

**Online Supplementary Materials C.** The learning dataset contains 70 viral proteins classified into 4 subcellular locations with respect to their host cells according to the experimental annotations. Both the accession numbers and sequences are given. None of the proteins has  $\geq 25\%$  sequence identity to any other in the same subset (subcellular location). See the text of the paper for further explanation.

---

(1) 14 Cytoplasm proteins

>P26669

MAVCIIDHDNIRGVIYFEPVHGKDKVLGSGVIGLKSGTYSLLIHRVYGDISQGCDSIGSPEI  
FIGNIFVNRYGVAYVYLDTDVNISTIIGKALSISKNDQRLACGVIGISYINEKIIHFLTI  
NENGV

>P34204

MKMEVIFKTDLRSSSQVVFHAGSLYNWFSVEIINSGRIVTTAIKTLTSTVKYDIVKSAHI  
YAGQGYTEHQAEWNMILHVLFEETESSASSESIEHEKNDNETNECTSSFETLFEQEPS  
SEEPKDSKLYMLAQKTVQHIEQYKAPDFNKVIRAHNFIQTIHGTPLKEEEEKEVRLMVI  
KLLKKNKLLSHLHLMF

>P21277

MSSVLKAFFERFTIEQELQDRGEEGSIIPPETLKSARKVVFVINTPNPTTRYQMLNFCLRIIC  
SQNRRASHRVGALIALFSLPSAGMQNHIRLADRSPQAQIERCEIDGFEPGTYRLIPNARA  
NLTANEIAAYALLADDLPPTINNGTPYVHADVELQPCDEIEQFLDRCYSVLIQAWVMVCK  
CMTAYDQPAGSADRRFAKYQQGRLEARYMLQPEAQRLIQTAIRKSLVVRQYLTFFELQLA  
RRQGLLSNRYAMVGDIGKYIENSGLTAFFLTLKYALGTKWSPLSLAAFTGELTKLRSLM  
MLYRDIGEQARYLALLEAPQIMDFAPGGYPLIFSYAMGVGSVLDVQMRNYTYARPFLNGY  
YFQIGVETARRQQGTVDNRVADDLGLTPEQRNEVTQLVDRLARGRGAGIPGGPVNPFVPP  
VQQQQPAAVYADIPALEESDDGDEDGGAGFQNGVQVPAVRQGGQTFRAQPLQDPIQAQ  
LFMPLYPQVSNIPNNRIIRSIASGGWKTKIYYDTTRMVILNKMQGANTETLSQTIPIKTH  
SCKWATGMSK

>P26589

MSLDRLKLNVDVSNKDSLLSNCKYSVTRSTGDVTSVSGHAMQKALARTLGMFLLTAFNRCE  
EVAEIGLQYAMSLGRDDSIKILREAGYNVKCVDTQLKDFTIKLGKEYKIQVLDIVGID  
AANLADLEIQARGVAKELKTGARLPDNRHDAPDCGVIVLCIAALVVSKLAAGDRGGLD  
AVERRALNVLKAEKARYPNMEVKQIAESFYDLFERKPYIDVFITFGLAQSSVRGGSKVE  
GLFSGLFMNAYGAGQVMLRWGLLAKSVKNIMLGHASVQAEMEQQVEVVEYEAQKQGGEAGF  
YHIRNPKASLLSLTNCNFTSVVLGNAAGLGIIGSYKGAPRNRELFDAAKDYAERLKDN  
NVINYSALNLTAEERELISQQLNIVDDTPDDDI

>O57285

MDQDALISKEDSEVEREASGGRESLSDVIGFLDAVLSSEPTDIGGDRSWLHNTINTLQRP  
GSTHRVKGEGEVSTSTQDNRSGEESRVSGGTSEPEAEAHARNVDKQNIHWATGRGAS  
TDSVPQDLGNRDSGILEDPPNEGGYPRSGAEDENREMAANPDKRGEDQAEGLPEEIRRS  
APLPDEREGRADNNGRGVEPGSPHSARVTGVLVIPSELEEAVLQRNKRRPANSRSRSLT  
PVVVPSTRSPPPDHDNSTRSPPRKPPTTQDEHTNPRNTPAVRIKDRRPPTGTRSAPDRPT  
DGYPTHSPETDATKKGIEENTSSMKEMATLLTSLGVIQSAQEFESSRDASYVFAKRALK  
SANYAEMAFNVCGLILSAEKSFANRVDENKQLLKQIQESVESFRDIYKRFSEYQKEQNSL  
LMSNLSTLHIITDRGGKTDNPDSPTRSPSVFAKTENKTKATRFDPSPMETMGDMRYKPD  
LREDEFREEIRNPVYQERDTEPRASNARLLPSREKPTIHSCLKLVISSPLSRAEKAAAYV  
KSLSKCKTDQEVKAVMELVEEDIESLTN

>O57286

MAGLLSTFDTFSSRRSESIKNSGGGAVIPGQRSTVSVFVLGPSVTDDADKLSIATTTFLAH  
 SLDTDKQHSQRGGFLVSLLAMAYSSPELYLTTNGVNADVYVIYNIKDKPKRTKTDFIV  
 KTRDMEYERTTEWLFGPMVNSPLFQQRDAADPDTLQIYGPAACLGAIIVQVWIVLVK  
 AITSSAGLRKGFNRLEAFRQDGTVKGALVFTGETVEGIGSVMRSQQSLVSLMVETLVTM  
 NTARSDLTLEKNIQIVGNYIRDAGLASFMNTIKYGVETKMAALTLSNLRPDINKLRSLI  
 DTYLSKGPRAFFICILKDPVHGEFAPGNYPALWSYAMGVAVVQNKAMQQYVTGRTYLDME  
 MFLLGQAVAKDAESKISSALEDELGVTDTAKERLRHHLANLSGGDGAYHKPTGGGAI EVA  
 LDNADIDLEPEAHTDQDARGWGGDSGDRWARSMGSGHFITLHGAERLEEEETNDEDVSDIE  
 RRIARRLAERRQEDATTHEDEGRNNGVDHDEEDDAAAAAGMGGI

>P03557

MENIEKLLMQEKILMLELDLVRAKISLARANGSSQQGELSLHRETPEKEVAVHSALVTFT  
 PTQVKAIPEQTAPGKESTNPLMASILPKDMNPVQTGTRLAVPSDFLRPHQGIPIPOKSEL  
 SSTVPLRAESGIQHPHINYVYNGPHAGIYDDWGCTKAATNGVPGVAHKKFATITTEAR  
 AAADAYTTRQQTDRLNFIKGEAQLPKPSFAEALTSPPKQKAHWLTLGTTKPSDDPAPKE  
 ISFAPEITMDDFLYLYDLVRKFDGEGDDTMFTTDNEKISLNFNRKNANPQMVREAYAAGL  
 IKTIYPSNNLQEKYLPKVKDAVKFRFTNCIKNTEKIDIFLKIRSTIPVWTIQGLLHKPR  
 QVIEIGVSKKVIPTESKAMESRIQIEDLTELVKTGEQFIQSLRLNDKKKIFVNMVEHD  
 TLVYSKNIKETDSEDQRAIETFQQRVISGNLLGFHCPAICHFIMKTVEKEGGAYKCHHCD  
 KGKAIQDASADEGTTDKSGPPPTRSI VEKEDVPNTSSKQVD

>Q00946

MSILEKITSSPSECAEHLTNKDSCLSKKIQKELTSFLEKKETLGCDSESCVITHPAVKAY  
 AQQKGLDLSKELETRFKAPGPRNNTGLLTNFNIDETLQRWAIKYTKFFNCPFSMDDFERV  
 HYKFNQVDMVKVYKGEELQYVEGKVVKRPCNTFGCVLNTDFSTGTGKHVVAIFVDMRGDC  
 WSIEYFNSAGNSPPGPVIRWMERVKQQLLKIHTVKTAVTNIRHORSQTECGPYSLFYI  
 RARLDNVSIAHFISARITDEDMYKFRTHLFRIA

>P01126

MDFLTNLRFTEGISEPYIEIFEQPRQRGTRFRYKCEGRSAGSIPGEHSTDNNKTFPSIQI  
 LNYFGVKVIRTTLVTKNEPYKPHPHDLVGGKCRDGYEAEFGPERQVLSFQNLGIQCVKK  
 KDLKESISLRISKINPNFNPVEEQLHNIDEYDLNVVRLCFQAFPLPDEHGNYTLALPPLIS  
 NPIYDNRAPNTAELRICRVNKNCGSVKGGDEIFLLCDKVQKDDIEVRFVLGNWEAKGSFS  
 QADVHRQVAIVFRTPPFLGDITEPITVKMQLRRPSDQAVSEPVDFRYLPDEEDPSGNKAK  
 RQRSTLAWQKPIQDCGSAVTERPKAAPIPTVNPEGKLLKKEPNMFSPTLMLPGLGTLSSSQ  
 MYPACSQMPTQPAQLGPGKQDTLHSCWQQLYSPSPSASSLLSLHSHSSFTA EVPQPGAQG  
 SSSLPAYNPLNWPDEKNSSFYRNFGNTHGMGAALVSAAGMQSVSSSSIVQGTHQASATTA  
 SIMTMPRTPGEVPPFLRQVVG

>P06447

MDGQESSQNPSDILYPECHLNSPIVRGKIAQLHVLLDVNQPYRLKDDSIINITKHKIRNG  
 GLSPRQIKIRSLGKALQRTIKDLDRYTFEPYPTYSQELLRLDIPEICDKIRSVFAVSDRL  
 TRELSSGFQDLWLNIFKQLGNIEGREGYDPLQDIGTIPEITDKYSRNRWYRPFLTWFSIK  
 YDMRWMQKTRPGGPLDTSNSHNLECKSYTLVTYGDLVMILNKLTTLTGTYILTPELVLMYC  
 DVVEGRWNMSAAGHLDKKSIGITSKGEELWELVDSLFSLGEIYNVIALLEPLSLALIQ  
 LNDPVIPLRGAFMRHVLTTELQTVLTSRDVYTDAAEDTIVESLLAIFHGTSIDEKAEIFSF  
 FRTFGHPSLEAVTAADKVRAHMYAQKAIKLTLYECHAVFCTIIINGYRERHGGQWPPCD  
 FPDHVCLELRNAQGSNTAISYECVDNYTSFIFGKFRKFI EPQLDEDLTIYMKDKALSPR  
 KEAWDSVYPDSNLYYKAPESSEETRRLIEVFINDENFNPEEIIINYVESGDWLKDEEFNISY  
 SLKEKEIKQEGRLFAMTYKMRAVQVLAETLLAKGIGELFRENGMVKGEIDLLKRLTTLS  
 VSGVPRTDVYNNKSSSEKRNEGMEKNKNSGGYWDEKKRSRHEFKATDSSTDGYETLSCFL

TTDLKKYCLNWRFESTALFGQRCNEIFGFKTFNWMHPVLERCTIYVGDYCPVADRMHR  
 QLQDHADSGIFIHNPRGGIEGYCQKLWTLISISAIHLAAVRVGVRSAMVQGDNQAIAVT  
 SRVVAQTYKQKKNHVEEITKYFGALRHVMFVGVGHELKLNETIISSKMFVYSKRIYYDG  
 KILPQCLKALTKCVFVWSETLVDENRSACSNISTSIAKAIENGYSPILGYCIALYKTCQQV  
 CISLGMTINPTISPTVRDQYFKGNWLRCAVLI PANVGGFNYMSTSRCFVRNIGDPAVAA  
 LADLKR FIRADLLDKQVLYRVMNQEPGDSSFLDWASDPYSCNLPHSQSITTI IKNITARS  
 VLQESP NPLLSGLFTETSGEEDLN LASFLMDR KVILPRVAHEILGN SLTG VREAIAGMLD  
 TTKSLVRASVRKGGLSYGILRRLVNYDLLQYETL TRTLRKPVKDNIEY EYMC SVELAVGL  
 RQKMWIHLTYGRPIHGLETDPDLELLRGIFIEGSEVCKLCRSEGADPIYTW FYLPDNIDL  
 DTLTNGCPAIRIPYFGSATDERSEAQLGYVRNLSKPAKAAIRIAMVYTWAYGTDEISWME  
 AALIAQTRANLSLENLKL LTPVSTSTNLSHRLKDTATQMKFSSATLVRASRFITISNDNM  
 ALKEAGESKDTNLVYQQIMLTGLSLFEFNMRYKKGSLGKPLIHLHLNNGCCIMESPQEA  
 NIPPRSTLDLEITQENNKLIYDPDPLKDVLELFSKVRDVVHTVDMTYWSDDEVIRATSI  
 CTAMTIADTMSQLDRDNLKEMIALVNDDVNSLITEFMVIDVPLFCSTFGGILVNQFAYS  
 LYGLNIRGREEI WGHVVRILKDTSHAVLKVLSNALSHPKIFKRFWNAGVVEPVYGP NLSN  
 QDKILLALSVC EYSVDLFMHDWQGGVPLEIFICDNPDVADMRRSSFLARHLAYLCSLAE  
 ISRDGPRLES MN SLERLESLSYSLELTFLLDDPVLRYSQLTGLVIKVF PSTLTYIRKSSIK  
 VLRTRGIGVPEVLEDWDPEADNALLDGI AAEIQQNIPLGHQTRAPFWGLRVSKSQVLR LR  
 GYKEITRGEIGRSGVGLTLPFDGRYLSHQRLRFGINSTSCLKALELTYLLSPLVDKDKDR  
 LYLGE GAGAMLS CYDATLGPCINYYNSGVYSCDVNGQRELNIYPAEVALVGK KLNNTSL  
 GQRVKVLFNGNPGSTWIGNDECEALIWNELQNSSIGLVHCDMEGGDHKDDQVVLHEHYSV  
 IRIAYLVGDRDVVLI SKIAPRLGTDWTRQLSLYLRYWDEVNLI VLKTSNPASTEMYLLSR  
 HPKSDI IEDSKTVLASLLPLSKEDSIKIEKWILIEKAKAHEWVTRELREGSSSSGMLRPY  
 HQALQTFGFEPNLYKLSRDFLSTMNIADTHNCMIAFN RVLKDTIFEWARITESDKRLKLT  
 GKYDLYPVRDSGK LKTI SRRLVLSWISLSMSTRLV TGSFPDQKFEARLQLGIVSLSSREI  
 RNLRVITKTLDRFEDI IHSITYRFLTKEIKILMKILGAVKMFGARQNEYTTVIDDGS LG  
 DIEPYDSS

>P68460

MKNVLIIFGKPYCSICENVSDAVEELKSEYDILHVDILSFFLKDGDSSMLGDVKRGT LIG  
 NFAAHL SNYIVSIFKYNPQTKQMAFVDINKSLDFTKTKDKSLVNLEILKSEIEKATYGVWP  
 PVTE

>P08353

MARRRRHRGPRRPRPPGPTGAVPTAQSQVTSTPNSEPAVRSAPAAAPPPPPASGPPPPSCS  
 LLLRQWLHVPE SASDDDDDDWPDSPPPEPAPEARPTAAAPRPRSPPPGAGPGGGANPSH  
 PPSRPFRLPPRLALRLRVTAEHLARLRLRRAGGEGAPEPPATPATPATPATPATPATPAT  
 PATPATPATPARVRFSPHVRVRHLVWASAARLARRG SWARERADRARFRRRVAEAEAVI  
 GPCLGPEARARALARGAGPANSV

>P15628

MEQAIAKQLRNQQKQLDQQIEQQKQTLNSLISRREEVTKGIQTLELLSGTTLVDPDQOIHEA  
 LQPAVSESTKGNQEIKTWKQIFDEEHELTREEKIDATVGQKKKAYVIFDGPWKGIYQDWH  
 IVKSKVNAQPYRYKGYNSLEEAKLAHKQAYAEVTKADEVKVEKTMKSF AKNNVAEKIHGL  
 HKSGPKELTEAEFYRNWKMIT EWTEESANLGFYPDCSKQVKAVFFIGADPHLLSSFYQSG  
 LISYIYLQEDEGQKGAISKATSQ LPKELRRTCQQYQLSFAKTREFYLAIQSTYPVFDEEK  
 MLVPAKHLVKLGISASSYPENKIVKTNFNFSLFINSVDKLYNYIRQYGTTIKGFVKMLKT  
 QLCLAVCLIRDQAEESKIMVMEFELDISTLTGIFSNLPKELK KATCEKMHRYKSHLCE  
 CEINFPELSETMNVSNDEKRSTKSVSSDEINLSAENDGYQHS

>P03170

MSWALEMADTFLDTMRVGPRTYADVRDEINKRGREDREAARTAVHDPERPLLRSPGLLPE  
IAPNASLGVARRRTGGTVTDSRNPVTR

(2) 8 Extracell proteins

>Q98157

MDTKGILLVAVLTALLCLQSGDTLGASWHRPDKCCLGYQKRPLPQVLLSSWYPTSQLCSK  
PGVIFLTRGRQVCADKSKDWVKKLMQQLPVTAR

>P16948

MLIFSVMYLGLLL LAGASALPNGLSPRNNAFCAGFGLSCKWECWCTAHGTGNELRYATAA  
GCGDHLKSSYYDARAGHCLFSDDLNRNQFYSHCSSLNNMNSCRSLSKRTIQDSATDTVDLG  
AELHRDDPPPTASDIGRGRPRPVMCQCVDTTNGGVRLDAVTRAACSIDSFIDGYYTEK  
DGFCRAKYSWDLFTSGQFYQACLRYSHAGTNCQPDPQYE

>P15879

MKGLRKSILCLVLSAGVIAPVTSGMIQSPQKCYAYSINQKAYSNTYQEFTNIDQAKAWGN  
AQYKKGYSKSEKEAIVSYTKSASEINGKLRQNKGVINGFP SNLIKQVELLDKSFNKMKT  
PENIMLFRGDDPAYLGTEFQNTLLNSNGTINKTAFEKAKAKFLNKDRLEYGYISTSLMNV  
SQFAGRPIITKFKVAKGSKAGYIDPISAFAGQLEMLLPRHSTYHIDDMRLSSDGKQIIIT  
ATMMGTAINP

>P60172

MEGLSLLQLPRDKFRKSSFFVWVILFQKAFSMP LGVVTNSTLEVTEIDQLVCKDHLAST  
DQLKSVGLNLEGSVSTDIP SATKRWGF RSGVPPQVVS YEAGEWAENCYNLEIKKPDGSE  
CLPPPPDGVRGFPRCRYVHKAQGTGPCPGDYAFHKDGAFFLYDRLASTVIYRGNFAEGV  
IAFLILAKPKETF LQSPPIREAAANYTENTSSYYATSYLEYEIENFGAQHSTTLFKINNNT  
FVLLDRPHTPQFLFQLNDTIQLHQQLSNTTGKLIWTL DANINADIGEWAFWENKKISPNN  
YVEKSCLSKLYRSTRQKTMMRHRRELQREESPTGPPGSIRTWFQRIPLGWFHCTYQKGKQ  
HCRLRIRQKVEE

>O10367

MGVKSALFIMAVFAAANVQYVLAACAETGAVCVHSDECCSGACSPVFNVC

>Q90121

MQIINVVYSFLFAAAMPLPVVHSLGINCRGSSQCGLSGGNLMVRIRDQACGNQGQTWCPGE  
RRAKVCGTGNSISAYVQSTNNCISGTEACRHLTNLVNHGCRVCGSDPLYAGNDVSRGQLT  
VNYVNSC

>O73559

MKSVLYSYILFLSCIIINGRDIAPHAPSNGKCKDNEYNRHNLCCLS CPPGTYASRLCDSK  
TNTNTQCTPCGSGTFTSRNNHLPACLSCNGRCD SNQVETRSCNTTTHNRICECAPGYCYLL  
KGSSGCKACVSQTKCGIGYGVSGHTSTGDVVCSPCGLGTYSHTVSSADKCEPVPSNTFNY  
IDVEINLYPVNDTSTRTTTTGLSESI STSELITMNHKDCDPVFRDGYF SVLNKVATSG  
FFTGENRYQNISKVCTLNFEIKC NNKDSSSKQLTKTKNDDGIMPHSETVTLVGDCLSSVD  
IYILYSNTNTQDYETDTISYHVG NVLDVDSHMPGSCDIHKLITNSKPTRF

>P20494

MSMKYLMLLFAAMIIRSFADSGNAIETTLP EITNATTDIPAIRLCGPEGDGYCLHGDCIH  
ARDIDGMYCRC SHGYTGIRCQHVVLDVYQRSENPNNTTTSYIPSPGIMLVLVGIIIIITCC  
LLSVYRFTRRTNKLPLQDMVVP

(3) 25 Nucleus proteins

>O55653

MTGSTIAPTTDYRNTTATGLKSALNLPQVHAFVNDWASLGMWWFSIALMFVCLIIIMWLIC  
CLKRRRARPPPIYRPIIVLNPHNEKIHRLDGLKPCSLLLQY

>P24109

MLDMETPLKEPESSLGSCNEPSSRTSGQDATTQELAKLGEEILSQLYQPLEECDNSCYCK  
RCCYHCQLCFLKKGLGICYDRKGRRRRTPKKAKAHSSSSASDKSISTRTRNSQPAKKQKKT  
LEATVETDPGLGR

>P26668

MATRPSSLVDSLEDEEDPQTLRRERSGSPRPRKI PRNALTQPV DQLLKD L RKNPSMISDP  
DQRTGREQLSNDELIKKLVTELAENSMIEAEV RGT LGDISARIEAGFESLSALQVETIQ  
TAQRCDHSDSIRILGENIKILDRSMKTMETMKLMMEKVDLLYASTAVGTSAPMLPSHPA  
PPRIYPQLPSAPTAEWDII

>P34051

MFVTRLEHTTQFLPQACKLEDEVLVLYAIYLN GFDYTLPRFVKREVQVNADGFVRFDYN I  
KMFDFARFTDMTQTTPEDIDDYINLTRINSLSDHDLKMLQLVCRDRWYKGDVARLRRI LQ  
QKDVDDLKVFVCNVMWERAYEDHYTLGQQLSIRITTKLIQSGLDFKHQPDTTAPVSVRGW  
EDATFEKYLQSITSISEVIKRHVFSKKYICLEVAASYWSDAVESLQQENFRIILNSKTPY  
VLLIEMDDDKNSMVYLRKLAHLLENKIVNLLFVTDVEFYFKNGNFMFYLYNSLKFYYYCL  
KNKFAFENV DKEIFFLLYTI VALEWFNGGHLNSFTLEKSALYNPLELSTRRLNSIKRAAQ  
HNRVINCDSEINMDYIRGKRVRTGAHYGKRVVNF DLN QSL

>P04296

METKPKTATTIKVPPGPLGYVYARACPSEGI ELLALLSARSGSDVAVAPLVVGLTVESG  
FEANVAVVVGSRRTGLGGTAVSLKLT PSHYSSSVYVFHGGRHLDPSTQAPNLTRLCERAR  
RHF GFSDYTPRPGDLKHETTGEALCERLGLDPDRALLYL VVTEGFKEAVCINNTFLHLGG  
SDKVTIGGAEVHRIPVYPLQLFMPDFSRVIAEPFNANHRSIGEKFTYPLPFFNRPLNRL L  
FEAVVGPAVALRCRNVDAVARAAHLAFDENHEGAALPADITFTA FEASQGKTPRGGRD  
GGGKGAAGGFEQRLASVMAGDAALALESI VSMVDFDEPPTDISAWPLFEGQDTAAARANA  
VGAYLARAAGLVGAMVSTNSALHLTEVDDAGPADPKDHSKPSFYRFFLVPGTHVAANPQ  
VDREGHVVPGFEGRPTAPLVGGTQEFAGEH LAMLCGFSPALLAKMLFYLERCDGAVIVGR  
QEMDVFRYVADS NQTDVPCNLCTFDTRHACVHTTLMRLRARHPKFASAARGAIGVFGTMN  
SMYSDCDVLGN YAAF SALKRADGSETARTIMQETYRAATERVMAELETLYVDQAVPTAM  
GRLETIITNREALHTVVNNVRQVVDREVEQLMRNLVEGRNFKFRDGLGEANHAMSLTLDP  
YACGPCPLLQLLGRSNLAVYQDLALSQCHGVFAGQSVEGRNFRNQFPVLRRRVMDMFN  
NGFLSAKTLTVALSEGA AICAPSLTAGQTAPAESSFEGDVARVTLGFPKELRVKSRV LFA  
GASANASEAAKARVASLQSAYQKPKRVDILLGPLGFLKQFHAAIFPNGKPPG SNQPNP  
QWFWTALQRNQLPARLLSREDIETIAFIKKFSLDYGAINFINLAPNNVSELAMYYMANQI  
LRYCDHSTYFINTLTAIIAGSRRPPSVQAAAAWSAQGGAGLEAGARALMDA VDAH PGAWT  
SMFASCNLLRPVMAARPMVVLGLSISKYYGMAGNDRVFQAGNWASLMGGKNACPLLI FDR  
TRKFVLACPRAGFVCAASSLGGGAHESLCEQLRGI ISEGGA AVASSV FVATVKS LGPRT  
QQLQIEDWLALLEDEYLSEEMMELTARAL ERNGEWSTDAALEVAHEAEALVSQ LGNAGE  
VFNF GDFGCEDDNATPFGGPGAPGPAFAGRKRAFHGDDPFGEGPPDKKGDLTLDML

>P03204

MESFEGQGDSRQSPDNERGDNVQTTGEHDQDPGPGPPSSGASERLVPEESYSRDQQPWGQ  
SRGDENRGWMQRI RRRRRRRAALSGHLLDTE DNVPWLP PHDITPYTARNIRDAACRAVK  
QSHLQALSNIILDSGLDTQHILCFVMAARQLQDIRRGPLVAEGGVGWRHWLLTSPSQSW  
PMGYRTATLRTLTPV PNRV GADSI MLTATFGCQNAARTLNTFSATVWTPPHAGPREQERY  
AREAEVRF LR GKWQRRYRIYDLIELCGSLHHIWQNL LQTEENLLDFVRFMGVMSSC NNP  
AVNYWFHKTI GNFKPYYPWNAPPNENPYHARRGIKEHVIQNAFRKAQIQGLSMLATGGEP  
RGDATSETSSDEDTGRQGS DVELESSDDEL PYIDPNMEPVQQRPMFVSRVPAKKPRKLP  
WPTPKTHPVKRTNVKTSDRSDKAEAQSTPERPGPSEQSSVTVEPAHPTPVEMPMVILHQ P  
PPVPKPVKPTPPPSRRRRGACVVYDDD VIEVIDVETTEDSSSVSQPNKPHRKHQDGFQ

RSGRRQKRAAPPTVSPSDTGPPAVGPPAAGPPAAGPPAAGPPAAGPPAAGPPAAGPRILA  
PLSAGPPAAGPHIVTPPSARPRIMAPPVVRMFMRRERQLPQSTGRKPQCFWEMRAGREITQ  
MQQEPESSHLQSATQPTTPRPSWAPSVCALSVMDAGKAQPIESSHLSSMSPTQPI SHEEQP  
RYEDPDAPLDLSLHPDVAAQPAPQAPYQGYQEPPAPQAPYQGYQEPPPPQAPYQGYQEPP  
AHGLQSSSYPGYAGPWTPRSQHPCYRHPWAPWSQDPVHGHTQGPWDPRAPHLPPQWDGSA  
GHGQDQVSQFPHLQSETGPPRLQLSLVPLVSSSAPSWSPPQPRAPIRPIPTRFPPPPMPL  
QDSMAVGCDSSTACPSMPFASDYSQGAFTPLDINATTPKRPRVEESSHGPARCSQATAE  
AQEILSDNSEISVFPKDAKQTDYDASTESELD

>O09802

MDPALRSYHQRLRLYTPVAMGINLAASSQPLDPEGPIAVTPRPPPIRPPSSGKAPHPEAPRR  
SPNWATAGEVDVGDDELIAISDERGPPRHDRPPLATSTAPSPHPRPPGYTAVVSPMALQAV  
DAPSLFVAWLAARWLRGASGLGAVLCGIAWYVTSIARGA

>P03206

MMDPNSTSEDVKFTPDYQVFPVQAFDQATRVIYQDLGGPSQAPLPCVLWVLPPEPLPQGG  
LTAYHVSTAPTGSWFSAQPPAPENAYQAYAAPQLFPVSDITQNNQTNQAGGEAPQPGDNS  
TVQTAADVVFACPGANQGQQLADIGVPQAPVAAPARRTRKPPQPESSLEECDSELEIKRY  
KNRVASRKCRAKFKQLLQHYREVAAAKSSSENDRLRLLLKQMCPSLDVDSIIPRTPDVLHE  
DLLNF

>P17473

MASQRSDFAPDLYDFIESNDFGEDPLIRAASAAEEGFTQPAAPDLLYGSQNMFGVDDAPL  
STPVVVI PPPSPAPEPRGGKAKRSPSAAGSGGPPTPAAAAQPASPAPSPAPGLAAMLKMV  
HSSVAPGNRRATGSSSPGGGDAADPVALDSDTETCPGSPQPEFPSSASPGGGSPAPRVR  
SISISSSSSSSSSMDEDDQADGAGASSSSSSSSSDSDSDEGGEEETPRPRHSQNAAKTPS  
AAGSPGPSSGGDRPAAGAATPKSCRSGAASPGAPAPAPASAPAPSRPGGGLPPGARILE  
YLEGVREANLAKTLERPEPPAGMASPPGRSPHRLPKDQRPKSALAGASKRKRANPRPRPQ  
TQTQAPAEAEAPQTAVWDLDMNSSQATGAAAAAASAPAAAASCAPGVYQREPLLTSPGDPW  
PGSDPPPMGRVRYGGTGDSRDGLWDDPEIVLAASRYAEAQAPVPVFPVPEMGDSTKQYNAL  
VRMVFE SREAMSWLQNSKLSGQDQNLAQFCQKFIHAPRGHGSFITG SVANPLPHIGDAMA  
AGNALWALPHAAASVAMSRRYDRTOQSFILOSLRRAYADMAYPRDEAGRPDSLAAVAGCP  
AQAAAAAASQQQPEAPAPSVRVREAYTRVCAALGPRRKAAAAAAPGTRAPRPSAFRLRE  
LGDACVLACQAVFEALLRLRGASAVPGLDPSEIPSPACPPEALCSNPAGLETAALSLEYE  
LRDLVERARLLGDSDPHRLGSDRLRAVRAVLVARTVAPLVRYNAEGARARASAWTVT  
QAVFSIPLVGGMLGEAVSLLAPPTRSQQPSSSSPGGEPFSGSAAAEGLQTLPLWPTV  
PGKQSATVPSSHSQSPQHSQSGGAGATTATCCRATQTNARSRGQHQHPQKARSPQAAAS  
PAHLSQEAMPGSSSDDRAIHGRPRGKSGKRRSEPLEPAAQAGASASFSSSARGYDPSGPV  
DSPPAPKRRVATPGHQAPRALGMPAEGPDRRGFRVPRGDCHTPRPSDAACAAYCPPE  
LVAELIDNQLFPEAWRPALTFDPQALATIAARCSGPPARDGARLGELAASGPLRRRAAWM  
HQIPDPEDVKVVVLYSPLQDEDLGGLPASRPGGSRREPLWSDLKGGLSALLAALGNRIL  
TKRSHAWAGNWTGAPDVSAALNAQGVLLLSTGDLAFTGCVEYLCLRLGSAARRKLLVLDVAVS  
TEDWPQDGPASISQYHIYMRAALTPRVACAVRWPGERHLSRAVLTSSSTLFGPGLFARAEAA  
FARLYPDSAPLRLCRSSNVAYTVDTRAGERTRVPLAPREYRQRVLPDYDGCKDMRAQAE  
LGFHDPDFEEGAAQSHRAANRWGLGAWLRPVYLACGRRGAGAVEPSELLIPELLSEFCRV  
ALLEPDAEAEPLVLPITEAPRRRAPRVDWEPGFGSRSTSVLHMGATELCLPEPDELEID  
GAGDVELVVEHPGPSGVAQALRRAPIKIEVVSDEDEDGGDWCNPYLS

>P03246

MEAWECLEDFSAVRNLEQSSNSTSWFWRFLWGSSQAKLVCRKEDYKWEFEELLKSCGE  
LFDLNLGHQALFQEKVIKTLDFSTPGRAAAAVAFLSFIKDKWSEETHLSGGYLLDFLAM

HLWRAVVRHKNRLLLLSSVRPAIIPTEEQQQQEEARRRRQEQSPWNPRAGLDPRE

>P03191

METTQTLRFKTKALAVLSKCYDHAQTHLKGGLVQVNLSSVNYGGPRLAAVANAGTAGLIS  
FEVSPDAVAEWQNHQSPEEAPAAVSFRNLAYGRTCVLGKELFGSAVEQASLQFYKRPQGG  
SRPEFVKLTMEYDDKVKSHHTCALMPYMPASDRLRNEQMIGQVLLMPKTASSLQKWAR  
QQGSGGVKVTLNPDLYVTTYTSGEACLTLDYKPLSVGPYEAFTGPPVAKAQDVGAVEAHVV  
CSVAADSLAAALSICRIPAVSVPILRFYRSGIIAVVAGLLTSAGDLPLDLSVILFNHASE  
EAAASTASEPEDKSPRVQPLGTGLQQRPRHTVSPSPSPPPPPRTPTWESPARPETPSPAI  
PSHSSNTALERPLAVQLARKRTSSEARQKQKHPKKVKQAF

>P52639

MSFHASLLREEETPRPVAGINRTDQSLKNPLLGTEVSFCLKSSSLPHHVRLGQIKARNL  
ASCDYLLFRQVVLPEVYPIGVLIRAAEAILTVIVSAWKLDHMTKTLYSSVRYALTNP  
VRAQLELHIAQYRIVGQVSYSREADIGPKRLGNMSLQFIQSLVIATIDTTSCLMTYNHFL  
AAADTAKSRCHLLIASVVQALWEQGSFLDHIINMIDIIDSINLPHDDYFTI IKSIFPYS  
QGLVMGRHNVSVSDFASVFAIPELCPQLDSSLKLLQLDPLVLLMVSSVQKSWYFPEIR  
MVDGSREQLHKMRVELETPQALLSYGHTLLSIFRAEFIKGYVSKNAKWPPVHLLPGCDKS  
IKNARELGRWSPAFDRRWQLFEKVILRIADLDMPDFNDIVSDKAI I SSRDWWFEYNA  
AAFWKYGERLERPPARSGPSRLVNALIDGRLDNI PALLEPFYRGAVEFEDRLTVLVPKE  
KELKVKGRFFSKQTLAIRIYQVVAEAAALKNEVMPYLKTHSMTMSSTALHLLNRLSHTIT  
KGDSEVINLDYSSWCNGFRPELQAPICRQLDQMFNCGYFFRTGCTLPCFTTFI IQDRFNP  
PYSLSGEPVEDGVTCVGTKTMGEGMRQKLWTILTSCWEI IALREINVTFNILGQGDNQT  
I I IHKSASQNNQLLAERALGALYKHARLAGHNLKVEECWVSDCLYEYGGKLF FRGVPVPG  
CLKQLSRVTDSTGELFPNLYSKLACTSSCLSAAMADTSPWVALATGVCLYLI ELYVELP  
PAI IQDESLTTLCLVGPSIGGLPTPATLPSVFFRGMSDPLPFQLALLQTLIKTTGVTCS  
LVNRVVKLRIAPYDWSLVTDPSTLNIAQVYRPERQIRRWIEEAIATSSHSSRIATFFQ  
QPLTEMAQLLARDLSTMMPLRPRDMSALFALSNVAYGLSI IDLFQKSSTVVSASQAVHIE  
DVALESVRYKESI IQGLLDTTEGYNMQPYLEGCTYLAQQLRRLTWGRDLVGVTMPFVAE  
QFHPHSSVGAKELYLDAI IYCPQETLRSHHLTTRGDQPLYLGSNTAVKVQRGEITGLAK  
SRAANLVKDTLVLHQWYKVRKVTDPHLNLTLMARFLLKGYTSDARPSIQGGTLTHRLPSR  
GDSRQGLTGYVNILSTWLRFSDDLHSHFSKSSDDYTIHFQHVFTYGCYADSVIRSGGVI  
STPYLLSASCKTCFEKIDSEEFVLACEPQYRGAEWLISKPVTVPEQITDAEVEFDPCVSA  
SYCLGILIGKSFLVDIRASGHDIMEQRTWANLERFSVSDMQKLPWSIVIRSLWRFLIGAR  
LLQFEKAGLIRMLYAATGPTFSFLMKVFQDSALLMDCAPLDRLSPRINFHSRGDLVAKLV  
LLPFINPGIVEIEVSGINSKYHAVSEANMDLYIAAAKSVGKPTQFVEETNDF TARGHHH  
GCYSLSWSKSRNQSQVLKMMVVRKLLKCVLYIYPTVDPVALDLCHLPALTI I LVLGGDPA  
YYERLLEMDLCGAVSSRVDIPHSLAARTHGRFAVGPDAGPGVIRLDRLESVCYAHPCLEE  
LEFNAYLDSELVDISDMCCLPLATPCKALFRPIYRSLQSFRLALMDNYSFVMDLIMIRGL  
DIRPHLEEFDELLEVGGHILGQPVLVEVVYVGVVRKRPVLRHPWSADLKRITVGGRAP  
CPSAARLRDEDCQGSLLVGLPAGLTQLLIID

>P03192

MFSCKQHLSLGGACVFCLGLLASTPFIWCFVFANLLSLEIFSPWQTHVYRLGFPTACLMVA  
LWTLVPAKHAVRAVTPAIMLNIASALIFFSLRVYSTSTWVSAPCLFLANLPLLCLWPRLA  
IEIVYICPAIHQRFFELGLLLACTIFALSVVSRALEVSAVFMSPFFIFLALGSGSLAGAR  
RNQIYTSGLERRRSIFCARGDHSVASLKETLHKCPWDLAISALTVLVVCMIVLHVHAE  
VFFGLSRYLPLFLCGAMASGLYLGHSSIIACVMATLCTLTSVVVYFLHETLGLPLGKTVL  
FISIFVYVYFSGVAALSAAMRYKLLKVFVNGPLVHLRVVYMCCFVFTFCEYLLVTFIKS

>P08028

MSDRRQNRKTPDEQRKANALI INENIEAYIAICKEVGLNGDEMLILENGIAIEKAIRICC  
 DGKYQEKREKKAREAQRADSNFNADSIGIRLVKRAGSGTNTITYHAVVELTSRSRIVQILK  
 SHWGNELNRAKIAGKRLGFSALFASNLEAIYQGRNAARRNGSAELFTLTQGAGIETRY  
 KWIMEKHIGIGVLIADAKGLINGKREGKRGVDANVKLRAGTTGSPLERAMQGIEKKAFFPG  
 PLRALARRVVKANYNDAREALNVIAEASLLLPQITNKMTMPWCMWLAARLTLKDEFTNF  
 CAYAGRRAFEVFNIAMEKIGICSFQGTIMNDDEIESIEDKAQVLMMACFGLAYEDFSLVS  
 AMVSHPLKLRNRMKIGNFRVGEKVSTVLSPLLRFTRWAAFAQRFALQANTSREGTQISNS  
 AVFAVERKITTDVQRVEELLNKVQAHEDPLQTLTKVREQISIIIGRNKSEIKEFLGSSM  
 YDLNDQEKQNPINFRSGAHPFFFEFDPDYNPIRVKRPKKPIAKRNSNISRLEEEGMDENS  
 EIGQAKMKPLDQLASTSSNIPGEN

>P20919

MAESKEARDQEMNLKEESKEEKRRNDWWKKDPQGPLESQWCRVLRQSLPEEKIPSQTCI  
 ARRHLGPGPTQHTPSRRDRWIRGQILQTEVLQERLEWRIRGVQQAAKELGEVNRGIWREL  
 YFREDQRGDFSAAWGGYQRAQERLWGEQSSPRVLRPGDSKRRRKHL

>Q02479

MKGIIPRKSSADAYVYLSPGVREHLFSLLVKYTSSQRDSPAGSSTKVAPDTPQTNFIVE  
 IQWVCHKILLEMEEQFGSLGGLVADINICLIWTLFRNYKHKHRMNNSDTGKSACAEYAQSV  
 VKHLTERMVYCTDKFFINSACSGVTVPQNLALVIASISQVCRNKCQGAWRRMGNGRRTL  
 DLGLQLVNTYNLLNACGAIDDKCAFIKLTFPYLNLETVYSPVHAASTGLHQMAISLYK  
 GQEKRKVPNATIYSNLVTQEKFALPEILLGEITDEGLLANKGPDLEKLLSEPQTILKLV  
 KLSVPSQFQVFMARWGDKLPSHLKDVCVSDSSQLPQQYENFKVWVWPQHQTWPNDETSALP  
 IASASLLSLQAENLPTSVQNSVAQETTGLPPQVPAGEMPPRGNCVPEDHGLLNSDAPDDL  
 SGKGDYDLFTSDDLPMSSGNVAIPVCEDEPLAPPCKKQRMPEPDPPLSTPGDNQWGDGAAS  
 AWIQSYMENEDAYLELILQGLYHLDEPPKLEDSIQVQDVSNSTNDPLEGPSRSQASCFED  
 TENIANPLDL

>P22379

MLVQEMLLSLYALLSKEGFRNKVPCLQETWLSTCLFTGLPAEPIPQTARQRRRRKQQLRT  
 RRAQLRELEGRILKQILDRGPDQLCQGVTNLALAEKSESS

>P06481

MTSRLSDPNSSARSMSVPLYPTASPVSVVEAYYSESEDEEAANDFLVRMGRQQSVLRRRRR  
 RTRCVGMVIACLLVAVLSGGFGALLMWLLR

>O36433

MSNMDIDGINTGTIDKTPEEITSGTSGTTRPIIRPATLAPPSNKRTRNPSPERATTSSEA  
 DVGRKTQKKQTPTEIKKSVYNMVVKLGEFYNQMMVKAGLNDDMERNLIQNAHAVERILLA  
 ATDDKKTEFQRKKNARDVKEGKEEIDHNKTGGTFYKMRDDKTIYFSPIRITFLKEEVKT  
 MYKTTMGSDGFSGLNHIMIGHSQMNDVCFORSKALKRVGLDPSLISTFAGSTLPRRSGAT  
 GVAIKGGGTLVAEAIRFIGRAMADRGLLRDIKAKTAYEKILLNLKKNCSAPQQKALVDQV  
 IGSRNPGIADIEDLTLLARSMVVVRPSVASKVVLPISIIYAKIPQLGFNVEEYSVMGYEAM  
 ALYNMATPVSILRMGDADKDSQLFFMSCFGAAYEDLRVLSALTGIEFKPRSALKCKGFH  
 VPAKEQVEGMGAALMSIKLQFWAPMTRSGGNEVGGDGGSGQISCSVPFAVERPIALSQQA  
 VRRMLSMNIEGRDADVKNLLKMMNDSMAKKTNGNAFIGKKMFQISDKNKTNPVEIPIKQ  
 TIPNFFFGRDTAEDYDDL

>Q05002

MACNRVTLAVASDSEISANGCSTIAQAVRRYSEAASNGFRACRFVSLDLQDCIVGIADDT  
 YVMGLHGNQTLFCNIMKFSDRPFMLHGWLVFNSNYLLEEFDVVFGKRGGGNVYTYTDQYL  
 CGADGKPVMSEDLWQFVDHFGENEEIIINGHTYVCAWLTKRKPLDYKRQNNLAIEEIEYV  
 HGDALHTLRNGSVLEMAKEVKTSSKVVLSDALDKLYKVFGSPVMTNGSNILEAFTKPVFI



SALVQCTCGTKSWSVGDWTGFKSSCCNVISNKLCVVPGNVKPGDAVITTTQQAGAGIKYFC  
GMTLKFVANIEGVSVWRVIALQSVDCFVASSTFVEEEHVNRMDTFCFNVRNSVTDECRLA  
MLGAEMTSNRRQVASGVIDISTGWFDVYDDIFAESKPWFVRKAEDIFGPCWSALASALK  
QLKVTTGELVRFVKSICNSAVAVVGGTIQILASVPEKFLNAFDVFTAIQTVFDCAVETC  
TIAGKAFDKVFDYVLLDNALVKLVTTKLGVRERGLNKVKYATVVVGSTEEVKSSRVERS  
TAVLTIANNYSKLFDEGYTVVIGDVAYFVSDGYFRLMASPNSVLTTAVYKPLFAFNVNM  
GTRPEKFPTTVTCENLESAVLFFVNDKITEFQLDYSIDVIDNEIIVKPNISLCVPLYVRDY  
VDKWWDFCRQYSNESWFEDDYRAFISVLDI TDAAVKAAESKAFVDTIVPPCPSILKVIDG  
GKIWNGVIKNVNSVRDWLKSLKLNLTQQGLLGTCAKRFKRWLGI LLEAYNAFLD TVVSTV  
KIGGLTFKTYAFDKPYIVIRDIVCKVENKTEAEWIELFPHNDRIKSFSTFESAYMPIADP  
THFDIEEVELLDAEFVEPGCGGILAVIDEHVFYKKDGVYYPNSNGTNILPVAFTKAAGKV  
SFSDDVEVKDIEPVYRVKLCFEFEDEKLDVCEKAI GK KIKHEGDWDSFCKTIQSALSVV  
SCYVNLPTYIYDEEGGNDLSLPMISEWPLSVQQAQQEATLPDIAEDVDQVEEVNSIF  
DIETVDVKHDVSPFEMPFEELNGLKILKQLDNNCWVNSVMLQIQLTGILDGDYAMQFFKM  
GRVAKMIERCYTAEQCIRGAMGDVGLCMYRLLKDLHTGFMVMDYKCSCTSGRLEESGAVL  
FCTPTTKAFPYGTCLNCNAPRMCTIRQLQGTIIFVQQKPEPVNPVSVFVVKPVCSSIFRGA  
VSCGHYQTNIYSQNLCDVDFGVNKIQPWTNDALNTICIKDADYNKVEISVTPIKNTVDT  
TPKEEFVVKELNAFLVHDNVAFYQGDVDTVNVGVDFDFIVNAANENLAHGGGLAKALDV  
YTKGKLQRLSKEHIGLAGKVKVGTGVMVECDLRIFNVVGPRKGKHERDLLIKAYNTINN  
EQGTPLTPILSCGIFGIKLETSLEVLLDVCNTKEVKVFVYTDTEVCKVKDFVSGLVNVQK  
VEQPKIEPKPVSVIKVAPKPYRVDGKFSYFTEDLLCVADDKPIVLF TDSMLT LDDRGLAL  
DNALSGVLSAAIKDCVDINKAIPSGNLIKFDIGSVVYMCVVPSEKDKHL DNNVQRCTRK  
LNRLMCDIVCTIPADYILPLVLSSLT CNVSFV GELKAAEAKVITIKVTEDGVNVHDVTVT  
TDKSFEQQVGV IADKDKDLSGAVPSDLNTSELLTKAIDVDWVEFYGFKDAVTFATVDHSA  
FAYESAVVNGIRVLKTS DNNCWNAVCIALQYSKPHFISQGLDAAWNKFVLGDVEIFVAF  
VYYVARLMKGDGDAEDTLTKLSKYL ANEAQVQLEHYSSCVECDAKFKNSVASINSAIVC  
ASVKRDGVQVGYCVHGIKYYSRVR SVRGRAIIVSVEQLEPCAQSRLLSGVAYTAFSGPVD  
KGHYTVYDTAKKSMYDGRFVKHDL SLLSVTSVVMVGGYVAPVNTVKPKPVINQLDEKAQ  
KFFDFGDFLIHNFVIFFTWLLSMFTLCKTAVTTGDVKIMAKAPQRTGVVLKRSLKYNLKA  
SAAVLKSKWLLAKFTKLLLLIYTLYSVLLCVRFGPFNFCSETVNGYAKSNFVKDDYCD  
GSLGCKMCLFGYQELSQFSHLDVVKHITDPLFSNMQPFIVMVL LLI FGDNYLRCFLLYF  
VAQMISTVGVFLGYKETNWF LHFIPFDVICDELLVTVIVIKVISFVRHVLFGCENPDCIA  
CSKSARLKRFPVNTIVNGVQRSFYVNANGGSKFCKKRRFFCVDCDSYGYGSTFITPEVSR  
ELGNITKTNVQPTGPAYVMIDKVEFENGFYRLYSCETFWRYNFDITESKYSCKEVFKNCN  
VLDDFIVFNNNGTNVTQVKNASVYFSQLL CRPIKLV DSELLSTLSVDFNGVLHKAYIDVL  
RNSFGKDLNANMSLAECKRALGLSISDHEFTSAISNAHRCDVLLSDLSFNNFVSSYAKPE  
EKLSAYDLACCMRAGAKVVNANVLT KDQTPIVWHAKDFNSLSAEGRKYIVKTSKAKGLTF  
LLTINENQAVTQIPATSIVAKQGAGDAGHSLTWLWLLCGLVCLIQFYLCFFMPYFMYDIV  
SSFEGYDFKYIENGQLKNFEAPLKCVRNVFENFEDWHYAKFGFTPLNKQSCP I VGVSEI  
VNTVAGIPSNVYL VGKTLIFTLQAAF GNAGVCYDIFGVTTPEKCI FT SACTRLEGLGNN  
VYCYNTALMEGSLPYSSI QANAYKYDNGNFIKLPEVIAQGFGRFTRVTIATKYCRVGEC  
VESNAGVCFGFDKWFVNDGRVANGYVCGTGLWNLVFNILSMFSSFSVAAMSGQILLNCA  
LGAF AIFCCFLVTKFRRMFGDLSVGVCTVVAVLLNNVSYIVTQNLVTMIAYAILYFFAT  
RSLRYAWIWCAAYLIAYISFAPWWLCAWYFLAML TGLLPSLLKLVSTNLFEGDKFVGT  
ESAAAGTFVIDMRSYEKLAN SISP EKLKSYAASYNRYKYYSGNANEADYRCACYAYLAKA  
MLDFSRDHNDILYTPPTVSYGSTLQAGLRKMAQPSGFVEKCVVRVCYGN TVLNGLWL GDI  
VYCPRHVIASNTTSAIDYDHEYSIMRLHNF SII SGTAF LGVVGATMHGVTLKIKVSQTNM

HTPRHSFRTLKSGEGFNILACYDGCAQGVFVNMRTNWTIRGSFINGACGSPGYNLKNGE  
VEFVYMHQIELGSGSHVGSSEFDGVMYGGFEDQPNLQVESANQMLTVNVVAFLYAAAILNGC  
TWWLKGEKLFVEHYNEWAQANGFTAMNGEDAFSILAAKTGVCVERLLHAIQVLNNGFGGK  
QILGYSSLNDEF SINEVVKQMFVNLQSGKTTSMFKSISLAFAGFFVMFWAELFVYTTTTIW  
VNPGLTTPFMILLVALSLCLTFVVKHKVFLQVFLLP SII VAAIQNCAWDYHVTKVLAEK  
FDYNVSVMQMDIQGFVNIFICL FVALLHTWRFAKERCTHWCTYLFSLI AVLTYTALYSYDY  
VSLLVMLLCAISNEWYIGAIIFRICRFGVAFLPVEYVSYFDGVKTVLLFYMLLGFVSCMY  
YGLLYWINRFCKCTLGVYDFCVSPAEFKYMVANGLNAPNGPFDALFLSFKLMGIGGPRTI  
KVSTVQSKLTDLKCCTNVVLMGILSNMNIASNSKEWAYCVEMHNKINLCDDPETAQELLLA  
LLAFFLSKHSDFGDLVDSYFENDSILQSVASSFVGMP SFVAYETARQEYENAVANGSS  
PQIIKQLKKAMNVAKAEFDRESSVQKKINRMAEQAAAAMYKEARAVNRKSKVVSAMHSL  
FGMLRRLDMSSVDTILNMARNGVPLSVIPATSAARLVVVVDPHDSFVKMMVDGFVHYAG  
VVWTLQEVKDNDGKNVHLKDVTKENQEILVWPLILTCERVVKLQNN EIMP GKMKVKATKG  
EGDGGITSEGNALYNNEGGRAFMYAVTTKPGMKYVKWEHDSGVVTVELEPPCRFVIDTP  
TGPQIKYLYFVKNLNLRGAVLGYIGATVRLQAGKQTEFVSNSHLLTHCSFAVDPAAAY  
LDAVKQGAKPVGNCVKMLTNGSGSGQAITCTIDSNTTQDTYGGASVCIY CRAHVAHPTMD  
GFCQYK GKWVQVPIGTNDPIRFCLENTVCKVCGCWLNHGCTCDRTAIQSF DNSYLNVRG  
SSAARLEPCNGTDIDYCVRAF DVYNKDASF IGKNLKSNCVRFKNVDKDDAFYIVKRCIKS  
VMDHEQSMYNLLKGCNAVAKHDFFTWHEGRTIYGNVSRQDLTKYTMMDL CFALRNFDKED  
CEVFKEILVLTGCCSTDYFEMKNWFDPIENEDIHRVYAALGKV VANAMLKCVAFCD EMLV  
KGVVGVLTLDNQLNGNFYDFGDFVLCPPGMGIPYCTSYSYMMPVMGMTNCLASECFMK  
SDIFGQDFKTFDLLKYDFTEHKEVLFNKYFKYWGQDYHPDCVDCHDEMCI LHCSNFNTLF  
ATTIPNTAFGPLCRKVFIDGVPVATAGYHFQKQLGLVWNKDVNTHSTRLTITELLQFVTD  
PTLIVASSPALVDKRTVCF SVAALSTGLTSQTVKPGHFNKEFYDFLRSQGFDEGSELTL  
KHFFFTQKGDAAIKDFDYRYNRPTMLDIGQARVAYQVAARYFDCYEGGCITSREVVVTN  
LNKSAGWPLNKFGKAGLYYESISYEEQDAIFSLTKRNILPTMTQLNLKYAISGKERARTV  
GGVSLLATMTTRQFHQKCLKSIVATR NATVVIGTTK FYGGWDNMLKNLMADVDDPKLMGW  
DYPKCDRAMPSMIRMLSAMILGSKHVTCCTASDKFYRLSNE LAQVLTEVVYSNGGFYFKP  
GGTTSGDATTAYANSVFNIFQAVSSNINCVLSVNSSNCN FNVKKLQRQLYDNCYRNSNV  
DESFVDDFYGYLQKHFSMMILSDDSVVCYNKTYAGLGYIADISAFKATLYYQNGVFMSTA  
KCWTEEDLSIGPHEFCSQHTMQIVDENGKYYPDPDSRIISAGVFVDDITKTDAVILLE  
RYVSLAIDAYPLSKHPKPEYRKVFYALLDWVKHLNKT LNEGVL ESFSVTLLDEHESK FWD  
ESFYASMYEKSTVLQAAGLCVVCQSQT VLR CGDCLRRPMLCTKCA YDHVFGTDHKFILAI  
TPYVCNTSGCNVNDVTKLYLGGLNYCVDHKPHLSFPLCSAGNVFGLYKSSALGSMDIDV  
FNKLSTSDWSDIRDYKLANDAKESLRLFAAE TVKAKEESVKSSYAYATLKEIVGPKELLL  
LWESGKAKPPLNRNSVFTCFQITKDSKFQVGEFVFEKVDYGS DTVTYKSTATTKLVPGML  
FILTSHNVAPLRAPTMANQEKYSTIYKLHPSFNVS DAYANLVPYYQLIGKQRITTIQGP  
GSGKSHCSIGIGVYYPGARIVFTACSHA AVDSLCAKAVTAYSVDKCTRIIPARARVECYS  
GFKPNNNSAQYVFSTVNALPEVNADIVVVDEVSMCTNYDLSVINQRISYKHIVYVGD PQQ  
LPAPRVLISKGVMEPIDYNVVTQRMCAIGPDVFLHKCYRCPAEIVNTVSELVYENKFVPV  
KEASKQCFKIFERGSVQVDNGSSINRRQLDVVKRF IHKNSTWSKAVFISPYN SQNYVAAR  
LLGLQTQTVD SAQGEYDYVIFAQTS DTAHACNANRFNVAITRAKKGIFCIMS DRTLFD A  
LKFFEITMTDLQSESSCGLFKDCARNPIDLPPSHATTYLSLSDRFKTSGDLAVQIGNNNV  
CTYEHVISYMGFRFDVSMPSHSLFCTRDFAMRHVRGWLGM DVEGAHVTDGDNVGTNVPLQ  
VGF SNGVDFVAQPEGCVLTNTGSVVKPVRARAPPGEQFTHIVPLLRKGQPWSVLRKRIVQ  
MIADFLAGSSDVLVFLWAGGLELTTMRYFVKIGAVKHCQCGTVATCYNSVSN DYCCFKH  
ALGCDYVYNPYVIDIQWGYVGSLS TNHHAICNVHRNEHVASGDAIMTRCLAVYDCFVK

VDWSITYPMIANENAINKGGRTVQSHIMRAAIKLYNPKAIHDIGNPKGIRCAVTDKAWYC  
 YDKNPINSNVKTLEYDYMTHGQMDGLCLFWNCNVDMYPEFSIVCRFDTRTRSTLNLEGVN  
 GGSLYVNNHAFHTPAYDKRAMAKLKPAPFFYYDDGSCEVVHDQVNYVPLRATNCITKCN  
 GGAVCSKHANLYRAYVESYNI FTQAGFNIWVPTTFDCYNLWQTFTEVNLQGLENI AFNVV  
 NKGSVFGADGELPVAISGDKVFRDGNNTDNLVFNKTSLPTNIAFELFAKRKVGLTPPLS  
 ILKNLGVVATYKFLWDYEAEERPLTSFTKSVCGYTDFAEDVCTCYDNSIQGSYERFTLST  
 NAVLFSATAVKTGGKSLPAIKLNFGLMNGNAIATVKSEDGNIKNINWFVYVRKDGPVDH  
 YDGFYTQGRNLQDFLPRSTMEEDFLNMDIGVFIQKYGLEDFNFEHVYVYGDVSKTTLGGLH  
 LLISQVRLSKMGILKAEFVAASDITLKCTVTYLNDPSSKTVCTYMDLLLDLDFVSVLKS  
 LDLTVVSKVHEVIIDNKPWRWMLWCKDNAVATFYPQLQSAEWKCGYSMPGIYKTQRMCL  
 PCNLVNYGAGLKLPSGIMFNVVKYTQLCQYFNSTTLCVPHNMRVLHLGAGSDYGVAPGTA  
 VLKRWLPHDAIVVDNDVVDYVSDADFSVTGDCATVYLEDKFDLLISDMYDGRTKAIDGEN  
 VSKEGFFTYINGFICEKLAIGGSAIAKVTEYSWNKKLYELVQRFSFWTMTFCTSVNTSSSE  
 AFVVGINYLGDFAQGPFI DGNI IHANYVFWRNSTVMSLSYNSVLDLSKFNCKHKATVVVQ  
 LKDS DINEMVLSLVRSGKLLVRGNGKCLSF SNHLVSTK

>P12978

MPTFYALHGGQTYHLIVDTS LGNPSLSVIPSNPYQEQLSDTPLIPLTIFV GENTGVPP  
 PLPP  
 DVGHGPLASAMRMLWMANYIVRQSRGDRGLILPQGPQTAPQARLVQPHVPLRPTAPTIL  
 SPLSQPRLTTPPQPLMMPRPTPTPLPPATLTVPPRPTPTLPTPLLLTVLQRPTELQP  
 TSPPRMHLPLVHVPDQSMHPLTHQSTPNDDPSPEPRSPVFNIPPMPLPPSQLPPAA  
 PAQPPPGVINDQQHLHPLSGPPWWPPICDPPQPSKTQGGSRGQSRGRGRGRGRGRGKGS  
 RDKQRKPGGPWRPEPNTSSPSMPELSPVLGLHQGGAGDSPTPGPSNAAPVCRNSHTATP  
 NVSPIHEPESHNSPEAPILFPDDWYPPSIDPADLDESWDYIFETTESPSSDEDYVEGPSK  
 RPRPSIQ

>Q06906

MWGALILLILLVFLFYLVYNGKLNLSL TESSPLAQSSDSVQVDPQTEQLNVKLGNNKM  
 TYMRVAHGDNKVSQVYVAEKPM SMDDIEKQGNARVGANSLF IGTVYDQGVRSNAPGASN  
 DVTVTRTTANFDVKEYKNMFI VVKGLPPAKMTKEDNMLCF TVDGLHVCLVDANAAPLSER  
 VFARLPPSACTLVYTRNSAAQQLLLENGFTV VNAEHTAFLKNHKSREL

>P03093

GAALTLLGLDIATVSEAAAATGFSVAEIAAGEAAAAIEVQLASVATVEGLTTSEAI AAI  
 LIPQAYAVISGAPAAIAGFAALLQTVTVGVSVAQVGYRFFSDWDHKVSTVGLYQQPGMAV  
 DLYRPDDYYDILFPGVQTFVHSVQYLDPRHWGPTLFNAISQAFWRVIQNDIPRLTSQELE  
 RRTQRYLRDSLARFLEETTWTVINAPVNWYNSLQDYYSTLSPIRPTMVRQVANREGLQIS  
 FGHTYDNIDEADSIIQQVTERWEAQSQSPNVQSGEFIEKFEAPGGANQRTAPQWMLPLLLG  
 LYGSVTSALKAYEDGPNKKKRKLSRGSSQTKGT SASAKARHKRRNRSSR

>P03508

MDPNTVSSFDILLRMSKMQLESSSGDLNGMITQFESLKL YRDSLGEAVMRMGDLHSLQN  
 RNEKWREQLGQKFEEIRWLIIEEVRHKLKITENSFEQITFMQALHLLLEVEQEIRTF SFQL  
 I

>Q74125

MNGRADEEGLQRKQRLIRLLHQTNYPYQGLGTARQRNRNRNRNRKQHWRQLVALANSIYTF  
 PDPPADSPLDRAIQRLQGLTIQELPDPPTNLPESSESTNNNQGLAETYNSLPAIWVRVDP  
 RSAPGPKDYERDSCERVERLVGGNGTDRQNGTCSK KDQAGGRTCPPVRGSGINRETL

(4) 23 Plasma membrane proteins

>P03705

MKMPEKHDLAAI LAAKEQIGAILAFAMAYLRGRYNGGAFTKTVIDATMCAI IAWFIRD  
LLDFAGLSSNLAYITSVFIGYIGTDSIGSLIKRFAAKKAGVEDGRNQ

>P16849

MDTIIHNSTRNNTPPHINDTCNMTGPLFAIRTTEAVLNTFII FVGGPLNAIVLITQLLTN  
RVLGYSTPTIYMTNLYSTNFLTTLTVLPFIVLSNQWLLPAGVASCKFLSVIYYSSCTVGF  
TVALIAADRYRVLHKRTYARQSYRSTYMILLTTLWLAGLIFSVPAAVYTTVMHHDANDTN  
NTNGHATCVLYFVAEEVHTVLLSWKVLLTMVWGAAPVIMMTWFYAFFYSTVQRTSQKQRS  
RTLTFVSVLLISFVALQTPYVSLMIFNSYATTAWPMQCEHLTLRRTIGTLARVPHLHCL  
INPILYALLGHDFLQRMQRQCFRGLLDRRAFLRSQQNQQRATAETNLAAGNNSQSVATSLD  
TNSKNYNQHAKRSVSFNFPSGTWKGGQKTASNDTSTKIPHRLSQSHHNLG

>P07975

MFFSLLLVLGLTEAEKIKICLQKQVNSSFSLHNGFGGNLYATEEKRMFELVKPKAGASVL  
NQSTWIGFGDSRTDKSNSAFPVSADVS AKTADKFRFLSGGSLMLSMFGPPGKVDYLYQGC  
GKHKVFYEGVNWS PHAA INCYRKNWTDIKLNFQKNIYELASQSHCMSLVNALDKTIPLQV  
TAGTAGNCNNSFLKNPALYTQEVKPS ENKCGKENLAFFTLPTQFGTYECKLHLVASCYFI  
YDSKEVYNKRGCDNYFQVIYDSFGKVVGGLDNRVSPYTGNSGDTPTMQCDMLQLKPGRY  
VRSSPRFLLMPERSYCFDMKEKGPVTAVQSIWGKGRESYAVDQACLSTPGCMLIQKQKP  
YIGEADDHGDQEMRELLSGLDYEARCISQSGWVNETSPFTEKYLLPPKFGRCPLAAKEE  
SIPKIPDGLLIPTSGTDTTVTKPKSRIFGIDDLIIGVLFVAIVETGIGGYLLGSRKESGG  
GVTKESA EKGF EKIGNDIQILKSSINIAIEKLNDRISHDEQAIRDLTLEIENARSEALLG  
ELGIRALLVGNISIGLQESLWELASEITNRAGDLAVEVSPGCWIIDNNICDQSCQNFIF  
KFNETAPVPTIPPLDTKIDLQSDPFYWGSSGLAITATISLAALVISGIAICRTK

>Q86917

MNYTLSTVSSATMYNSSSNITTIATTIISTILSTISTNQNNVTPSTYENTTTISNYTTA  
YNTTYYSDDYDDYEVSIVDIPHCDGVDTTSFGLITLYSTIFFLGLFGNIIVLTVLRKYK  
IKTIQDMFLLNLTLSDLIFVLVFPFNLYDSIAKQWSLGDCLCKFKAMFYFVGFYNSMSFI  
TLMSIDRYLAVVHPVKSMPIRTKRYGIVLSMVVWIVSTIESFPIMLFYETKKVYGITYCH  
VFYNDNAKIWKLFINFEINIFGMIIPLTILLYCYKILNTLKTSQTKNKAIKMVFLIVI  
CSVLFLLPFSVTVFVSSLYLLNVFSGCMALRFVNLAVHVAEIVSLCHCFINPLIYAFCSR  
EFTKLLRLRTTSSAGSISIG

>P19106

MKTIIALS YIFCLVLGQDFPGNDNSTATLCLGHHAVPNGTLVKTITNDQIEVTNATELVQ  
SSSTGKICNPNHRILDGIDCTLIDALLGDPHCDGFQNETWDLFVERSKAFSNCYPYDVPD  
YASLRSLVASSGTLEFISEGFTWTGVTQNGGSNACKRGPDSGFFSRLNWLYKSGSTYPVL  
NVTMPNNDNFDKLYIWGVHHPSTDQEQTSLYVQASGRVTVSTKRSQQTII PNIGSRPWR  
GLSSRISIIYWTIVKPGDILVINSNGNLIAPRGYFKMRTGKSSIMRSDAPIGTCISECITP  
NGSIPNDKPFQNVNKITXGACPKYVKQNTLKLATGMRNVPEKQTRGLFSAIAGFIENGWE  
GMIDGWYGFRHQNSEGTQPADLKSTQAAIDQINGKLN RVIEKTNEKFHQIEKEFSEVEG  
RIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMKNLFEKTSRQLRENAEDMGN  
GCFKIYHKCDNACIGSIRNGTYDHDVYRDEALNNRFQIKGVELKSGYKDWILWISFAISC  
FLLCVLLGFIMTACQKGNIRCNICI

>P35253

MKTTCFLISLILIQGTKNLPILEIASNNQPQNVDSVCSGTLQKTEDVHLMGF TLSGQKVA  
DSPLEASKRWA FR TGVPK NVEYTEGEEAKTCYNI SVTDPSGKSLLLDPPTNIRDYPKCK  
TIHHIQGNPHAQGIALHLWGAFFLYDRIASTMYRGKVFTEGNIAAMIVNKT VHKMIF  
RQGQGYRHMNLTSTNKYWTSSNGTQTNDTGCFGALQEYNSTKNQTCAPSKIPPPLPTARP  
EIKLTSTPTDATKLN TTDPS DDEDLATS GSGSGEREPHTTSDAVTKQGLSSTMPPTPSP

QPSTPQQGGNNTNHSQDAVTELDKNNTTAQPSMPPHNTTTISTNNTSKHNFSTLSAPLQN  
 TTNDNTQSTITENEQTSAPSITTLPTGNPTTAKSTSSKKGPATTAPNTTNEHFTSPPPT  
 PSSTAQHLVYFRKRKRSILWREGDMFPFLDGLINAPIDFDPVNTKTIFFDESSSSGASAE  
 DQHASPNI SLTLSYFPNINENTAYS GENENDCDAELRIWSVQEDDLAAGLSWIPFFGPGI  
 EGLYTAVLIKNNLVCRLRLRLANQTAKSLELLLRVTTEERTFSLINRHAI DFLLRWG  
 TCKVLGPDCCIGIEDLSKNISEQIDQIKKDEQKEGTGWGLGGKWWTSDWGVLTNLGILL  
 LSI AVLIALSCICRIFTKYIG

>P05780

MSLLTEVETPIRNEWGCRCNDSSDPLVIAANIIGILHLILWILDRLFFKCIYRRFKYGLK  
 RGPSTEGVPESMREEYRKEQQNAVDVDDGHFVNIELE

>P06817

MNNAFNFCTNINPITHIRGSIITICVSLIVILIVFGCIAKIFINKNNCTNNVIRVHKRI  
 KCPDCEPFCNKRDDISTPRAGVDIPSFILPGLNLSEGTPN

>P52382

MEKETKSLAWPATAEFYGWVFISSIQCTMVLITVRFNSFKVGREYAVFTFAGMSFNCF  
 LLPKMGLLSGHWSLPRDFCAILLYIDDFSIFSSWSLVFMAIERINHFCYSTPLL NENS  
 KALAKVCFPIVWIIISGVQALQMLN NYKATALQNETPQCFLAFLRSGYDMWMLVYSVMIP  
 VMLVFIYIYSKNFMLLKDELSTVTTYLCIYLLGTIAHLPKAGLSEIESDKIFYGLRDI  
 MALPVLKVVYIIPVMAYCMACDDHTVPVRLCSIWLVNLCKKCFSTRREKESDLEVGKML

>Q38008

MLYNIMLEVAKGDYITILFALILDFDITGFLKAWKWKVTD SWTGLKGVIKHTLTFIFYF  
 VAVFLTYIHAMAVGQILLVIINLYALSIMENLAVMGVFI PKFM TARVQEELQKYTAQLD  
 AGKDLLEEFKGEKK

>P09723

MSLRGQVQIARSVFLLRIYILIWVQCLILMSVCAFCWLVLPHRLEQLFSSVRLTSLCLMI  
 SIVCLGLLRWAEPNFPKNVWILLTYTLLTSVAVTASGFHF SHRSVIYAMVATVTLFCFLT  
 LATYLFARDVELQRSLLTGASTLILLFAVFSLFPEAVSEILVMIAGLAVIVTSVVC DTQ  
 DILHDIEYESYIPGALCLYMDL MYLFVSVLYFMPSEPGSAHTAQTVAATAAASPQFVS

>P69539

MKKS LVLKASVAVATLVPMLSFAAEGDDPAKAAFDSLQASATEYIGYAWAMVVVIVGATI  
 GIKLFKKFTSKAS

>P03621

MKAMKQRIAKFSPVASFRNL CIAGSVTAATSLPAFAGVIDTSAVE SAITDGQGMKAIGG  
 YIVGALVILAVAGLIYSMLRKA

>Q98146

MAAEDFLTIFLDDDES WNETLNMSGYDYSGNFSLEVSVCEMTTVVPYTWNVGILSLIFLI  
 NVLGNGLVTYIFCKHR SRAGAI DILLGICLNSLCLSI SLLAEVLMFLFPNIISTGLCRL  
 EIFFYYLYVYLDIFSVVCVSLVRYLLVAYSTRSWPKQSLGWVLTSAALLIALVLSGDAC  
 RHRSRVVDPVSKQAMCYENAGNMTADWRLHVRTVSVTAGFLLPLALLILFYALTWCVVRR  
 TKLQARRKVRGVIVAVVLLFFVFCFPYHVLNLLDTLLRRRWIRDSCYTRGLINVLAVTS  
 LLQALYSAVVPLIYSCLGSLFRQRM YGLFQSLRQSFMSGATT

>Q65984

MIVLILCLLLFSYNSVIC TSNND CVQGNVTQLPGNENI IKDFLFHTFKEEPSVVVGGYYP  
 TEVWYNCSRSATTTAYKDFSNIHAFYFDMEAMENSTGNARGKPLL VHVHGDVPSIIYIS  
 AYRDDVQPRPLLKHGLLCITKNKI IDYNTFTSAQWSAICLGDDRKIPFSVIPTDNGTKIF  
 GLEWNDYVTAYISDRSHHLNINNNWFNNVTILYSRSSATWQKSAAYVYQGVSNFTYYK  
 LNN TNGLKSYELCEDYEYCTGYATNVFAPT VGGYIPHGFSFNWFMRTNSSTFVSGRFVT

NQPLLVNCLWPVPSFGVAAQQFCFEGAQFSQCNGVSLNNTVDVIRFNLNFTALVQSGMGA  
 TVFSLNTTGGVILEISCYNDTVSESSFYSGEISFGVTDGPRYCFALYNGTALKYLGTLF  
 PSVKEIAISKWGHFYINGYNFFSTFPIDCISFNLTGDSGAFWTIAYTSYTDALVQVENT  
 AIKKVTYCNSHINNIKCSQLTANLQNGFYVPASSEVGLVNKSVLLPSFYSHTSVNITID  
 LGMKRSGYGQPIASTLSNITLPMQDNNTDVYICIRSNRFSVYFHSTCKSSLWDDVFNSDCT  
 DVLYATAVIKGTGTCFSDKLNLYLTFNKFCLSLNPVGANCKFDVAARTRTNEQVVRSLY  
 VIYEEGDNIVGVPSDNSGLHDLVHLHDSCTDYNIYGITGVGIIRQTNSTLLSGLYYTSL  
 SGDLLGFKNVSDGVIYSVTPCDVSAHAAVIDGAIVGAMTSINSELLGLTHWTTTPNFYYY  
 SIYNYTNERTRGTAIDSNDVDCEPIITYSNIGVCKNGALVFINVTHSDGDVQPISTGNVT  
 IPTNFTISVQVEYIQVYTPVSIDCSRYVCNGNPRCNKLLTQYVSACQOTIEQALAMGARL  
 ENMEIDSMLFVSENAKLKASVEAFNSTETLDPYKEWPNIGGSWLGLKDI LPSHNSKRK  
 YRSAIEDLLFDKVVTSGLGTVDEYKRCCTGGYDIADLVCAQYNGIMVLPGVANDDKMAM  
 YTASLAGGITLGSLLGAVSIPFAIAVQARLNYVALQTDVNLKNQQILANAFNQAIGNIT  
 QAFGKVNDAIHQTSQGLATVAKVLAKVQDVVNTQGQALSHLTLQLQNNFQAISSSISDIY  
 NRLDELSADAQVDRITGRILTALNAFVSQTLTRQAEVRASRQLAKDKVNECVRSQSQRFG  
 FCGNGTHLFLSLANAAPNGMIFHTVLLPTAYETVTAWSGICASDGDRTFGLVVKDVQLTL  
 FRNLDDKFYLTPRTMYQPIVATSSDFVQIEGCDVLFVNATVIDLPSIIPDYIDINQTVQD  
 ILENFRPNWTVPELPLDIFNATYLNLTGEINDLEFRSEKLNHTTVELAILIDNINNTLVN  
 LEWLNRIETYVKWPWYVWLLIGLVVIFCIPILLFCCCSTGCCGCIGCLGSCCHSICSRRO  
 FESYEPIEKV

>P68326

MAKLTGMFSAAILLMAICSTAIYRGEHMSMYLNASSEFAVYPTDQSLVLVGHLLFLDGL  
 QRLPTTNYSGLIELIHYNYSSVCYTVIQTISYESCPRVANNAFRSCLHKTSKHYHDYFRV  
 NASVETNVLLNITKPQPTDSGAYILRVKLDHAPTADVFGVSAFVYDLKSKTVPDPMPPTQ  
 TVEPTTSYVSTPTYDYTDVTTETESTSTSTQQAMTSTQTPSATWGTQLTTELPTNETVV  
 IGQEALLCHWFQPSTRVPTLYLHLLGRTGNLPELVLLVEDSEFLRTTSPAHRPSASPADG  
 DDFKQTNSTSLKARNKIVAMVVIPTACVLMLLLVVVGAIINGAVRKHLLSCASRIYRSG  
 QGGASAAERRRLTCGPTLAASSESLADDTTSSPPTPKPSKTKLETDPMEQLNRKLEAI  
 KEES

>P03759

MRNRIMPGVYIVIPYVIVSICYLLFRHYIPGVSFSAHRDGLGATLSSYAGTMIAILIAA  
 LTFLIGSRTRRLAKIREYGYMTSVVIVYALSVELGALFFCGLLLLSSISGYMIPTIAIG  
 IASASFIHICILVFQLYNLTREQE

>P04849

MGTIIQFLVVSCLLAGAGSLDPAALMQIGVIPTNVRQLMYYTEASSAFIVVKLMPTIDSP  
 ISGCNITSISSYNATVTKLLQPIGENLETIRNQLIPTRRRRRFAGVVI GLAALGVATAAQ  
 VTAVALVKANENAAAILNLKNAIQKTNAAVADVQATQSLGTAVQAVQDHINSVSPA  
 TAANCKAQDAIIGSILNLYLTELTTIFHNQITNPALSPITIQALRILLGSTLPTVVEKSF  
 NTQISAAELLSGLLTGQIVGLDLYMQMVIKIELPTLTVQPATQIIDLATISAFINNQE  
 VMAQLPTRVMVTGSLIQAYPASQCTITPNTVYCRYNDAQVLSDDTMAQLQGNLTRCTFSP  
 VVGSFLTRFVLFDGIVYANCRSMLCKCMQPAAVILQPSSSPVTVIDMYKCVSLQLDNLRF  
 TITQLANVTYNSTIKLESSQILSIDPLDISQNLAAVNKSLSDALQHLAQSDTYLSAITS  
 TTTSVLSIIAICLGLGLILILLSSVVVWKLITIVVANRNRMENFVYHK

>P08714

MTRLPIILLLLISLVYATPPFPQTSKKIGDDATLSCNRNNTNDYVVMASAWYKEPN  
 KSDVLYFDNYTKDKISYDSPYDDLVTITIKSLTARDAGTYVCAFFMTSTTNDTDKVDYE  
 EYSTEIVNTDSESTIDIILSGSTHSPETSSEKPEDIDNFCSSVFEIATPEPITDNVED

HTDVTYTSDSINTVSASSGESTTDETPPEITDKEEDHTVTDTVSYTTVSTSSGIVTTKS  
TTDDADLYDTYNDNDTVPPTTVGGSTTSISNYKTKDFVEIFGITALIILSAVAIFCITYY  
IYNKRSRKYKTENKV

>P28081

MSPQRDRTNAFYKDNPHPKGSRIVINREHLMIDRPYVLLAILFVMFLSLIGLLAIAGIRL  
HQAAIHAEIHKSLSTNLDVTNSIEHQVKDVLTPFKIIGDEVGLRTPQRFTDLVKFISD  
KIKFLNPDREYDFRDLNWCINPPERIKLDYDQYCADVAEELMNALVNSTLLETRTTNQF  
LAVSKGNCSGPTTIRGQFSNMSTSLLDLYLSRGYNVSSIVTMTSQGMYGTYLVEKPNLS  
SKRSELSQLSMYRVFEVGVIRNPGLGAPVVFHMTNYFEQPVSNDLNMCVALGEFKLAALC  
HREDSITIPYQSGKGVSFQLVNLGVWKSPTDMQSWIPLSTDDPVIDRLYLSSHRGVIAD  
NQAKWAVPTTRTDDKLRMETCFQQACKGKIQALCENPEWAPLKDNRIPSYGVLSVDLSPT  
VELKIKIASGFGPLITHGSGMDLYKSNHNNVYWLTIIPPMKNLALGVINTLEWIPRFKVS  
NLFTVPIKEAGKDCHAPTYLPAEVDGDVKLSSNLVILPGQDLQYVLATYDTSRVEHAVVY  
YVYSPGRSFSYFYPFRLPIRGVPIELQVECFDQKLWCRHFCVLANSSESGGHITHSGMV  
GMGVSCTVTREDGTNRRQSC

>P09717

MGLRFPTATQRQIVFRRLFDSGNNDYDEAAVVAVLGVVHRFEVVVRIAGLLLFQISAAV  
AVLGSFSLVFPATLKSRLGPFCHVWVAPEVLLLVPVASALFVYFRYERPVLAQRNRHPR  
CRRPFRQLVLLLAGLLAHIPALGVTACQEPREVLTSFVLTTLVITLLCAEVVVICRDNCT  
LSDQFTLINGVWVFLVNLIVFTRPWTWPLRLLLGFYSTVGLIFAGHFSQQVLFVRHV  
LMPRDVAHTSLQLFITFISLFFLILRIRNCQDLLSDLRLLLELPSSDAMTLPNDLSSHASS  
STPLSTLSP

>P09718

MRREKGFQVPTDGTVIYVPPGIQETRLATRSLAWVDCRVALHTYGAVGWQLAGLTALLS  
AFCYAAPATWFHHSRLCLTESSPSLVFVIPVTSVIFIHICYETSHPSNIGVLLFYTLLHVP  
PLIVICLCLDGTLVISAALFTLLAFLSCTGVALLAPERTVRRQIVVHALITLFTAIIV  
VILRRGWSWCFKIVLSFSVLITCLAVSHFHEAALAVRYETPLERALLAAVKVFLSLVFTL  
LMVLRIMTLRFTLQTYFSSDKL

>P09719

METVSTQRETASSETERTREAASAETTDATFRSLEEGSTISSRYSETASTVSEDAVCWLR  
RTAIVMRVYGLLTLETAFSVLISALVWLGYPVSLGYKCSDDPSPLLSCTPVLVLGALELT  
DHRHPSNGLVFALYVALLSFTTAGLNLCATAPIGISSLILTWTLFVACNGVAWEHRLSSV  
WRDALFTSTLLTVMVSVLASTYTWLHKTLCLYTVFVGCILAVLFQDVRYIATKMPVSHV  
IRSSLILYATETLIYHTTLLMLTPVWWSARWDQMF SYLAKLGTYYHYLIDNGTLSVILNT  
TTATFQSKAA