

Pro tyto sekvence:

>sekvence1

MLEICLKLVLGCKSKKGLSSSSSSCYLEEALQRPVASFEPQGLSEAARWNSKENLLAGPSE
NDPNLFWALYDFVAGDNTLSITKGEKLRVLGYNHNGEWCEAQTKNGQGWVPSNYITPVN
SLEKHSWYHGVPVSRNAEYLLSSGINGSFLVRESESSPGQRSISLRYEGRVYHYRINTAS
DGKLYVSSSESRFNTLAEVHHHSTVADGLITTLHYPAKRNKPTVYGVSPNYDKWEMERT
DITMKHKLGGGQYGEVYEGVWKKYSLTVAVKTLKEDTMEVEEFLEKAAVMKEIKHPNLVQ
LLGVCTREPPFYIITEFMTYGNLLDYLRECNRQEVNAVVLVLYMATQISSAMEYLEKKNFI
HRDLAARNCLVGENHLVKVADFLSRLMTGDTYTAHAGAKFPIKWTAPESLAYNKFSIKS
DVWAFGLLWEIATYGMSPYPGIDLSQVYELLEKDYRMERPEGCPEKVYELMRACWQWNP
SDRPSFAEIHQAFETMFQESSISDEVEKELGKQGVAVSTLLQAPELPTKTRTSRRAAE
HRDTTDVPMPHSGKQGESDPLDHEPAVSPLLPRKERGPPEGGLNEDERLLPKDKKTNLF
SALIKKKKKTAPTTPKRSSSFREMDGQPERRGAGEEEGRDISNGALAFTPLDTADPAKSP
KPSNGAGVPNGALRESGGSGFRSPHLWKKSSLTSSRLATGEEEGGGSSSKRFLRSCSAS
CVPHGAKDTEWRSVTLPRDLQSTGRQFDSSTFGGHKSEKPALPRKRAGENRSQVTRGTV
TPPPRLVKKNEEADEVFKDIMESSPGSSPPNLTTPKPLRRQVTVAPASGLPHKEEAGKGS
ALGTPAAAEVPTPTSKAGSGAPGGTSKGPAAESRVRRHKHSSESPGRDKGKLSRLKPAPP
PPPAASAGKAGGKPSQSPSQEAAGEAVLGAKTKATSLVDAVNDAKPSQPGEGLKPKPVL
PATPKPQSAKPSGTPISPAPVPSLPSASSALAGDQPSSTAFIPLISTRVSLRKRTRQPE
RIASGAIKGVVLDSTEALCLAISRNSEQMASHSAVLEAGKNLYTFCVSYVDSIQQMRNK
FAFREAINKLENNLRELQICPATAGSGPAATQDFSKLLSSVKEISDIVQR

>sekvence2

MGCGCSSHPEDDWMENIDVCENCHYPIVPLDGKGTLLIRNGSEVRDPLVTYEGSNPPASP
LQDNLVIALHSYEP SHDGLGFEKGEQLRILEQSGEWWKAQSLTTGQEGFIPFNFAKAN
SLEPEPWFFKNLSRKDAERQLLAPGNTHGSFLIRESESTAGSFSLSVRDFDQNGQEVVKK
YKIRNLNNGGFYISPRITFPGLHELVRHYTNASDGLCTRLSRPCQTQKPKQPWWEDEWEV
PRETLKLVRLGAGQFGEVWMGYNGHTKVAVKSLKQGSMSPD AFLAEANLMKQLQHQR
VRLYAVVTQEPYIITEYMEGSLVDFLKTSPGIKLTINKLLDMAAQIAEGMAFIEERNY
IHRDLRAANILVSDTLCKIADFLARLIEDNEYTAREGAKFPIKWTAPEAINYGTFTIK
SDVWSFGILLTEIVTHGRIPYPGMTNPEVIQNLERGYRMVRPDNCPEELYQLMRLCWER
PEDRPTFDYLRVLEDFFTATEGQYQPQP

>sekvence3

MPSETPQAEVGP TGCPHRSGPHSAKGSLEKGSPEDEKAKEPLWIRPDAPSRTCWQLGRPA
SESPHHHTAPAKSPKILPDI LKKGIDTPMVRINKIGKKFGLKCELLAKCEFFNAGGSVKD
RISLRMIEDAERDGTLPKGD TIEPTSGNTGIGLALAAAVRGYRCIIVMPEKMSSEKVDV
LRALGAEIVRTPTNARFDS PESHVGVAVRLKNEIPNSHILDQYRNASNPLAHYDTTAD E I
LQQCDGKLDMLVASVGTGGTITGIARKLKEKCPGCRIGVDPEGSILAEPEELNQTEQTT
YEVEGIGYDFIPTVLDRTVVDKWFKSNDEEAFTFARMLIAQEGLLCGGSAGSTVAVAVKA
AQELQEGQRCVILPDSVRNYMTKFLSDRWMLQKGFLEEDLTEKKPWWWHLRVQELGLS
APLTVLPTITCGHTIEILREKGFQDQAPVVD EAGVILGMVTLGNMLSSLLAGKVQPSDQVG
KVIYKQFKQIRLTDTLGRLSHILEMDHFALVVHEQIQYHSTGKSSQRQMVFGVVT AIDL
NFVAAQERDQK

zjistěte následující informace:

1. o jaký protein se jedná (vhodné nástroje znáte)
2. z jakých domén se skládá (vhodné nástroje jsou např. InterPro, <http://www.ebi.ac.uk/interpro/>, nebo PFAM, <http://pfam.sanger.ac.uk/>, případně přímo informace v záznamu sekvence), pro každou sekvenci najdete pět dalších proteinů v lidském genomu, které ji obsahují.
3. jestli je protein po odštěpení prvních 9 aminokyselin na N-koncovém glycinu myristylován (vhodný nástroj je např. <http://web.expasy.org/myristoylator/>)
4. jestli je protein N-glykosylován (vhodný nástroj <http://www.cbs.dtu.dk/services/NetNGlyc/>)
5. jestli je protein O-glykosylován (vhodný nástroj <http://www.cbs.dtu.dk/services/NetOGlyc/>)
6. jestli je protein fosforylován (<http://www.cbs.dtu.dk/services/NetPhos/>, <http://www.cbs.dtu.dk/services/NetPhosK/>)