

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence	intracellular	extracellular	cytosol	nucleus	mitochondrion	endoplasmic reticulum	golgi apparatus	p n
P12111	CO6A3_HUMAN	Homo sapiens	Collagen alpha-3(VI) chain	24.532706	T118;S1944;S2253;T2560	T433;S1225	30620550;28657654;38253038;38665916	MRKRRHLPLVAVFCLFLSGFPTTHA QQQADVKNGAADIIFLVDSSWTI GEEHFQLVREFLYDVVKS LAVGEND PFFALVQFNGNPHTEFLLNTYRTKQ EVLSHISNMSYIGGNTQTGKGLLEYI MQSHLTKAAGSRAGDGVVQVIVVLT DGHSKDGLALPSAELKSADVNVFAI GVEDADEGALKKEIASEPLNMHMFN LENFTSLHDIVGNLVCVHSSVSPE RAGDTETLKDITAQDSADIIFLDGNS NTGSVNFVILDFLVNLEKLPITGQ QIRVGVVQFSDEPRTMFLSDTYSTK AQVLGAVKALGFAGGELANIGLALD FVVENHFTTRAGGSRVEEGVQVLLVL ISAGPSSDEIRYGVVALKQASVFSFG LGAQAASRAELQHIATDDNLVFTVP EFRSFGDLQEKLLPYIVGVAQRHIVL KPPTIIVTQVIEVNRDIVFLVDGSSAL GLANFNAIRDFAIKVIQRLEIGQDLIQ VAVAQYADTVRPEFYFNTHPTKREVI TAVRKMKPLDGSALYTGSAIDFVRN NLFTSSAGYRAEGIPKLLVLITGGK SLDEISQPAQELKRSSIMAFIGNKG ADQAELEIAFDSSLVFPAEFRAAPL QGMLPGLLAPLRTLSGTPEVHSNKR DIIFLLDGSANVGKTNFPYVRDFVM NLVNSLDIGNDNIRVGLVQFSPTPV TEFSLNTYQTKSDILGHLRQLQLQG GSGLNTGSALSYYANHFTEAGGSR IREHVPQLLLLLTAGQSEDSYLQAAN ALTRAGILTFCVGASQANKAELEQIA FNPSLVYLMDDFSSLPALPQQLIQPL TTYVSGGVEEVPLAOPESKRDLFLF DGSANLVGQFPVVRDFLYKIIDELNV KPEGTRIAVAQYSDDVKVESRFEDEH QSKPEILNLVCRMKIKTGKALNLGYA LDYAQRIFVKSAGSRIEDGVLQFLV LLVAGRSSDRVDGPAASNKQSGVVP FIFQAKNADPAELEQIVLSPAFILAAE SLPKIGDLHPQIVNLLKSVHNGAPAP VSGEKDVVFLLDGSEGVRS GFPLLK EFVQRVVESLDVGQDRVRVAVVQYS DRTRPEFYLNYSYMNKQDVVNAVRO LTLLGGPTPNTGAALEFVLRNLVSS AGSRITEGVPQLLIVLTADRSGDDVR NPSVVVKRGGAVPIGIGGNADITEM QTISFIPDFAVAIPTRFQLGTVQQVIS ERVTLTRELSRLQPVLPSPGV GGKRDVFLIDGSQSAGPEFOYVRT LIERLDVLDVGFDTTRVAVIQFSDD PKVEFLLNAHSSKDEVQNAVQRLRP KGGQRQINVGNALEYVSRNIFKRPLG SRIEEGVPQFLVLISGKSDDEVDDP AVELKQFGVAPFTIARNADQEELVKI SLSPEYVFSVSTFRELPSLEQKLLTPI TTLTSEIQKLLASTRYPPPAVESDA ADIVFLIDSSGVRPDGFAHIRDFVS RIVRRLNIGPSKVRVGVVQFSNDVFP EFYLTKYRSQAPVLDAIRLRLRGGG PLNTGKALEFVARNLFVKSAGSRIED GVPQHLVLVLGKGSQDDVSRFAQVI RSSGIVSLGVGDRNIDRTELQITIND PRLVFTVREFRELPIEERIMNSFGP SAATPAPPVDTPPPSRPEKKAIDIV FLLDGSINFRRDSFQEVLRVSEIVD TVYEDGDSIQVGLVQYNSDPTDEFFL KDFSTKROIIDAINKVYKGGRHANT KVGLEHLRVNHVFPEAGSRLDQRPV QIAFVITGGKSVEDAQDVSLALTORG VKVFAVGVRNIDSEEVGKIASNSATA FRVGNVQELSELSEQVLETLHDAM HETLCPGVTDAAKACNLDVILGFDG SRDQNVFVAQKGFESKVDAILNRIS QMHRVSCSGGRSPTVRVSVVANTPS GPVEAFDFEYQPEMLEKFRNMRS QHPYVLTEDTLKVYLNKFRQSSPDS VKVVIHFTDGDGDLADLHRASEN RQEGVRLALVGLERVVNLRLMHL EFGRGFMYDRPLRLNLLDLDYELAE QLDNIAEKACCGVPCCKSGQRGRG	False	True	1.84	2.269	2.203	4.394	1.313	4

PIGSIGPKGIPGEDGYRGPYDDEGGP
GERGPPGVNGTQGFQGPCGQGRVK
GSRGFPGEKGEVGEIGLDGLDGEDG
DKGLPGSSGEKGNPGRRGDKGPRG
EKGERGDVGIRGDPGNPGQDSQER
GPKGETGDLGPMGVPRDGVPGGP
GETGKNGGFRGRRGPPAKGNKGGP
GQPFEGEQGTRGAQGPAGPAGPPG
LIGEQISGPRGSGGAAGAPGERGRT
GPLGRKGEPGEPGPKGGIGNRGP
ETGDDGRDGVGSEGRGKKGGERGF
PGYPGPKGNPGEPLNGTTGPKGIR
GRRGNSGPPGIVGQKGDGYPGPAG
PKGNRGDSIDQCALIQSIKDKCPCY
GPLECPVFTELAFAFDLTSEGVNQD
TFGRMRDVLIVNDLTIAESNCP
GARVAVVYNNNEVTTEIRFADSKRKS
VLLDKIKNLQVALTSKQSQLETAMS
FVARNTFKRVRNGFLMRKVAVVFFS
NTPTRASQLREAVLKLSDAGITPLF
LTRQEDRQLINALQINNTAVGHALV
LPAGRDLTDFLENVLTCHVCLDICNI
DPSCGFGSWRPSFRDRRAAGSDVDI
DMAFILDSAETTLFQFNEMKKYIAY
LVRQLDMSDPKASQHFARVAVVQ
HAPSESVDNASMPVVFSLTDYG
SKEKLVDFLSRGMQLQGTRALGSA
IEYTIENVFESAPNPRDLKIVVMLT
GEVPEQQLEEAQRVILQAKCKGYFF
VVLGIGRKNIKEVYTFASEPNDVFF
KLVKSTELNNEPLMRFGRLLPSFV
SSENAFYLSPIRQCDWFQGDQPT
KNLVKFGHKQVNPNNVTSSPTS
PVTITKPVITTKPVITTKPVITTKP
VTIINQPSVKPAAAKPAPAKPVAAP
VATKMATVRPPVAVKPATAAKPVA
KPAAVRPPAAAAAKPVATKPEVPRP
AAKPAATKPAATKPMVKMSREVQVF
EITENSAKLHWERAEPGPYFYDLTV
TSAHDQSLVLKQNLTVTDRVIGLL
AGQTYHVAVVCYLRQVRAHYHGSF
STKKSQPPPPQPARSASSSTINLMVS
TEPLALTETDICKLPKDEGTCRDFIL
KWYYDPNTKSCARFWYGGCGGNEN
KFGSQKECEKVCAPVLAKPGVISVM
GT