

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
P31629	ZEP2_HUMAN	Homo sapiens	Transcription factor HIVEP2	18.971071	S163;T251;S384;S392;S404;S412;T416;S420;S439;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	34019948;30379171;35132862;35254053;33214551;35289036;37340703;34846842;27655845;30620550;31492838;29351928;28657654	MDTGD TALGQKATSRSGETDKASGR WRQEQSAVIK MSTFGSHEGQRQPI EPEQIGNTAS AQLFGSGK LASPSEVV QQVAEKQYPPHRSPYSCQHLSL SFP QHSLPQGVMMHSTKPHQSLEGPPWL FPGPLPSVASEDLFFPPIHGHS GGYP RKKISSLNPAYSQYSQKSIEQAEEAH KKEHKPKKPGKYICPYCSRACAKPSV LKKHIRSHTGERPYPCIPCGFSFKTK SNLYKHRKSHAHAIKAGLVPFTESA VSKLDLEAGFIDVEAEIHSDGEQSTD TDESSLFAEASDKMSPGPPILDI SRGGYHGSLEESLGGPMKVPIIIPK SGIPLPNESSQYIGPDMLPNPSLNTK ADDSHTVKQKLALRLSEKKGQDSEP SLNLLSPHSGSTDSGYFSRSESAE QQISPPNTNAKSYEEIIFGKYCRLSPR NALSVTTSQERAAMGRKGIMEPLP HVNTRL DVKMFEDPVSQLIPSKGDV DPSQTSMLKSTKFNSESRQPQIIPSS IRNEGKLYPANFQGSNPVLEAPVD SSPLIRSN SVPTSSATNLTI PPSLRGS HSFDERMTGSDDV FYPGTVGIPPQR MLRRQA AFELPSVQEGHVEVEHHG RMLKGISSSSLKEK KLSPGDRVGYD YDVCRKPYKKWEDSETPKQNYRDIS CLSSLKHGGEYFMDPVVPLQGVPS MFGTTCENRRRKEKSVGDEEDTP MICSSIVSTPVGIMASDYDPKLQMQ EGVRS GFAMAGHENLSHGHTERFD PCR PQLQPGSPSLVSEESPSAIDSDK MSDLGGRKPPGNVISVIQHTNSLSR PNSFERSESAELVACTQDKAPSPSET CDSEISEAPVSP EWAPPGDGAESGG KPSPSQVQVQSYHTQ PRLVRQHNI QVPEIRVTEEPDKPEKEKEAQ SKEPE KPVVEEFQWPQRSETLSQLPAEKLPP KKKRLRLADMEHSSGESSFESTGTG LSRSPSQESNL SHSSFSMSFEREE TSKLSALPKQDEF GKHS EFLTVPAG SYLSVPGHHHQKEMRRCSSEQMP CPHPAEVPEVRSKSF DYGNLSHAPV SGAAASTVSPSRERKKCF LVRQASFS GSPEISQGEVGMDSVKQE QLEHL HAGLRSGWHHGPPAVLPPLQ QEDP GKQVAGPCPPLSSGPLHLAQPQIMH MDSQESLRNPLIQPTSYMTSKHLPE QPHLFPHQETIPFSPIQNALFQFQYP TVCMVHLPAAQQPPWWQAHFPHFPA QHPQKSYGKPSFQTEIHSSYPLEHV AEHTGKKPAEYAHTKEQTYPCYSGA SGLHPKNLLPKF PSDQSSKSTETPSE QVLQEDFASANAGSLQSLPGTVVPV RIQTHVPSYGSVMYTSISQILGQNSP AIVICKVDENMTQRTLVTNAAMQGI GFNIAQVLGQHAGLEKYPIWKAPQT LPLGLESSIPLCLPSTSDSVATLGGSK RMLSPASSLELFMETKQKQKRVKEEK MYGQIVEELSAVELTNSDIKKDL SRP QKPQLVRQGCASEPKDGLQSGSSSF SSLSPSSSQDYPSVSPSSREPFLPSK EMLSGSRAPLPQKSSGPSESKESS DEL DIDETASDMSMSPQSSSLPAGD

GQLEEEGKGHKRPVGM LVRMASAP
SGNVADSTLLLTD MADFQQLQFPS
LRTTTTVSWCFLN YTKPNYVQQATF
KSSVYASWCISSCNPNPSGLNTKTT
LALLRSKQKITA EIYTLAAMHRPGTG
KLTSSSAWKQFTQMKPDASF LFGSK
LERKLVGNILKERGKGD IHGDKDIGS
KQTEPIRIKIFEGGYKSNEDYVYVRG
RGRGKYICEECGIRCKKPSMLKKHIR
THTDVRPYVCKLCNFAFKTKGNLTK
HMKSKAHMKKCLELGVSM TSVDDT
ETEEAENLEDLHKA AEKHSMSST
DHQFSDAEESDGEDGDDNDDDDDED
EDDFDDQGD LTPKTRSRSTSPQPPR
FSSLPVNVGAVPHGVPSDSSLGHSS
LISYLVTLPSIRVTQLMTPSDSCEDT
QMTEYQRLFQSKSTDSEPKDRLDI
PSCMDEECMLPSEPSSSPRDFSPSS
HHSSPGYDSSPCRDNSPKRYLIPKG
DLSPRRHLSPRRDLSPMRHLSPRKE
AALREMSQRDVSPRRHLSPRRPVS
PGKDITARRDLSPRRERRYMTTIRAP
SPRRALYHNPPLSMGQYLQAEPIVL
GPPNLRRLGLQVPYFSLYGDQEGAY
EHPGSSLFPEGPN DYVFSHLPLHSQ
QQVRAPIPMVPVGGIQMVHSMPPAL
SSLHPSPTLPLMEGFEEKKGASGE
SFSKDPYVLSKQHEKRGP HALQSSG
PPSTPSSPRLLMKQSTSEDSL NATER
EQEENIQCTCKAIASLRIATEEAALLG
PDQPARVQEPHQNPLGSAHVSIRHF
SRPEPGQPCTSA THPDLHDGEKDNF
GTSQTPLAHSTFY SKSCVDDKQLDF
HSSKELSSSTEEKDP SSEKSQLH