

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
P51610	HCFC1_HUMAN	Homo sapiens	Host cell factor 1	91.184852	T262;T405;T408;S411;T413;S419;S427;T441;T446;T457;T458;T459;T460;T466;S470;T480;T490;T495;T496;T498;T502;S507;T515;T517;S518;T529;S531;T535;S539;S540;T556;S562;S563;T566;S569;T574;T575;T579;T583;T586;T587;T588;T592;S597;S598;S603;T607;T612;T619;S620;S622;S623;T625;T627;S628;T629;T634;S638;T640;T642;T651;T652;T658;S666;S669;S674;S678;S685;T689;T694;S695;T698;S702;T703;T707;T712;T720;T726;S727;T733;T734;T737;T738;T739;S742;T746;T749;S754;S755;S757;S759;T760;T761;T765;T767;T771;S775;T779;T784;T787;S788;S789;S794;T797;T800;T801;T805;S806;T808;T823;T831;T848;T858;T861;S863;T870;T871;T877;T878;T881;T882;T885;T887;T889;S891;T892;S893;S901;T902;S905;T908;T911;T912;T915;T918;S920;S921;T927;T930;S932;T936;T937;T939;T945;T946;T948;T950;S955;T958;S967;S980;T995;T1031;S1058;T1061;S1062;T1063;T1143;S1150;T1153;S1162;S1221;S1234;S1238;T1239;T1243;S1245;S1246;T1273;T1335;T1386;S1398;T1400;T1485;T1486;T1488;T1491;S1497;T1743;S1749;T1752;S1758;T1760	S6;S411;S598;S666;S669;S1205;S1224;T1491;S1497;S1507;S1771;S1838	23301498;34105348;31373491;21295698;36240223;22661428;29351928;34931806;27114449;33214551;30059200;26305326;30397120;24311690;21158410;30620550;20676906;29237092;26075789;34065225;32522876;35254053;28657654;32092778;32870666;31637018;34725712;20305658;28584052;20068230;23335398;32574038;29249144;30379171;31492838;34846842;33465208;28604694;33103664;32119511;35289036;35138101;35083852;28510447;27655845;35132862;27056667;35164752;21740066;34019948	MASAVSPANLPVLLQPRWKRVVG WSGPVPRRHGHRAVAIKELIVVFG GGNEGIVDELHVYNTATNQWFIPAV RGDIPPGCAAYGFVCDGTRLLVFGG MVEYGKYSNDLYELOASRWEWKRL KAKTPKNGPPPCPRLGHSFSLVGNK CYLFGGLANDSEDPKNNIPRYLNDL YILELRPGSDVVAWDIPITYGVLPPPR ESHTAVVYTEKDNKSKLVIYGGMS GCRLGDLWTLDDTLTWNKPSLSGV APLPRSLHSATTIGNKMYVFGWVVP LVMDDVKVATHEKEWKCTNTLACL NLDTMAWETILMDTLEDNIPRARAG HCAVAINTRLYIWSGRDGYRKAWN QVCKDLWYLETEKPPPARVQLVR ANTNSLEVSWGAVATADSYLLQLQK YDIPATAATATSPTPNPVPSVPANPP KSPAPAAAAPAVQPLTQVGITLLPQA APAPPTTTTIQVLPVTPVSSISVPTAA RTQGVPAVLKVTGPQATTGTPLVTM RPASQAGKAPVTVTSLPAGVRMVVP TQSAQGTVIGSSPQMSGMAALAAA AATQKIPSSAPTLSVVPAGTIVKTM AVTPGTTTTLPATVKVASSPVMVSNPA TRMLKTAQAQVGTSSSATNTSTRPI ITVHKSGTVAQAQVVTTVVGGV TKTITLVKSPISVPGGASALISNLGKVM SVVQTKPVQTSAVTQASTGPVQTQII QTKGPLPAGTILKLVTSADGKPTTIIT TTQASGAGTKPTILGISSVSPSTTKPG TTTIKTIIPMSAIITQAGATGVTSSPGI KSPITIIITKVMSTSGTGAPAKIITAVPK IATGHGQQGVTVVVLKQAPGQPGTI LRTVPMGGVRLVTPVTVSAVKPAVTT LVVKGTTGVTTLGTVTGTVSTSLAGA GGHSTSASLATPITTLGTIATLSSQVI NPTAITVSAAQTTLTAAGGLTPTIT MQPVSQPTQVTLITAPSGVEAQPVH DLPVSILASPTTEQPTATVTIADSGQG DVQPGTVTLVCSNPPCETHETGTTN TATTTVVANLGGHPQPTQVQFVCDR QEAAAALVSTVGGQNGSVVRVCSN PPCETHETGTTNTATTATSNMAGQH GCSNPPCETHETGTTNTATTAMSSV GANHQRDARRACAAGTPAVIRISVAT GALEAAQGSKSCQCTRQTSATSTTM TVMATGAPCSAGPLLGPSMAREPGG RSPAFVQLAPLSSKVRLLSSSIKDL AGRHSHAVSTAAMTRSSVVGAGEPR MAPVCESLQGGSPSTTVVTALEALL CPSATVTQVCSNPPCETHETGTTNT ATTSNAGSAQRVCSNPPCETHETGT THTATTATSNGGTGQPEGGQPPAG RPCETHQTTSTGTTMSVSVGALLPD ATSSHRTVESGLEVAAPSVTPQAGT ALLAPFPPTQVCSNPPCETHETGTT HTATTVTSNMSSNQDPPPAASDQGE VESTQGDSVNITSSAITTTVSTLTR AVTTVTQSTPVPGPSVPPPEELQVSP GPRQQLPPRQLLQSASTALMGESAE VLSASQTPELPAAVDLSSTGEPSSGQ

