

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
P31629	ZEP2_HUMAN	Homo sapiens	Transcription factor HIVEP2	15.941634	S163;T251;S384;S392;S404;T416;S420;T441;T442;T443;S500;S511;S545;S549;S924;S1050;S1057;S1062;T1274;T1278;S1283;S1286;S1320;S1325;S1328;T1377;T1380;S1945;S1961;T1974;S1977;S2157	S819;S950;S955;S1048;S1443;S1447;S2118;S2297;S2301;S2429;S2431	31492838;29351928;34019948;30620550;34846842;27655845;35254053;35132862;33214551;35289036;30379171;28657654	MDTGD TALGQKATSRSGETDKASGR WRQEQSAVIKMFSTFGSHEGQRQPI EPEQIGNTASQLFGSGKCLASPSEVV QQVAEKQYPPHRPSPYSCQHSLSFP QHSLPQGVMMHSTKPHQSLEGPPWL FPGPLPSVASEDLFFPPIHGHSGGYP RKKISSLNPAYSQYSQKSIEQAEAAH KKEHKPKKPGKYICPYCSRACAKPSV LKKHIRSHTGERPYPCIPCGFSFKTK SNLYKHKRSHAHAIKAGLVPFTESA VSKLDLEAGFIDVEAEIHSDEQSTD TDESSLFAEASDKMSPGPPILPLDIA SRGGYHGSLEESLGGPMKVPILIPK SGIPLPNESSQYIGPDMLPNPSLNTK ADDSHTVKQKLALRLSEKKGQDSEP SLNLLSPHSGSTDGSGYFSRSESAE QQISPPNTNAKSYEEIIFGKYCRLSPR NALSVTTSQERAAMGRKGIMEPLP HVNTRLDVKMFEDPVSQLIPSKGDV DPSQTSMLKSTKFNSESRQPQIIPSS IRNEGKLYPANFQGSNPVLEAPVD SSPLIRSNSVPTSSATNLTIPPSLRGS HSFDERMTGSDDFYFPGTVGIPPQR MLRRQA AFELPSVQEGHVEVEHHG RMLKGISSSSLKEKKLSPGDRVGYD YDVCRKPYKKWEDSETPKQNYRDIS CLSSLKHGGEYFMDPVVPLQGVPS MFGTTCENRRRKEKSVGDEEDTP MICSSIVSTPVGIMASDYDPKLQMQ EGVRS GFAMAGHENLSHGHTERFD PCRQQLQPGSPSLVSEESPSAIDSDK MSDLGGRKPPGNVISVIQHTNSLSR PNSFERSESAELVACTQDKAPSPSET CDSEISEAPVSPWAPPGDGAESGG KPSPSQVQVQSYHTQPRLVQRHNI QVPEIRVTEEPDKPEKEKEAQSKPE KPVEEFQWPQRSETLSQLPAEKLPP KKKRLRLADMEHSSGESSFESTGTG LSRSPSQESNLSHSSSFSMSFEREE TSKLSALPKQDEFKHSFELTVPAG SYLSVPGHHHQKEMRRCSSEQMP CPHPAEVPEVRSKSFYDGNLSHAPV SGAAASTVSPSRERKCCFLVRQASFS GSPEISQGEVGMDSVKQEQLHL HAGLRSGWHHGPPAVLPPLQQEDP GKQVAGPCPLSSGPLHLAQPMH MDSQESLRNLIQPTSYM TSKHLPE QPHLFPHQETIPFSPIQNALFQFYYP TVCMVHLPAAQQPPWWQAHFPHFPA QHPQKSYGKPSFQTEIHSSYPLEHV AEHTGKKPAEYAHTKEQTYPCYSGA SGLHPKNLLPKFPSDQSSKSTETPSE QVLQEDFASANAGSLQSLPGTVVVP RIQTHVPSYGSVMYTSISQILGQNSP AIVICKVDENMTQRTLVTNAAMQGI GFNIAQVLGQHAGLEKYPIWKAPQT LPLGLESSIPLCLPSTSDSVATLGGSK RMLSPASSELEFMETKQKRVKEEK MYGQIVEELSAVELTNSDIKKDLSRP QKPQLVRQGCASEPKDGLQSGSSSF SSLSPSSSQDYPSVSPSSREPFLPSK EMLSGSRAPLPQKSSGPSESKESS DEL DIDETASDMSMSPQSSSLPAGD

GQLEEEGKGHKRPVGMIVRMASAP
SGNVADSTLLLTDMAFDQILQFPS
LRTTTTVSWCFLNYTKPNYVQQATF
KSSVYASWCISSCNPNSGLNKTIT
LALLRSKQKITAIEIYTLAAMHRPGTG
KLTSSSAWKQFTQMKPDASFJGSK
LERKLVGNILKERGKGDHGDKDIGS
KQTEPIRIKIFEGGYKSNEDYVYVRG
RGRGKYICEECGIRCKKPSMLKKHIR
THTDVRPYVCKLCNFAFKTKGNLTK
HMKSKAHMKKCLELGVSMTSVDDT
ETEEAENLEDLHKAEEKHSMSSIST
DHQFSDAEESDGEDGDDNDDDDDED
EDDFDDQGDLPKTRSRSTSPQPPR
FSSLPVNVGAVPHGVPSDSSLGHSS
LISYLVTLPSIRVTQLMTPSDSCEDT
QMTEYQRLFQSKSTDSEPKDRLDI
PSCMDEECMLPSEPSSSPRDFSPSS
HHSSPGYDSSPCRDNSPKRYLIPKG
DLSPRRHLSPRRDLSPMRHLSPRKE
AALREMSQRDVSPRRHLSPRRPVS
PGKDITARRDLSPRRERRYMTTIRAP
SPRRALYHNPPLSMGOYLQAEPIVL
GPPNLRRLPQVPYFSLYGDQEGAY
EHPGSSLFPEGPNDYVFSHLPLHSQ
QQVRAPIMVPVGGIQMVHSMPPAL
SSLHPSPTLPLMEGFEEKKGASGE
SFSKDPYVLSKQHEKRGPHALQSSG
PPSTPSSPRLLMKQSTSEDSLATER
EQEENIQCTCKAIASLRIATEEAALLG
PDQPARVQEPHQNPLGSAHVSIRHF
SRPEPGQPCTSATHPDLHDGEKDNF
GTSQTPLAHSTFYKSCVDDKQLDF
HSSKELSSSTEEKDPSSEKSQLH