

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
Q86YZ3	HORN_HUMAN	Homo sapiens	Hornerin	15.391251	NaN	S659;S661;S890;S993;S1008;S1463;S1478;S1712;S1714;S1829;S1831;S1933;S1948;S2299;S2301;S2403;S2418;S2652;S2654	34229054;23271734;26374642;26853435;28510447;34019948	MPKLLQGVITVIDVFYQYATQHGEYD TLNKAELKELLENEFHQILKNPNDP DTVDIILQSLDRDHNKKVDFTEYLL MIFKLVQARNKIIGKDYCQVSGSKLR DDTHQHQQEEQEETEKEENKROESS FSHSSWSAGENDSYSRNVRGSLKP GTESISRRLSFQRDFSGQHNSYSGQ SSSYGEQNSDSHQSSGRGQCGSGS GQSPNYGQHGSGSGQSSSNDTHGS GSGQSSGFQHKSSSGQSSGYSQH GSGSGHSSGYGQHGSRSQSSRGE RHRSSSGSSSSYGQHGSGSRQSLGH GRQGSGSRQSPSHVRHGSGSGHSS SHGQHGSGSSYSYRGRHYESGSGQT SGFGQHESGSGQSSGYSKHGSGSG HSSSQQHGSTSGQASSGQHGSS SRQSSSYGQHEASARHSSGRGQHSS GSGQSPGHGQRGSGSGQSPSSGQH GTGFRSSSSGYPVSGSGYSSGFGH HESSEHSSGYTQHGSGSGHSSGH GQHGSRSGQSSRGERQSSAGSSSS YGQHGSGSRQSLGHSRHGSGSGQS PSPSRGRHESGSRQSSSYGPHGYGS GRSSSRGPYESGSGHSSGLGHQESR SGQSSGYGQHGSSSGHSSSTHGQHG STSGQSSSCGQHGATSGQSSSHGQ HGSGSSQSSRYGQQGSGSGQSPSRG RHGSDFGHSSSYGQHGSGSGWSSS NGPHGSVSGQSSGFGHKSGSGQSS GYSQHGSGSSHSSGYRKHGSRSGQ SSRSEQHGSSSGLSSSYGQHGSGSH QSSGHGRQGSQSGHSPSRVRHGSS SGHSSSHGQHGSGTSCSSSCGHYES GSGQASGFGQHEGSGGQGYSQHGS ASGHFSSQGRHGSTSGQSSSSGQH DSSSQSSSYGQHEASHHASGRG RHGSGSGQSPGHGQRGSGSGQSPS YGRHGSGSGRSSSSGRHGSGSGQSS GFGHKSSSGQSSGYTQHGSGSGHS SSYEQHGSRSGQSSRSEQHGSSSGS SSSYGQHGSGSRQSLGHGQHGSGS GQSPSPSRGRHGSGSGQSSSYGPYR SGSGWSSSRGPYESGSGHSSGLGH RESRSGQSSGYGQHGSSSGHSSSTH GQHGSTSGQSSSCGQHGASSGQSS SHGQHGSGSSQSSGYGRQGSQSGQ SPGHGQRGSGSRQSPSYGRHGSGS GRSSSSQHGSGGLGESSGFGHHES SSGQSSSYQHGSGSGHSSGYGQH GSRSGQSSRGERHGSSSGSSSHYGQ HGSGSRQSSGHGRQGSQSGHSPSR GRHGSGLGHSSSHGQHGSGSGRSS SRGPYESRSGHSSVFGQHEGSGGHS SAYSQHGSGSGHFCSQGQHGSTSG

QSSTFDQEGSSTGQSSSYGHRGSGS  
SQSSGYGRHGAGSGQSPSRGRHGS  
GSGHSSSYGQHGSQSGWSSSSGRH  
GSGSGQSSGFGHHSSSWQSSGCT  
QHGSQSGHSSSYEQHGSRSGQSSR  
GERHSSSSGSSSYGQHGSQRSL  
GHGQHGSQSGQSPSPSRGRHGSGS  
GQSSSYSPYGSQSGWSSSRGPYESG  
SSHSSGLGHRESRSGQSSGYGQHG  
SSSGHSSTHGQHGSTSGQSSCGQ  
HGASSGQSSSHGQHGSGSSQSSGY  
GRQSGSGQSPGHGQRGSGSRQSP  
SYGRHGSGSRSSSSGQHGSGLGES  
SGFGHHSSSQSSSYQHGSQSG  
HSSGYGQHGRSGQSSRGERHGSS  
SRSSRYGQHGSQRQSSGHGRQ  
SGSGQSPSRGRHGSLGHSSSHGQ  
HGSGSRSSSRGPYESRSGHSSVFG  
QHESGSGHSSAYSQHGSQSGHFCS  
QGQHGSTSGQSSSTFDQEGSSTGQSS  
SHGQHGSQSSQSSSYGQQGSGSGQ  
SPSRGRHGSGSGHSSSYGQHGSQSG  
GWSSSSGRHGSGSGQSSGFGHHES  
SSWQSSGYTQHGSQSGHSSSYEQH  
GSRSGQSSRGEQHGSSSGSSSYGQ  
HGSGSRQSLGHGQHGSQSGQSPSP  
SRGRHGSGSGQSSSYGQPYGSQSGW  
SSSRGPYESGSGHSSGLGHRESRSG  
QSSGYGQHGSSSGHSSTHGQHGS  
SGQSSSCGQHGASSGQSSSHGQHG  
SGSSQSSGYGRQSGSGQSPGHGQ  
RGSGSRQSPSYGRHGSGSRSSSSG  
QHGPGLGESSGFGHHSSSQSSSY  
SQHGSGSGHSSSYGQHGRSGQSS  
RGERHSSSSGSSRYGQHGSQRQS  
SGHGRQSGSGHSPSRGRHGSGSG  
HSSSHGQHGSGSGRSSSRGPYESRS  
GHSSVFGQHESGSGHSSAYSQHGS  
GSGHFCSQGQHGSTSGQSSSTFDQE  
GSSTGQSSSHGQHGSGSSQSSSYGQ  
QGSGSGQSPSRGRHGSGSGHSSSY  
GQHGSQSGWSSSSGRHGSGSGQSS  
GFGHHSSSWQSSGYTQHGSQSGH  
SSSYEQHGSRSGQSSRGERHGSSSG  
SSSYGQHGSQRQSLGHGQHGSQ  
SGQSPSPSRGRHGSGSGQSSSYSPY  
GSGSGWSSSRGPYESGSGHSSGLG  
HRESRSGQSSGYGQHGSSSGHSSST  
HGQHGSTSGQSSSCGQHGASSGQS  
SSHGQHGSGSSQSSGYGRQSGSG  
QSPGHGQRGSGSRQSPSYGRHGSG  
SGRSSSSGQHGSGLGESSGFGHHE  
SSSGQSSSYQHGSQSGHSSSYGQH  
GSRSGQSSRGERHGSSSGSSSHYGQ  
HGSGSRQSSGHGRQSGSGQSPSR  
GRHGSLGHSSSHGQHGSGSGRSS  
SRGPYESRLGHSSVFGQHESGSGHS

SAYSQHSGSGHFC	SQGQHG	STSG
QSSTFDQEGS	STGQSS	SYGHRGSGS
SQSSGYGRHG	GAGSGQ	LSHGRHGS
GSGQSSSYGQ	HGSGSG	QSSGYSQH
GSGSGQDGYSY	CKGGSN	HDGGSSG
SYFLSFP	SSTSPY	EYVQEQR
		CYFYQ