

UniprotKB ID	Entry name	organism	full name	oglcnacscore	oglcnac sites	phosphorylation sites	PMIDS	sequence
P21675	TAF1_HUMAN	Homo sapiens	Transcription initiation factor TFIID subunit 1	5.263779	NaN	S137;S307;S1669;S1672;S1778;S1781;S1799;S1826;S1697;S1700;S1700;S1702	34019948	MGPGCDLLLRTAATITAAAIMSDTDS DEDSSAGGGPFFSLAGFLFGNINGAGQ LEGESVLDDCECKKHLAAGLGLGLGS LITELTANEELTGTGALVNDGEWV RSTEDAVDYSINEVAEDESRRYQQ TMGSLQPLCHSDYDEDDYDADCEDI DCKLMPPPPPPGPMKKDKDQDSIT GEKVDFSSSSDSESEMGPQEQATQAE SEDGKLTPLAGIMQHDATAKLLPSVT ELFPEFRPGKVLRLRFGPGKNVPS VWRSARRKRRKKHRELIIQEEQIQEV ECSVESEVSQKSLWNYDYAPPPPE QCLSDDEITMMAPVESKFSQSTGDI DKVTDTKPRVAEWRYGPARLWYDM LGVPEDEGSGFDYGFKLKTEHEPVI KSRMIEEFRKLEENNGTDLLADENF LMVTQLHWEDDIIWDGEDVKHKG KPQRASLAGWLPSSMTRNAMAYNV QQGFAATLDDDKPWYSIFPIDNEDL VYGRWEDNIIWDAQAMPRLLEPPVL TLDPNDENLILEIPDEKEEATSNSPS KESKKESSLKSRILLGKTGVIKEEP QQNMSQPEVKDPWNLSNDEYYYPK QQGLRGTFGGNIIQHSIPAVELRQPF FPTHMGPIKLRQFHRPPLKKYSFGA LSQPGPHSVQPLLKHIKKAKMREQ ERQASGGGEMFFMRTPDQDLTGKDG DLILAEYSEENGPLMMQVGMATKIK NYYKRKPGKDPGAPDCKYGETVYCH TSPFLGSLHPGQLLQAFENLFRAPI YLHKMPETDFLIIRTRQGYIYRELVDI FVVGQQCPLFEVPGPNSKRANTHIR DFLQVFIYRLFVWKSADRPRRIRMEDI KKAFFSHSESSIRKRLKLCADFKRTG MDSNWWVLKSDFRLPTEEEIRAMV SPEQCCAYYSMAAEQRLKDAGYGE KSFFAPEEENEEDFQMKIDDEVRTA PWNTRAFIAAMKGGKCLEVTGVAD PTGCGEGFSYVKIPNKPTQKDDKE PQPVKKTVTGTDADLRRLSLKNAKQ LLRKFGVPEEEIKKLSRWEVIDVVRT MSTEQARSGEGPMSKFARGSRFSV AEHQERYKEECQRIFDLQNKVLSST EVLSTDTDSSSAEDSDFEEMGKNIE NMLQNKKTSSQLSREREEQERKEL QRMLLAAGSAASGNNHRDDDTASV TSLNSSATGRCLKIYRTFRDEEGKEY VRCETVRKPAVIDAYVRIRTTKDEEFI RKFALFDEQHREEMRKERRRIQEQL

RRLKRNQEKEKLKGPPEKKPKMK
ERPDLKLCGACGAIGHMRTNKFCP
LYYQTNAPPSNPVAMTEEQEELEK
TVIHNDNEELIKVEGTKIVLGKQIE
SADEVRRKSLVLKFPKQQLPPKKR
RVGTTVHCDYLNRPKHSIHRRRDP
MVTLSILESIINDMRDLPNTYPFHT
PVNAKVVKDYYKIITRPMDLQTLREN
VRKRLYPSREEFREHLELIVKNSATY
NGPKHSLTQISQSMLDLCDEKLKEK
EDKLARLEKAINPLDDDDQVAFSFI
LDNIVTQKMMAVPDSWPFHHPVVK
KFVPDYKVVIVNPMDLTIRKNISKH
KYQSRESFLDDVNLILANSVKYNGP
ESQYTKTAQEIVNVCYQTLTEYDEHL
TQLEKDICTAKEAALEEAELESLDPM
TPGPYTPQPPDLYDTNTSLSMSRDA
SVFQDESNMSVLDIPSATPEKQVTQ
EGEDGDGLADEEEGTQQPQASVL
YEDLLMSEGEDDEEDAGSDEEGDN
PFSAIQLSESGSDSDVGSGGIRPKQP
RMLQENTRMDMENEESMMSYEGD
GGEASHGLEDSNISYGSYEEDPKS
NTQDTSFSSIGGYEVSEEEEEDEEEE
QRSGPSVLSQVHLSEDEEDEDSEDFHS
IAGDSLDSDE