

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
Q8BRH4	KMT2C_MOUSE	Mus musculus	Histone-lysine N-methyltransferase 2C	19.12418	T1833;S1836;S3746	S28;S46;S89;S113;S847;S1294;S1983;S2822;T2824;S4027;S4260	29187734;36852467	MSSEEDRSAEQQPPPPAPPEEPGAP APSPAAADKRPRGRPRKDGASPFQR ARKKPRSRGKSTVEDEDSMDGLETT ETENIVETEIKEQSVVEEDAETVDS KQPVSAQRVSVSEESANSLVSVGVE AKISEQLCAFCYCGEKSSLGQGD LQFRVTPGLTLPWKDQPSNKDIDDNS SGTCEKIQNYAPRKQGRQKERPPQ QSAVSCVSVSTQTACEDQAGKLWDE LSLVGLPDAIDVQALFDSTGTCWAH HRCVEVSLGICQMEEP LLVNVDKA VVSGSTERCAFCCKHLGATIKCCEEK CTQMYHYPCAAGAGTFQDFSHFFLL CPEHIDQAPERSKEDANCAVCDSPG DLLDQFFCTTCGQHYHGMCLDI AVT PLKRAGWQCPECKVCQNCQSGED SKMLVCDTCDKGYHTFCLQPMKS VPTNGWKCKNCRICIEGTRSS TQW HHNCLICDTCYQQQDNLC PFCGKC YHPELQKDM LHCNMCKR WVHLEC DKPTDQEL DSQLKEDYIC MYCKHLG AEIDPL HPGNEVEMPEL PTDYASGM EIEGTE DEVVLEQTVN KDVSDHQC R PGIVPDVQVY TEEPQKSNPL ESPDT VGLITSE SDNKMNPDL ANEIAHEV DTEKTE MLSKGRHVCE EDQNE DRM EVTENIE VLPHQTIVPQ EDLLSEDS EVASKELSP PKSAPETAAP EALLSPH SERLSCKE PLLTERVQEE MEQKEN SEFSTGC VDFEMTLAV DSCDKDSS C QGDYVEL PAEEESTF SSATDLNKA DVSSSTL CSDLPSCD MLHGYPPAF NSAAGS IMPTTYSV TPKIGMGK PAIT KRKFS PGRPRSKQ GAWSNHNT VSP PSWAPD TSEGREIF KPRQLSGS AIWS IKVGRG SGFPGKRR PRGAGLSG RGG RGRSKL KSGIGAVV LPGVSAAD ISSN KDEEEN SMHNTVV LFSSDKFT LQ QDMCV VCGSFGQ GAEGRLL ACSQC GQCYHPY CVSIKIKV VLSKGW RCLC CTVCEAC GKATDPG RLLLCDD CDIS YHTYCL DPPLQTV PKGGWCK WCV WCRH CGATSAG LRCEWQ NNYTQCA PCASLSS CPVCCRN YREEDL LILQCRQ CDRWMH AVCQNLN TEEEVEN VADI GFDCSM CRPYMP VSNVPSS DCCDS SLVAQI VTKVKE LDPKTYT QDGVCL TESGMS QLQSLT VTAPRR KRTKPKL KLKIIN QNSVAV LQTPDI QSEHSR D GEMDDS REGELM DCDGKSE SSPER EAGDDE TKGIEG TDAIKR KRKPYR P GIGGF MVRQRS RTGQGK AKRSV VRK DSSGS ISEQLP SRDDG WREQL PDTL VDEPVS VAENTD KIKKR YRKRK NKL EETFP AYLQEA FFGKDL LDTSR QNK

LSVDNLSEDAAQLSFKTGFLDPSSD
PLLSSSSTSAPGTQGTADDPLADIS
EVLNTDDDDILGIISDDLAKSVDHSDI
GPTTADASSLPQPGVSQSSRPLTEEQ
LDGILSPELDKMVTDGAILGKLYKIP
ELGGKDVEDLFTAVLSPATTQPAPLP
QPPPPQLLPMHNQDVFSRMLPMN
GLIGSPHLPHNSLPPGSGLGTFPPI
AQSPYTDVRDKSPAFNAIASDPNSS
WAPTTPSMEGENDTLSNAQRSTLK
WEKEEALGEMATVAPVLYTNINFPN
LKEEFPDWTRVKQIAKLWRKASSQ
ERAPYVQKARDNRAALRINKVQMSN
DSMKRQQQDSIDPSSRIDSDFKD
PLKQRESEHEQEWKFRQMRQKSK
QAKIEATQKLEQVKNEQQQQQQ
QQQQQQQLASQHLLVAPGSDTPS
SGAQSPFTPQAGNGNVSPAQTFHKD
LFSKHLPGTPASTPSDGVFVKPQPPP
PPSTPSRIPVQESLSQSQNSQPPSPQ
MFSPGSSHSRPPSPVDPYAKMVGTP
RPPPGGHSFPRRNSVTPVENCVPLS
SVPRPIHMNETSATRPSARDLCASS
MTNSDPYAKPPDTPRPMMDQFSK
PFSLPRSPVISEQSTKGLTTGTSDH
FTKPSPRTDAFQRQRLDPYAGPSLT
PAPLGNGPFKTPHPPPSQDPYGSVS
QTSRRLSVDPYERPALTPRPVDFNS
HSQSNDPYSHPLTPHPAMTESFTH
ASRAFPQPGTISRASQDPYSQPPGT
PRPLIDSYSQTSGTARSNDPYSQPP
GTPRPNTIDPYSQQPPTPRPSPQDM
FVSSVANQRHTDPYTHHLGPPRPGI
SVPYSQPPAVPRPRTSEGFTRPSSAR
PALMPNQDPFLQAAQNRVPLPGPL
IRPPDTCSTPRPPGPRIDTFTHAS
SSAVRDPYDQPPVTPRPHSEFGTS
QVVHDLVDRPVPVPGSEGNFSTSSNLP
VSSQGGQFSSVSQLPGPVPTSGGTD
TQNTVNMSQADTEKLRQRQKLEII
LQQQQQKKIASRQEKGPQDTAVVPH
PVPLPHWQPESINQAFTRPPPPYPGS
TRSPVIPPLGPRYAVFPKDQRGYPYP
EVAGMGMRPHGFRFGFPAGHGP
MPSQDRFHVPPQIQGSGIPPHRRP
MSMEMPRPSNNPPLNPNVGLPQHFP
PPQGLPVQQHNILGQAFIELRHRAP
DGRSRLPFAASPSVIESPSHPRHGN
FLPRPDFGPRHTDPIRQPSQCLSN
QLPVHPNLEQVPPSQEQGHPAHQ
SSIVMRPLNHPLSGEFSEAPLSTSTP
AETSPDNLEIAGQSSAGLEEKLDSD
DPSVKELDVKDLEGVEVKDLDDDEL
ENLNLDTEDGKGDDLDLTDNLETN
DPNLDDLLRSGEFDIIAYTDPDLG
DKKSMFNEELDLNVPIDDKLDNQC
ASVEPKTRDQGDKTMVLEDKDLPQR
KSSVSSEIKTEALSPYSKEEIQSEIKN

HDDSRGDADTACSQAASAQTNHSD
RGKTALLTTDQDMLEKRCNQENAG
PVVSAIQGSTPLPARDMNSCDITGS
TPVLSLLSNEKCDDSDIRPSGSSPP
SLPISPSTHGSSLPPTLIVPPSPLLDN
TVNSNVTVVPRINHAFSQGVVNPNG
FIQQSSVNHNLGTGKPTNQTVPPLT
NQSSTMSGPQQLMIPQTLAQQNRE
RPLLEEQPLLLQDLLDQERQEQQQ
QRQMAMIRQRSEPFPPNIDFDAIT
DPIMKAKMVALKGINKVMAQNSLG
MPPMVMSRPFPMGPSVAGTQNNND
GQTLVPQAVAQDGSITHQISRPNPP
NFGPGFVNDQRKQYEEWLQETQQ
LLQMQQKYLEEQIGAHRKSKKALSA
KORTAKKAGREFPEEDAEQLKHVTE
QQSMVQKQLEQIRKQKQKEHAELIED
YRIKQQQQQQCALAPPILMPGVQP
QPPLVPGATSLTMSQPNFPMVPQQL
QHQQHTAVISGHTSPARMPSLPGW
QSNASASHLPLNPPRIQPPIAQLSLK
TCTPAPGTVSSANPQNGPPPRVEFD
DNNPFSESFQERERKERLREQQER
QRVQLMQEVDQRALQQRMEMEQ
HCLMGAELANRTPVSQMPFYGSDR
PCDFLQPPRPLQOSPQHQQQIGPVL
QQQNVQQGSVNSPPNQTFMQTNE
QRQVGPPSFVPDSPSASGGSPNFHS
VKPGHGNLPGSSFQQSPLRPFPTIL
PGTSPVANSNVPCGQDPAVTQGGQNY
SGSSQSLIQLYSDIPEEKGKKRTRK
KKKDDDAESGKAPSTPHSDCAAPLT
PGLSETTSTPAVSSPELPOQRQEP
VEPVPVPTPNVSAGQPCIESENKLPN
SEFIKETSNNQQTHVNAEADKPSVET
PNKTEEIKLEKAETQPSQEDTKVEEK
TGNKIKDIVAGPVSSIQCPSHPVGT
TTKGDGTGNELLKHLLKNKASSLLT
QKPEGTLSSDESSTKDGLIEKQSPA
EGLQTLGAQMGGGFGGNSQLPKT
DGASENKKQRSKRTQRTGEKAAPRS
KKRKKDEEEKQAMYSSDSFTHLKQ
QNNLSNPPTPPASLPPTPPPMACQK
MANGFATTEELAGKAGVLSHEVAR
ALGPKPQLPFRPQDDLARAIQGP
KTVDVPASLPTPPHNNHEELRIQDH
YGDRDTPDSFVPSSSPEVVGVEVN
KYPDLSLVKEEPPVPSPIIPILPSIS
GKNSESRRNDIKTEPGTLFFTSFPGS
SPNGPRSGLISVAITLHPTAAENISSV
VAAFSDLLHVRIPNSYEVSNAPDVPP
MGLVSSHVRNPSLEYRQHLLLRGPP
PGSANPPRLATSYRLKQPNVFPPTS
NGLSGYKDSHGPAEGASLRPQWC
CHCKVVILGSGVRKCKDLTFVNKG
SRENTKRMEKDIVFCSNNCFILYSS
AAQAKNSDNKESLPSLPQSPMKEPS
KAFHQYSNNISTLDVHCLPQFQEKV

SPPASPPISFPPAFAEAAKVESKPDELK VTVKLKPRLRTPVVGLEDRCRPLNKK WRGMKWKKWSIHIVIPKGTfkPPCE DEIDEFLKKGTLKPDVPKDCRKC CFCHEEGDGLTDGPARLLNLDL WVHLNCALWSTEVYETQAGALINVE LALRRGLQMKCVFCHKTGATSGCH RFRCTNIYHFTCATKAQCMFFKDKT MLCPMHKPKGIHEQQLSYFAVFRRV YVQRDEVRQIASIVQRGERDHTFRV GSLIFHTIGQLLPQMQAFHSPKALF PVGYEASRLYWSTRYANRRCRYLCSI EEKDGRPVFVIRIVEQGHEDLVLSDS SPKDVWDKILEPVACVRKKSEMLQL FPAYLKGEDLFGLTSAVARIAESLP GVEACENYTFRYGRNPLMELPLAVN PTGCARSEPKMSAHVKRFVLRPHTL NSTSTSKSFQSTVTGELNAPYSKQFV HSKSSQYRRMKTEWKSNNVYLARSRI QGLGLYAARDIEKHTMVIEYIGTIIRN EVANRKEKLYESQNRGVYMFMRMDN DHVIDATLTGGPARYINHSCAPNCV AEVVTFFERGHKIISSNRRIQKGEELC YDYKFDDEDQHKIPCHCGAVNCRK WMN
--