

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence	intracellular	extracellular	cytosol	nucleus	mitochondrion	endoplasmic reticulum	golgi apparatus	plasma membrane	extracellular region
P04052	RPB1_DROME	Drosophila melanogaster	DNA-directed RNA polymerase II subunit RPB1	25.496522	NaN	NaN	33925313	MSTPTDSKAPLRQVKRVQFGILSPD EIRRMVSVTEGGVQFAETMEGGRPKL GGLMDPRQGVDRSRTSRCQTCAGNM TECPGHFGHIDLAKPVFHIGFITKTIK ILRCVCFYCCKMLVSPHNPKIKEIVM KSRGQPRKRLAYVVDLCKGKTICEG GEDMDLTKENQQPDPNKKPGHGCG GHYQPSIRRTGLDLTAEWKHKQNEGS QEKKIVVSAERVWEILKHITDEECFI LGMDFKYARPDWMIIVLVPVPLAV RPAVVMFGAAKNQDDLTHKLSDIK ANNELRKNEASGAAAHVIQENIKML QFHVATLVDNDMPGMPRAMQKSG KPLKAIKARLKGKEGRIRGNLMGKR VDFSARTVITPDNLRIDQVGVRSIA QNLTFPELVTPFNIDRMQELVRRGN SQYPGAKYIVRDNGERIDLRFPKSS DLHLQCGYKVERHLRDDLVIFNRQ PTLHKMSMMGHRVKVLPWSTFRM NLSCTSPYNADFDGDEMNLHVPOS METRAEVENIHITPRQIITPQANKPV MGIVQDTLTAVRKMTKRDVFITREQ VMNLLMFLPTWDAKMPQPCILKPR PLWTGKQIFSLIIPGNVNMIRTHSTH PDEEDEGPYKWISPGDTKVMVEHGE LIMGILCKKSLGTSAGSLLHICFLEL GHDIAGRFGYNIQTVINNWLLEFEGH SIGIGDTIADPOTYNEIQQAIIKAKDD VINVIQKAHNMELEPTPGNTRLQTF ENKVNRIKNDARDKTGGSACKSLTE YNNLKAMVVSGSKGSNINISQVIAC VGQONVEGKRIPYGFRRKRTLPHFIK DDYGPESRQFVENSYLAGLTPSEFYF HAMGGREGLIDTAVKTAETGYIQR LIKAMESVMVNYDGTVRNSVGLIQ LRYGEDLGCGLVEFQNMPTVKLSN KSFQKRFKFDWSNERLMMKVFDD VIKEMTDSEAIQEAWEWDRLVSD RDSLRLQIFPNGESKVVLPCLNLRMI WNVQKIFHINKRLPTDLSPIRVIKGV KTLLERCVIVTGNDRISKQANENATL LFQCLIRSTLCTKYVSEEFRLSTEAF EVLVGEIETRFQQAQANPGEMVGA LAAQSLGEPATQMTLNTFFHAGVSS KNVTLGVPRLKEIINISKKPKAPSLTV FLTGGAARDAEKAKNVLCRLEHTTL RKVTANTAIYDPPQRTVISEDQEF VNVYEMPDQDPTRISPWLLRIELDR KRMTDKCLTMEQIAEKINVGFGEDL NCIFNDDNADKLVLRIRIMNNEENK FQDEDEAVDKMEDDMFLRCIEANM LSDMTLQGIEAIGKVMHLPTDSK KRIVITETGEFKAIGEWLLETDTGSM MKVLSEKRVDPVRTSSNDICEIFQVL GIEAVRKSVEKEMNAVLFYGLYVN YRHLALLCDVMTAKGHLMAITRHI NRQDTGALMRCSEETVDVLMDDAA AHAETDPMRGVSENIIMGQLPKMG TGCFDLLLLDAEKCRFGIEIPNTLGN MLGGAAMFIGGGSTPSMTPMTPW ANCNTPRYFSPGHVSAMTPGGPSF SPSAASDASGMSPSWSPAHGSSPS SPGSPMSPYFPASPSVSPSYSPSPN YTASSPGGASPNYSPSPNYSPSPSPL YASPRYASTTPNFNPQSTGYSPSSSG YSPTSPVYSPTVQFQSSPSFAGSGSN IYSPGNAYSPSSSNYSPNSPSYSPTSP SYSPSSPSYSPTSPCYSPSPSYSPS PNYTPVTPSYSPSPNYSPSPQYSPA SPAYSQTGVKYSPTSPYSPSPSYD GSPGSPQYTPGSPQYSPASPKYSPTS PLYSPSSPQHSNQSPTGTYSAT SPRYSPNMSIYSPSSTKYSPTSPYTP TARNYSPTSPMYSPTAPSHYSPTSPA YSPSSPTFEESD	None	None	None	None	None	None	None		