

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
A2AKX3	SETX_MOUSE	Mus musculus	Probable helicase senataxin	17.394869	S1369;T2455	S102;S640;S870;S871;S872;S938;S1002;S1004;S1318;S1472;T1474;T2450	36852467;34418053	MSTCCWCTPGGSSTIDVLKRYASSTGSSEFQTADEDLCYCLECVAEYHRA RDEVPPFLHEVLWELETLRVSHFEK SMKAEAEEDDDLYIVDNNNGEEQLF DCSGQDFENKLRVPLFEILKYPYLLL HERVNELCVEALCRMEQNNCSFQV FDKYPGIYFLVHPNEMVRRWAILT ARNLGKVDRDDYDLQEVLTCLFKV IELGLLESPDIYTSSVLEKGLILLPA HMYDTTNYKNYWLGCMLLTILEEQ AMDSL L LGS DK Q NDFMQSILHTME KQSDDDSMDFPWPALHCFMVILDR LGSKVWGQLIDPIEAFQTIINNESYN REIQNIRNSSIRT KLEPEPHFDDMVT CSQIVYNFNPEKTKKDSGWRSAICP DYCPNMYEEMETLANVLQSDIGQD MRVHNSTFLWFIPFVQSLMDLKD L GVAYIVEVIHHL YSEVKDVLNQTDAV CDKVTEFFILILISVIELHRNKKCLHL LWVSSQQWVEAVVKCAKLPTTAFVR SCEKSPGSTSRGAAIMSSLALHSVQS NSVQLACVQLIRGLLKEGYQLGQQT LCKRFWDKLNLF L RGNLSL GWQLT GOETHELQ MCLKQIRNIKFKMPQY STFGDSTSTFKTPPSFKEESDKIDRK HKKNIYCLENCSPVSSKEPMKADTH RVL MKVNTTEENFKQHYIDLNEEE QEPLPAELCLKQKSEALFSESAEQ VKISAEKSGKESSSYAPSNSTSRNGP EWGCDRGVIMSAHSLTDSSSDFME QVSTSNEDVSLKDG SVGKTSKPSFK LQKDEICAKLSHVIKKQIRKSTLVDNI IDLENTAISDLENCSGTDGGALKED SIGHNVPSDPVLDKHEEQKSONSS LFKKEIKSEELDNSSDDEDKLOIQE GRADDDLVSFTEVTDTLVKAPCEGH VKMVVESRDKEMRESTALTSNLVE GQVPHDSSKPLVAGRQIDLCNITLIS QTTVIQFPSGLSKQNSFQLQKGDKR CLTANQNSAATCRGQVIVISDSDEEE DEDEDERS S SEENIKQSKACIGKDC SEHRSLAVNASVEKQLVKEEERYPV EFEDSESQVFEFESSSEVFSVWQDH KIDSKNSLQGEQKSYVTHVADSTNN NLGCGDSVSEEVVRNKAEGVKEHA GPHSSVSAEEFCKTGVKKPKRKRYD KVTAEDPQRPSSSVGTDQLPDRRDL TESDLKSADMGMATPSSSVERDSTI LQKSTKSRTHSKPVRKVPASKATKKT HSDTRRGQSKSSCYISCRTPAIVPP

KKLRQCPEPTSTVEKLGKAPRKA
ELSQRSLECIQLRDHGKTVGVVDA
PKKAKLISPQTLSEIKNNKLLTSQDL
QFQRLMRSRSHKRDYKNTDTV
RVSRIVQGSVDLEADSDHRVS
EPLAISNEKQLAKCMLSKEVAEASS
DPWVTGITCLVNQCESRVLSSGGVPT
DVVMVSASEDPVDGGAVTVQVGEV
ASVAAAEPASSSDTDDDDNLFLTQH
DPQDMDLCSQLENKTIIVAHKKDTV
QREDSLSRPQLESLSITCKYKDCVE
TTKNQGEYCPRHSEAKAADGLFRK
PGLPLSVARPLRPTTTKIFSSSSASRT
ANLSKSLESTTLQQSALKNKSSGAQ
PNLKVTPSSMGSQKPVAEVKSLCN
IFHFQTPSSSSKQSKLTFSENRPST
AASPVNILLPSQSIFDTFIKEVLKWKY
QMFLNFDKCGAPTSQCQISRPVPV
RFQDCAEYFNVFLPLIILNAFETVAQ
EWLSSPNKENFYQLQLRKFADYKK
YWEFLIYLNESLAKQLHPKENDLV
FLAPEKSYMDRHGMQDCSHYYCGY
VHKFRRTSVMRSGKAECSLCIQTQD
TLPASVKNLTRCIVISSLVTTQRKLKA
MSLLSSRNQLARAVLNPNPMDFCT
KDLLTTTTERIVAYLKDFNEDQKAI
ETAYAMVKHSPSVAKICLIHGPPGTG
KSKTIVGLLYRLLTENQRKGHSDEN
FNAKIKQNRVLCAPSNAAVDELMMK
KIILEFKEKCKDKKNPLGNCGDINLV
RLGPEKSINTEVLKFSLDSQVNHMM
KKDLPSHIQEMLRKEILDAQLDELS
RQRALCRGGREMQRQELDEHIAIVS
KERQELASKIKEVQGRPQRAQNTIIL
ESHVICCTLSTSGGLLESFAFRGQGG
VPFSCVIVDEAGQSCEVETLSPLIHR
CNKLILVGDPKQLPPTVISMKAQEYGG
YDQSMMARFCKLLEENVEQNMIGR
LPVLQTLTIQYRMHPDICLFPSNYVYN
KNLKTNRLTESIRCSSEWPFQPYLVF
DVGDGSERRDNDSYINVQEIKLVME
IIKLIKEKRKDISFRNIGIITHYKAQKT
MIQKDLEKEFDKKGPAEVDTVDAFQ
GRQKDCIIVTCVRASAVQGSIGFLAS
LQRLNVTITRAKYSLFILGHLRTLME
NQHWYELIQDAQKRGAIKTSDPNY
RHDAMKILKLPVLRSLTHPPATA
PEAPRPQGGPLSNRLDSGLATTSFA
ASLYHTPSDVTSTKGPRLQDRLR
DPRLLRRLDAEAKGTFLKDPQPVSP
QLPGVVHLLGEPGFPVVFQDLGFVV
PPSTAIVAPLGSHRSPMQAEPHPAHP
AAAASSTSKRKYSDDAGLSHKREPR

