

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
P11087-2	CO1A1_MOUSE	Mus musculus	Isoform 2 of Collagen alpha-1(I) chain	9.218204	NaN	S161;S260;S776	36288343;34887587	MFSFVDLRLLLLLGATALLTHGQEDI PEVSCIHNGLRVPNGETWKPEVCLI CICHNGTAVCDDVQCNEELDCPNP QRREGECCAFCPPEEYVSPNSEDVGV EGPKGDPGPQGPRGPVGPGRDGIP GQPGLPGPPGPPGPPGPPGLGGNFA SQMSYGYDEKSAGVSVPGMPGSPGP RGLPGPPGAPGPQGFQGPPEPGEPE GGSGPMGPRGPPGPPGKNGDDGEA GKPRPGERGPPGPQGARGLPGTAG LPGMKGHRGFSGLDGAKG DAGPAG PKGEPGSPGENGAPGQMGRGLPG ERGRPGPPGTAGARGNDGAVGAAGP PGPTGPTGPPGFP GAVGAKGEAGPQ GARGSEGPQGV RGEPPGPAGAAG PAGNPGADGQPGA KANGAPGIAGA PGFPGARGPSGPQGPSGPPGPKGNS GEPGAPGNKGD TGAKGEPGATGVQ GPPGPAGEEGKRGARGEPGPSGLPG PPGERGGPGRGFP GADGVAGPKGP SGERGAPGAPGKGS PGEAGRPGEA GLPGAKGLTGS PGPDPGKTGPPG PAGQDGRPGPAGPPGARGQAGVMG FPGPKGTAGEPGKAGERGLPGPPGA VGPAGKDGEAGA QGAPGPAGPAGER GEQGPAGSPGFQGLPGPAGPPGEAG KPGEQGVPGDLGAPGSPGARGERGF PGERGVQPPGPAGPRGNNGAPGN DGA KGDTGAPGAPGSQ GAPGLQGM PGERGAAGLPGPKGDRGDAGPKGA DGSPGKD GARGLTGPIGPPGPAGAP GDKGEAGPSGPPGPTGARGAPGDRG EAGPPGPAGFAGPPGAPGAPGAPGP VGPAGKNGDRGETGPAGPAGPIGPA GARGPAGPQGRGDKGETGEQGDR GIKGHRGFSGLQGGPSGSPGEEQG PSGASGPAGPRGPPGSAGSPGKDGL NGLPGPIGPPGPRGRTGDSGPAGPP GPPGPPGPPGPPSGGYDFSFLPQPPQ EKSQDGGRYRADDANVVRDRDLE VDTTLKSLSQIENIRSP EGSRK NPA RTCRLK MCHSDWKS GEYWIDPNQ GCNLDAIKVYCNMETGQTCVFPTQP SVPQKNWYISPNPKEKKHVWFGES MTDGFPF EYGSESDPADVAIQ LTF LRLMSTEASQNITYHCKNSVAYMDQ QTGNLKKALLLQGSNEIELRGEGNS RFTYSTLV DGCTSH TGTWGKT VIEYK TTKTSRLPIIDVAPLDIGAPDQEFGLD

