

| UniprotKB ID | Entry name  | organism     | full name                   | oglcnacscore | oglcnac sites | phosphorylation sites | PMIDS    | sequence  |
|--------------|-------------|--------------|-----------------------------|--------------|---------------|-----------------------|----------|---|
| Q80X19       | COEA1_MOUSE | Mus musculus | Collagen alpha-1(XIV) chain | 16.794757    | NaN           | NaN                   | 37453647 | MMIWQCKMRDWLILAFLLAAACFCTI<br>VRGQVAPPTRLRYNVISHDSIQISWK<br>APRGKFGGYKLLVAPASGGKTNQM<br>NLQNTATKAIIQGLLPEQNYTVQLIA<br>YYKDKEKPAQGGQFRIKDLEKRKDP<br>TKPKVKVVDKGNNGSKPTSPEEVKFF<br>CETPAIADIVILVDGWSWSIGRFNFRLV<br>RNFLLENLVTAFNVGSEKTRIGLAQY<br>SGDPRIEWHLNAFNKDEVIDAVRS<br>LPYKGGNTLTGLALNFIFENSFKPEA<br>GSRSGVSKIGILITDGKSQDDIIPPSR<br>NLRESGVELFAIGVKNADLSELQEIA<br>SEPDSTHVYNAEFDLMHTVVESLT<br>RTVCSRVEEQDKEIKASALATIGPPT<br>ELITSEVTARFVMVNWTSQSPGKVEK<br>YRVVYYPTRGGKPEEVVVDGVSSTV<br>LKNLMSSTEYQIAVFAVSAHTASEGL<br>RGAETTLALPMASDLELYDVTENSM<br>RVRWDAVPGATGYLILYAPLTEGLAG<br>DEKEMKIGETHTDIELSGLFPNTEYT<br>VTVYAMFGEEASDPATGQETTLPLT<br>PPRNLRISNVGSNSARLTWDPASGK<br>ISGYRIVYTSADGTEINEVEVDPITTF<br>PLKGLTPLTEYSIAIFSIYEEGQSLPLV<br>GEFTTEEVPAQQYLEIDEVKTDSFRV<br>TWHPLSAEEGQHKLWIPVYGGKT<br>QEVDLKEEQDSYVIEGLDPGTEYEVS<br>LLAVLDDGSESEVVTAVGTTLDDFW<br>TEAPTAIEPTSPVTSVLQTGIRNLVVD<br>DETATSLRVSWDISDSNVEQFRVY<br>LKAQGDPMEEVVGTVMPGVQNSL<br>LLKALLPDTEYKVTVTPVYTVGEGVS<br>VSAPGKTLPSGGPQNLRVSEEWYNR<br>VRITWDPPSGPVKGYRIVYKPVSVPG<br>QTLETFVGADINTIVMTNLLSGMDY<br>NVKIFASQASGFSALTGLVQTLFLG<br>VTDLQANQVEMTSLCARWQIHRHA<br>TAYRIVLESLODTQAQESTVGGGVN<br>RHCIFYGLQPDSEYKISVYTKLQLEG<br>PSVSIMQKTQSLPTEPPTFPPTIPPAK<br>EVCKAAKADLVFMVDGWSWSIGDDN<br>FNKIINFLYSTVGALDKIGADGTQVA<br>MVQFTDDPRTEFKLDSYKTKETLLD<br>AIRHISYKGGNTKTGKAIKHVRDTLF<br>TSDSGTRRGIPKIVVITDGRSQDDV<br>NKISREMQADGFNIFAIGVADADYS<br>ELVQIGSKPSSRHVFFVDDDFDAFKKI<br>EDELITFVCETASATCPMVHKDGV<br>LAGFKMMEMFGLVEKDFSAVEGVS |

|  |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|
|  |  |  |  |  |  |  |  | <p>MEPGTFNLFPCYQIHKDALVSQPTK<br/> YLHPEGLPSDYTMSFLFRILPDTPQE<br/> PFALWEILNKNSEPLVGIILDNGGKT<br/> LTYFNLDYTGDFQTVTFEGPDIRKM<br/> FYGSFHKLHVVVSKTLAKVVVDCKE<br/> VGQKAINASANITSDGVEVLGRMVR<br/> SRGPNNGNSAPFQLQMFDIVCSTSWA<br/> SKDRCCLPGLRDEESCPDLPRSCS<br/> CSETNEVALGPAGPPGGPGLRGPKG<br/> QQGEQGPKGPEGPRGETGPAGPQG<br/> PPGPQGSPGSLIQGMPGMPGDKGD<br/> KGDAGLPGPQGVPGVGVSPGRDQSP<br/> GQRFPGKDGSSGPPGPPGIPGPA<br/> PGVPGITGSMGPQALGPPGVPGAK<br/> GERGERGDLQSQAMVRAVARQVCE<br/> QLIQSHMARYTAILNQIPSQSSSIRTI<br/> QGPPGEPGRPGSPGTPGEGQPPGTP<br/> GFPGNAGVPGTPGERGLTGVKGEKG<br/> NPGIGTQGPRGPPGPAGPSGESRPG<br/> SPGPPGSPGPRGPPGHLGVPGPQGP<br/> SGQPGYCDPSSCSAYGVGVSHPDQP<br/> EFTPVQDEQEAMDLSAGI</p> |
|--|--|--|--|--|--|--|--|--|