	UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
0	P15822	ZEP1_HUMAN	Homo sapiens	Zinc finger protein 40	31.045117	S523;S667;T699;S710;S717;T718;S804;T837;S1550;S1608;S1610;S1633;S1749;S1753;T1805;S1863;S1864;S2220;S2227;S2295;T2325;S2587	S141;T429;S476;S479;S492;S495;S571;S577;S670;S681;S1036;S1051;S1091;S1158;S1161;S1180;T1268;S1735;S1740;S1749;S1753;S1884;S2033;S2327;S2599;S2669;S2682	31637018;20068230;33214551;30059200;31492838;20305658;28657654;32574038;29237092;34019948;30620550;23301498	MPRTKQIHPRNLRDKIEEAQKELNG AEVSKKEILQAGVKGTSESLKGVKRK KIVAENHLKKIPKSPLRNPLQAKHKQ NTEESSFAVLHSASESHKKQNYIPV KNGKQFTKQNGETPGIIAEASKSEES VSPKKPLFLQQPSELRRWRSEGADP AKFSDLDEQCDSSSLSSKTRTDNSE CISSHCGTTSPSYTNTAFDVLLKAME PELSTLSQKGSPCAIKTEKLRPNKTA RSPPKLKNSSMDAPNQTSQELVAES QSSCTSYTVHMSAAQKNEQGAMQS ASHLYHQHEHFVPKSNQHNQQLPG CSGFTGSLTNLQNQENAKLEQVYNI AVTSSVGLTSPSSRSQVTPQNQQMD SASPLSISPANSTQSPPMPIYNSTHV ASVVNQSVEQMCNLLLKDQKPKKQ GKYICEYCNRACAKPSVLLKHIRSHT GERPYPCVTCGFSFKTKSNLYKHKK SHAHTIKLGLVLQPDAGGLFLSHESP KALSIHSDVEDSGESEEEGATDERQ HDLGAMELQPVHIIKRMSNAETLLK SSFTPSSPENVIGDFLLQDRSAESQA VTELPKVVVHHVTVSPLRTDSPKAM DPKPELSSAQKQKDLQVTNVQPLSA NMSQGGVSRLETNENSHQKGDMN PLEGKQDSHVGTVHAQLQRQQATD YSQEQQGKLLSPRSLGSTDSGYFSR SESADQTVSPPTPFARRLPSTEQDSG RSNGPSAALVTTSTPSALPTGEKALL LPGQMRPPLATKTLEERISKLISDNE ALVDDKQLDSVKPRRTSLSRRGSIDS PKSYIFKDSFQFDLKPVGRRTSSSSD IPKSPFTPTEKSKQVFLLSVPSLDCLP ITRSNSMPTTGYSAVPANIIPPPHPLR GSQSFDDKIGTFYDDVFVSGPNAPV PQSGHPRTLVRQAAIEDSSANESHV LGTGQSLDESHQGCHAAGEAMSVR SKALAQGPHIEKKKSHQGRGTMFEC ETCRNRYRKLENFENHKKFYCSELH GPKTKVAMREPEHSPVPGGLQPQIL HYRVAGSSGIWEQTPQIRKRRKMKS VGDDEELQQNESGTSPKSSEGLQFQ NALGCNPSLPKHNVTIRSDQQHKNI QLQNSHIHLVARGPEQTMDPKLSTI MEQQISSAAQDKIELQRHGTGISVIQ HTNSLSRPNSFDKPEPFERASPVSF QELNRTGKSGSLKVIGISQEESHPSR DGSHPHQLALSDALRGELQESSRKS PSERHVLGQPSRLVRQHNIQVPEILV TEEPDRDLEAQCHDQEKSEKFSWP QRSETLSKLPTEKLPPKKKRLRLAEI EHSSTESSFDSTLSRSLSRESSLSHT SSFSASLDIEDVSKTEASPKIDFLNK AEFLMIPAGLNTLNVPGCHREMRRT ASEQINCTQTSMEVSDLRSKSFDCG SITPPQTTPLTELQPPSSPSRVGVTG HVPLLERRRGPLVRQISLNIAPDSHL SPVHPTSFQNTALPSVNAVPYQGPQ LTSTSLAEFSANTLHSQTQVKDLQA ETSNSSSTNVFPVQQLCDINLLNQI HAPPSHQSTQLSLQVSTQGSKPDKN

									SVLSGSSKSEDCFAPKYQLHCQVFT SGPSCSSNPVHSLPNQVISDPVGTD HCVTSATLPTKLIDSMSNSHPLLPPE LRPLGSQVQKVPSSFMLPIRLQSSVP AYCFATLTSLPQILVTQDLPNQPICQ TNHSVVPISEEQNSVPTLQKGHQNA LPNPEKEFLCENVFSEMSQNSSLSE SLPITQKISVGRLSPQQESSASSKRM LSPANSLDIAMEKHQKRAKDENGAV CATDVRPLEALSSRVNEASKQKKPIL VRQVCTTEPLDGVMLEKDVFSQPEI SNEAVNLTNVLPADNSSTGCSKFVV IEPISELQEFENIKSSTSLTLTVRSSP APSENTHISPLKCTDNNQERKSPGV KNQGDKVNIQEQSQQPVTSLSLFNI KDTQQLAFPSLKTTTNFTWCYLLRQ KSLHLPQKDQKTSAYTDWTVSASNP NPLGLPTKVALALLNSKQNTGKSLY CQAITTHSKSDLLVYSSKWKSSLSKR ALGNQKSTVVEFSNKDASEINSEQD KENSLIKSEPRRIKIFDGGYKSNEEY VYVRGRGRGKYICEECGIRCKKPSM LKKHIRTHTDVRPYHCTYCNFSFKT KGNLTKHMKSKAHSKKCVDLGVSV GLIDEQDTEESDEKQRFSYERSGYD LEESDGPDEDDNENEDDDEDSQAE SVLSATPSVTASPQHLPSRSSLQDPV STDEDVRITDCFSGVHTDPMDVLPR ALLTRMTVLSTAQSDYNRKTLSPGK ARQRAARDENDTIPSVDTSRSPCHQ MSVDYPESEEILRSSMAGKAVAITQS PSSVRLPPAAAEHSPQTAAGMPSVA SPHPDPQEQKQQITLQPTPGLPSPHT HLFSHLPLHSQQQSRTPYNMVPVG GIHVVPAGLTYSTFVPLQAGPVQLTIP AVSVVHRTLGTHRNTVTEVSGTTNP AGVAELSSVVPCIPIGQIRVPGLQNLS TPGLQSLPSLSMETVNIVGLANTNM APQVHPPGLALNAVGLQVLTANPSS QSSPAPQAHIPGLQILNIALPTLIPSV SQVAVDAQGAPEMPASQSKACETQP KQTSVASANQVSRTESPQGLPTVQR ENAKKVLNPPAPAGDHARLDGLSK MDTEKAASANHVKPKPELTSIQGQP ASTSQPLLKAHSEVFTKPSGQQTLSP DRQVPRPTALPRRQPTVHFSDVSSD DDEDRLVIAT