

UniprotKB ID	Entry name	organism	full name	oglnacscore	oglnacsites	phosphorylation sites	PMIDS	sequence	intracellular	extracellular	cytosol	nucleus	mitochondrion	endoplasmic reticulum	golgi apparatus	plasma membrane	extracellular region
Q5XG71	UTP20_MOUSE	Mus musculus	Small subunit processome component 20 homolog	25.019509	S2536	S788;S2640	34418053	MKPKPLSHKTENTYRFLTFAERLGN VNIDIHRIDRTASYDEDVETYPFEAL LKWRELNLTTEHFGKGYKEVIDKQCS FNQLVYHQNEIVQSLKTHLQIRNSL AYQPLLDLVVQLARDLQTDYFPHFE DFFLTITSILETQDTELLEWAFTSLSY LYKYLWRLMVKDMSKIYSLYSTLLA HKKLHIRNFAAESFTFLMRKVS DKN ALFNLMF LDLNEHPEKVEG V GQLLF EMCKGVRNMFH S CTGQALKLL LQK LGPVTETETQLPWILVGETLKTMAKS SVVYIYKEHFGVFFDCLQESLLELHN KVTEANCENSEQMRRLLETYLIVV KHGSGSKITRPADVCGVLEALQTAS LSTSCRKTLDDVVSALLAENVSLPE TLIKETVEKVFESKFERRSVLDFSEV MFAMKQFEQLFLPSFLYIENCFLM DNSVVSDEALAILAKLILHKAPPPTA GSMAIEKYPLVFSQQT V GSYLQQRK ADSKRRRKEQFPVLSHLLSIVQLPPNK DATYLSRSWAALVVLPHLRPLEKEK TISLVS CFIESLFLAVDRGSFGK GHL FVLCQAVNTLLSLEESSELLHLVPVG RVKHLVLTSPTEPSVLLADLYYQRL ALCGCKGPLSEEALMELFPKLQANI STGVSKIRLLTIRILNHFDIRLPVSME DDGLSERQSAFAILRQAE L PAVT VSD YREKLLHLRKL R HDV VQ GAVPQGR L QEVPLRYLLGMLYVNF SALWD PVIE LISSHAYGMENKQFWNV CYEHLEK AASHAEKELHKDVRDEESTGDESW EQTQEGDVG DLYQQQLALKTD C RER LDHTNFRFLWRALAKFPERVEPRS RELSPLFLRFINNEYPADLQVAPTQ DLRKKGRGAVAE EEEEE E PAAGEDE ELEEEAVPTEDAPQKKTRRAAAKQ LIAHLQVFSKFSNPRALYLESKLYEL YLQLLHQDQAVQKITLDCIMTYRH PHILPYRENLRLLDRSFKEIIVHF NISEDNTVVKAAHRADLFPILMRILY GRMKNKTGSKTOGKSASGTRMAIVL RFLACTQPEEIQFLDLLSEPVKHFK DGDCCSAVIQAVEDLDVSKVLPVGR QHGVNLSLEVVLKNISHLISTYLPKIL QILLCMTATVSHILDQREKIQLRFIN PLKNLRRLGIKMVTIDIFLDWESYQF KAEEIDAVFHGT VWPQICRLGSESQ YSPTPLK LISIWSRNARYFP LLAKQ KPGHPEYDILT N VFAVLSAKNLSEAT ASIIIMDIVDDLNLNLPDFQFTEAVPSL PVTGCVYADVAEDTEPVT VGGRLVL PHVPAIQYLSKTTISAEKVKKKKNR AQVSKELGILSKISKFMKDREQCSLL ITLLPFLLRGNVAQDTELDILVTVQ NLLQHCLHPAHFLRPLAKLFSVIKN KLSRQLLCTVFMQSD FESRLKYITDI VKLN AFDKRHLDDINFDVRFSAFQT ITSNIKAMQTVDADYLIAMHNCFY NMEIGDMSLSDNASICTLSIIKRLAA LNVTEKEYKEIIHRTLLEKLRKGLKS QTESVQHDYTLILSCLIQTFPNQLEF KDLVQLTHCHDPEMDF FENMKHIQ IHRRARALKKLAKQLLEGGVVLSSKS LQNYIMPYAMAPILDEKMLKHENITI AATEVIGAICRHLSWPAYVYVYLKHF HVLQSGQINQKLAVSLLVIVLEAFHF DYKTLEEQMGNVKNEENTVEMAEL LEPEAMEVEDMDEAGKEQASERLS DSKEALGAPEAAASEGTVAKEQECIS KSVSFLPRNKEELERTIQTIQGAITG DILPRLHKCLASATKREEEHKLVKSK VVNDEEVVRVPLAFAMVKLMRSLPR EVMEANLPSILLKVCVLLKNRAQEIR DIARSTLSKIHEDLG VHFLOVVLKELQ TTLVRGYQVHVLTFTVYVTLQLGLSSK LQVGDLD SCLHIMTEIFNH E LFGAL	None	None	None	None	None	None	None		

AEEKEVKQILSKVMEARRSKSYDSY  
EILGKFGKQVTKLILPLKEILQNTT  
SLKLARKVHETLRRRIAGLIVNPDMT  
ADALLLSYGLVSENLPLLTEKEKKP  
AAPVPDARLPPQSCLLLPATPVRRGGP  
KAVVNNKKTNMHIFIESGLRLLHLSL  
KTSRIKSSSEHVLEMLDPFVSVLINC  
LGAQDVKITGALQCLIWVLRFPPLPS  
IASKAEQLTKHLFLLKNYARVGAAR  
GQNFHLVVNCFKCVTIVVKKVKSHQ  
ITEKQLQVLLAYAEEDIYDTSRQATA  
FGLLKAILSRRKLLVPEIDDIMRKVSKL  
AISAQNEPARVQCRCQVFLKYILDYPL  
GEKLRPNLEFMLAQLNYEHETGRES  
TLEMIAYLFETFPQGLLHEHCGMFF  
IPLCLMMVNDDSAMCKRMAISMAIK  
SLLSKVDREKKDWLFGVTSWFEEK  
KRLNRQALACGLFVESEGVDFER  
RLGTLPPVIEKEIDPENFKDIEBETEE  
KAADRLLFGFLTMRKLIKECSIIHF  
TKPSETLSKIWSHVSHLRHPSWV  
WLTAAQIFGLLFASCQPEELIQKWK  
GKTKKKTSDPIAVRFLTSDLGQKM  
KSISLASCHQLHSKFLDESLGEQVV  
KNLLFIAKVLYLLEESGNKRGEVKD  
SEEQDTLADALAREAAEEKAGAGGK  
MESNREKKEEPSKPATLMWLIQKLS  
RMAKLEAAYSARNPLKRTCIFKFLGA  
VAVDLGVDRVKPYLPLIAPLRELN  
STFAEQDPVLKNLSQEIHELLKLVG  
LESFSLAFASVQKQASEKRALKRKR  
KALEFVTNPDIAAKKKLKKHKNKSE  
AKKRKIEFLRPGYKAKRQKSHSLRDL  
AMVE