

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
P24928	RPB1_HUMAN	Homo sapiens	DNA-directed RNA polymerase II subunit RPB1	22.632334	S1534;S1619;S1621	S27;S217;T1840;S1843;S1845;S1847;S1849;S1850;T1854;S1857;T1860;S1861;T1863;S1864;T1867;S1868;T1870;T1874;S1875;T1877;S1878;T1881;S1882;T1885;T1894;S1896;S1899;S1906;T1909;T1912;S1913;T1915;T1916;S1917;T1919;S1920;T1923;T1926;S1927;T1929;T1930;S1931;T1933;S1934	34019948;34725712;26807597;37217939;22605332;22371499;26853435;29351928	MHGGGPPSGDSACPLRTIKRVQFGV LSPDELKRMSVTEGGIKYPETTEGG RPKLGGLMDPRQGVIERGTGRCQTCA GNMTECPGHFGHIELAKPVFHVGF VKTMKVLRVCVFFCSKLLVDSNNPK IKDILAKSKGQPKKRLTHVYDLCKGK NICEGGEEMDNKFGVEQPEGEDDL TKEKGHGGCGRYQPRIRRSGLLEYA EWKHVNEDSQEKKILLSPERVHEIF KRISDEECFVLGMEPRYARPEWMIV TVLPVPPLSVRPVAVVMQGSARNQDD LTHKLADIVKINNQLRRNEQNGAAA HVIAEDVKLLQFHVATMVDNELPGL PRAMQKSGRPLKSLKQRLKGEGRV RGNLMGKRVDFSARTVITPDPNLSI DQVGVPRSIAANMTFAEIVTPFNIDR LOELVRRGNSQYPGAKYIIRDNGDRI DLRFHPKPSDLHLQTYKVERHMC DGDIVIFNRQPTLHKMSMMGHRVRI LPWSTFRLNLSVTTYPYNADFDGDEM NLHLPOSLETRAEIQELAMVPRMIV TPQSNRPVMGIVQDTLTAVRKFTKR DVFLERGEVMNLLMFLSTWDGKVP QPAILKPRPLWTGKQIFSLIIPGHINC IRTHSTHPDDEDSGPYKHISPGDTKV VVENGELIMGILCKKSLGTSAGSLVH ISYLEMGHDITRLFYSNIQTVINNW LIEGHTIGIGDSIADSKTYQDIQNTIK KAKQDVIEVIEKAHNNELEPTPGNT LRQTFENQVNRILNDARDKTGSSAQ KSLSEYNNFKSMVVS GAKGSKINIS QVIAVVGQQNVEGKRIPFGFKHRTL PHFIKDDYGPESRGFVENSYLAGLTP TEFFFHAMGGREGLIDTAVKTAETG YIQRRLIKSMESVMVKYDATVRNSIN QVVQLRYGEDGLAGESVEFQNLATL KPSNKAFEKKFRFDYTNERALRRTL QEDLVKDVLSNAHIQNELEREFERM REDREVLRVIFPTGDSKVVLPCNLLR MIWNAQKIFHINPRLPSDLHPIKVV GVKELSKKLVIVNGDDPLSRQAQEN ATLLFNIHLRSTLCSRRMAEEFRLSG EAFDWLLGEIESKFNQAIHPGEMV GALAAQSLGEPATQMTLNTFFHYAGV SAKNVTLGVPRLKELINISKKPKTPS LTVFLLGQSARDAERAKDILCRLEHT TLRKVTANTAIYYDPNPQSTVVAEDQ EWVNVYYEMPDPFDVARISPWLLRVE LDRKHMTDRKLTMEQIAEKINAGFG DDLNCIFNDDNAEKLVLIRIMNSD ENKMQEEEEVVDKMDDDVFLRCIE SNMLTDMTLQIEQISKVYMHLPTQ DNKKKIIITEDGEFKALQEWILETDG VSLMRVLSEKDVPVRTTSDNDIVEIF TVLGIEAVRKALEREYHVISFDGSY VNRYRLALLCDTMTCRGHLMAITR HGVNRQDTGPLMKCSFEETVDVLM EAAAHGESDPMKGVSENIMLGQLA

PAGTGCFDLLLLDAEKCKYGMEIPTNI
PGLGAAGPTGMFFGSAAPSPMGGISP
AMTPWNQGATPAYGAWSPVSGSM
TPGAAGFSPAASDASGFSPGYSPA
WSPTPGSPGSPGPSSPYIPSPGGAMS
PSYSPTSPAYEPRSPGGYTPQSPSYS
PTSPSYSPTSPSYSPTSPNYSPTSPSY
SPTSPSYSPTSPSYSPTSPSYSPTSPS
YSPTSPSYSPTSPSYSPTSPSYSPTSP
SYSPTSPSYSPTSPSYSPTSPSYSPTS
PSYSPTSPSYSPTSPSYSPTSPNYSPT
SPNYTPTSPSYSPTSPSYSPTSPNYTP
TSPNYSPTSPSYSPTSPSYSPTSPSYS
PSSPRYTTPQSPTYTPSSPSYSPSSPSY
SPASPKYTPTSPSYSPPSPEYTPSPK
YSPTSPKYSPTSPKYSPTSPYSPPTP
KYSPTSPYSPYSPVYTPSPKYSPTS
PTYSPYSPKYSPTSPYSPYSPKYSTY
SPTSPGYSPYSPYSLTSPAISPDDSD
EEN