA-IV--Q---C--Y-I-Q---VL-D--G-L-IV--I-----Q--IA--PP--.L-A-	238
GRI_AGM677	---C---RWGA--V---Q---CLSY-V-Q---VI-D--G-L-I--V-----T-RPP--.L-A-	228
TAN_AGM17	---T---R-GA-IV--Q---CLSY-I-Q---VI-D--G---II--V-----QX-LT--PP--.L-A-	234
SAB_SAB1C	---I---K--A--V-----A---I-Y-I-Q---A--E--G-L-IV-DV-----D--LR--PPQQ--.PAQ-	239
SYK_SYK	--EA-NS-K-DASIV-L-Q--T--FI-Y---G---A--D--G---VI-DI---G-A--LE--QPQQ--.A-QA	251

HIV1 GAG

	<- MHR	
most-likely	QMREPRGSDIAGTTSTLQEQIGWMT...SNPPIPVGEIYKRWIILGLNKIVRMYSVPSILDIRQGPKEPFRDY	296
A_U455	-----V-----G-----D-R-----	292
B_HXB2R	-----N-----T-----	296
C_UG268	-----N-----K-----	291
D_ELI	-----A-----V-----	296
F_BZ163B	-I-----Q-----V-----	292
G_LBV217	-I-D-----R-----	294
H_VI557	-----A-----D-----	296
O_ANT70	-I---T-----Q---H-T-R.PNQ.----D-RK--V---M-K-----K-----	295
O_MVP51	-I---T-----Q---I-T-R.GA-.S----D-RK--V---M-K-----	295
CPZGAB	-L-----T-. . .A-----DV-R--V-----V---C-----	304
CPZANT	-L---T-----V---MQ--STPQQ-GGV---D-----M---V--X-----E-K-----	315
A_ROD	-L-----VE---Q--FR.PQ-.V---N-R---QI-Q-C---N-TN---K-----QS-	298
B_EHOA	-L-----VE---Q--YR.PQ-.V---N-R---Q-Q-C---N-TN---K-----QS-	293
C_2238	-L-D-----VE---Q--YR.AQ-.V---N-R---Q-Q-C---N-TN---VK-----QS	260
D_FO784	-L-D-----VE---Q--YR.QQ-.----N-R---Q-Q-C---N-TN---K-----T-QS	260
SD_MM251	-L---S-----SVD---Q--YR.QQ-.----N-R---Q-Q-C---N-TN---VK-----QS-	297
STM_STM	-L---S-----SPE---Q--YR.QQ-.----N-R---Q-Q-C---N-N---VK-----T-QS-	299
VER_AGM3	-L-D-----V---LE-IY..TA--RVD--A-R-----Q-C-K-N--V-----A-K--	309
GRI_AGM677	-L-D-T-----SI---E-TF..NA--R-D--AQ-RK-V---Q-V-Q--N-QKV-----Q--	299
TAN_AGM17	-L-D-----SV---E-TF..NAT-KVD--R--RGXV--R-Q-C-K-N-I-V-----A-G--K--	305
SAB_SAB1C	VL-D-Q-----I---E-T-R.AQ-.AVN--N--G-----Q-C-K-N-N---K-----K--	310
SYK_SYK	GL-D-SA-----SIA--E-I-R.QN.-.VQ-----R-----Q-C-QV-N-----K-----I-K--	322

	MHR ->	p24 \ / \ /	
most-likely	VDRFYKTLRAEQATQEVKNWMTETLLVQNANPDCKTILKALGPGATLEEMMTACQVGGPGHKARVLAEAMSQ		369
A_U455	---F-----D-----S-R-----		365
B_HXB2R	---S-----A-----		369
C_UG268	---F-----D-----D-----R-----S-----		364
D_ELI	---S-D-----Q-----S-----		369
F_BZ163B	---F-----G--D-----Q-----		338
G_LBV217	---F-----D-----D-----S-----		367
H_VI557	---F-A-----G--D-----R--Q--SI-----K-----		369
O_ANT70	-----Q--S-----V-----T-----AT		368
O_MVP51	-----S-----Q-----E-----V-----T--KI-----AS		368
CPZGAB	-----S-----D-----Q-----S-----M		377
CPZANT	-----I---S-P-A---I-----H-----T-S---L-----A-----AS		388
A_ROD	-----S---TDPA---Q-----LV-G--MNP---L-----Q---LM---LKE		371
B_EHOA	-----S---TDPA---Q---I-----LV-G--MNP---L-----I---Q---LM---LKE		366
SD_MM251	-----S---TDAA---Q---I-----LV-G--VNP---L-----Q---LM---LKE		370
STM_STM	-----S---DPS---R---I-----LV-G--MNP---L-----Q---LM---LKE		372
VER_AGM3	-----AI---SG---Q---S-I-----V--G--MHP---L-----SY--K-M--M-QN		382
GRI_AGM677	-----A---P-D---Q---I-----L--G--MNP---LI-----Q---KLMV-M--N		372
TAN_AGM17	-----QA--R--TP-D--X-----I-----H-LV-R--MHP---L-----I-----KLM---Q-		378
SAB_SAB1C	-----A---TDPA---QS--I-----V--G--MNP---L-----I--AQ---LM---TA		383
SYK_SYK	-----HC-----DPS--G-L-Q---I-----E-RQ---M.VKP---LQ-----L---KLM---VM		394

	'p2'	\ / p7	<- Zn motif ->	<- Zn motif	
most-likely	V.TN.....	ANIMMQ.....	RG.N.FRGQRKT...	IKCFNCGKEGHIARNCRAPRKKGCWCKGEGHQMK	423
A_U455	-.QQ.....	TS-----	-----P-R.....	-----L-K-----	418
B_HXB2R	-.--S.....	-T-----	-----N--I...V-----	-----T-----	424
C_UG268	A.N-----	I-----	-S-.--K-PKRI.....	-----L-----R-----	418
D_ELI	A.--SV....	TTA-----	-----K-P--I.....	-----K-----R-----L-	425
G_LBV217	A.SGTA...	TA-----	-----K-.--K-P--N.....	-----D-----L-----R-----	423
H_VI557	-N-----	-VM-----	-----K-.--K--RI.....	-----L-----L-----	424
O_ANT70	A.QQDLKGGYTAVF--	Q.NPIRKG.....	-----Q-----	428
O_MVP51	A.QQDLKGGYTAVF--	Q.NPNRKG.....	-----K-----R-----Q-----	428
CPZGAB	-Q-QG....	R-DVFF-....	K-.Q.GA-PKRK.....	-----L---K---R---R---Q-----	434
CPZANT	A.N-AQ...	GTAVFL-....	-----GN--GKRP....L-----	-----T---K---R---R---Q---L-	446
A_ROD	-IGP.....	-P-PFA.....	AA.Q....--A...F--W-----	S-Q---RQ-----P--I-T	423
B_EHOA	AL-P.....	STNPFA.....	AA.Q.P-AGKR....VT-W---A--T--Q-K--RQ-----	QQ--I-S	421
SD_MM251	ALAP.....	VP-PFA.....	AA.Q.K--P--P.....	W-----S-Q---RQ-----MD-V-A	425
STM_STM	-FQP.....	DPLPFA.....	AA.Q.QQ-R-T....V--W---A--T-KQ-KG--RQ-----	QQ---A	426
VER_AGM3	M.QS.....	QN---Q-..G.Q--RPRPP...V--Y---F--MQ-Q-PE--MR-L---P--LA-			436
GRI_AGM677	G.Q.....	-V-.....	V-PQ.KK-P-GP....L-----	F--MQ-E-K---QIK-F---I--MA-	424
TAN_AGM17	M.QG.....	VN-I-GHSGKG--	R.G--PPR.....	--K--QI--VQKD-PRGGPIK-L---PR-MA-	434
SAB_SAB1C	AFQQQT...	VG--FV-QGA.RP--	P.LG-RGRPLNPN--Y---P--L--F-K--RQ-----	SPD----	450
SYK_SYK	A.QQ.....	SVN-V.....	Q-.P.SK-RSM.....	Y---QI--MQKD-KK-L-AK-FN---T--LAR	446

	pol cds ->		VpR		
	->	p7 \ /	'p1'	\ / p6	<-binding->
most-likely	DCT.....	ERQANFLGKIWPSH.	KGRPGNFL.....	QSRP.....	EPTAPPA.
A_U455	-----	-----	N-----	P-----	-----
B_HXB2R	-----	-----	Y-----	-----	E-----
C_UG268	-----	-----	PN-----	-----	E-----
D_ELI	-----	-----	R-----	-----	-----
G_LBV217	-----	-----	-----	N-----	S-----
H_VI557	-----	-----	SPN-----	-----	-----
O_ANT70	--RN.....	GK-----	Y--PG.GT---YV.....	-----RPA....	H-S---M.
O_MVP51	--KN.....	G-----	Y--PG.GT---YV.....	-----KQV....	S-S---M.
CPZGAB	-----	G--V-----	G--R.S-----	V.....	N-T-----I.
CPZANT	N-PAT....	NTGKV----	PT-TWWGC----	V.....	QKEEVV....
A_ROD	N-P.....	D--G---LGPWG..	-K.-R--PVA.....	QVPQGL....	T----V.
B_EHOA	K-P.....	-----G---FGPWG..	-K.-R--PVQAPQGVPSAPPMPNPAFGMTPQGA....	I-S----	-----
SD_MM251	K-P.....	D--G---LGPWG..	-K.-R--PMA.....	QVHQGL....	T----E.
STM_STM	K-P.....	-----VG---FGPWG..	-K.-H--PMA.....	QIPQGL....	T----E.
VER_AGM3	--RG.....	-----V---YGRWMG.AK.	-R--P.....	AATLGV....	-----PP
GRI_AGM677	--K.....	NG-----	YGHWGG.AK.-R--V.....	Y-GDTVGL-----	ME
TAN_AGM17	--R.....	GG-----	RRMPAPX.GSK-R--.....	E-GGA....	V----IA
SAB_SAB1C	--QK.....	-----V---FGPWGR.GK.	-R--P.....	LTSI....	R----ME
SYK_SYK	A-RQPKRNQGPVVA-----	G-GVS.RRP-A--P.....	-----	V-S.....	S---LE

	\ / minor	
most-likely	ESFGFGEE..TTPSQKQ.EQKDK.....
A_U455	-----I--M--K..M-SPA--L--R.....
B_HXB2R	-----RS-V-T.--P---PI-----
C_UG268	PTA.....	PPA--R-.G.--PP--P--R.....
D_ELI	-----I-----
G_LBV217	-----L-----IA--P---P.....
H_VI557	-----M--PP--R--A.....
O_ANT70	-----EEVK-Q-N.....--GGPN.....
O_MVP51	-----EAVKEQ-N...Q---G-D..E.....
CPZGAB	-----Y-YQ--...EKS-.K-EG.....
CPZANT	-----IYQE.-HKR-QKGL-G.-E.....
A_ROD	-----DPAVDLL-KYMQQGKR-R--RERPYPKVEDLL
B_EHOA	-----DPAEEMLKNYMQLGK--K-NRERPYPK.....
SD_MM251	-----DPAVDLLKNYMQLGKQ-R-SRE-PYK.....
STM_STM	-----DPAADLLRSYMQLGK--R-SRKTPYK.....
VER_AGM3	S.....	PYDPAKLLQYADKQQLR--RK-PPAVN.....
GRI_AGM677	T.....	AYDPAKLLQYAEKG-RLR-EREQTRKQ.....
TAN_AGM17	AHGFPFG.....	PPVAGAYDPAKLLGQYAKKGDQLRR--E.....
SAB_SAB1C	RDYSRPEENWY.....	ADRPPTRGPPDDPATALLKQYAVQGKR.....
SYK_SYK	DIEDGPWLTWSAQMSQQAQAKAQNSPSKPPPTNREVLSPK--S-K.--T.....	

HIV1 GAG

VpR		
<-binding->		/ p6 terminus (80%)
most-likelyE..LYPLASLNSL.....KQ	492
A_U455-..QT--V--K--FGNDPLS-	493
B_HXB2R-..-T--R--FGNDPSS-	500
C_UG268-..P-I--.....S	488
D_ELI-..-T--K--FGNDPLS-	500
G_LBV217-..-T--.....S	487
H_VI557-..S..P-T--.....S	489
O_ANT70-..-F--K--FGT...D-	498
O_MVP51-..-F--K--FGT...D-	498
CPZGAB-..SS--PT--K--FGSDPSS-	508
CPZANT-..-P-SY--K--FGK...D-	522
A_ROD	HLEQGGETPYREPPT..EDL-.H---FGK...D-	522
B_EHOAEVT..EDL-.H---FGE...D-	519
SD_MM251EVT..EDL-.H---FGG...D-	506
STM_STMEVT..EDL-.H---FGE...D-	507
VER_AGM3PDWT-.G-S.....FGE...D-	521
GRI_AGM677KEK-V.EDV.--S--FGG...D-	513
TAN_AGM17-L.EDY.....--.....SS	518
SAB_SAB1CSPQQSPYE..A-S.--R--FGE...D-	554
SYK_SYKS..-S--FGE...D-	554

****COBBLER sequence from MOTIF****

>GAG A_ROD, with embedded consensus blocks

```

mgarnsvlrgkkadelerirlrpggkkkyrlkhivwaankldRFALNPGLLETKEGCQKI
LQQLQPALPTGSENKSLFNTVCviwcihaeekvkdtegakqivrrhlvaetgtaekmps
tsrptapssekgnypvqhvgnnyVHQPLSPRTLNAWVKVVEEKKFGPEVVPMFQALSEG
CTPYDINQMLNCVDHQGAMQIIKEVINEEAAEWDRTHPipgplpagqlreprgsdiagt
tstveeqiqwmfrpqnPIPVGNIYRRWIIILGLQKCVRMYPVSVILDIKQGPKEPFRDYVD
RFYKTLRAEQtdpavknwmtqtllvqnanpdcklvkglgmptleemltacqvgvggq
karlmaealkevigpapipfaaaqqrkafkcwncgkeghsarqcraprqqgkcgkpg
imtncpdrqagflglgpgwkkprnfpvaqvpqgltpappvdpavdllekymqqgkrqre
qrerpykevtedllhlegetpyrepptedllhlnslfgkde

```

HIV1 GAG CONSENSUS

p17 -> /<- nls ->/
 /<- membrane binding ->/
 CONSENSUS-A mGARaSVLsggkLDawekIrLRPgGkKkYrLkHlvwAsreLerFaLnPslLeTaegcqqimeQlqsalkT 70
 CONSENSUS-B -----e-r-----k--i-----v-g---s---R--lg---ps-q- 70
 CONSENSUS-C -----i-r-----?-----h-Mi-----g---s---k--ik---P--Q- 69
 CONSENSUS-D -----?-----?-----i-----G---s---k--ig---P-iq- 68
 CONSENSUS-F -----i-g---s---rk-Ig---ps-Q- 70
 CONSENSUS-G -----?-----?--?-----?-----G---T-----??--P?--Q- 63
 CONSENSUS-H -----?-----?-----?-----?--?--?--L-?I---P--- 64
 CONSENSUS-O ---?---T-S---?---?---S---?---C---?---?E?LLQ---EP--- 62
 CONSENSUS-CPZ ---?---?---?---?---?---M?---?---?---?---?---?---K?---?---P?---? 42

/<- nls ->/
 CONSENSUS-A g?eElkSLfNtvatLycvHqrIdvkdTKeAldkiEeiqnKskqk?????tqgaaA..?T.gs?.sskv 126
 CONSENSUS-B -s---r--y-----e-----E-----k-.....a-----??d-.n-??--q- 128
 CONSENSUS-C -T---r--?-----??-e-r-----E---?Q-----k-.ad?.-k.....- 120
 CONSENSUS-D -s-----?-----e-e-----e-m-E---k-----a---t-.D-.rn...-Q- 125
 CONSENSUS-F -S---r--y---v--f---vE-----L--E---q-.....-?--.dK-.....- 123
 CONSENSUS-G -T--?--?--?--?--?--e-----eEV-Ka-kn-Q-.....-?--??..e?.-n...-q- 110
 CONSENSUS-H -T---Q---LL-?-----?-----?--?--?--?Q???.-----T?.DK.???.??-? 106
 CONSENSUS-O -S??-?-W-AI?V-W---N-??I?--QQ-IQ-LK-V.M?-RKS...A-AAKE.....-..?RQ? 106
 CONSENSUS-CPZ ?S??-?-?V-W-?--?---?---?---?---?---?---?---?---?---?---?---?---? 61

p17 \ / p24
 CONSENSUS-A ???S?NYPIVQNaqqQm?hQ?lSPrTLnAwkVkiEekaFspEVIpMfSaLSEGATpQdLNmMLNiVgGH 190
 CONSENSUS-B ...-----l---V--ai-----v-----T---T--- 194
 CONSENSUS-C ...-----L---v--ai-----t-----t---t--- 186
 CONSENSUS-D ...-----L---V--ai-----t-----T--- 191
 CONSENSUS-F ...-----l---V--i-----T---T--- 188
 CONSENSUS-G ...-----i-----v-----t---T--- 174
 CONSENSUS-H ...-----?V--AI-----V-----A---?--- 170
 CONSENSUS-O ...?-----?-----V--AI-----AV---N--I---M-----??Y-I-T---AI--- 168
 CONSENSUS-CPZ ---??---?---?---?---?---?---?---?V---?---?---?---?---T---A---?--- 107

CONSENSUS-A QAAMQLKdItInEAAewDr?HPVhAgPippgQmREPrGSDIAGtTStlqEqigwmTs...NPPiPVGdI 256
 CONSENSUS-B -----e-----l-----a-----n...-----e- 261
 CONSENSUS-C -----l-----?a-----a----- 252
 CONSENSUS-D -----E-----l-----A-----?-----e- 257
 CONSENSUS-F -----L---q-----i-----q-----v---e- 255
 CONSENSUS-G -----I---?Q-----I---?-----R---...---e- 239
 CONSENSUS-H -----?-----A---?----- 233
 CONSENSUS-O -G-L-V--EV---?---T--P??--L---I---T---Q---?---T-R.??-?--- 229
 CONSENSUS-CPZ -G--V--EV-----L--T---??-?---L--?-----?---?---?---?---?---?--- 160

/<- MHR ->/
 CONSENSUS-A YkrwiIlgLNKIVRMYPvSILDirQgPKPEFrdYVdrFfKtLRAeqAtQeVKnwMTeTLLvQNaNPDCK 326
 CONSENSUS-B -----t-----Y-----s----- 331
 CONSENSUS-C -----k-----D-----d----- 322
 CONSENSUS-D -----Y-----s-d----- 327
 CONSENSUS-F -----g---D----- 325
 CONSENSUS-G -----?---?---D----- 307
 CONSENSUS-H -----?---?---D----- 299
 CONSENSUS-O -RK--V---M-K-----?-----Y-----?----- 297
 CONSENSUS-CPZ -?--?--?---V--?--?---?---?---Y--?---S--?--?---?---?--- 218

Zn-motif
 p24 \ / \ / 'p2' \ / p7 /<-
 CONSENSUS-A sILraLg?gAtLeEMMTacQgVggPgHKArvLAEAmSqv...q??n??iMmQrGnf.rgqkr?iKCFN 384
 CONSENSUS-B T--K---Pa-----...tn-s.at?-----.-n-rKtv--- 394
 CONSENSUS-C T-----P--s-----s-----a...nn.--...-s---.K-p--iv--- 383
 CONSENSUS-D t--K---P?-----s-----a...tn.s-ta-----.K-prki--- 390
 CONSENSUS-F T--K---P-----a...TN.-?a-.-ks--.K--R-iv--- 386
 CONSENSUS-G T--?---P-----?-----A...SG.-A-A.-?---K??-.K-P??-?--- 360
 CONSENSUS-H ?--?---SI-----?---?---?---.TN.-?A.-?---K---.K--R-I?--- 353
 CONSENSUS-O Q--K?--P?-----V-----T---?---A?AQQDLKGGYTA.VF---QN.P?R-G--- 358
 CONSENSUS-CPZ ?--K-----?---?---?---?---?---?---?---.Q.-?-.VF?--?--?G?--?---?--- 262

HIV1 GAG CONSENSUS

	pol cds ->	
	Zn-motif ->/ /<-Zn-motif ->/ p7 \/\ 'pl' \/\ p6	
CONSENSUS-A	CGkEGHlArNCrAPrKkGCwKCgkEGHqMkDCT.?e.rQANFlgkiwpSsKG.RPgNFpQsRp.....	443
CONSENSUS-B	-----i-k-----h-----l-----???????	453
CONSENSUS-C	-----i-----?-----L-----???????	440
CONSENSUS-D	-----i-k-----h-----l-----	449
CONSENSUS-F	-----i-k-----r-----n-----L-----	445
CONSENSUS-G	-----?-----?-----?-----H-----L-?-?.....	414
CONSENSUS-H	-----?-----?-----?-----?-----L-----	406
CONSENSUS-O	-----I-?-?-?-Q-----?..NG?-----Y--PGGT.----YV-????.....	411
CONSENSUS-CPZ	-----?-K--R--R--Q----?--??-??V-----??-??-?-V-?????.....	306
	vpr binding	
	/<-->/ \/\ (minor) (minor) \/\ /<- ->/	
CONSENSUS-AEptAppAE.....?f?gmgeeit.s?...pkqeqkd..??ke??ppl?slKSlFGNDpIS	484
CONSENSUS-B	??..???-----e-.....s-.rf---t-tps????q---pi-...--lY?--a--r-----s-	499
CONSENSUS-C	??????-S-.rF.--t-.pa....--p-??-?-?--t-----x	480
CONSENSUS-DS-.F-----Ps....q-----??--ly.--a-----	494
CONSENSUS-Fs-.F?-----PS.....egly--a----	482
CONSENSUS-G-?-----?..??-?..?S....--P??....--LY?-----	440
CONSENSUS-HS-.F--M-.P-....--?--?--?--?-----	436
CONSENSUS-O?-S--M-.....-?VK.?Q....EN-?-G...-LY.-FA-----T-Q\$	444
CONSENSUS-CPZ-I-.....-Y.??Q--?K.?-....?--?????....??L---?-----?--?--	332
	p6 terminus	
	/ (80%)	
CONSENSUS-A	Q	485
CONSENSUS-B	-\$	500
CONSENSUS-D	-	495
CONSENSUS-CPZ	-	333