

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
D3ZN95	HCFC1_RAT	Rattus norvegicus	Host cell factor 1	30.406554	T1148	S6;S411;S598;S666;S669;S1204;S1223;T1500;S1506;S1516;S1782;S1849	18683930;19692427;17496889;19322778;15340146	MASAVSPANLPVALLQPRWKRVRVW WSGPVPRPRHGHRAVAIKELIVVFG GGNEGIVDELHVYNTATNQWFIPAV RGDIPPGCAAYGFVCDGTRLLVFGG MVEYGKYSNDLYELQASRWEWKRL KAKTPKNGPPPCPRLGHSFSLVGNK CYLFGGLANDSEDPKNNIPRYLNDL YILELRPGSGVVAWDIPITYGVLPPPR ESHTAVVYTEKDNKKSCLVIYGGMS GCRLGDLWTLDIETLTWNKPSLSGV APLPRSLHSATTIGNKMYVFGGWVP LVMDDVKVATHEKEWKCTNTLACL NLDTMAWETILMDTLEDNIPRARAG HCAVAINTRLYIWSGRDGYRKAWNN QVCKDLWYLETEKPPPPARVQLVR ANTNSLEVSWGAVATADSYLLQLQK YDIPATAATATSPTPNPVSPVPANPP KSPAPAAAAPAVQPLTQVGITLVPQA AAAPPSTTTIQVLPVPGSSISVPTAA RTQGVPAVLKVTGPQATTGTPLVTM RPASQAGKAPVTVTSLPASVRMVVP TQSAQGTVIGSNPQMSGMAALAAAA AATQKIPPSSAPTIVLSPAGTTIVKTV AVTPGTTTTLPATVKVASSPVMVSNPA TRMLKTAQAQVGTSSAANTSTRPI ITVHKS GTVTVAQQAQVTTVGGV TKTITLVKSPISVPGGSALISNLGKVM SVVQTKPVQTSAVTGQASTGPVTQII QTKGPLPAGTILKLVTSADGKPTTIIT TTQASGAGTKPTILGISSVSPSTTKPG TTTIKTIPMSAIITQAGATGVTSSPGI KSPITIIITKVM TSGTGAPAKIITAVPK IATGHGQQGVTVVVLKGAPGQPGTI LRTVPMGGVRLVTPVTVSAVKPAVTT LVKGTGVTTLGTVTGTVSTSLAGA GAHSTSASLATPITTLGTIATLSSQVI NPTAITVSAAQTTLTAAGGLTPTIT MQPVSQPTQVTLITAPSGVEAQPVH DLPVSILASPTTEQPTATVTIADSGQG DVQPGVTTLVCSNPPCETHETGTTN TATTVVANLGGHPQPTQVQFVCDR QEAAASLVTSAVGQQNGNVVRVCS NPPCETHETGTTNTATTATSNMAGQ HGCSNPPCETHETGTTSTATTAMSS MGTGQQRDARRATNPTVVRIIVAP GALERAQGTVKPPCQTQQTNMTSTT MTVQATGAPCSAGPLLRSVALETG SHSPAFVQLALPSVRVGLSGPSSKDV PTGRQPETYHTYTTNTPTTARSIMVA GELGTARVVPTSQYDCLQASSPSST MTMTALEALLCPSATVTQVCSNPPC ETHETGTTNTATTSNAGSAQRVCSN

PPCETHETGTTHTATTATSNGGAGQ
PEGGQQPASGHPCETHQTTSTGTT
MSVSVGALIPDATPSHGTLESGLVV
AVPTVTSQAGATLLASFSTQRVCSNP
PCETHETGTTHTATTVTSNMSSNQD
PPPAASDQGEVASTQGDSTNITSASA
ITTTVSSTLPRAVTTVTQSTPVPGPSV
PPPEELQVSPGPRQQLPPRQLQSAS
TPLMGESAEVLSASQTPELQAAVDL
SSTGDPSSVQEPTTSAVVATVVVQPP
QPTQSEVDQLSLPQELMAEAQAGTT
TLMVTGLTPEELAVTAAAEAAAQAA
ATEEAQALAIQAVLQAAQQAVMGTG
EPMDTSEAAAAVTQAELGHLSAEGQ
EGQATTIPIVLTQQELAALVQQQQQL
QEAQAQAQQQHHLPTREALPADSL
NDPSIESNCLNELASAVPSTVALLPS
TATESLAPSNTFVAPQPVVVVASPAK
MQAAATLTEVANGIESLGVKPDLPPP
PSKAPIKKENQWFDVGVKGTSMVMV
THYFLPPDDAVQSDDDSGTVPDYNQ
LKKQELQPGTAYKFRVAGINACGRG
PFSEISAFKTCLPGFPGAPCAIKISKS
PDGAHLTWEPPSVTSGKIEYSVYLAI
QSSQASGEPKSSTPAQLAFMRVYCG
PSPSCLVQSSLSNAHIDYTTKPAIIF
RIAARNEKGYGPATQVRWLQETSKD
SSGTKPASKRPMSSPEM