

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
Q03164	KMT2A_HUMAN	Homo sapiens	Histone-lysine N-methyltransferase 2A	13.754781	T259;S540;T616;S764;S775;S777;S778;S781;S783;S925;S926;S1325;S1858;S2340;S3050;T3051;T3210;S3238;S3241;S3267;S3273;T3353	S136;S142;S153;S197;S518;S680;T840;S926;S1056;S1837;T1845;S1858;S2098;T2147;S2151;S2201;T2525;S2611;S2796;S2955;S3036;T3372;S3511;S3515;S3527	34019948;34725712;32119511;37217939;35254053;21740066;20305658;33214551;35138101;35289036;29351928;23301498;35132862;32574038;30379171	MAHSCRWRFPARPGTTGGGGGGGR RGLGGAPRQRPALLPPGPPVGGG GPGAPPSPPAVAAAAAAGSSGAGVP GGAAAASAASSSSASSSSSSSSASS GPALLRVGPGFDAALQVSAIQTNLR RFRAVFGESGGGGGSEDEQFLGF GSDEEVVRVSPTRSPSVKTSRKRPRG RPRSGSDRNSAILSDPVFSPLNKSE TKSGDKIKKKDSKSIEKKRGRPTFP GVKIKITHGKDISELPGNKEDSLKK IKRTPSATFQQATKIKKLRAGKLSPL KSKFKTGKQLIGRKGVQIVRRRGRPP STERIKTPSGLLINSELEKPOKVRKD KEGTPPLTKEDKTVVROSPRIKPV IIPSSKRTDATIAKQLLQRAKKAQK KIEKEAAQLQGRKVKTVQKNIRQFI MPVVSAISSRIKTPRRFIEDYDPP IKIARLESTPNSRFSAPSCGSSEKSSA ASQHSSQMSSDSSRSSSPVDTSTD SQASEEIQVLPERSDTPVHPLPI SQSPENESNDRRSRRYSVRSRFGS RTTKLSTLQSAQQQTSSSPPPPLL TPPPPLQPASSISDHTPWLMPTIPL ASPFLPASTAPMQGKRKRSILREPTFR WTSCLKHSRSEPQYFSSAKYAKEGLIR KPIFDNFRPPPLTPEDVGFASGFSAS GTAASARLFSPLHSGTRFDMHKRSP LLRAPRFTPSEAHSRIFESVTLPSNR TSAGTSSSGVSNRKRKRKRVFSPIRSE PRSPSHSMRTRSGRLSSSELSPLTPP SSVSSLSISVSPLATSALNPTFTFPS HSLTQSGESAENQRPRKQTSAPAE PFSSSSPTPLFPWFTPGSQTERGRN KDKAPEELSKDRDADKSVEKDKSRE RDREREKENKRESRKEKRRKGEIQ SSSALYPVGRVSKKVVGEDVATSSS AKKATGRKKSSSHDSGTDITSVTLGD TTAVTKILIKKGRGNLEKTNLDLGP TAPSLEKEKTLCLSTPSSSTVKHSTS SIGSMLAQADKLPMTDKRVASLLKK AKAQLCKIEKSKSLKQTDQPKAQGQ ESDSSSETSVRGPRIKHVCRRAAVALG RKRAVFPDDMPTLSALPWEEEREKIL SSMGNDKSSIAGSEDAEPLAPPIKP IKPVTRNKAPQEPVKKGRRSRRCG QCPGCQVPEDCGVCTNCLDKPKFG GRNIKKQCKMRKCQNLQWMPK AYLQKQAKAVKKEKSKTSEKKS KESSVKNVVDSSQKPTPSAREDP PKKSSSEPPRKPVEEKSEEGNVSAP GPESKQATTPASRKSSKQVSQPALVI PPQPPTTGPPRKEVPKTTTSEPKKKQ PPPPESGPEQSKQKQVAPRPSIPVKQ KPKEKEKPPPVNKQENAGTLNILSTL SNGNSSKQKIPADGVHRIRVDFKED CEAENVWEMGGGLGILTSVPITPRVV CFLCASSGHVEFVYCCVCEPFHFKF CLEENERPLEQLENWCCRCKFC HVCGRQHQAATKQLLECNKCRNSYH PECLGPNYPTKPTKKKVVICTKCV RCKSCGSTTPGKGWDAQWSDHDFSL CHDCAKLFAGKGFCLCDKCYDDD DYESKMMQCGKCDRWVHSCENL SDEMYEILSNLPESVAYTCVNCTER HPAEWRLALEKELQISLKQVLTALL

NSRTTSHLLRYRQAAKPPDLNPETE
ESIPSRSSPEGDPDPVLETSKQDDQ
QPLDLEGVKKRMDQGNYSVLEFSD
DIVKIIQAAINSDDGGQPEIKKANSMV
KSFIRQMERVFPWFVSVKKSFRWEP
NKVSSNSGMLPNAVLPPSLDHNYA
QWQEREENSHTEQPPLMKKIIPAK
PKGPGEPDSTPLHPPTPILSTDRS
REDSPELNPPPGIEDNRQCALCLTY
GDDSANDAGRLLYIGQNEWTHVNC
ALWSAEVFEDDDGSLKNVHMAVIR
GKQLRCEFCQKPGATVGCCLTSCTS
NYHFMCSRAKNCVFLDDKVVYQCR
HRDLIKGEVVPENGFEVFRRVFVDF
EGISLRRKFLNGLEPENIHMIGSM
TIDCLGILNDLSDCEDKLFPIGYQCS
RVYWSTTDARKRCVYTKIVECRPPV
VEPDINSTVEHDENRTIAHSPTSFTE
SSSKESQNTAEIISPPSPDRPPHSOT
SGSCYYHVISKVPRIRTPSYSPQRSP
GCRPLPSAGSPTTTHEIVTVGDPDLL
SSGLRSIGSRRHSTSSLSPOKSLRI
MSPMRTGNTYSRNNVSSVSTGTAT
DLESSAKVVDHVLGPLNSSTSLGQN
TSTSSNLQRTVVTVGNKNSHLDGSS
SSEMKQSSASDLVSKSSSLKGEKTK
VLSKSSSEGAHNVAYPGIPKLAQV
HNTTSRELVNSKIGFAEPSSVSFSS
KEALSFPHLHLRGQRNDRDQHTDS
TQSANSSPDEDTEVKTLKLSGMSNR
SSIINEHMGSSSRDRRQKGGKSCKE
TFKEKHSSKSFLEPGQVTTGEEGNL
KPEFMDEVLTPEYMGQRPCNNVSS
DKIGDKGLSMPGVPKAPPMQVEGSA
KELQAPRKRTVKVTLTPLKMNESQ
SKNALKESSPASPLQIESTSPTEPISA
SENPGDGPVAQSPNNTSCQDSQS
NNYQNLVQDRNLMLPDGPKQED
GSFKRRYPRRSARARSNMFFGLTPL
YGVRSYGEEIDIPFYSSSTGKKRGRS
AEGQVDGADDLSTSDDEDLYYNFT
RTVISSGGEERLASHNLFREEEQCD
LPKISQLDGVDDGTESDTSVTATTRK
SSQIPKRNGKENG TENLKIDRPEDA
GEKEHVTKSSVGHKNEPKMDNCHS
VSRVKTQGGDSLEAQLSSLESSRRV
HTSTPSDKNLLDTYNTELLKSDSDN
NNSDDCGNILPSDIMDFVLKNTPSM
QALGESPESSSELLNLGEGGLDS
NREKDMGLFEVFSQQLPTTEPVDSS
VSSSISAEQFELPLELPSDLVLTTR
SPTVPSQNPSRLAVISDSGEKRVITIT
EKSVASSESDPALLSPGVDPTPEGH
MTPDHFIQGHMDADHISSPPCGSVE
QGHGNNQDLTRNSSTPGLQVPVSP
TVPIQNQKYVPNSTDSFGPSQISNAA
VQTPPHLKPATEKLIVVNQNMQPL
YVLQTLPNGVTQKIQLTSSVSTPSV
METNTSVLGPMGGGLTTLTGLNPSL
PTSQSLFPSASKGLLPMSSHQHLHS
FPAATQSSFPPNISNPPSGLLIGVQP
PPDPQLLVSESSQRTDLSTTVATPSS
GLKKRPISRLQTRKNKKLAPSTPSN
IAPSDVSNM TLINFTPSQLPNHPSL
LDLGLSNTSSHRTVPNIIKRSSSIM
YFEPAPLLPQSVGGTAATAAGTSTIS
QDTSHLTSGSVSGLASSSVLNVVS
MQTTTTPTSSASVPGHVTLTNPRL

