

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
P98088	MUC5A_HUMAN	Homo sapiens	Mucin-5AC	13.589521	NaN	NaN	37217939	MSVGRRLALLWALALALACTRHTG HAQDGSSESSYKHHHPALSPIARGPS GVPLRGATVFP SLRTIPVVRASNPAH NGRVCSTWGSFHYKTFDGDVFRFP GLCNYVSEHCGAAYEDFNIQLRRS QESAAPTLSRVLKVDGVVIQLTKG SVLVNGHPVLLPFSQSGVLIQSSSY TKVEARLGLVLMWNHDDSLLELD TKYANKTCGLCGDFNGMPVSELLS HNTKLTPEFGNLQKMDDPTDQCQ DPVPEPPRNCSTGFGICEELLHGQL FSGCVLVDVGSYLEACRQDLCFCE DTDLLSCVCHTLAEYSRQCTHAGGL PQDWRGPDFCPQKCPNMQYHEC RSPCADTCSNQEHSRACEDHCVAG CFCPEGTVLD DIGQTGCVPVSKCAC VYNGAAYAPGATYSTDCTNCTCSGG RWSCQEVPCPGTCSVLGGAHFSTFD GKQYTVHGDCSYVLTKPCDSSAFTV LAELRRCGLTDSETCLKSVTSLDGA QTVVVIKASGEVFLNQIYTQLPISAA NVTIFRPSTFFIIAQTSGLQLNLQLV PTMQLFMQLAPKLRGQTCGLCGNF NSIQADDFRTL SGVVEATAAAFFNTF KTQAACP NIRNSFEDPCSLSVENEK YAQHWCSQLTDADGPFGRCHAAVK PGTYYSNCMFDT CNCERSEDCLCAA LSSYVHACA AKGVQLGGWRDGVCT KPM TTCPKSMTYHYHVSTCQPTCRS LSEGDITCSVGFIPVDGCICPKGTF DDTGKCVQASNCPCYHRGSMIPNG ESVHDSGAICTCTHGKLS CIGGQAPA PVCAAPMVFFDCRNATPGDTGAGC QKSCHTLDMTCYSPQCVPGVCPD GLVADGEGGCITAEDCPCVHNEASY RAGQTIRVGCNTCTCDSRMWRCTD DPCLATCAVYGDGHYLTFDQSYSF NGDCEYTLVQNHCGGKDSTQDSFR VVTENVPCGTTGTTCSKAIKIFLGGF ELKLSHGKVEVIGTDESQEVPTIRQ MGIYLVVDTDIGLVLLWDKKT SIFIN LSPEFKGRVCGLCGNFDDIAVNDFA TRRSVVDVLEFGNSWKLSPSCPD ALAPKDPCTANPFRKSWAQKQCSIL HGPTFAACHAHVEPARYYEACVNDA CACDSGGDCECFCTAVAAYA QACHE VGLCVSWRTPSICPLFCDYINPEGQ CEWHYQPCGVPCLRNPRGDCL RDVRGLEGCYKCPPEAPIFDEDKM QCVATCPTPLPPRCHVHGKSYRPG

AVVPSDKNCQSCLECTERGVECTYKA
EACVCTYNGQRFHPGDVIYHTTDGT
GGCISARCGANGTIERRVYPCSPPTP
VPPTTFSFSTPPLVVSSTHTPSNGPS
SAHTGPPSSAWPTTAGTSPRTRLPTA
SASLPPVCGEKCLWSPWMDVSRPG
RGTDSGDFDTLENLRAHGVRVCEP
RSVECRAEDAPGVPLRALGQRVQCS
PDVGLTCRNREQASGLCYNQIRVQ
CCTPLPCSTSSSPAQTTPTTSTKTE
TRASGSSAPSSTPGTVSLSTARTPA
PGTATSVKKTFTSPPPPVPATSTSS
MSTTAPGTSVVSSKPTPEPSTSSCL
QELCTWTEWIDGSYPAPGINGGDFD
TFQNLRDEGYTFCESPRSVQCRAES
FPNTPLADLGQDVICSHTEGLICLNK
NQLPPICYNYEIRIQCCETVNVCRDI
TRLPKTVATTRPTPHPTGAQTQTTFT
THMPSASTEQPTATSRGGPTATSVT
QGTHHTLVTRNCHPRCTWTKWFDV
DFPSPGPHGGDKETYNNIIRSGEKIC
RRPEEITRLQCRAKSHPEVSIEHLGQ
VVQCSREEGLVCRNQDQQGPFKMC
LNYEVRVLCCEPRGCHMTSTPGST
SSSPAQTTPTTSTTETQASGSSAP
SSTPGTVSLSTARTTPAGTATSVKK
TFSTPSPPPPVPATSTSSMSTTAPGTS
VVSSKPTPEPSTSSCLQELCTWTE
WIDGSYPAPGINGGDFDTFQNLRDE
GYTFCESPRSVQCRAESFPNTPLADL
GQDVICSHTEGLICLNKNQLPPICYN
YEIRIQCCETVNVCRDITRPPKTVATT
RPTPHPTGAQTQTTFTTHMPSASTE
QPTATSRGGPTATSVTQGTHHTPVTR
NCHPRCTWTTWFDVDFPSPGPHGG
DKETYNNIIRSGEKICRRPEEITRLQ
RAKSHPEVSIEHLGQVVQCSREEGL
VCRNQDQQGPFKMLNYEVRVLC
ETPKGCPVTSTPVTAPSTPSGRATSP
TQSTSSWQKSRTTLVTTSTTSTPQT
STTYAHTTSTTSAPTARTTSAPTTRTT
SASPASTTSGPGNTSPVPTTSTISAP
TTSITSAPTSTTSAPTSTTSGPGTT
PSPVPTTSITSAPTSTTSAPTSTTS
ARTSTTSATTTSRISGPETTPSPVPT
TSTTSATTTSTTSAPTSTTSAPTSTT
TSSPQTSTTSAPTSTTSGPGTTPSPV
PTTSTTSAPTTRTTSAPKSSTTSAATT
STTSGPETTPRPVPTTSTTSSPTTSTT
SAPTSTTTSASTTSTTSGAGTTPSPVP
TTSTTSAPTSTTTSAPISSTTSATTTST
TSGPGTTPSPVPTTSTTSAPTSTTSG

PGTTPSAVPTTSITSAPTTSTNSAPIS
STTSATTTSRISGPETTPSPVPTASTT
SASTTSTTSGPGTTPSPVPTTSTISVP
TTSTTSASTTSTTSASTTSTTSGPGTT
PSPVPTTSTTSAPTTSTTSAPTTSTIS
APTTSTTSATTTSTTSAPTPRRTSAPT
TSTISASTTSTTSATTTSTTSATTTSTI
SAPTTSTTLSPTTSTTSTTITSTTSAPI
SSTTSTPQSTTSAPTTSTTSGPGTTS
SPVPTTSTTSAPTTSTTSAPTRRTTSV
PTSSTTSTATTSTTSGPGTTPSPVPTT
STTSAPTRRTTSAPTTSTTSAPTTSTT
SAPTSSTTSATTTSTISVPTTSTTSVP
GTTSPVPTTSTISVPTTSTTSASTTS
TTSGPGTTPSPVPTTSTTSAPTTSTTS
APTTSTISAPTTSTPSAPTTSTTLAPT
STTSAPTTSTTSTPTSTTSSPQTSTT
SASTTSITSGPGTTPSPVPTTSTTSAP
TTSTTSAATTSTISAPTTSTTSAPTT
TTSASTASKTSGLGTTPSPIPTTSTTS
PPTTSTTSASTASKTSGPGTTPSPVPT
TSTIFAPRTSTTSASTTSTTPGPGTTP
SPVPTTSTASVSKTSTSHVSISKTHS
QPVTRDCHLRCTWTKWFDIDFSPG
PHGGDKETYNNIIRSGEKICRRPEEI
TRLQCRAESHPEVSIHLGQVVQCS
REEGLVCRNQDQGGPFKMCLNYEV
RVLCCETPKGCPVTSTPVTAPSTPSG
RATSPTQSTSSWQKSRTTTLVTTSTT
STPQSTTSAPTTSTTSAPTTSTTSAP
TTSTTSTPQTSISSAPTSSTTSAPTSS
TISARTTSIISAPTTSTTSSPTTSTTSA
TTTSTTSAPTSSTTSTPQTSKTSAA
STTSGSGTTPSPVTTTSTASVSKTST
HVSVSKTTHSQPVTRDCHPRCTWTK
WFDVDFPSPGPHGGDKETYNNIIRS
GEKICRRPEEITRLQCRASHPEVSI
EHLGQVVQCSREEGLVCRNQDQGG
PFKMCLNYEVRVLCCETPKGCPVTS
TSVTAPSTPSGRATSPTQSTSSWQKS
RTTTLVTSSITSTTQSTTSAPTTSTT
PASIPSTTSAPTTSTTSAPTTSTTSAP
TTSTTSTPQTTTSSAPTSSTTSAPTT
TISAPTTSTISAPTTSTTSAPTASTTSA
PTSTSSAPTTNTTSAPTTSTTSAPITS
TISAPTTSTTSTPQSTISSPTTSTTST
PQTSTTSSPTTSTTSAPTTSTTSAPTT
STTSTPQTSISSAPTSSTTSAPTASTIS
APTTSTTSFHTTSTTSPPTSSTSTPQ
TSKTSAASTTSGSGTTPSPVPTTST
ASVSKTSTSHVSVSKTTHSQPVTRD
CHPRCTWTKWFDVDFPSPGPHGGD

KETYNNIIRSGEKICRRPEEITRLQCR
AESHPEVSIIEHLGQVVQCSREEGLV
CRNQDQQGPFKMCCLNYEVRVLCCE
TPKGCPVTSTPVTAPSTPSGRATSPT
QSTSSWQKSRTTTLVTTSTTSTPQTS
TTSAPTTSTIPASTPSTTSAPTTSTTS
APTTSTTSAPTHRRTSGPTTSTTLAPT
TSTTSAPTTSTNSAPTTSTISASTTSTI
SAPTTSTISSPTSSTTSTPQTSKTSAA
TSSTTSGSGTTPSPVPTTSTTSASTTS
TTSAPTTSTTSGPGTTPSPVPSTSTTS
AATTSTTSAPTRRTSAPTSSMTSGP
GTTTPSPVPTTSTTSAPTTSTTSGPGTT
PSPVPTTSTTSAPITSTTSGPGSTPSP
VPTTSTTSAPTTSTTASTASTTSGPG
TTPSPVPTTSTTSAPTRRTSASTAST
TSGPGSTPSPVPTTSTTSAPTRRTTTPA
STASTTSGPGTTPSPVPTTSTTSASTT
STISLPTTSTTSAPITSM TSGPGTTPS
PVPTTSTTSAPTTSTTASTASTTSGP
GTTTPSPVPTTSTTSAPTTSTTASTAS
TTSAGPGTSLSPVPTTSTTSAPTTSTTS
GPGTTPSPVPTTSTTSAPTTSTTSGP
GTTTPSPVPTTSTTPVSKTSTSHLSVS
KTHSQPVTSDCHPLCAWTKWFDV
DFPSPGPHGGDKETYNNIIRSGEKIC
RRPEEITRLQCR AESHPEVNIEHLGQ
VVQCSREEGLVCRNQDQQGPFKMC
LNYEVRVLCCETPRGCPVTSVTPYGT
SPTNALYPSLSTSMVSASVASTSVAS
SSVASSVAYSTQTCFCNVADRLYPA
GSTIYRHRDLAGHCYALCSQDCQV
VRGVSDCPSTTLPPAPATSPSISTSE
PVTELGCPNAVPPRKKGETWATPNC
SEATCEGNNVISLRPRTCPRVEKPTC
ANGYPAVKVADQDGCCHHYQCQCV
CSGWGDPHYITFDGTYTFLDNCTY
VLVQQIVPVYGHFRVLVDNYFCGAE
DGLSCPRSIILEYHQDRVVLTRKPVH
GVMTNEIIFNKNVSPGFRKNGIVV
SRIGVKMYATIPELGVQVMFSGLIFS
VEVPFSKFANNTGQCGTCTNDRK
DECRTPRGTVVASCSEMSGLWNVSI
PDQPACHRPHTPTTVGPTTVGSTT
VGPTTVGSTTVGPTTPPAPCLPSPIC
QLILSKVFEPCHTVIPLLFYEGCVFD
RCHMTDLDVVCSLELYAALCASHD
ICIDWRGRTGHMCPFTCPADKVYQP
CGPSNPSYCYGNDASL GALPEAGPI
TEGCFCEPMTLFTSAQVCVPTGC
PRCLGPHGEPVKVGHTVGMDCQEC
TCEAATWTLTCRPKLCPLPPACPLPG

								FVPVPAAPQAGQCCPQYSCACNTSR CPAPVGCPEGARAIPTYQEGACCPVQ NCSWTVCSINGTLYQPGAVVSSSLC ETCRCELPGGPPSDAFVWSCETQICN THCPVGFHEYQEESGQCCGTCVQVA CVTNTSKSPAHLFYPGETWSDAGNH CVTHQCEKHQDGLVVTTKKACPPL SCSLDEARMSKDGCCRFCPPPPPPY QNQSTCAVYHRSLIIQQGCSSEPV RLAYCRGNCGDSSSMYSLEGNTVE HRCQCCQELRTSLRNVTLHCTDGSS RAFSYTEVEECGCMGRRCPAPGDTQ HSEAEPEPSQEAESGSWERGVPVS PMH
--	--	--	--	--	--	--	--	--