

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
P16356	RPB1_CAEEL	Caenorhabditis elegans	DNA-directed RNA polymerase II subunit RPB1	10.590252	NaN	NaN	30250452	MALVGVDFQAPLRIVSRVQFGILGPE EIKRMSVAHVEFPEVYENGGPKLGG LMDPRQGVIDRRGRCMTCAGNLTD CPGHFGHLELAKPVFHIGFLTITLKI LRCVCFYCGRLLIDKSAPRVLEILKK TGTNSKKRLTMIYDLCKAKSVCEGA AEKEEGMPDDPDDPMNDGKKVAGG CGRYQPSYRRVGIDINAEWKKNVNE DTQERKIMLTAERVLEVFQQITDEDI LVIGMDPQFARPEWMICTVLPVPPL AVRPAVVTFGSAKNQDDLTHKLSDII KTNQQLQRNEANGAAHVLTDDVR LLQFHVATLVDNCIPGLPTATQKGG RPLKSIKQRLKGKEGRIRGNLMGKR VDFSARTVITADPNLPIDTVGVPRTIA QNLTFPEIVTPFNVDKLQELVNRGD TQYPGAKYIIRENGARVDLRYHPRAA DLHLQPGYRVERHMKDGDIIVFNRQ PTLHKMSMMGHRVKILPWSTFRMN LSVTSPYNADFDGDEMNLHLPQSLE TRAEIEEIAMVPRQLITPQANKPVMG IVQDTLCAVRMMTKRDVFIDWPFM MDLLMYLPTWDGKVPQPAILKPKPL WTGKQVFSLIIPGNVNLRTHSTHP DSEDSGPYKWISPGDTKVIIHGELL SGIVCSKTVGKSAGNLLHVVTLELGY EIAANFYSHIQTVINAWLIREGHTIGI GDTIADQATYLDIQNTIRKAKQDVVD VIEKAHNDLEPTPGNTRQTFENK VNQILNDARDRTGSSAQKSLSEFNN FKSMVVS GSKGSKINISQVIACVGQQ NVEGKRIPFGFRHRTLPHFIKDDYGP ESKGFVENSYLAGLTPSEFFFHAMG GREGLIDTAVKTAETGYIQRRLIKAM ESVMVNYDGTVRNSLAQMVLRYG EDGLDGMWVENQNMPTMKPNNAV FERDFRMDLTDNKFLRKNYSEDVV REIQESEDGSLVESEWSQLEEDRRL LRKIFPRGDAKIVLPCNLQRLIWNAQ KIFKVDLRKPVNLSPLHVISGVRELS KKLIIVSGNDEISKQAQYNATLLMNI LLRSTLCTKNMCTKSKLNSEAFDWL LGEIESRFQQAIAQPGEMV GALAAQ SLGEPATQMTLNTFHYAGVSAKNVT LGVPRLKEIINVSKTLKTPSLTVFLTG AAAKDPEKAKDVLCKLEHTTLKKVT CNTAIYYDPDPKNTVIAEDEEWVSIF YEMPDHDL SRTSPWLLRIELDRKRM VDKCLTMEIADRIHGGFGNDVHTI

