

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
P08775	RPB1_MOUSE	Mus musculus	DNA-directed RNA polymerase II subunit RPB1	15.111155	NaN	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	24968300;29187734	MHGGGPPSGDSACPLRTIKRVQFGV LSPDELKRMDSVTEGGIKYPETTEGG RPKLGGLMDPRQGVIERTRGRCQTCA GNMTECPGHFGHIELAKPVFHVGF VKTMKVLRVCVFFCSKLLVDSNNPK IKDILAKSKGQPKKRLTHVYDLCKGK NICEGGEEMDNKFGVEQPEGDEDL TKEKGHGGCGRYQPRIRRSGLELYA EWKHVNEDSQEKKILLSPERVHEIF KRISDEECFVLGMEPRYARPEWMIV TVLPVPLSVRPAVVMQGSARNQDD LTHKLADIVKINNQLRRNEQNGAAA HVIAEDVKLLQFHVATMVDNELPGL PRAMQKSGRPLKSLKQRLKGKEGRV RGNLMGKRVDFSARTVITPDPNLSI DQVGVPRSIAANMTFAEIVTPFNIDR LQELVRRGNSQYPGAKYIIRDNGDRI DLRFHPKPSDLHLQTYKVERHMC DGDIVIFNRQPTLHKMSMMGHRVRI LPWSTFRLNLSVTTTPYNADFDGDEM NLHLPQSLETRAEIQELAMVPRMIV TPQSNRPVMGIVQDTLTAVRKFTKR DVFLERGEVMNLLMFLSTWDGKVP QPAILKPRPLWTGKQIFSLIIPGHINC IRTHSTHPDDEDSGPYKHISPGDTKV VVENGELIMGILCKKSLGTSAGSLVH ISYLEMGHDITRLFYSNIQTVINNW LIEGHTIGIGDSIADSKTYQDIQNTIK KAKQDVIEVIEKAHNNELEPTPGNT LRQTFENQVNRILNDARDKTGSSAQ KSLSEYNNFKSMVVSGAKGSKINIS QVIAVVGQQNVEGKRIPFGFKHRTL PHFIKDDYGPESRGFVENSYLAGLTP TEFFFHAMGGREGLIDTAVKTAETG YIQRRLIKSMESVMVKYDATVRNSIN QVVQLRYGEDGLAGESVEFQNLATL KPSNKAFEKKFRFDYTNERALRRTL QEDLVKDVLSNAHIQNELEREFERM REDREVLRVIFPTGDSKVLPCLNLLR MIWNAQKIFHINPRLPSDLHPIKVVE GVKELSKLVIVNGDDPLSRQAQEN ATLLFNIHLRSTLCSRRMAEEFRLSG EAFDWLLGEIESKFNQAIHPGEMV GALAAQSLGEPATQMTLNTTFHYAGV SAKNVTLGVPRLKELINISKPKTPS LTVFLLGQSARDAERAKDILCRLEHT TLRKVTANTAIYYDPNPQSTVVAEDQ EWNVYYEMPFDVARISPWLLRVE LDRKHMTDRKLTMEQIAEKINAGFG

DDLNCIFNDDNAEKLVLRLRIMNSD  
ENKMQEEEEVVDKMDDVFLRCIE  
SNMLTDMTLQIEQISKVYMHLPQT  
DNKKKIITEDGEFKALQEWILETDG  
VSLMRVLSEKDVPVRTTSNDIVEIF  
TVLGIEAVRKALEREYHVISFDGSY  
VNYRHLALLCDTMTCRGHLMAITR  
HGVNRQDTGPLMKCSFEETVDVLM  
EAAAHGESDPMKGVSENIMLGQLA  
PAGTGCFDLLLLDAEKCKYGMEIPTNI  
PGLGAAGPTGMFFGSAPSPMGGISP  
AMTPWNQGATPAYGAWSPSVGSGM  
TPGAAGFSPSAASDASGFSPGYSPA  
WSPTPGSPGSPGPSSPYIPSPGGAMS  
PSYSPTSPAYEPRSPGGYTPQSPSYS  
PTSPSYSPTSPSYSPTSPNYSPTSPSY  
SPTSPSYSPTSPSYSPTSPSYSPTSPS  
YSPTSPSYSPTSPSYSPTSPSYSPTSP  
SYSPTSPSYSPTSPSYSPTSPSYSPTSP  
PSYSPTSPSYSPTSPSYSPTSPNYSPT  
SPNYTPTSPSYSPTSPSYSPTSPNYTP  
TSPNYSPTSPSYSPTSPSYSPTSPSYS  
PSSPRYTQSPTYTPSSPSYSPSSPSY  
SPTSPKYTPTSPSYSPSSPEYTPASPK  
YSPTSPKYSPTSPKYSPTSPYSPPTP  
KYSPTSPYSPYTPYTPSPKYSPTSP  
PTYSPYSPKYSPTSPYSPYSPKYSTY  
SPTSPGYSPYSPYSLTSPAISPDDSD  
EEN