

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
Q8TCU4	ALMS1_HUMAN	Homo sapiens	Alstrom syndrome protein 1	11.375847	S951;T1136;T1142;S1143;S1236;T1242;S1266;T1275;T1276;T1282;S1283;S1300;S1303;T1373;T1512;T1560;S1645;S1662;S1708;S1786;T1796;T1802;S1805;T2553;T2613	S464;S1189;S2143;S2466;S2632;S2805	31492838;29351928;34019948;31373491;30620550;35254053;35132862;30059200;30379171;33214551	MEPEDLPWPGELEEEEEEEEEEEEEEE EEEEAAAAANVDDVVVVVEVEVEEA GRELDSDSHYGPHLESIDDEEDDEE AKAWLQAHPRILPPLSPPQHRYSE GERTSLEKIVPLTCHVWQQIVYQGN SRTQISDTNVVCLLETTAQRGSGDDQ KTESWHCLPQEMDSSQTLDTSQTR FNVRTEDTEVTFDPSLEEGILTQSEN QVKEPNRDLFCSPLLVIQDSFASPDL PLLTCLTQDQEFAPDSLFFHQSELSFA PLRGIPDKSEDTEWSSRPSEVSEALF QATAEVASDLASSRFSVSVQHPLIGST AVGSQCPFLPSEQGNNEETISSVDE LKIPKDCDRYDDLCSYMSWKTRKDT QWPENNLADKQVSVATSFIDITDEN IATKRSDHFDAAARSYGQYWTQEDSS KQAETYLTKGLQKVESDVITLDGL NENAVVCSERVAELQRKPTRESEYH SSDLRMLRMSPDTPKAPKHLKAG DTSKGGIAKVTQSNLKSGITTPVDS DIGSHLSLSLEDLSQLAVSSPLETTT GQHTDTLNQKTLADTHLTEETLKVT AIPPEADQKTATPTVLSSSHSHRGKP SIFYQQGLPDSLHTEALKVSAAPGL ADQTTGMSTLTSTSYSHREKPGTFY QQELPESNLTEEPLEVSAAPGPVEQ KTGIPTVSSSTSHSHVEDLLFFYRQTL PDGHLTDQALKVSAVSGPADQKTGT ATVLSSTPHSHREKPGIFYQQEFADS HQTEETLTKVSATPGPADQKTEIPAV QSSSYSQREKPSILYPQDLADSHLPE EGLKVS AVAGPADQKTGLPTVPSSAY SHREKLLVFYQQALLDLSHLPEEALK VSAVSGPADGKTGTPAVTSTSSASS LGEKPSAFYQQTLPNSHLTEEALKV SIVPGPDQKTGIPSAPSSFYSHREK PIIFSQQTLPDFLPPEEALKVSAVSVL AAQKTGTPTVSSNSHSHSEKSSVFY QQELPDSDLPRESLKMSAIPGLTDQ KTVPTPTVPSGFSHREKPSIFYQQE WPDSYATEKALKVSTGPGPADQKTEI PAVQSSSYPQREKPSVLYPQVLSDSH LPPEESLKVSAFPGPADQMTDTPAVP STFYQREKPGIFYQQTLPESHLPKE ALKISVAPGLADQKTGTPTVTSTSYS QHREKPSIFHQALPGTHIPEEAQK VSAVTGPGNQKTWIPRVLSTFYQR EKPGIFYQQTLPGSHIPEEAQKVS LGPADQKTGTPTPTSASYSHTKPGI FYQQVLPDNPTEEALKISVASEPVD QTTGTPAVTSTSYSQYREKPSIFYQQ SLPSSHLTEEAKNVSAVPGPADQKT VIPILPSTFYSHTEKPGVFYQQVLP HPTEEALKISVASEPVDQTTGTPTVT STSYSQHTKPSIFYQQSLPGSHLTE EAKNVSAVPGGDRKTGIPTLPSTFY SHTEKPGSFYQQVLPVSHLPEEAL VSAVPGVDQITGTPTVTSPPSSSFG KPIVYKQAFPEGHLPEESLKVSVAP GPVQTTGAPTITSPSYSQHRAKSGS FYQLALLGSQIPEEALRVSSAPGPAD QTTGIPTITSTSYSGEKPVNYKQAF PDGHLPEEALKVSIIVSGPTEKTDIP

AGPLGSSALGEKPITFYRQALLDSPL
NKEVVKVAAPGPADQKTETLPVHS
TSYSNRGKPVIFYQQTLSDSLPEEA
LKVPPVPGPDAQKTETPSVSSLSYS
REKPIVFYQQALPDELSTQEALKVSA
VPQPADQKTGLSTVTSSFFYSHTEKP
NISYQQELPDSLHTEALKVSINVPGP
ADQKTGVSTVTSTSYSHREKPIVSYQ
RELPHFTEAGLKILRVPGPADQKTGI
NILPSNSYPQREHSVISYEQELPDLT
EVLTKAIGVPGPADQKTGIQIASSSSY
SNREKASIFHQELPDVTEALNVF
VVPQGDRKTEIPTVPLSYSRREKP
SVISQQELPDSLHTEALKVSIPVIPA
EQKTGIPIGLSSSYSHSHKEKLIKISTV
HIPDDQKTEFPAATLSSYSQIEKPKIS
TVIGPNDQKTPSQTAFHSSYSQTVKP
NILFQQQLPDRDQSKGILKISAVPEL
TDVNTGKPVSLSSSYFHREKSNIFSP
QELPGSHVTEDEVLVKSTIPGPAGQKT
VLPTALPSSFSHREKPDIFYQKDLDP
RHLTEDALKISSALGQADQITGLQTV
PSGTYSHGENHKLKLVSEHVQRLIDNL
NSSDSSVSSNNVLLNSQADDRVIN
KPESAGFRDVGSEEIQDAENSAKTL
KEIRTLLEAENMALKRCNFPAPLA
RFRDISDISFIQSKKVVCFKESSTGV
SNGDLLHRQPFTTEESPSSRCIQKDIG
TQTNLKCRRGIENWEFISSTTVRSPL
QEAESKVSMALEETLRQYQAAKSVM
RSEPEGCSGTIGNKIIIPMMTVIKSDS
SSDASDGNCSWDSNLPELESVS
DVLLNFFPYVSPKTSITDSREEEGVS
ESEDGGGSSVDSLAAHVKNLLQCES
SLNHAKELRNAEEEESRVRAHAWN
MKFNLAHDCGYSISELNEDDRRKEVE
EIKAEELFGHGRTDLSKGLQSPRGM
GCKPEAVCSHIIIESHEKGFRTLTS
EHPQLDRHPCAFRSAGPSEMTRGR
QNPSSCRAKHVNLSASLDQNNSHF
KVVNSLQKSHSPFQNFIPDEFKIS
KGLRMPFDEKMDPWLSELVEPAFV
PPKEVDFHSSSQMPSPPEPMKKFTTS
ITFSSHRHSCISNSSVVKGVTEGS
QCTGASVGVFNHSHFTEEQNPPRDLK
QKTSSPSSFKMHSNSQDKEVTILAE
GRRQSQKLPVDFERSFQEEKPLERS
DFTGSHSEPSTRANCSNFKEIQISD
NHTLISMGRPSSTLGVNRSSSRLGV
KEKNVTITPDLPCIFLEQRELFEQS
KAPRADDHVRKHHSPSPQHODYVA
PDLPCIFLEQRELFEQCKAPYVDHQ
MRENHSPLPQGQDSIASDLPSPISE
QCQSKAPGVDDQMNKHHFPLPQGQ
DCVVEKNNQHKPKSHISNINVEAKF
NTVVSQAPNHCTLAASASTPPSNR
KALSCVHITLCPKTSKLDSTGLDER
FHSLDAASKARMNSEFNFDLHTVSS
RSLEPTSKLLTSKPVAQDQESLGLG
PKSSLDFQVVQPSLPDSNTITQDLKT
IPSQNSQIVTSRQIQVNISDFEGHSN
PEGTPVFADRLPEKMKTPLSAFSEKL
SSDAVTQITTESPEKTLFSSEIFINAE
DRGHEIIEPGNQKLRKAPVKFASSSS
VQQVTFSRGTDGQPLLLPYKPSGST
KMYVYVQLRQIPSPDSKSDTTVESS

HSGSNDAIAPDFPAQVLGTRDDDL
ATVNIKHKEGIYSKRVTKASLPVGE
KPLQENADASVQVLITGDENLSDK
KQQEIHSTRAVTEAAQAKEKESLQK
DTADSSAAAAAEHSAQVGDPEMKN
LPDTKAITQKEEIHRKKTVPPEAWPN
NKESSLQINIEESECHSEFENTTRSVF
RSAKFYIHHPVHLPDQDICHESLG
KSVFMRHSWKDFQHHHPDKHREH
MCLPLPYQNMDTKTDTYTRIKSLSI
NVNLTGNKEVMDTTKSQVRDYPKHN
GQISDPQRDQKVTPQTTQHTVSLN
ELWNKYRERQRQRQPELQDRKEL
SLVDRDLRLAKILQNPITHSLQVSES
THDDSRGERSVKEWSGRQQORNKL
QKKRFRKSLEKSHKNTGELKSKVL
SHHRAGRSNQIKIEQIKFDKYLSKQ
PGFNYSNTSSDCRPSESELLTDTT
TNILSGTTSTVESDILTQTDREVALH
ERSSSVSTIDTARLIQAFGHERVCLS
PRRIKLYSSITNQRRYLEKRSKHSK
KVLNTGHPLVTSEHTRRRHIQVANH
VISSDSISSASSFLSSNSTFCNKQN
DVGITFPTSSSEAKLEENSDVTSWS
EEKREEKMLFTGYPEDRKLKKNKK
NSHEGVSWFVPVENVESRSKKENV
PNTCGPGISWFEPITKTRPWREPLRE
QNCQGQHLGGRGYLAGPGREAGR
LLRPFVRATLQESLQFHRPDFISRS
ERIKRLKLVQERKLSMLQTERDAL
FNIDRERQGHQNRMCPLPKRVFLAI
QKNKPISKKEMIQRSKRIYEQLEPEV
KKREEKRSYKSYRLRAQLYKRR
VTNQLLGRKVPWD