

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
P58107	EPIPL_HUMAN	Homo sapiens	Epiplakin	16.051874	T1624	T34;T1543;S2422;S2508;S2718;S2958;S3044;S3249;S3489;S3575;S3783;S4023;S4109;S4317;S4557;S4645;S4850	33214551;33465208;26374642;23301498;30397120	MSGHTLPLPLPVPGTNSTEQASVPRAMAATLGAGTPPRPQARSIAGVYVEASGQAQSVYAAMEQGLLPAGLGQALLEAQAATGGLVDLARGQLLPVSKALQQLVGLGLELKEKLLAAERATTGYDPYGEKLLALFQAIGKEVVDRALGQSWLEVQLATGGLVDPAQGVLVAPEPACHQGLLDRETWHKLSELEPGTGLRFLDPNTLERLTYHQLLERCVRAPGSGLALLPLKITFRSMGGAVSAAELLEVGILDEQAVQGLREGRLAAVDVSARAEVRRYLEGTGSVAGVLLPEGHKKSFQAAATEHLLPMGTALPLLEAQAATHLVDPIQGQRLWVDEAVRAGLVSPELHEQLLVAEQAVTGHHDPFSGSQIPLFQAMKKGLVDRPLALRLDAQLATGGLVCPARRRLPLEAALRCGCLDEDTQRQLSQAGSFSFGTHGGLRYEQLLALCVTDPETGLAFLPLSGGPRGGEPQGPPFIKYSTRQALSTATATVSVGKFRGRPVSLWELLFSEASSEQRAMLAQYQEGTSLVEKLAAKLSATLEQAAATARVTFSGLRDTPVTPGELLKAEIIDQDLYERLEHGQATAKDVGSLASVQRYLQGTGCIAGLLPQSQRERLSIYEARCKGLLRPGTALILLEAQAATGFIDPKANKGHSVEEALRAAVIGPDVFAKLLSAERAVTGYTDPYTGQQISLQAMQKGLIVREHGIRLLEAQIATGGVIDPVHSHRVPVDVAYRRGYFDQMLNLILLDPSDDTKGFFDPNTHENLTYLQLLERCVRDPETGLYLLPLSSTQSPLVDSATQQAQFNLLLSVKYGRFQGRVSAWELINSEYFSEGRRRQLRRYRQREVTLGQVAKLLEAETQRQADIMLPALRSRVTVHQLLEAGIIDQQLLDQVLGAGTISPEALLLMDGVERRYL CGLGAVGGVRLPSGQRLSLYQAMRQKLLGPRVALALLEAQAATGTIMDPHSPELSVDEAVRRGVVGPPELYGRLKRAEGAIAGFRDPFSGKQVSVFQAMKKGLIPWEQAARLLEAQVATGGIIDPTSHHHLMPVAIQRGYVDQEMETALSSSETFPTPDGQGRTSYAQLLEECPRDETSGLHLLPLPESAPALPTEEQVQRSLQAVPGAKDGTSLWDLSSCHFTEEQRRGLEDVQEGRTTVPQLLASVQRWVQETKLLAQARVMVPGPRGEVPAVWLLDAGIITQETLEALAQGTQSPAQVAEQPAVKACLWGTGCVAGVLLQPSGAKASIAQAVRDGLLPTGLGQRLLEAQVASGFLVDPLNNQRLSVEDAVKVLVGVRELSEQLGQAERAAAAGYPDPYSRASLSLWQAMEKGLVPQNEGLPLLQVQLATGGVVDPV

HGVHLPQAAACRLGLLDTQTSQVLT
AVDKDNKFFFDPSARDQVTYQQLRE
RCVCDSETGLLLLPLPSDTVLEVDD
HTAVALRAMKVPVSTGRFKGCSVSL
WDLLESEYVGADKRRELVALCRSGR
AAALRQVVS AVTTLVEAAERQPLQAT
FRGLRKQVSARDLFRACLISRKTLDE
LSQGT TTVKEVAEMDSVKRSLEGGN
FIAGVLIQGTQERMSIPEALRRHILRP
GTALV LLEAQAATGFIIDPVENRKL
VEEAFKAGMFGKETYVKLLSAERAV
TGYTDPYTGQQISLFQAMQKDLIVRE
HGIRLLEAQIATGGIIPVHSHRVPV
DVAYRCGYFDEEMNRILADPSDDTK
GFFDPNTHENLTYLQLLERCVEDPE
TGLYLLQIIKKGENYVYINEATRHLV
QSRTAKMRVGRFADQVVSFWDLLS
SPYFTEDRKRELIQEYGAQSGGLEKL
LEIITTTIETETQNGIKVAAIRGEV
TAADLFNSRVIDQKTLHTLRVGRGTG
GQALSTLECVKPYLEGSGCIAGVTVP
STREVM SLHEASRKELIPAAFATWL
LEAQAATGFLDPCTRQKLSVDEAV
DVGLVNEELRERLLKAERAATGYRD
PATGDTIPLFQAMQKQLIEKAEALRL
LEVQVATGGVIDPQHHRHRLPLETAY
RRGCLHKDIYALISDQKHMRRKRFVD
PNTQEKVSYRELQERC RPQEDTGWL
LFPVNKAARDSEHIDDETRRALEAE
QVEITVGRFRGQKPTLWALLNSEYV
TEEKKLQLVRMYRTHTRRALQTVAQ
LILELIEKQETSNKHLWFQGIRRQIT
ASELLSSAITEEMLQDLETGRSTTQ
ELMEDDRVKRYLEGTSCIAGVLVPA
KDQPGRQEKMSIYQAMWKGVL RPG
TALV LLEAQAATGFVIDPVRNLRLSV
EEAVAAGVVGGEIQEKLLSAERAVTG
YTDPYTGQQISLFQAMQKDLIVREH
GIRLLEAQIATGGVIDPVHSHRVPVD
VAYRRGYFDEEMNRVLADPSDDTK
GFFDPNTHENLTYVQLLRRCVDPDP
TGLYMLQLAGRS AVHQLSEELRCA
LRDARVTPGSGALQGQSVSWELLF
YREVSEDRRQDLLSRYRAGTLTVEEL
GATLTSLLAQAQAQARAEAEAGSPR
PDPREALRAATMEVKVGR LRGRAVP
VWDV LASGYVSRAAREELLA EFGSG
TLDLPALTRRLTAHIEEAEAPGARPO
LQDAWRGPREP GPAGRGDGD SGRS
QREGQGEGETQEAAAAAAAAAARRQ
EQTLRDATMEVQRGQFQGRPVSVW
DVL FSSYLSEARRDELLAQHAAGAL
GLPDLVAVLTRVIEETEERLSKVSFR
GLRRQVSASELHTSGILGPETLRDLA
QGTKTLQEVTEMDSVKRYLEGTSCI
AGVLVPAKDQPGRQEKMSIYQAMW
KGVLRPGTALV LLEAQAATGFVIDPV

RNLRLSVEEAVAAGVVGGEIQEKLLS
AERAVTGYTDPYTGGQISLFFQAMQK
DLIVREHGIRLLEAQIATGGVIDPVHS
HRVPVDVAYQRGYFDEEMNRVLAD
PSDDTKGFFDPNTHENLTYVQLLRR
CVPDPDTGLYMLQLAGRGSVHQLS
EELRCALRDARVTPGSGALQGQSVS
VWELLFYREVSEDRRQDLLSRYRAG
TLTVEELGATLTSLLAQQAQARAEA
EAGSPRPDPREALRAATMEVKVGR
RGRAVPVWDVLAGSYVSRAAREELL
AEFGSGTLDLPALTRRLTAIIIEAEEA
PGARPQLQDAWRGPREPGPAGRGD
GDSGRSQREGQGEGETQEAAAAAR
RQEOTLRDATMEVQRGQFQGRPV
VWDVLFSSYLSEARRDELLAQHAAG
ALGLPDLVAVLTRVIEETEERLSKVS
FRGLRCQVSASELHTSGILGPETLRD
LAQGTKTLQEVTEMDSVKRYLEGT
CIAGVLPKADQPGRQEKMSIYQAM
WKGVLRPGTALVLEAQAATGFVIDP
VRNLRLSVEEAVAAGVVGGEIQEK
LSAERAVTGYTDPYTGGQISLFFQAM
QKDLIVREHGIRLLEAQIATGGVIDP
HSHRVPVDVAYRRGYFDEEMNRVL
ADPSDDTKGFFDPNTHENLTYVQLL
RRCVDPDTGLYMLQLAGRGSVH
QLSEELRCALRDARVTPGSGALQGQ
SVSVWELLFYREVSEDRRQDLLSRY
RAGTLTVEELGATLTSLLAQQAQAR
AEAEAGSPRPDPREALRAATMEVKV
GRLRGRAVPVWDVLAGSYVSRAARE
ELLAEFGSGTLDLPALTRRLTAIIIEA
EEAPGARPQLQDAWRGPREPGPAG
RGDGDGRSQREGQGEGETQEAAA
ATAAARRQEOTLRDATMEVQRGQF
QGRPVSVWDVLFSSYLSEARRDELL
AQHAAGALGLPDLVAVLTRVIEETEE
RLSKVSFRGLRRQVSASELHTSGILG
PETLRDLAQGTKTLQEVTEMDSVKR
YLEGTSCIAGVLPKADQPGRQEK
SIYQAMWKGVLRPGTALVLEAQA
TGFVIDPVRNLRLSVEEAVAAGVVG
GEIQEKLLSAERAVTGYTDPYTGGQI
SLFFQAMQKDLIVREHGIRLLEAQIAT
GGVIDPVHSHRVPVDVAYRRGYFDE
EMNRVLADPSDDTKGFFDPNTHEN
LTYVQLLRRRCVDPDTGLYMLQLAG
RGSVHQLSEELRCALRDARVTPGS
GALQGQSVSVWELLFYREVSEDRRQ
DLLSRYRAGTLTVEELGATLTSLLAQ
AQQAARAEAEAGSPRPDPREALRAA
TMEVKVGRRLRGRAVPVWDVLAGSY
VSRAAREELLAEFGSGTLDLPALTRR
LTAIIIEAEEAPGARPQLQDAWRGPR
EPGPAGRGDGDGRSQREGQGEGET
QEAAAATAAARRQEOTLRDATMEV

QRGQFQGRPVSVWDVLFSSYLSEAR
RDELLAQHAAGALGLPDLVAVLTRVI
EETEERLSKVSFRGLRRQVSASELH
TSGILGPETLRDLAQGKTLQEVTEM
DSVKRYLEGTSCIAGVLVPAKDQPG
HQEKMSIQAMWKGVL RPGTALVL
LEAQAATGFVIDPVRNLRLSVEEAVA
AGVVGGEIQEKLLSAERAVTGYTDPY
TGQQISL FQAMQKDLIVREHGIRLLE
AQIATGGVIDPVHSHRVPVDVAYRRG
YFDEEMNRVLAHPSDDTKGFFDPN
THENLTYVQLLRRCVDPDTGLYML
QLAGRGS AVHQLSEELRCALRDARV
MPGSGALQGSVSVWELLYREVSE
DRRQDLLSRYRAGTLTVEELGATLTS
LLAQAAQARAEEAEAEAGSPRPDPR
EALRAATMEVKVGR LRGRAVPVWD
VLASGYVSGAAREELLA EFGSGTLDL
PALTRRLTAIIIEAEEAPGARPQLQD
AWRGPREPGPAGRGDGD SGRSQRE
GQGEGETQEAAAAARRQE QTLRDAT
MEVQRGQFQGRPVSVWDVLFSSYL
SEHRDELLAQHAAGALGLPDLVAV
LTRVIEETEERLSKVSFRGLRRQVSA
SELHTSGILGPETLRDLAQGKTLQE
VTEMDSVKRYLEGTSCIAGVLVPAK
DQPGRQEKMSIQAMWKGVL RPGT
ALV LLEAQAATGFVIDPVRNLRLSVE
EAVAAGVVGGEIQEKLLSAERAVTGY
TDPYTGQQISL FQAMQKDLIVREHGI
RLLEAQIATGGVIDPVHSHRVPVDVA
YRRGYFDEEMNRVLADPSDDTKGFF
DPNTHENLTYLQLLQRATLDPETGL
LFLSLSLQ