

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
Q61329	ZFHX3_MOUSE	Mus musculus	Zinc finger homeobox protein 3	15.818779	T552	S426;T428;S535;S573;S1207;S1600;S2634;S2795;S2804;S2900;S2904;S3434;S3443;S3457;S3616;S3700	36852467	MEGCDSPVVSGKDNGCGIPQHRQW TELNSAHLDPKPSMEQPTGESHGP LDSLRAFNFNERLADSSTSAGPPAEP SKEVSCNECSASFSSLQTYMEHHCP GTHPPPALREESASDTSEEGEEESD VENLAGEIVYQPDGSAYIVESLSQLA QSGAACGSSSGSGAVPSLFLNSLPG VGGKQGDPSCAAPVYPQIINTSHIAS SFGKWFEGSDPAFPNTSALAGLSPV LHSFRVFDVRHKSNDYLNDSGSAK SSCVSKDVPNNVDLSKFDGFVLYGK RKPILMCFLCKLSFGYVRSFVTHAV HDHRMTLSEEERKLLSNKNISAIQ IGKDKPEPLVSFLEPKNKNFQHPLVS TGNLIGPGHSFYGKFSGIRMEGEEA LPAAAAAGPEQPQAGLLTPSTLLNLG GLTSSVLKTPITSVPLGPLASSPTKSS EGKDSGAAEGDKQESGGHQDCFSE KVEPAEEEEEEEEEEEEEEEEEEEE EEEEEEEEEEEEASKGLFPNDLEEL EDSPSEESGPPAGGTTKKDLALSNS SISNSPLMPNVLQTLRGPASTTSNS ASNFVVFDGANRRSRLSFNSEGVRA NVAEGRRLDFADESANKDSATAPEP NESTEGDDGGFVPHHQHAGSLCEL GVGESPSGSGVECPKCDTVLGSSRS LGGHMTMMHSRNSCKTLKCPKCN WHYKYQQTLEAHMKEKHPEPGGSC VYCKSGQPHPRLARGESYTCGYKPF RCEVCNYSTTTKGNLSIHMQSDKHL NNMQNLQNGGGEQVFSHSAGAAA AAAAAAAAAANIGSSWGAPSPTKPKT KPTWRCEVCDYETNVARNLRIHMTS EKHMHNMMLLQQNMTQIQHNRH LGLGSLPSPAEAELYQYYLAQNMNL PNLKMDSTASDAQFMMSGFQLDPT GPMMAAMPALVGGEIPLDMRLGGG QLVSEELMNLGESFIQTNDPSLKL QCAVCNKFTTDNLDMLGLHMNVER SLSEDEWKAVMGDSYQCKLCRYNT QLKANFQLHCKTDKHVQKYQLVAHI KEGGKANEWRLKCAIGNPVHLKC NACDYTNSLEKLRRLHTVNSRHEAS LKLYKHLQQHESGVEGESCYHCVL CNYSTKAKLNLIQHVRSMKHQRSES LRKLQRLQKGLPEEDEDLGQIFTIRR CPSTDPEEPVEDAEGPSEASADPEEL AKDQSGSSEEGQSKRAASSSQAEKE LTDSPATTKRTSFPGSSETPLSSKRP

KASEEIKPEQMYQCPYCKYSNADVN
RLRVHAMTQHSVQPLLRCPLCQDM
LNNKIHLQLHLTHLHSVAPDCVEKL
IMTVTAPEMVMPSSMFLPAAAARD
GNSTLEEVGKQPEASEDPGKNILPP
ASMEHGGDLKPTSADPSCGREDSG
FLCWKKGCNQVFKTSATLQTHFNE
VHAKRPQLPVS DRHVYKYRCNQCSL
AFKTIEKLQLHSQYHVIRAATMCCLC
QRSFRTFQALKKHLETSHLELSEADI
QQLYGGLLANGDLLAMGDPTLAED
HTIIVEEDKEEESDLEDKQSPTGSDS
GSVQEDSGSEPKRALPFRKGNFTM
EKFLDPSRPYKCTVCKESFTQKNILL
VHYNSVSHLHKLKRALQESATGQPE
PTSSPDNKPFKCNVCNVAYSQSSTL
EIHMRVSLHQTAKAAKLEAASGNS
NGTGNSGGVLSSTSPSPVGSSGAN
NTFTATNPSSAAMAPSVNALSQVPP
ESVVMPLGNPISANIASPSEPKEAN
RKKLADMIASRQQQQQQQQQQQQ
AQTLAQAQAQVQAHLQQLQQA
LIQSOLFNP TLLPHFPM TTTETLLQLQ
QQQHLLFPFYIPSAEFQLNPEVSLPV
TSGALTGTGSGPGLLEDLKVQVQIPQ
QSHQQILQQQQQSQLSLSQSHSAL
LQPSQHPEKKNKVVIKEDKESQRE
REGPEGAEGNTGPQESLPDASKAKE
KKDLAPGGGSEGTMLPPRIASDARG
NATKALLENFGFELVIQYNENKQKA
QKKNKGAEQGGESLEKLECDSCGK
LFSNILILKSHQEHVHQNYFFPKQLE
RFAKQYREHYDKLYPLRPQTPEPPPP
PPPPPPPLPTAPPQASAPAIPASAP
PITSPTIAPAQSVPLTQLSMPMELPI
FSPLMMQTMPLQTLPAQLPPQLGPV
EPLPADLAQLYQHQLNPTLLQQQNK
RPRTRITDDQLRVLRQYFDINNSPSE
EIKEMADKSGLPQKVIKHWFRNTL
FKERQRNKDSPYNFSNPPITSLEELK
IDSRPPSPEPQKQEYWGSKRSSRTR
FTDYQLRVLQDFFDANAYPKDDEFE
QLSNLLNLPTRVIVVWFQNAQKAR
KNYENQGEKDGGERRELTNDRIYRT
SNLNYQCKKCSLVFQRIFDLIKHQK
KLCYKDEDEEGQDDSQNEDSMDA
MEILTPTSSSCSTPMPSQAYSTPAPS
AAAANTAPSAFLQLTAETDELATFNS
KAEASDEKPKQADPPSAQPNQTQEK
QGQPKPEMQQLEQLEQKT NAPQP
KLPQPAAPSLPQPPPQAPPQCPLPQ
SSPSPSQLSHLPLKPLHTSTPQQLAN

LPPQLIPYQCDQCKLAFPSFEHWQE
HQQLHFLSAQNQFIHPQFLDRSLD
MPFMLFDPSNPLLASQLLSGAIPQIP
ASSATSPSTPTSTMNTLKRKLEEKAS
ASPGENDSGTGGEFPQRDKRLRTTI
TPEQLEILYQKYLLDSNPTRKMLDHI
AHEVGLKKRVVQVWFQNTRARERK
GQFRAVGPAQAHRRCPFCRALFKAK
TALEAHIRSRHWHEAKRAGYNLTLS
AMLLDCDGGGLQMKGDIFDGTSFSH
LPPSSSDGQGVPLSPVSKTMELSPRT
LLSPSSIKVEGIEDFESPSMSSVNLN
FDQTKLDNDDCSSVNTAITDTTTGD
EGNADNDSATGIATETKSSAPNEGL
TKAAMMAMSEYEDRLSSGLVSPAPS
FYSKEYDNEGTVDYSETSSLADPCSP
SPGASGSAGKSGDGGDRPGQKRFRT
QMTNLQLKVLKSCFNDRPTPTMLE
CEVLGNDIGLPKRVVQVWFQNARAK
EKKSKLSMAKHFGINQTSYEGPKTE
CTLCGIKYSARLSVRDHIFSQQHISK
VKDTIGSQLDKEKEYFDPATVRQLM
AQQELDRIKKANEVLGLAAQQQGM
FDNAPLQALNLPPTYPALQGIPPVLL
PGLNRPSLPGFTPANTALTSPKPNL
MGLPSTTVPSGLPTSGLPNKPSSAS
LSSPTPAQATMAMAPQPPPQPPQP
PPVQPPPPPPAAQQIPAPQLTPQQQR
KDKDGEKGKEKEKAHKGKGEPLPVP
KKEKGEAPPAGTGTISAPLPAMEYAV
DPAQLQALQAALTSDPALTLLTSQFLP
YFVPGFSPYYAPQIPGALQSGYLQPM
YGMGLFPYSPALSRPLMGLSPGSL
LQQYQQYQQSLQEAIQQQQQQQQQ
QQQQQQQQQRQLQQQQQQQQQKV
QQQQQQQQQPKASQTPVPQGAASP
DKDPAKESPKPEEQKNVPRELSPLL
PKPPEEPEAESKSASADSLCDFIVP
KVQYKLVCRKCQAGFGDEEAARSHL
KSLCCFGQSVVNLQEMVLHVPTGS
GGGGGGGGGGGGGGGGGGSYHCLACES
ALCGEEALSQHLESALHKHRTITRA
ARNAKEHPSLLPHSACFPDPSTAST
SQSAAHSNDSPPPPSAAPSSSASPH
ASRKSWPPVGSRASAAKPPSFPPLSS
SSTVTSSSCSTSGVQPSMPTDDYSEE
SDTDLSQKSDGPASPVEGPKDPSCP
KDSGLTSVGTDTFRL