	UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence	intracellular	extracellular	cytosol	nucleus	mitochndrion	endoplasmic reticulum	golgi apparatus	plasma membrane	extracellula region
0	A0A087WS16	A0A087WS16_MOUSE	Mus musculus	von Willebrand factor A domain-containing protein 1	27.83319	NaN	NaN	39927985;37507081	MRKHRHLPLVAVFSLLLSGIATTHA QQHGVMEVNKRDIVFLVDGSSSLGP SNFNAIRDFVTRVIQRLEIGQDLVQV SVAQYADTVKPEFYLNSYTNKRDAIT AVRKMRALNGSALYTGSSLDFVRNN LFTSSAGHRAAEGVPKLLVLITGGKS LDEVSQPAQELKRGSIMALAVGSKA ADEDELKEIAFDSSLVFIPAEFRPAPL QNMLPSLMAPLRTLTGTTEESKRDI LFLFDGSVNVLGQFPAVRDFLYRIIE ELDVKPDGTRVAIAQFSDDVRLESRF SEHQTKAEILNLVKKMKLKTGKALN LGYALDYALRNIFVRSAGSRIEDNVQ QFLVLLVAGRSSDAVAGPASSLKQR GVVPFIFQAKNANPSELEQIVLSPAFI LAAESLPKIGDLQSQIVSLLKAEQGS GPVSGEKDVVFLIDGSEGVRSGFPLL KDFVQRVVESLDVGPDRVRVALVQY SDRTRPEFYLNSHMDQQGVISAIRR LTLLGGPTPNTGAALEFVLRNILTSS TGSRIAEGVPQLLIVLTAEPSGDDVR GPSVVLKQGGAVPIGIGIGNADISEM QTISFIPDFAVAIPTFRELGTIQQVISE RVIQLNREELSSLKPILTPSTGAGVG SKKDVVFLIDGSRNAGPEFQYIRTLI ERIVEYLDIGFDTTRVAVIQFSEDSK MEFPLNAHFSKDEVQNAVRRLRPK GGSQVYIGNALEYVLKNIFQRPLGSR IEEGVPQFLVLISSGKSDDEVDDSAV ELKQFGVAPLTIARHTDQEELVKISL SPEYVYSVSTFRELPRLEQKLLTPITT LTSQQIHQILASTRYPPSVVESDAADI VFLIDSSDAVKPDGIAHIRDFVSRIVR RLNIGPSKVRIGVVQFSNDVFPEFYL KTHKSQSSVLEAIRRLRFKGGSPLNT GRALEFVARNLFVKSAGSRIEDGVP QHLVLFLGGKSQDDVARHAQVISSS GIVSLGIGDRNIDRTDLQTITNDPRL VFTVREFRELPNIEERVMLSFGPSGA TPQPPGVDLPSPSRPEKKKADIVFLL DGSINFRRDSFQEVLRFASEIVDTVY EDGDSIRVGLVQYNSDPTDEFFLRD FSTKRQIIDAINKVVYKGGRHANTRV GIEHLLRNHFVPEAGSRLDERVPQIA FVITGGKSVEDAQDVSLALTQKGVK VFAVGVRNIDSEEVGKIASNSATAFR VGSVQELSELSETVLETLHDAMHET LCPGVTDVSKACNLEVILGFDGSRD QNVFVSQKGLESKVDIILNRISQIQRI SCSGNQLPTVRVSVMANTPSGPVEA FDFAEYQPELFEKFRNMRSQRPYVL TADTLKLYQNKFRQSSPDTVKVVIHF TDGADGDMADLYRASEELRQAGAQ ALILVGLERVANLERLMHLEFGRGF MYDRPLRLNLLDLDYELAEQLDNIA EKACCGVPCKCSGERGDRGPIGSIGP KGISGEDGYRGYPGDEGGPGERGPP GVNGTQGFQGCPGQRGVKGSRGFP GEKGELGEIGLDGLDGEEGDKGLPG SSGEKGSPGRRGDKGPKGDKGERG DVGIRGDPGDSGRDSQQRGPKGETG DIGPMGLPGRDGIPGSPGDPGKDGG SGRRGPAGAKGNRGGPGQPGFEGE QGTRGSQGPPGPIGPPGLIGEQGIPG PRGGGGTAGAPGERGRTGPLGRKGE PGEPGPKGSIGNRGPRGETGDDGRD GVGSEGRRGKKGERGFPGYPGPKGT PGEPGADGPPGPKGIRGRRGNSGPP GATGQKGDPGYPGPSGHKGNRGDS VDQCALIQSIRDKCPCCYGPLECPVF PTELAFALDTSEGVTQDTFSRMREV LLGIVGDLTIAESNCPRGARVAVVTY NNEVTTEIRFADSKKKSALLDSIQNL QVALTSKQQSLETAMSFVARNTFKR VRSGFLMRKVAVFFSNKPTRASPQL REAVLKLSDAGITPLFLTSQEDRQLI NALQINNTAVGHALVLPARRDLTDF LKNVLTCHVCLDICNIDPSCGFGSW RPSFRDRRAAGSDVDIDLAFILDSSE ATTLFQFNEMKKYIGYVIRQLDLSPD	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	None

								PKASQHFARVAVVQQSTYESVDNAS VPPVKVEFSLTDYGAKEKLLDFLSRR MTQLQGTMGLGNAIEYTIENIFESAP NPRDLKIMVLMLTGDMQRQQLEEA QRAILQAKCKGYFFVVLGIGRKVNIK EVYSFASEPNDVFFKFVDKSTELNE EPLMRFGRLLPSFVSSENAFYLPPDL RKQCDWFQGDQPVKNGVKFGHKQI NPPHTANSSLTSKVVTTMKPVTTTK PTAIVNLPPAKPAPARPAPAQPVLAK PDPAKPAQARPAPAKPASAKLVPPQP VHVQPAPAQTASVRPAPAKPAPPQP AAAKPVPAKPAVPAQPAPPQPAAAKP VPAKPAVPAQPAPPQPAAAKPVPAKP AVPAQPAAAQPMPAQPVLTKSAAVK PASANKPVAAKPVATNTATATARPAL AAKPAAAKPAATRPLAAAIRPVATKP EAPRQQAKPAATKPATTKPLARVSRE VQVSEVTENSARLHWERPEPSSSFF YDLTVTSAHDQSLVLRQNLTVTDRVI GGLLAGQLYHVVVVSYLQSQVRAIY QGSFNTKKTQPPPLQAAHRASSSTI NLMVNTEPLFLTKTDICKLSRDAGT CVDFKLLWHYDLESKSCKRFWYGG CGGNENRFHSQEECEKMCSPDLLV								