	UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
0	P15822	ZEP1_HUMAN	Homo sapiens	Zinc finger protein 40	39.007362	T331;S332;S523;T527;S588;S667;T699;S706;S710;T715;S717;T718;T724;S804;S827;T837;T845;S881;T1330;T1366;T1416;S1550;S1608;S1610;S1632;S1633;S1744;S1749;S1753;T1805;S1863;S1864;S1913;T1918;S1919;S2220;S2227;T2228;S2295;T2325;S2327;T2586;S2587;S2590;T2597;S2659;S2661	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	34019948;35132862;35254053;31637018;20305658;29237092;33214551;20068230;30059200;35289036;34846842;30620550;37340703;23301498;31492838;32574038;28657654	MPRTKQIHPRNLRDKIEEAQKELNGAEVSKKEILQAGVKGTSESLKGVKRKKIVAENHLKKIPKSPLRNPLQAKHKQNTEESSFAVLHSASESHKKQNYIPVKNGKQFTKQNGETPGIIAEASKSEESVSPKKPLFLQQPSELRRWRSEGADPAKFSDLDEQCDSSSLSSKTRTDNSECISSHCGTTSPSYTNTAFDVLLKAMEPELSTLSQKGSPCAIKTEKLRPNKTARSPPKLKNSSMDAPNQTSQELVAESQSSCTSYTVHMSAAQKNEQGAMQSASHLYHQHEHFVPKSNQHNQQLPGCSGFTGSLTNLQNQENAKLEQVYNIAVTSSVGLTSPSSRSQVTPQNQQMDSASPLSISPANSTQSPPMPIYNSTHVASVVNQSVEQMCNLLLKDQKPKKQGKYICEYCNRACAKPSVLLKHIRSHTGERPYPCVTCGFSFKTKSNLYKHKKSHAHTIKLGLVLQPDAGGLFLSHESPKALSIHSDVEDSGESEEEGATDERQHDLGAMELQPVHIIKRMSNAETLLKSSFTPSSPENVIGDFLLQDRSAESQAVTELPKVVVHHVTVSPLRTDSPKAMDPKPELSSAQKQKDLQVTNVQPLSANMSQGGVSRLETNENSHQKGDMNPLEGKQDSHVGTVHAQLQRQQATDYSQEQQGKLLSPRSLGSTDSGYFSRSESADQTVSPPTPFARRLPSTEQDSGRSNGPSAALVTTSTPSALPTGEKALLLPGQMRPPLATKTLEERISKLISDNEALVDDKQLDSVKPRRTSLSRRGSIDSPKSYIFKDSFQFDLKPVGRRTSSSSDIPKSPFTPTEKSKQVFLLSVPSLDCLPITRSNSMPTTGYSAVPANIIPPPHPLRGSQSFDDKIGAFYDDVFVSGPNAPVPQSGHPRTLVRQAAIEDSSANESHVLGTGQSLDESHQGCHAAGEAMSVRSKALAQGPHIEKKKSHQGRGTMFECETCRNRYRKLENFENHKKFYCSELHGPKTKVAMREPEHSPVPGGLQPQILHYRVAGSSGIWEQTPQIRKRRKMKSVGDDEELQQNESGTSPKSSEGLQFQNALGCNPSLPKHNVTIRSDQQHKNIQLQNSHIHLVARGPEQTMDPKLSTIMEQQISSAAQDKIELQRHGTGISVIQHTNSLSRPNSFDKPEPFERASPVSFQELNRTGKSGSLKVIGISQEESHPSRDGSHPHQLALSDALRGELQESSRKSPSERHVLGQPSRLVRQHNIQVPEILVTEEPDRDLEAQCHDQEKSEKFSWPQRSETLSKLPTEKLPPKKKRLRLAEIEHSSTESSFDSTLSRSLSRESSLSHTSSFSASLDIEDVSKTEASPKIDFLNKAEFLMIPAGLNTLNVPGCHREMRRTASEQINCTQTSMEVSDLRSKSFDCGSITPPQTTPLTELQPPSSPSRVGVTGHVPLLERRRGPLVRQISLNIAPDSHLSPVHPTSFQNTALPSVNAVPYQGPQLTSTSLAEFSANTLHSQTQVKDLQAETSNSSSTNVFPVQQLCDINLLNQIHAPPSHQSTQLSLQVSTQGSKPDKN

									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