

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
Q62059	CSPG2_MOUSE	Mus musculus	Versican core protein	24.474581	NaN	S2585;S2586	33300544	MLINMKGILWMCSTLLLTHALHQA KMETSPPVKGSLSGKVVLPCHFSTL PTLPPNYNTSEFLRIKWSKMEVDKN GKDIKETTFLVAQNGNIKIGQDYKGR VSVPTHPDDVGDASLTMVKLRASDA AVYRCDMVYGIEDTQDTMSLAVDGV VFHYRAATSRYTLNFAAAQACLDI GAVIASPEQLFAAYEDGFEQCDAGW LSDQTVRYPIRAPREGCYGDMMGKE GVRTYGFSPQETYDVYCYVDHLDG DVFHITAPSKFTFEEAEACTSRDAR LATVGELQAAWRNGFDQCDYGWLS DASVRHPVTVARAQCGGGLLGVRTL YRFENQTCFPLPDSRFDAYCFKPKQ NISEATTIEMNILAETSSPSLSKEPH MVPDRATPVIPLATELPIFTTHFPPA GNIVNSEQKSVVYSQAITGRLATESP TTTRNTINSWDLNDSLASGSGPLGM PDISEIKEEELRSTTVISQHATGSQAV ITEDTQTHESVSQIEQIEVGPLVTSM EITNHISLKLPEKNKTPYESTEVTL EHTTEMPTVSASPELATTSHYGFTL REDDREDRDLTVRSQSTRVFSQIPE VITVSKTSEDITYSQLGDLESISTSTI TMLGTDRLIDKEKEPKTNGKVTE EFGSQPTTTFPSOHLTEVELLPYSG DTTSVEGISTVIYPSLQTDVTQGRER TETPRPELKKDPYTVDEIPEKVTKDP FIGKTEEVFSGMPLSTSSSESSVERT ESVSPALTIEKLTGKPTTEARDVEEMT TLTRLETDTVTKSDKDVTRVHLTHSTL NVEVTVSKWPGDEDNSTSKPLPST EHAGFTKLPVPLSTIGINGKDKKEIPS FTDGGGEYTLFPDGTPKPLEKVSEE DLASGELTVTFHTSTSIGSAEKSASG EPTTGDRFLPTTSTEDQVINATAEGS ALGEDTEASKPLFTGPPFVHTSDVEE LAFVNYSSQOEPTYVDISHTSPLSII PKTEWSVLETSVPLEDEILGKSDQDI LEQTHLEATMSPGALRITGVSQGET QEEPQTPGSPFPPTFSSTAVMAKETTA FEEGEGSTYTPSEGRLMTGSERVPG LETTVPVGTSYPPGAITDQEVEMDTM VTLMSTIRPTVVSSTESEVIYEAEGSS PTEFASTLRPFQTHVTQLMEETTEE GKKASLDYTDLGSGLFEPRALEPKF PSTPSDISVFTAIDSLHRTPLSPSS FTTEQRVFEESSEKTTGDILPGESV TQHPVTTLIDIVAMKTESDIDHMTSK PPVTQTPRPSVVERKTTSKTQELSTS TPAAGTKFHPDINVYIIEVRENTGR LSDMIVSGHPIDSESKEEPCSEETD PLHDLFAEILPELPDSFEIDIHSEED EDGEEDCVNATDVTTTPSVQYINGK QLVTTVPKDEAAEARRGQYESVAP SQNFPDSSATDTHQFILAETESST MQFKKSKEGTELEITWKPETYPET PDHVSSGEPDVFPTLSSHDKTTRW SESITESPNNLENPVHKQPKPVPLFP EESSGEGAIEQASQETILSRATEVAL GKETDQSPTLSTSSILSSSVSVNVLE EEPLTLTGISQTDISMSTIESWVEIT PSQTVKFSSESSAPIIEGSGEVEENK NKIFNMVTDLPQRDPTDLSPLDMS KIMITNHHIYPATIAPLDSKLPSPDA RPTTVWNSNSTSEWVSDKSFEGRK KKENEDEEGAVNAAHQGEVRAATE RSDHLLLTPELESSNVDASSDLATW EGFILETTPTESEKEMANSTPVFRET IGVANVEAQPFHSSSSHPRVQEEL TTLSGNPPSLFTDLGSGDASTGMELI TASLFTLDLESETKVKKELPSTPSPS VEISSSFEPTGLTPSTVLDIEIAGVMS QTSQKTLISEISGKPTSQSGVRDLYT GFPMGEDFSGDFSEYPTVSYPTMKE	None	None	None	None	None	None	None		

ETVGMGGSDDERVRDTQTSSSIPTT
SDNIYPVPDSKGPSTVASTTAFPW
EVMSSAEGSGEQLASVRSVGPVLP
LAVDIFSGTESPYFDEEFEEVAVTE
ANERPTVLPATAASGNTVDLTENGYE
VNSTMSLDFPQTMPEPSKLSKPEV
NLDKQEIGRETVTKEKAQGQKTFES
LHSSFAPEQILETQSLIETEFQTS
SMLTTLKTYITNKEVEEEMSIAM
STPGPGIKDLESYTTTPEAPGKSHS
SATALVTEGAARSVLMDSSTQEEE
SIKLFQKGVKLTNKNESNADLSFSG
GSGGALPLPPTTSVNLDMKQIISTL
YAETSHMESLGTSLGDKMEDHER
MEDVSSNEVRMLISKISISQDSTEA
LDTTSLHTGTTEPPTSTLFPVKLMDL
ERSPKQDPSGGKRPKTHRPQMS
GLISNENSASAEEGATSPTAFLPQ
TYSVEMTKHFAPSESQPSDLFNVNS
GEGSGEVDLTLVYTSGTQASSQG
DSMLASHGFLEKHPEVSKTEAGATD
VSPTASAMFLHHSEYKSLYPTSTLP
STEPYKSPSEGIEDGLQDNIQFEGST
LKPSRRKTTESIIDLKEDSKDLGLT
ITESAIVKSLPELTSKNIIDIDHTKP
VYEYIPGIQTDLDPEIKLESHGSSEES
LQVQEKYEGAVTSLPTEESFEGSGD
ALLAGYTQAIYNESVTPNDGKQAEDI
SFSFATGIPVSTETELHTFFPTASTL
HIPSKLTTASPEIDKPNIEAISLDDIFE
SSTLSDGQAIADQSEVISTLGHLEKT
QEEYEEKYGGPSFQPEFFSGVGEV
LTDPPAYVIGSTYLIAQTLTELPNVV
RPSDSTHYTEATPEVSSLAEKSPQIP
SSPPVYVDNGVSKFPEVPHSAQP
VSTVTSSQKSIESPFKEVHANIEETIK
PLGGNVHRTEPPSMSRDPALDVS
ESKHKLELETSPTKETSQDFPN
KAKDHIPGETVGMLAGIRTTSEPV
TADDMELGGATQPPHSASAAFRVET
GMVPQPIQEPERPTFPSLEINHET
HTSLFGESILATSEKQVSQKILDNSN
QATVSTLDLHTAHALSPFSILDNSN
ETAFLIGISEESVEGTAVYLPGLDCK
TNPCLNGGTCYPTETSIVCTCAPGYS
GDQCELDLDFECHSNPCRNATCVD
GFNTFRCLCLPSYVVALCEQDTETC
DYGWHKFGQCYKFAHRRRTWDA
ERECLQGAHLTSILSHEEQMFVNR
VGHDYQWIGLNDKMFHDFRWD
GSALQYENWRPNQPSFFSAGEDC
VVIWHENGQWINDVPCNYHLTYC
KKGTVACGQPPVENAKTFGKMKP
RYEINSLIRYHCKDGFIRHLPTIRCL
NGRWAMPKITCMNPSAYQRTYSK
KYLKNSSAKDNSINTSKHEHRWSR
RQETRR