

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
P55200	KMT2A_MOUSE	Mus musculus	Histone-lysine N-methyltransferase 2A	15.818779	S3045	S134;S140;S151;S195;S516;S678;T837;S923;S1053;S1839;T1847;S1860;S2100;T2148;S2152;S2202;S2560;S2607;S2792;S2951;S3032;T3369;S3510;S3523	36852467	MAHSCRWRFPARPGTTGGGGGGGR RGLGGAPRQVRPALLLPPGPQAGGG GPGAPPSPPAVAAAAAGSSGAGVPG GAAAASAASSSSSSSSSSSSSSSSG PALLRVGPGFDAALQVSAAGTNLRR FRAVFGESGGGGGSGEDEQFLGFGS DEEVRVRSPTSPSVKASPRKPRGRP RSGSDRNPAILSDPSVFSPLNKSETK SADKIKKKDSKSIEKKRGRPPTFPGV KIKITHGKDIAELTQGSKEDSLKKVK RTPSAMFQQATKIKKL RAGKLSPLKS KFKTGKLGQIGRKG VQIVRRRGRPPST ERIKTPSGLLINSELEKPKVVRKDKE GTPPLTKEDKT VVRQSPRRIKPVRIIP SCKRTDATIAKQLLQRAKKG AQKKIE KEAAQLQGRKVKTQVKNIRQFIMPV VSAISSRIIKTPRRFIEDEDYDPPMKI ARLESTPNSRFSATSCGSSEKSSAAS QHSSQMSSDSSRSSSPSIDTTSDSQ ASEEIQALPEERSNTPEVHTPLPISQ SPENESNDRRSRRYSMSERSFGSRA TKKLPTLQSAPQQQTSSSPPPPLLTP PPPLQPASGISDHTPWLMPTIPLAS PFLPASAAPMQGKRKSILREPTFRW TSLKHSRSEPQYFSSAKYAKEGLIRK PIFDNFRPPPLTPEDVGFASGFSASG TAASARLFSPLHSGTRFDIHKRSPIL RAPRFTPSEAHSRIFESVTLPSNRTS SGASSGVSNRKRKRKRVFSPIRSEPR SPSHSMRTRSGRLSTSELSPLTPPSS VSSLSIPVSPLAASALNPTFTFP SHS LTQSGESTEKNQRARKQTSALAEPF SSNSPALFPWFTPGSQTEKGRKKDT APEELSKDRDADKSVEKDKSREDR EREKENKRESRKEKRKKGSDIQSSS ALYPVGRVSKEKVAGEDVGTSSSAK KATGRKKSSSLDSGADVAPVTLGDT TAVKAKILIKKGRGNLEKNNLDLGA APSLEKERTPCLSAPSSSTVKHSTSSI GSMLAQADKLPMTDKRVASLLKKAK AQLCKIEKSKSLKQTDQPKAQGQES DSSETSVRGPRIKHVCRRRAVALGR KRAVFPDDMPTLSALPWEEREKILS SMGNDDKSSVAGSEDAEPLAPPIKPI KPVTRNKAPQEPPVKKGRRSRRCGQ CPGCQVPEDCGICTNCLDKPKFGGR NIKKQCKMRKCQNLOWMPSKASL QKQTKAVKKEKSKTTEKESKES TAVKSPLPAQKAAPPREEPAPKKS

SSEPPPRKPVEEKSEEGGAPAPAPAP
EPKQVSAPASRKSSKQVSQPAAVVPP
QPPSTAPQKKEAPKAVPSEPKKKQPP
PPEPGPEQSKQKKVAPRPSIPVKQKP
KDKEKPPPVSQENAGTLNILNPLS
NGISSKQKIPADGVHRIRVDFKEDCE
AENVWEMGGLGILTSVPITPRVVCFL
CASSGHVEFVYCQVCCPEPFHKFCLE
ENERPLEDQLENWCCRCKFCHVC
GRQHQATKQLLECNCRNSYHPEC
LGPNYPTKPTKKKKVWICKCVRCK
SCGSTTPGKGWDAQWSDFLCHD
CAKLFAGNFCLCDKCYDDDDYES
KMMQCGKCDRWVHSCESLSGTE
DEMYEILSNLPESVAYTCVNCTERH
PAEWRLALEKELQASLKQVLTALLN
SRTTSHLLRYRQAAKPPDLNPETES
IPSRSSPEGPDPPVLTEVSKQDEQQP
LDLEGVKKRMDQGSYVSVLEFSDDI
VKIIQAAINS DGGQPEIKKANS MVKS
FFIRQMERYFPWFVSVKSRFWEPN
KVSNNSGMLPNAVLPPSLDHNYAQ
WQEREESHTEQPPLMKKIIPAPKP
KGPGEPSPTPLHPPTPPILSTDRSR
EDSPELNPPPIDDNRQCALCLMYG
DDSANDAGRLLYIGQNEWTHVNCA
LWSAEVFEDDDGSLKNVHMAVIRG
KQLRCEFCQKPGATVGCCLTSCTSN
YHFMCSRANKCVFLDDKKVYCQRH
RDLIKGEVVPENGFVFRRVFVDFE
GISLRRKFLNGLEPENIHMMIGSMT
IDCLGILNDLSDCEDKLFPIGYQCSR
VYWSTTDARKRCVYTCKIMECRPPV
VEPDINSTVEHDDNRTIAHSPSSFID
ASCKDSQSTAAILSPPSPDRPHSQTS
GSCYYHVISKVPRIRTPSYSPTQRSPG
CRPLPSAGSPTPTTHEIVTVGDPLLS
SGLRSIGSRRHSTSLSPLRSKLRIM
SPVRTGSAYSRSVSSVPSLGTATDP
EASAKASDRGGLSSSANLGHSAAPP
SSSSQRTVGGSKTSHLDGSSPSEVK
RCSASDLVPKGSLVKGEKNRTSSSK
STDGSAHSTAYPGIPKLTQVHNATP
GELNISKIGSFAEPSTVPFSSKDTVSY
PQLHLRGQRSDRDQHMDPSQSVKP
SPNEDGEIKTLKLPGMGHRPSILHE
HIGSSSRDRRQKGKKSSKETCKEKH
SSKSYLEPGQVTTGEEGNLKPFAFAD
EVLTPGFLGQRPCNNVSSEKIGDKV
LPLSGVPKGQSTQVEGSSKELQAPR
KCSVKVTPMKMEGENQSKNTQKES
GPGSPAHIESVCPAEPVSASRSPGAG

PGVQSPNNTLSQDPOSNNYQNL
EQDRNLMIPDGPKQEDGSFKRRYP
RRSARARSNMFFGLTPLYGVRSYGE
EDIPFYSNSTGKKRGRSAEGQVDG
ADDLSTSEDDLYYNFTRTVISSGG
EERLASHNLFREEEQCDLPKISQLD
GVDDGTESDTSVTATSrkSSQIPKRN
GKENGTEENLKIDRPEDAGEKEHVIK
SAVGHKNEPKLDNCHSVSRVKAQG
QDSLEAQLSSLESSRRVHTSTPSDK
NLLDTYNAELLKSDSDNNNSDDCG
NILPSDIMDFVLKNTPSMQALGESP
ESSSELLTLGEGGLDSNREKDIGL
FEVFSQQLPATEPVDSSVSSSISAE
QFELPLELPSDLSVLTRRSPVPSQN
PSRLAVISDSGEKRVTITEKSVASSE
GDPALLSPGVDPAPEGHMTDPHFIQ
GHMDADHISSPPCGSVEQGHGNSQ
DLTRNSGTPGLQVPVSPTVPVQNK
YVPSSTDSPGPSQISNAAVQTPPHL
KPATEKLIVVNQNMQPLYVLQTLPN
GVTQKIQLTSPVSSTPSVMTNTSVL
GPMGSLTTLTGLNPSLPPSPSLFPP
ASKGLLSVPHHQHLHSFPAAAQSSF
PPNISSPPSGLLIGVQPPDPQLLGE
ANQRTDLTTVATPSSGLKKRPI SRL
HTRKNKKLAPSSAPSNIAPSDVVS
MTLINFTPSLSNHPSLDLGSLNP
SSHRTVPNIKRKSGIMYFEQAPLL
PPQSVGGTAATAAGSSTISQDTSHT
SGPVSALASGSSVLNVVSMQTTAAP
TSSTSVPGHVTLANQRLGTPDIGSI
SHLLIKASHQSLGIQDQPVALPPSSG
MFPQLGTSQTPSAAAMTAASSICVL
PSSQTAGMTAASPPGEAEEHYKLQR
GNQLLAGKTGTLTSQRDRDPDSAPG
TQPSNFTQTAEAPNGVRLEQNKTLP
SAKPASSASPGSSPSGQQSGSSVP
GPTKPKPKAKRIQLPLDKGSGKHK
VSHLRTSSEAHIPHRDTPAPQPSVT
RTPRANREQQDAAGVEQPSQKECG
QPAGPVAALPEVQATQNPANEQENA
EPKAMEEEESGFSSPLMLWLQEQ
KRKESITERKPKKGLVFEISSDDGFQ
ICAESIEDAWKSLTDKVQEARSNARL
KQLSFAGVNGLRMLGILHDAVVFLI
EQLAGAKHCRNYKFRFHKPEEANE
PPLNPHGSARAEVHLRQSAFDMFN
FLASKHRQPPEYNPNDEEEVQLK
SARRATSMDLPMRFRHLKKTSK
EAVGVYRSPIHGRGLFCKRNIDAGE
MVIEYAGNVIRSIQTDKREKYYDSKG

