

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
Q03172	ZEP1_MOUSE	Mus musculus	Zinc finger protein 40	20.468652	T332;S1277;S1282;T1694;S1706;S1771;T1790;S1865;S1910;S1912;S2647	S141;T430;S477;S493;S496;S665;S676;S1030;S1045;S1148;T1253;S1721;S1726;S1735;S1739	36852467;34678516;35822049;36084651	MPRTKQIHPRNLRDKIEEAQKELNG AEVSKKEVLEAGVKGTSSELKGVKR KKIVAENHLKKIPKSLRNPLQTKHK QNTTEPPFSVLPASASESHKKHNCVP AKQGRQFTKQNGETPGMTAESSES GDLVSPKKTSSPHQRSELRRWRSEG SDPTRLGSLDGQRDSSSSSSKARTD NSECSSPCCSTTPPSYTSSTAFDVLLK AMEPELSTLSQKGSSCAIKTEKLRPN KTVRSPSKLKNSSLDAPNATSPDLVV ESPCPPCTSYPVHVAQTQKSEQVAA QCVSHLYSSQDHLVPKLSQQNQQLP GHLGFTGSLTNLHTLESTKLEPIYNT AVTSTVGLTSPSTRQTVPHPHQMD SVSPLSVSPASSTQSPPGPIYSSAHVA SVVSQSVEQMCSLLLRDQKPKKQGK YICEYCNRAKAPSVLLKHIRSHTGE RPYPCVTCGFSTKSNLYKHKKSH AHTIKLGLVLOPEAGGLFLSQECPKA LSVHSDIEDSGESDEEGLADGRQNN PCVKDLQPVQTMKTVSNPESLPKLP SNSDHVVRGFSQDRPDSQAPTEL PKVVVHPVSMPLKTDCLQVANPNP ELPSPQSPRDLHVASILSHSASVSSL EMDESCHQKGDVIQSEGKPDHSHG TAHAQLQRQATEDPQEQGKLLLS PRSLGSTDGYSRSEADQAVSPPT PFARTFPTMDPPAKNGGAPGPRIS APAPSALATGEKSSVVTQMRPPLA TKTLEERISKLISDNEALVDDKQLDS VKPRRTSLRREGSIDSPPKSYIFKDSF QFDLKPMGRRTSSSSDIPKSPFTPTE KSKQVFLLSVPSLDCLPITRSNSMPT TGYSaipaniipppslrgsqsfddki GTYDDVFVSGPNPMPPSGHRPL VRQAAVEDSTASESHVPGSQSVDE SCQGCPSSEAGPVQSKAAQTPHLE KKKSHQGRGTMFECETCRNRYRKL ENFENHKKFYCSELHGPKTKAAVRE AEHGPAAGGAQPQVLHYRVAAPTAV WEQTPQIRKRRKMKSVGDEEDLQP HESGRSPESADALQLQVPVGAAPSP SKHTSATAADQAHRGVQLVARGPER ALPLKQCPMVEQQLNAAAQDKMEV KRQGGISVIQHTNSLSRPNFSDKP EPLEGGITFSLQELGRTGMPGALKVI GMAPEEGHPPQDAMHQTALSHNLR GEPRESARKIPSERVVLGQPLRLVRQ HNIQVPEILVTEEPDRDLEAQSHDE EKSEKFTWPQRSETLSKLPTEKLPPK KKRLRLAEIEHSSTESSFESTLSRSL SRESSLSHASSFSASLDLEDISKVEL APKIDFPSKAEFLLIPLGSNTLSVPGS HREMRAASEQISCVPTLMEVSDFR SKSFDCGSIAPSHVVPALVESQPSYS PSAVGGTAHVPLLERRRGPLIRQISL NIASDHLSPGSAALQTIIVLPSVNT VPFQAPRLPDMAADCPAHTVHPQA LAKDLQAEISSSSSTDFPPQQLFGA HLLNKTNTSLSHQNTPLPLPVSAQG GKPDAAAPTACVSTGEGSFAPKYQL

QCQAFSTSDQGCSAPLRSSPNQVLP
QAGADPCPASEAPPAKAADPMKPC
PLPPLLEGLPRDEVLQKQLPSFVLPV
PQKRNVTVDCFTPVTSLPQILVTQDL
PNMPICQTNQSLVPVSEEQNSMPKS
QNYLQNASPTPEKELACKTVLPEVG
QSVPVSESSPTVQKVSVGRLSPQQE
SSASSKRMLSPANSLDIAMEKHQKR
AKDENGAVCSTNIRALELPSSRANE
SHKQKPVLRQLCTTEPLEGALE
QGACSASGRSSNKAANLTQVLPTDS
LSSRPSTFAVTDHVNELQEFKNTKV
STSLTPTVGSFPAPSESACVLPKSID
NNQEKGSFGRHEENKVIQGRPPL
VSGLSLVSSSDTQQPSFSLKTATSF
TWCYLLRQKALPLAQNQKTSAYTG
WTVSSSNPNPLGLPTKVALSLLNSK
QKTGKSLYCQAITTHSKSDLVYSSK
WKNNLSKRALGNQKATVVEFSNKD
DSEINSEQDKENSLIKSEPRRIKIFD
GGYKSNEDYVYVRGRGRGKYICEEC
GIRCKKPSMLKKHIRTHTDVVRPYHC
SYCNFSFKTKGNLTKHMKSKAHSK
KCVDLGVSGLIDEQDTEESDEKQR
FCGERSGYDLEESDGPDEDDNDNE
EDDDDSQAESGLSAAPSVTASPQHL
PSRSLQDPPGSVEEELRVSSCFSGV
HTDPMDILPRALLTKMTVLSTVQSS
PNRTDLPKARQSTEKDEHEQAPPA
DTPRSPGHQLSVHSESDVLRSPAA
GNPAAGSPGAAVQDSSVGLPPAVAQ
LNPQPAARISSSVSPHPDSQDQKQOI
ILQPPPGLPSPQTHLFSHLPLHSQQQ
SRTPYNMVPVGGIHVVTAGLTYSTFV
PIQAGPMQLTIPAVSVIHRTVGTSGD
TITEASGSPNRPTGVAELSSVPCIP
GQIHVPGLQNLSPPALQSLTSLGME
TVNLVGLANATVGPQGHPPGLALNA
VGLQVLANAPAQSSPAPPAHIQGLQI
LNIALPTLIPSVGPVAVGTTGTPETTA
PNSKAMELQMPAGQGHSAPPEPQGS
PEGPQETPQTVSGPSADHARPEDST
KMDTKKGPSAGHVLPGRSPAQAQPA
PTPEALQKVATSAPPSLPTDRAAPRP
PVPHRQPIVHFSDVSSDDDEDRLVIA
T