

Online Supplementary Materials D. The testing dataset contains 42 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 others in the same subcellular location of either the dataset here or the learning dataset in the Online Supplementary Materials C. See the text of the paper for further explanation.

(1) 4 Cytoplasm proteins

>P31352

MQHPTQYPDARLSSPIILDQCDLLARSLGLYSHYSHNPCLRNCRI PHHIYRLRNSTALKT
 FLQNC SILTVPFHSIWDHILTSIQYDAINHVDVDFKYLLPSELVKYANWDNEFLKAYLNKI
 LGLDHVFSASARSQCEDFSPKENPYYWGMLLLVHLSQLARRIKGQRGSLRSNWKF IGTDL
 ELFGIADVFVIFKVPVKTIIRNAVSLQASKPGLRIWYRDQNLTPYLCDDDEFIVSVASYECF
 IMIKDVFIERYNTWEICARAWLESDGADYPPDVDLGE LYNQGDQIIAMYLEDGFKLIKH
 LEPLCVSCIQTHGIFTPRKYWFQSQMIKSYDELHDLNLKLQISDNKAECAQNFIKTIVQ
 AKLTPQQYCELFSLQKHWHGHPVLYNDVALDKVKKHAQSTKILKPKVMFETF CVFKFIVAK
 NHYHSQGSWYKTT HDLHLTPYLRQHIVSNSFPSQAEIYQHLWEWYFVEHEPLFSTKIISD
 LSIFIKDRLTAVNQECWDSVDFDRSVLGYNPPVRFQSKRVPEQFLGQADFSLNQILEFAEK
 LEYLAPSYRNF SFSLKEKELNIGRTFGKLPYRVRNVQTLAEALLADGLAKAFPSNMMVVT
 EREQKEALLHQASWHNSASIGENAIVRGASFVTDLEKYNLAFRYEFTTRHFIDYCNRCY
 G VKNLFDWMHFLIPLCYMHVSDFYSPPHCVTEDNRNNPPDCANAYHYHLGGIEGLQOKLWT
 CISCAQITLVELKTKLKLKSSVMGDNQCITTL SLPIDAPNDYQENAE LNAARVAVELA
 ITTGYSGIFLKP EETFVHSGFIYFGKKQYLNGVQLPQSLKTMARCGPLSDSIFDDLQGS
 L ASIGTSFERGTSETRHIFPSRWIASFHSM LAINLLNQNHLGFPLGFNIDISCFKKPLTFS
 EKLIALITPQVLGGLSFLNPEKLFYRNISDPLTSGLFQLKNALEFLEKEELFYILISKKP
 GLADASDFVMNPLGLNVPRSKEIITFLRQTVRENITITSONRIINSLFHIGSDLEDQRVC
 EWLSSNPVMSRFAADIFSRTPSGKRLQVLGYLEGTRTLLASRTISLTTEGTMLMKLREL
 TRNRWKSWSFSYIDALDDD LSESEKFTCTVDVANFLRAYSWSDVLK GKRLIGATLPCLLE
 QFEVKWINLSEDLREQFNLS SDSKSTINLLPYDCKELRLEGSNDTELNYVSCALDRKV VQ
 KHPSVNRLAWTIGNRAPYIGSRTE DKIGYPPLRVNCP SAALKEAIE MVSRLLWVTQGTAD
 REKLLIPLLNSRVNLDYQTVLNF LPTHYSGNIVHRYNDQYGQHSFMANRMSNTSTRAIIS
 TNTLGKYAGGGQAAIDSNIIFQNTINLGVAVLDIALSLAKLSSASNVTFRMLNKCCTR
 H VPSEYLYFDKPLDVLN KYMDNELVYDNDPLCSGIKGRLGRVSRSTLSLSLNVSDIGSYD
 FPTIAAWTLGETIVGSIFSDESSQSTDPISSGCTKTFVTHFLVYPVESIFYAFGANLIVE
 SL SLSRIKSIKNLSDLTFLISSTIRNLSHRSLRILQSTFRHELVLTRLAHHIPLISMLG
 GSAGEKSSSDAVRLF LTASYQNF INNFSCLMKGQSSLPVWLYFPSEGQQLKPILKILQR
 LSDLLSPDKIQKRKILADTCCPIGSFWVYPSKSTRTNHYASLNYWRDKANKVKNTPF
 SH LINCSPPEFSSHTSSVSSNQV TNSKYIVYPENITEINARTRLINYGSTALQGM
 DTKMPL SEQNLVENC RPSEGIRFKDNQKITKHDQRCEREESPQOMFPEDNMQTPAH
 IHSSSPFQI LIKSLDAHEDFDASKIILNSEINNLNLT EYTLNTKLLTTPRTEILDTSPL
 QSSRYSSTS RERSLLSREQASYLYVDCSNIP S ISLDPGFRS MSDQNQVQMLINTYKRD
 LHACFDSNQFC RFTGVVSSMHYKLYDLLPPGK LKKAICLAEGEGSGARLLLKWKETD
 YLFFNTLATDSQQE AEILSGRVI PRMLYNIDRLSALLE SRRLILNNLTIQITDITNPL
 WLDSVIQYLPEDSDIL TMDAETTKDE TREQLYKTIVNIWTRTSPNIPKISIIKVFL
 LDYEGTLFLMKNAIQYYGQV QLKKPYSSNAKNSEWYLCCGKRRIQRLQIDFSDQV
 GIFLICKAMSRQRQAIPYWLKHIEK NYPASLHEFFLTLGFPSLESSFCHRYTIP
 FSEGKALFHKVQSYVRQ GKQHLHSLMLDYEN NSPLLDLRNH FICSLRGKITKYND
 ILKLNLVIKAVEKGNWSQLVEILPNMHSVCIVHV

DHECSGCEKRLLLKLD FIRNTKIAEQKLLNRVIGYILFFPFGLFKSGSLRA

>Q80872

MAVAYADKPNHF INFPLTHFQGFVLNYKGLQFQILDEGVDCIKQTAPHISLTMLDIQPED
YKSVDVAIQEVIDDMHWGDFQIKFENPHILGRCIVLDVKGVEELHDDL VNYIRDKGCVA
DQSRKWIGHCTIAQLTDAALS IKENVD FINSMQFNKITINPSSPARLEIVKLGAEKKDG
FYETIVSHWMGIRFEYTSPTDKLAMIMGYCCLDVVRKELEEGDLPENDDDAWFKLSYHYE
NNSWFFRHVYRKS FHRKACQNLDCNCLGFYESPVEED

>P16913

MNFQGLVLT DNCKNQWVVGPLIGKGGFGSIYTTNDNNYVVKIEPKANGSLFTEQAFYTRV
LKPSVIEEWKKS HNIKHVGLITCKAFGLYKSINVEYRFLVINRLGADLDAVIRANNRRLP
KRSVMLIGIEILNTIQFMHEQGYSHGDIKASNIVLDQIDKNKLYLVDYGLVSKFMSNGEH
VPFIRNPNKMDNGTLEFTPIDSHKGYVVSRRGDLET LGYCMIRWLGGILPWTKIS ETKNC
ALVSATKQKYVNN TATLLMTSLQYAPRELLQYITMVNSLTYFEENYDEF RHILMQGVVY

>P89079

MAESLYAFIDSPGGIAPVQEGTSNRYTFFCPESFHIPP HGVVLLHLKVSVLVPTGYQGRF
MALNDYHARDILTQSDVIFAGRRQELTVLLFNHTDRFLYVRK GHPVGTLLLERVIFPSVK
IATLV

(2) 2 Extracellular proteins

>P15240

MLKRSLFLTLVLLLLLFSFSSITNEVSASSSFDKGYKKGDDASYFEPTGPYLMVNVTGVD
GKRNELLSPRYVEFP IKPGTTLTKEKIEYYVEWALDATAYKEFRVVELDPSAKIEVTTYD
KNKKKEETKSFPI TEKGFVVPDLSEHIKNPGFNLITKVVIEKK

>O40633

MTFRKTSLVLLLLLSIDCIVKSEITSAQTTPRCLAANNSFPR SVMVTL SIRNWNTSSKRAS
DYYNRSTSPWTL YRNEDQDRYPSVIWEAKCRYLGCVNADGNVDYHMNSVPIQQEILVVRK
GHNPCPNSFRLEKMLVTVGCTCVTPIVHNV

(3) 20 Nucleus proteins

>P23091

MASELAMSGSDLPTSPLAMEYVND FDLMKFEVKKEPVETDRIISQCGRLIAGGSLSSSTPM
STPCSSVPPSPSFSAPSPGSGTDQKTHLEDY YWMTGYPQQLNPEALGFSPEDA VEALINS
SHHPLPGAFDGYARGQQLAAAAGGSVPAEEMGSAAAVVSAVIAAAAAQGGAPHYHHHHHH
PHHGGGGGGGGPHGAAPGSAPPSSASSSAAGSGGGGGGGGGAGGLHHPHGGGGGGGG
LHFDDRFSDEQLVTMSMRELNRQLRGVSKEEVIRLKQKRRTLKNRGYAQSCRFKRVQQRH
VLESEKNQLLQQVEHLKQEI SRLVRERDAYKEKYEKLVSNGFRENGSSSDNPSSPEFFMY
PRESSTVM

>P11675

MADDLDFDIETEGNFSQLLAAAAAAEEEGIASGPDGGSQGSRRRGSSGEDLLFGPGGLF
SDDAAEAEAAVLAAAAGATRPPRPPSAQQQRHARRGSGEIVVLDDDEDEEED EPGSPAAGS
PGRALHQGSEHGHLVLGPRSRAGSGPRPPTPAALAAAEAGAPGGPGRSSPSAASPSSSS
GSSGSSGSPGPSAAPRRWSPARGDPVGEPPGAARPRTPAPPAQPAAVAAA PARRGPASPA
SPAAGPVSAPGGGAPSGGGDRGRHHHQHREPLLDEPAAARRLDPRPLGARSPVSSNPNS
SSSSTTTVAVEPVARGPEKDEDGLGLAGDGGAPLQRQPRRRRAGEGALRRGRGFSSSSSG
GSDSDLSPARSPSAPRAPAAAAAARRSASSSSSSSSSSSSSSSSSEGEDEGVRPGAPLA
RAGPPSPPPAPAAAPRPSASSASATSSSAAASPAPAPEPARPPRRKRRSTNNHLSLMADG
PPPTDGPLLTPLGEPWPGSDPPADGRVRYGGAGDSREGLWDEDDVRQAAAARYRAAAGPVP
VFIPEMGDSRKQHEALVRLIYSGAAGEAMSWLQNPQMOPDQRFNQFCQRRVHAPHGHGS
FITGSVTPPLPHIGDAMAAQDPLWALPHA VSAVAMSRRYDRTQKTFILQSLRRAYADMAY

PGRAADPRAGEATVEALCARVRAAFAAAQPGRVPRELADACVLACRGVLERLLPCPLRLP
 APARAPAALGPACLEEVTAALLALRDAIPGAGPAERQQAADSVALVARTVAPLVRYSDG
 ARAREAAWTYAAALFAPANVAGARLAEAAAARGPAPGLPPLWPEQPGLVVPAPAPAA
 AGAPSGLPGSGPSSPASTKSGSSTKSSSGTKSGLSGSSGYASSPAAGPDPAPERKKRR
 APGARRPGDGEDEGLSGSALRGDGHGRDDEEDRGPRRKRSLGLGPAPDPAPALVSS
 SSSSSSEDDRLRRPLGPMPEHPAPDGGFRRVPAGETHTPRPSAAALAAAYCPPEVARALVD
 QEVFPELWRPALTFDPAALAHIAARRGAAGAPLRRRAAWMRQIADPEDVRVVVLYDPLPH
 EELCAEPAEGAPRPAWDPRRGGLSALLAAFAHRLCTPDSHAWAGNWTGRPDIGRLNAQGV
 LLLSARDLGFAGAVEYLC SRLGAARRRLIVLDTIEDWPADGPAVGDYHVYVRARLDPAAQ
 CAVRWPECRELRAAVLDSSSIVGPACFARVEASFARLHPGAEPLRLCRQDNVRYTVSTRA
 GPRTVPPLPPRAYRERVLPTVDGCKDMARQRSALGLGDPDFDAGAAF GHRAANRWGLGAP
 LRPV FVSCGRRGLAELRGPEGLPAELRAFCAAALLEPDAEAAPLV LAPGALAAAGAPP AV
 RWDFAPFETS VRAAAGGAVETHRPSGARGPGEDGEDSAAVEIVGFRGGDRPRGPLGP IK
 VEAISDDEEAEDAGNPYLLLR

>P03116

MANDKGSNWDSGLGCSYLLTEAECESDKENEPEGAGVELSVESDRYDSQDEDFVDNASVF
 QGNHLEVFQALEKKAGEEQILNLKRKVLGSSQNSSGSEASETPVKRRKSGAKRRLFAENE
 ANRVLTPLOVQGEGERQELNEEQAI SHLHLQLVKSKNATVFKLGLFKSLFLCSFHDI TR
 LFKNDKTTNQQWVLAVFGLAEVFFEASFELLLKQCSFLQMQRSHEGGTCAVYLICFN TA
 KSRETVRNLMANTLNVREECLMLQPAKIRGLSALFWFKSSLSPATLKHGALPEWIRAQT
 TLNESLQTEKFDGFTMVQWAYDHKYAEESKIAYEYALAAGSDSNARAFLATNSQAKHVKD
 CATMVRHYLRAETQAL SMPAYIKARCKLATGEGSWKSILTFNYQNIELITFINALKLWL
 KGIPKKNCLAFIGPNTGKSMCLNSLIHFLGGSVLSFANHKS HFWLASLADTRAALVDDA
 THACWRYFDTYLRNALDGYPVSIDRKHKA AVQIKAPPLLVTSNIDVQAEDRYLYLHSRVQ
 TFRFEQPCTDESGEQPFNITDADWKSFFVRLWGRDLIDEEEDSEEDGDSMRTFTCSARN
 TNAVD

>P06921

MENLSERFNALQDQLMNIYEAAEQTLQAQIKHWQTLRKEPVLLYYAREKGVTRLGYQPVP
 VKAVSETKAKEAIAMVLQLESLOTSDFAHEPWTLVDTSIETFRSAPEGHFKKGPLPVEVI
 YDNDPDNANLYTMWTVYYMDADDKWHKARSGVNHIGIYYLQGTFKNYVLFADDAKRYG
 TTGEWEVKVNKETVFAPVTSSTPPGSPGGQADTNTTPATPTTSTTAVDSTSRQLTTSKQP
 QQTETRGRRYGRRPSSKSRRSQTQQRRSRSRHSRSRSRSRSRSKSQTHTRSTTRSSTSL
 TKTRALTSRSRSRGRSPTTCRRGGGRSPRRRSRSPSTSSSCTTQRSQRARAESSTTRGAR
 GSRGSRGSRGGRRRRGRSSSSSSPAHKRSRGGSAKLRGVSPGEVGGSLRSVSSKHTGR
 LGRLL EEARPPVIIVKGAANTLKNVRNRAKIKYMGLFRSFSTTWSWVAGDGTERLGRPR
 MLISFSSYTQRRDFDEAVRYPKGVDKAYGNLDSL

>P08546

MFFNPYLSGGVTGGAVAGGRRQRSQPGSAQGSGRPPQKQFLQIVPRGVMFDGQTGLIKH
 KTGRPLPMFYREIKHLLSHDMVWPCPWRETLVGRVVGP IRFHTYDQTDVLFDFSPENVS
 PRYRQHLVPSGNVLRFFGATEHGYSICVNVFGQRSYFYCEYSDTDRLREVIASV GELVPE
 PRTPYAVSVTPATKTSIYGYGTRPVPDLQCVSISNWTMARKIGEYLLEQGFVYEV RVDP
 LTRLVIDRRIITTFGWCSVNRDWRQQGRASTCDIEVDCDVSDLVAVPDDSSWPRYRCLSF
 DIECMSGEGGFPCA EKSDDIVIQISCVCYETGGNTAVDQGI PNGNDGRGCTSEGVIFGHS
 GLHLFTIGTCGQVGPVDVYEF PSEYELLGLFMLFFQRYAPAFVTGYNINSFDLKYILTR
 LEYLYKVD SQRFCKLP TAQGGRFLLHSPAVGFKRQYAAAFPSASHNNPASTAATKVYIAG
 SVVIDMYPVCMAKTNSPNYKLN TMAELYLQRKDDLSYKDI PRCFVANAEGRAQVGRYCL
 QDAVLVRDLFNTINFHYEAGAIARLAKIPLRRVIFDGQQIRIYTSLLDECACRDFILPNH

YSKGTTVPETNSVAVSPNAAI I STA AVPGDAGSVAAMFQMS PPLQSAPSSQDGVSPGSGS
 NSSSSVGVFSVSGSGSSGGVGSNDNHGAGGTA AVSYQGATVFEPEVGYNDP VAVFDFAS
 LYP S I IMAHNL CYSTLLVPGGEYPVDPADV SVTLENGVTHR FVRASVRVSVLSELLNKW
 VSQRRAVRECMRECQDPVRRMLLDKEQMALKVTCNAFYGFTGVVNGMMPCLPIAASITRI
 GRDMLERTARFIKDNFSEPCFLHNFFNQEDYVVG TREGDSEESSALPEGLETSSGGSNER
 RVEARVIYGD TDSVFRFRGLTPQALVARGPSLAHYVTACLFVEPVKLEFEKVFVSLMMI
 CKKRYIGKVEGASGLSMKGVDLVRKTACEFVKGVTRDVL SLLFEDREVSEAAVRLSRLSL
 DEVKKYGVPRGFWRILRRLVQARDDL YLHRVRVEDLVLSVLSKDISLYRQSNLPHIAVI
 KRLAARSEELPSVGD RVFYVLTAPGVRTAPQSSDNGDSVTAGVVSRSDAIDGTDDDADG
 GGVEESNRGGEPAK KRARKPPSAVCNYEVAEDPSYVREHGVP I HADKYFEQVLKAVTNV
 LSPVFPGETARKDKFLHMLVLRRLHLEPAFLPYSVKAHECC

>P54094

MNALQEDTPPGPSTVFRPPTSSRPLETPHCREIRIGIAGITITLSL CGCANARAPTLRSA
 TADN SESTGFKNV PDLRTDQPKPPSKKRSCDPSEYRVSELKESLITTTPSRPRTARRCIR
 L

>P36720

MDDNTGTEGGACSESERAGGW FIVEAIVDRRTGDPISSDDDEE EDEAGEDFVDFIDDTRS
 LGDGQEVAQELFQQQTAADDDVAVQTVKRKFAPSPYFSPVCEQASIEHELSPRLDAIKLG
 RQSAKAKRRLFELPDSGYGQTQVDTESGPKQVQGSSETQDGRQDDDEG SVVQSTLDTGNQ
 NGRQNNDEGSGRNVGEHGSQEEERAGGDGEESDLQSTSTGKGAGGVVEILRASNKATLL
 GKFKEQFGLGYNELIRHFKSDRTSCADWVVCVFGV FCTVAEGIKTLIQPLCDYAHIQVLP
 CQWGMTVLMLVRYKRAKNRET VAKGLSTLLNVPESQMLIEPPKLRSGPAALYWKTS MSS
 CSDVYGETPEWIVRQTMVGHAMEDAQFSLSEMVQWAYDHDITDESTLAYEYALIADTDSN
 AA AFLSSNCQAKYLKDACTMCRHYKRGEQARMSMSEWIWFRGDKVQGDGDWKPIVQFLRY
 QDVEFI PFLCAFKTF LQGVPKKSCLV FYGPADTGKSYFCMSLLRFLGGAVISYANSSSHF
 WLQPLSEAKIGLLDDATSQCWNYIDTYLRNALDGNQICVDRKHRALLQLKCPLLIT T NI
 NPLTDERWKFLRSRLQLFTFKNPFVTTQGEPMYTLNDQNWKCFRRLWARLSLTDPEDE
 EEHGNPSEPFRCVPGQNARTI

>P03264

MASREEEQRETTPERGRGAARRPPTMEDVSSPSPPPP RAPPKKRLRRRLESEDEEDSS
 QDALVPRTPSPRPSTSTADLAIASKKKKKRPSPKPERPPSPEVIVDSE EEREDVALQMGV
 FSNPPVLIKHGKGGKRTVRRLLNEDDPVARGMRTQEEKEESEAESESTVINPLSLPIVSA
 WEKGMEARALMDKYHVDNDLKANFKLLPDQVEALAAVCKTWLNEEHRGLQLTFTSNKTF
 VTMMGRFLQAYLQSF AEVTYKHHEPTGCALWLHRCAEIEGELKCLHGSIMINKEHVIEMD
 VTSENGQRALKEQSSKAKIVKNRWGRNVVQISNTDARCCVHDAACPANQFSGKSCGMFFS
 EGAKAQVAFKQIKAFMQALYPNAQTGHGHLMLPLRCECNSKPGHAPFLGRQLPKLTPFAL
 SNAEDLDADLISDKSVLASVHHPALIVFQCCNPVYRNSRAQGGGPNCDFKISAPDLLNAL
 VMVRS LWSENFTELPRMVVPEFKWSTKHQYRNVSLPVAHSDARQNPFD F

>P0C012

MVKSRSVSYRSVMSGVGEERVPSAFTILASWG WTFAPQNHDPGASPNTTPIESIAGTAPDA
 HVGPLDGE PDRDAISPLTSSVAGDPPGADGPYVTFDTL FMVSSIDELGRRQLTDTIRKDL
 RLSLAKFSI ACTKTSSFSGTAARQRKRGA PPQRTCVPRSNKSLQMFVLCKRANAAQVREQ
 LRAVIRSRKPRKYYTRSSDGR LCPAVPVFVHEFVSSEPMRLHRDNVMLSTEPD

>P03211

MSDEGPGTGPGNGLGEKGD TSGPEGSGSGSPQRRGGDNHGRGRGRGRGRGGRPGAPGGS
 GSGPRHRDGVRRPQKR PSCIGCKGTHGGTGAGAGAGGAGAGAGGGAGAGGGAGAGGGAGAGGG
 AGGAGAGGGAGAGGGAGGAGAGGGAGAGGGAGAGGGAGAGGGAGAGGGAGAGGGAGAGGG

AGGAGAGGGAGGAGGAGAGGGAGAGGAGGAGAGGGAGGAGGAGGAGAGGAGAGGAGAGGAGAGGA
 GAGGAGAGGAGGAGAGGAGGAGAGGAGGAGAGGGAGGAGAGGGAGGAGAGGAGGAGAGGAGAGGA
 GGAGAGGAGGAGAGGGAGAGGAGAGGGGRGRGSGGRGRGSGGRGRGSGGRGRGSGRRRGRGRER
 ARGGSRERARGRGRGRGEKRPRSPSSQSSSSGSPRRRPPGRRPFFHFPVGEADYFEYHQE
 GGPDPGEPDVPVGAIEQGPADDPGEGPSTGPRGQGDGGRRKKGGWFGKHRGQGGSNPKFEN
 IAEGRLALLARSHVERTTDEGTWVAGVFVYGGSKTSLYNLRRGTALAI PQCRLTPLSRLP
 FGMAPGPGPQPGPLRESIVCYFMVFLQTHIFAEVLKDAIKDLVMTKPAPT CNIRVTVCSF
 DDGVDLPPWFPMPMEGAAAEGDDGDDGDEGGDGDEGEEGQE

>P28859

MDRNAVLYGVLEHRLPKWVELSDDTDLEPFFFSSVRYITAGSEDAIMI QALNLNTDEIVV
 FLVTNLNFMAL IPTVYIENPGIRQLIASTPISYRSPITVFNGLKWKWDCDLFVFGTMAA
 QKAFIKAGNSVLGSLGGNVYTYGDHVS NFDGNT PVLQNNLMCSHVYYTRYKTDVYAPWEF
 YYDQQRDQGYLMSLPAIIPRCKREGAFDIETIVHENAMDQDLNCQKFFKSEFRSMEE SQV
 LIQRFREAGVTGLPPSPFVGITQKLHEIVSISLVVCNYHKTGPKKKEYVYVYNTKKMENP
 MEMIPVEHLHL DASRIKFEACKNEFYMLLAFINRLRKS VNVLYVYNAQFDIQVIQORLRY
 YAFKQRAPRCKGHDDIPHEWGKALMEKWEAFLSVKPQLFKAQILMGQDILKANYLKLLE
 GIGSVLAQAKSTMAKMCTIKERIDSYRKMKDTVQNFKSHGFGCDIIDMMYVCKRKEFEAK
 DGSLNTVAQLIIKKFKPHKATPKIHKMDDITYDKLDGYRAGGTKIAECLINLIDSLLV
 IRIAKNLKPMEEYIYRQLACYNIDTAAHTRGVMNFCGFIQSTKVVEVSRNKARLDAGIVM
 ATDYIRNSLFTPETIPRRGGFVMAPLTGLFFARPTQCFELCLDFTSMYPSMMCDLNISPE
 TIVDSKTNRVGDYMGYDWSKIDQGF EKFTLVLRVDRTDPENPKLVRHTSDTSLSLKRYL
 RL RTEHKRALKQSSGSVAEYHNRLQNEMKICTNTHYGVSEHTCSLMIT TQGQH KIKLVNE
 FIKTLNRTGHS LFPNYGDTDSTMLYHPSDESETQLEDMVTLEDEMRAELREYMLKKS LSAE
 LVNRVKEKTKRTDTFVQSFLSDVETVLFDDMVEKLRLFSQGEVIEPFKDGGTWVVDPLT
 GIWMDCSTPFSELICKLEYENASSIGCHVAKMVSIGSTYLFKKISLYHVRVWRMCAD
 TDGSPSHLYFPVSLSRTRAKQRGDH

>P03126

MHQKRTAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREVYDFAFRDLCIV
 YRDGNPYAVCDKCLKFYSKI SEYRHYCYSLYGTTLEQQYNKPLCDLLIRCINCQKPLCPE
 EKQRHLDKKQRFHNIRGRWTGRCMSSCRSSRTRRETQL

>O56129

MTYPRRRYRRRRHRPRSHLGQILRRRPWL VHPRHRYRWRRNKNGIFNTRL SRTFGYTVKAT
 TVRTPSWAVDMMRFNIDDFVPPGGGTNKISIPFEYYRIRKVKVEFWPCSPITQGDRGVGS
 TAVILDDNFVTKATALTYDPYVNYSSRHTIPQPF SYHSRYFTPKPVL DSTIDYFQPNNKR
 TQLWLRLQTSRNV DHVGLGTAFENSIYDQDYNIRVTMYVQFREFNLKDPPLKP

>P03073

MDRVLSRADKERLLELLKLPRQLWGDFGRMQQAYKQOSLLLHPDKGGSHALMQELNSLWG
 TFKTEVYNLRMNLGGTGFQGSPPRTAERGTEESGHSPLHDDYWSFSYGSKYFTREW NDFE
 RKWDP SYQSPPKTAESSEQPDLCYEEPLLSPNPSSPTDTPAHTAGRRRNPCVAEPDDSI
 SPDPPTPVSRRRPRPAGATGGGGGGVHANGGSVFGHPTGGTSTPAHPPPYHSQGGSESM
 GGS DSSGFAEGSFRSDPRCESENE SYSQSCS QSSFNATPPKKAREDPAPSDFPSSLTGYL
 SHAIYSNKTFPAFLVYSTKEKCKQLYDTIGKFRPEFKCLVHYEEGMLFFLTMTKHRVSA
 VKNYCSKLCRSFLMCKAVTKPMECYQVVTAAPFQLITENK PGLHQFEFTDEPEEQKAVDW
 IMVADFALENNLDDPLLIMGYLDFAKEVPSCIKCSKEETRLQIHWKNHRKHAENADLFL
 NCKAQKTICQQAASLASRRLKLVETR S QLLKERLQQSLLRLKELGSSDALLYLAGVAW
 YQCLLEDFPQTLFKMLKLLTENVPKRRNILFRGPVNSGKTGLAAALISLLGGKSLNINCP
 ADKLA FELGVAQDQFVVC FEDVKGQIALNKQLQPGMGVANLDNLRTTWNGSVKVNLEKKH

SNKRSQLFPPCVCTMNEYLLPQTVWARFHMVLDFTCKPHLAQSLEKCEFLQRERIIQSGD
 TLALLLIWNFTSDVFDPIQGLVKEVRDQFASECSYSLFCDILCNVQEGDDPLKIDICDIA
 EYTVY

>P12977

MYIMYAMAIRQAIRDRRRNPASRRDQAKWRLQTLAAGWPMGYQAYSSWMYSYTDHQTTP
 FVHLQATLGTGGRRCHVTF SAGTFKLP RCTPGDRQWLYVQSSVGNIVQSCNPRYSIFFD
 YMAIHRSLTKIWEEVLTPDQRV SFMEFLGFLQRTDLSYIKSFVSDALGTTSIQTPWIDDN
 PSTETAQAWNAGFLRGRAYGIDLLRTEGEHVEGATGETREESEDTESDGDDEDLPCIVSR
 GGPVKRPPPIFIRRLHRLLLMRAGKRTEQGKEVLEKARGSTYGTPRPPV PKPRPEVPQSD
 ETATSHGSAQVPEPPTIHLAAQGMAYPLHEQHGMAPCPVAQAPPTPLPPVSPGDQLPGVF
 SDGRVACAPVPAPAGPIVRPWEPSLTQAAGQAFAPVRPQHMPVEPVVPTVALERPVPYK
 PVRPAPPKIAMQGPGETSGIRRARERWRPAPWTPNPPRSPSQMSVRDRLARLRAEAQVKQ
 ASVEVQPPQLTQVSPQQPMEGPLVPEQQMFPGAPFSQVADVVRAPGVPAMQPQYFDLPLI
 QPISQGAPVAPLRASMGVPPVPATQPQYFDIPLTEPINQGASAAHFLPQQPMEGPLVPE
 QWMFPGAALSQSVRPGVAQSQYFDLPLTQP INHGAPAAHFLHQPPMEGPWVPEQWMFQGA
 PPSQGTDVVQHQLDALGYTLHGLNHPGVVSPAVNQYHLSQAAFGLPIDEDESGEGSDTS
 EPCEALDLSIHGRPCPQAPEWPVQEEGGQDATEVLDLSIHGRPRPRTPEWPVQEGEGQNV
 TGPETRRVVVSAVVHMCQDDEFDLDQDPPDEA

>P08392

MASENKQRPGSPGPTDGPPTPSPDRDERGALGWAETE EGGDDPDHDPDHPDLDLDDARR
 DGRAPAAGTDAGEDAGDAVSPRQLALLASMVEEAVRTIPTDPAASPPRTPAFRADDDDG
 DEYDDAADAAGDRAPARGREREAPLRGAYPDPDRLSPRPPAQPPRRRRRHGRWRPSASST
 SSDSGSSSSSSASSSSSSSDEDEDDGDNAADHAREARAVGRGPSSAAPAAPGRTPPPPG
 PPLSEAAPKPRAAARTPAASAGRIERRRARA AVAGRDATGRFTAGQPRRVELDADATSG
 AFYARYRDGYVSGEPWPGAGPPPPGRVLYGGLGDSRPLWGAPAEAEARRRFEASGAPAA
 VWAPELGDAAQQYALITRLLYTPDAEAMGWLQNPVVPGDVALDQACFRISGAARNSSSF
 ITGSVARAVPHLYAMAAGRFGWGLAHAAAAVAMSRRYDRAQKGFLLTSLRRAYAPLLAR
 ENAALTGAAGSPGAGADDEGVAAVAAAAPGERAVPAGYGAAGILAALGRLSAAPASPAGG
 DDPDAARHADADDAGRRAQAGRVAVECLAACRGILEALAEFGDGLAAVPGLAGARPAS
 PPRPEGPAGPASPPPHADAPRLRAWLRELRFVRDALVLMRLRGDLRVAGGSEAAVAVR
 AVSLVAGALGPALPRDPRLPSSAAAAAADLLFDNQSLRPLLAASAAPDAADALAAAAAS
 AAPREGRKRKSPGPARPPGGGGPRPPKTKKSGADAPGSDARAPLPAPAPPSTPPGPEPAP
 AQPAAPRAAAAQARPRPVAVSRRPAEGPDP LGGWRRQPPGPSHTAAPAAAALEAYCSPRA
 VAELTDHPLFPVPWRPALMFDPRALASIAARCAGPAPAAQAACGGGDDDDNPHPHGAAGG
 RLFGLRASGPLRRMAAWMRQIPDPEDVRVVVLYSPLPGEDLAGGGASGGPPEWSAERGG
 LSCLLAALANRLCGPDTAAWAGNWTGAPDV SALGAQGVLLLSTRDLAFAGAVEFLGLLAS
 AGDRRLIVVNTVRACDWPADGPAVSRQHAYLACELLPAVQCAVRWPAARDLRRTVLASGR
 VFGPGVFARVEAAHARLYPDAPPLRLCRGGNVRYRVRTRFGPDTVPVMSPREYRAVLPA
 LDGRAAAASGTTDAMAPGAPDFCEEEAHSHAACARWGLGAPLRPVYVALGREAVRAGPARW
 RGP RRDFCARALLEPDDDAPPLVLRGDDDGPGALPPAPPGIRWASATGRSGTVLAAAGAV
 EVLGAEAGLATPPRREVVDWEGAWDEDDGGAFEGDGV L

>P06934

MSRSESRKNRGGREEILEQWVAGRKKLEELERDLRKT KKKLKKIEDENPWLGNIKGILGK
 KDKDGE GAPPKRARTDQMEVDSGPRKRPLRGGFTDKERQDHRRRKALENKKKQLSAGGK
 NLSKEEEEELRRLTEEDERRERRVAGPPVGGVIPLEGGSRGAPGGGFVPSLQGVPESPFS
 RTGGLDIRGNRGGFPWDILFPADPPFSPQSCR PQ

>P24739

MARRSGNPDEDLLKAVRI IKLLYQSSPCPNPRGSRQARKNRRRRWRARQRQIDSLSERIL
SDCLGRPAEPVPLQLPPIERLRLDCESESCGTSQTPQGTETGVGGPQISVSSAVLGS
TKN

>P56260

MLDMGRKEVKMLPKSSFSINNLVPEAVQSDNHSGHSHHNSHHPHHHHHHHHHPPPPQQ
PQRAAAAEEDDEEKAPLLLPPPAAGALEAAKAEALAGKGEAGAAAEELEEKAAEEKKG
AAEGGKDGESGKEGKNGKYEKPPFSYNALIMMAIRQSPEKRLTLNGIYEFIMKNFPYY
RENKQGWQNSIRHNLNLCFVKVPRHYDDPGKGNWMLDPSSDDVFIGGTTDKLRRRST
TSRAKLAFKRGARLTSTGLTFMDRAGSLYWPMSPFSLHHPRASSTLSYNGTASAYPSHP
MPYSSVLTQNSLGNHNSFSTSNGLSVDRLVNGEIPYATHHLTAAALAASVPCGLSVPCSG
TYSLNPCSVNLLAGQTSYFSPTSLTPQ

>Q01552

MPPKRRLVDDADAMEDQDLYEPPASLPKLPKFLQYTVGGSDPHPGIGHEKDIRQNAVAL
LDQSRDMFHTVTPSLVFLCLLIPGLHAAFVHGGVPRESYLSSTPVTRGEQTVVKTAKFYG
EKTTQRDLTELEISSIFSHCCSLIGVIGSSSKIKAGAEQIKKRFKTMMALNRPSHGE
TATLLQMFNPHEAIDWINGQPWVGSFVLSLLTDFESPGKEFMDQIKLVASYAQMTTYTT
IKEYLAECMDATLTIPVVAYEIRDFLEVSARKLKEEHADLFPFLGAIRHPDAIKLAPRSFP
NLASAAFYWSKKNPTMAGYRASTIQPGASVKETQLARYRREISRGEDGAELSGEISAI
MRMIGVTGLN

(4) 16 Plasma membrane proteins

>P03470

MNPQKIITIGSICMVVGIISLILQIGNIISIWISHSIQTGNQNHTGICNQGIITYNVVA
GQDSTSVILTGNSSLCPIRGWAIHSDNGIRIGSKGDVVFVIREPFISCSHLECRFFLTQ
GALLNDKHSNGTVKDRSPYRALMSCPVGEAPSPYNSRFESVAWSASACHDGMGWLTI
GSGPDNGAVAVLKYNGIITETIKSWRKKILRTQSECTCVNGSCFTIMTDGPNGLASYKIF
KIEKGVTKSIELNAPNSHYEECSYCPDTGKVMCVRDNWHGNSRNPWVSFDQNLDYQIGY
ICSGVFGDNPRPKDGPVSGCPVSADGANGVKGFSYRYGNGVWIGRTKSDSSRHGFEMIWD
PNGWTESTRFSVRQDVVAMTDRSGYSGSFVQHPPELTGLDCMRPCFWVELIRGRPEEETI
WTSGSIISFCGVNSDITVDWSWPDGAELPFT

>P16071

MAEKGKTNSSYWSTTRNDNSTVNTYIDTPAGKTHIWLLIATTMHTILSFIIMILCIDLII
KQDTCMKTNIMTVSSMNESAKTIKETITELIRQEVISRTINIQSSVQSGIPILLNKQSRD
LTQLIEKSCNRQELAQICENTIAIHHADGISPLDPHDFWRCVGEPLLSNPNISLLPGP
SLLSGSTTISGCVRLPSLSIGDAIYAYSSNLITQGCADIGKSYQVLQLGYISLNSDMYPD
LKPVISHTYDINDNRKSCSVIAAGTRGYQLCSLPTVNETTDYSSEGIEDLVFDILDKGGK
TKSHRYKNEDITFDHPFSAMYPSVSGSIIKIENLIFLGYGGLTTPLOGDTKCVTNRCANV
NQSVCNDALKITWLKQRQVNVLIRINNYLSDRPKIVVETIPITQNYLGAEGRLKLGKK
IYIYTRSSGWHSHLQIGSLDINNPMTIKWPHEVLSRPGNQDCNWNRCPRECISGVYTD
AYPLSPDAVNVAATTLYANTSRVNPTIMYSNTSEIINMLRLKNVQLEAAAYTTTSCITHFG
KGYCFHIVEINQTSNLTLQPLFKTSIPKICKITS

>P03655

MAVYFVTGKLGSGKTLVSVGKIQDKIVAGCKIATNLDLRLQNLQVGRFAKTPRVLRIPD
KPSISDLLAIGRGNSYDENKNGLLVLDECCTWFNTRSWNDKERQPIIDWFLHARKLGWD
IIFLVQDLSIVDKQARSALAEHVYCRRLDRITLFPVGTLYSLVTGSKMPLPKLHVGVVK
YGDSQLSPTVERWLYTGKNLYNAYDTKQAFSSNYDSGVYSYLTPLYLHGRYFKPLNLGQK
MKLTKIYLKFSRVLCLAIGFASAFYTSYITQPKPEVKKVVSQTYDFDKFTIDSSQRLNL
SYRYVFKDSKGLINSDDLQKQGYISITYIDLCTVSIKKGNSNEIVKCN

>P01115

MPAARAAPAADEPMRDPVAPVRAPALPRPAPGAVAPASGGARAPGLAAPVEAMTEYKLVV
 VGARGVGKSALTIQLIQNHFVDEYDPTIEDSYRKQVVIDGETCLLDILDTTGQEEYSAMR
 DQYMRTEGEGFLCVFAINNTKSFEDIHQYREQIKRVKDSDDVPMVLVGNKCDLAGRTVESR
 QAQDLARSYGIPYIETSATRQGVEDAFYTLVREIRQHKLRLKLNPPDESGPGCMSCKCVL
 S

>Q04463

MPRPRPRALRGSSPGWALVAAVAVGAALLMATLAVAAPPGRGPAARSPGLEDASVEPVE
 DGDYDEYDDEGHTPTDVPVGSPEPSPGDRPPRGPGGGSGRRRGSPGNGTRSAARRQLRES
 LRRIQAEYAASAFYVCPPTGATVVQFEEPRPCPDVAAGKNFTEGIAVVFKENIAPYKFT
 ATKYYKEITVSQWQGSRYLQLTGLYNDRAPVPFSEITDLINGKGRCSRSDVTYTRSQRV
 TAYDGDDEWGREVALVPAKTSTPNSRGWYTTDRVYAPNAHAGFYKTGTTVNCIVEEMEARS
 AFPYDSFVLATGEFVYASPFSGFSEDARRERNRYAPDRFRQVDGFYPRDLDSGQRAATPV
 VRNLLTPTFTVGVWDWPKRPNVCSVTKWQVVEEMVRAEYGSAFRFTSAALSATFTSNLT
 QYPPELIEHSDCVAREAAESIEAIYARRYNASHVRVGGVQYLLAAGGFLAFQPLLSNSL
 AEMYRREALLGRSGDLAAALAPPVAAAPASGAGPRGTISTTQTVEFARLQFTYDHIQKHV
 NEMLGRIAAAWCQLQNQELVLWNEARKLNPGAIASATVGTTRVGARMLGDVMAVSTCIPVS
 PDNVMQNSMRI PGDPKTCYARPLVSFRYTDEGELVEGQLGEDNEIRLEQNNVEPCTVGH
 KRYFVFGGGYVYFEEYAYSHQVSRADVPVSTFVDLNLTMLEDHEFLPLEVYTREIKDS
 GLLDYAEVQRRNQLHALRFSDIDRIMNDSANAALMAGLARFFDGMGDAGKAIGRAVLGVT
 EGLISVVSGVSSFLSNPFGALAVGLLVLAGLVAAFFAMRYIMRLRANPMRALYPLTTSGI
 KAERAAALGSGGDKGGAGDGAAGVEDFDEAKLEAARDMIRYMTLVSAMERTAHKAKKRGT
 SARISAHLTDMVLAQAQAEYQPLSKDDEDGADP

>P17504

MKAIIVLLMVVTSNADRICTGITSSNSPHVVKATATQGEVNVGTGVIPLTTTPTKSHFANLK
 GTKTRGKLCPKCLNCTDLVALGRPKCMGTIPSAKASILHEVKPVTSGCFPIIMHDKIR
 QLPNLLRGYEHIRLSTHNVINAETAPGGPYKVGTSVSGPCPNVTNGNGFFATMAWAVPKNDN
 NKTATNPLTVEVPYICTEGEDQITVWGFHSDNEAQMVKLYGDSKPKQKFTSSANGVTTHYV
 SQIGGFNPQAEDGGLPQSGRIVVDYMVQKSGKTGTITYQRGILLPQKQVWCASGRSKVIK
 SLPLIGERDCLHEKYGGLNKSPPYTGHEHAKAIGNCPIWVKTPLKLANGTKYRPPAKLLK
 ERGFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSEL
 EVKNLQRLSGAMDELHNEILELDEKVDLDRADTISSQIELAVLLSNEGIINSEDEHLAL
 ERKLLKMLGPSAVDIGNGCFETKHKCNQTCLDRIAAGTFNAGEFSLPTFDLNLITAASLN
 DDGLDNHTILLYYSTAASSLAVTLMIAIFIVYMVSRDNVSCSICL

>P69336

MLHVVPLEWTVVEVVPYLERLAVWLRASVLVAFQLTATVALSVLSWWMPPPVAELCERG
 RDDDPPLSHLSLVVPVVGCLFLLLRGPSIDRCPRKLPDLLAYCLPHALAFLLTLLMCQPS
 QAFVGAALLALAVDLSCLGASLLGCDPGASLRRLWLPVLSLLCATALGLWLLRAAPFF
 LGLHATLLTTLMLIHDLSLITCQSSFPEFQPSLRLYVENVALFIGMYHLLRLWLWSP

>P52380

MDTVIELSKLQFKGNASCTSTPTLKTARIMESAVTGITLTTSSIPMIIIVVTTMILYHRVA
 KHNATSFYVITLFAFDVLMWCVFMTVNRKQLFSFNRFQCQLVYFIYHAVCSYSISMLA
 I IATIRYKTLHRRKTESKTSSTGRNIGILLASSMCAIPTALFVKTNMKTGKCVVYI
 SSKKAYELFLAVKIVFSFIWGVLPMTVFSFFYVIFCKALHDVTEKKYKKTFFIRILLLS
 FLLIQIPYIAILICEIAFLYMPQNTCFWLARVEILQLIIRLMPQVHCFSNPLVYAFTGGE
 LRNRFTACFQSFPPKTLCTSTQKRKDSASEHDQNSKSKASVEKNQPL

>P09703

MTTSTNNQTLTQVSNMTNHTLNSTEIYQLFEYTRLGVWLMCIVGTFLNVLVITITILYRR
 KKKSPSDTYICNLAVADLLIVVGLPFFLEYAKHHPKLSREVVCGLNACFYICLAFAGVCF
 LINLSMDRYCVIVWGVELNVRVNNKRATCWVVIWILAVLMGMPHYLMYSHTNNECVGEF
 ANETSGWFPVFLNTKVNICGYLAPIALMAYTYNRMVRFIINYVGKWHMQTLHVLLVVVVS
 FASFWFPFNALFLESIRLLAGVYNDTLQNVIIIFCLYVGQFLAYVRACLNPGIYILVGTQ
 MRKDMWTTLRVFACCCVKQEIPYQDIDIELQKDIQRRAKHTKRTHYDRKNAPMESGEEEE
 LL

>P09721

MVQIQFHQGEPLGHKKEKPPPVSPPSPPPIRRVTVITKDEDTLRSVQHFLWMVRLYGTVV
 FQTSATIATTILFMLIPWRVTAPYLRDTPFWSTLLPCALRCHAYWLERRRRPGLTMLVM
 VYTTLTITISVSTIGLCFDRTVVIQAYVLSSMLCVWCTGLAWLMAWNMQRRLAAILCLLSFM
 LPILWLFIAVQSWEPYQRIILALTVSFIYGLKIVLIRDTLTVLYRSPSNCYTDGDLRLTA
 MLLYMDQVIMFLLVVVPLTAPIWYPNYAGALGRTAHWLWLFHK

>P09724

MRVISRARSACTWTSCTSLSPCSTSCPPSPAAPTLLRRRSLPQQRRRPSSSPNRRVRGVT
 TSPCPTRSLVYKRRVQRLCAETVATMQAQEANALLSRMEALEWFKKFTVWLRVYAI
 FIFQLAFSFGSGVFWLGFPPQNRNFCVENYSFFLTVLVPVCMFITTYTLGNEHPSNATVL
 FIYLLANSLTAAIFQMCSESRVLVGSYVMTLALFISFTGLAFLGGRDRRRWKCISCVYVV
 MLLSFLTLALLSDADWLQKIVVTLCAFSISFFLGILAYDSLMIFFCPCPNQCIRHAVCLY
 LDSMAIFLTLMLLSGPRWISLSDGAPLDNGTLTAASTTGKS

>P12555

MDRAVSQVALENDEREAKNTWRLVFRIAILLTIVVTLAISAAALAYSMEASTPSDLVGIPT
 TAISRTEEKITSALGSNQDVVDRIYKQVALESPLALLNTESTIMNAITSLSYQINGAANS
 SGGGAPIHDPDYIGGIGKELIVDDASDVTSFYPSAFQEHNFIPAPTTGSGCTRIPSFDM
 SATHYCYTHNVILSGCRDRSHSHQYLALGLVLRTSATGRVFFSTLRSINLDDTQNRKSCSV
 SATPLGCDMLCSKVTEETEEDYNSAIPTSMVHGRLGFDGQYHEKDLDVTTLFEDWVANYP
 GVGGSFIDNRVWFPVYGGKPNSPSDTAQEGKYVIYKRYNDTCPDEQDYQIRMAKSSYK
 PGRFGGKRVQQAISIKVSTSLGEDPVLTVPPNTVTLMGAEGRVLTVGTSHFFYQRGSSY
 FSPALLYPMTVSNKTATLHSPYTFNAFTRPGSVPCQASARCPNSCVTVGVYTDPYPLVYR
 NHTLRGVFGTMLDDEQARLNVPVSAVFDSISRSRITRVSSSSTKAAYTTSTCFKVVKTNKT
 YCLSLIAEISNTLFGFRIVPLLVEILKDDGVREARSGRLSQLQEGWKDDIVSPIFCDAKN
 QTEYRRELESYAASWP

>P25049

MENQREEIELTEIPETEKKRRTAEKLLSCSAETAEEKVSLCSEETTDTSSSSSSEQTPAP
 IEVNVNIQTSTYLPQNAATNLNSLYTSFEDARAQGGKGLVRHNSDDLKSFLEKYPDYRKP
 KRDLSESWDPGMPKPTLPPRPANLGASQASTVRRHVREQNFKQLRERKANEGKIVKDLKR
 LEYKVNIILCLVVILAIILLTGLSILFIRIKS

>P00545

RMPSPGPHYGASAEETPGPRPPLCPASSCCLPTEAMGPRALLVLLMATAWAHQGVPVIQPS
 GPVELVVEPGTTVTLRCVNGSVEWDGPISPHWNLDDLDPSSILTTNNATFQNTGTYHCTE
 PGNPRGGNATIHLYVKDPPARPWKVLAQEVTVLEGQDALLPCLLTDPALEAGVSLVRVRGR
 PVLRQTNYSFSPWHGFTIHKAKFIENHVYQCSARVDGRTVTSMGIWLKVQKDISGPATLT
 LEPAELVRIQGEAAQIVCSASNIDVNFVSLRHGDTKLTISQQSDFHDNRYQKVLTLNLD
 HVSFQDAGNYSCTATNAWGNHSASMVFRVVESAYSNLTSEQSLLQEVTVGEKVDLQVKVE
 AYPGLESFNWTYLGPFSQDKLDFVTIKDITYRYTSTLSLPRLRSESGRYSFLARNAGG
 QNALTFELTLRYPPEVRVTMTLINGSDTLLCEASGYQPQSVTWVQCRSHTDRCDSEAGLV
 LEDSHSEVLSQVPFYEVIVHSLLAIGTLEHNRTYECRAFNSVGNSSQTFWPI SIGAHTPL

PDELLFTPVLLTCMSIMALLLLLLLLLLLLLLYKYKQKPKYQVRWKIIESYEGNSYTFIDPTQL
PYNEKWEFPRNNLQFGKTLGTGAFGKVVEATAFGLGKEDAVLKVAVKMLKSTAHADKEA
LMSELKIMSHLGQHENIVNLLGACTHGGPVLVITEYCCYGDLLNFLRRQAEAMLGPSLSV
GQDPEAGAGYKNIHLEKKYVRRDSGFSSQGVDTYVEMRPVSTSSSNDSFSEEDLGKEDGR
PLELRDLLHFSSQVAQGMFLASKNCIHRDVAARNVLLTSGRVAKIGDFGLARDIMNDSN
YIVKGNARLPVKWMAPESEIFDCVYTVQSDVWSYGILLWEIFSLGLNPYPGILVNSKFYKL
VKDGYQMAQPAFAPKNIYSIMQACWALEPTRRPTFQQICSLQKQAQEDRRVPNYTNLPS
SSSRLLRPWQRTPPVAR

>P09716

MSPNSEATGTAWAPPPRPSRGVIMISSVSTNDVRRFLLCMRVYSTVAVQGTCTFLLCLG
LVLAFFHLKGTVFLCCTGFMPPLSLMVPTICLALLHGKRDEGSFTSPPSPGLLTIYSVLT
TLVIVASACSSSTLVTFSGLLACVLFSLCSCVTGLAGHNHRRWQVIVTLFVIGVIAFLI
ALYLQPVPLGHKFLFLGYAMALSFMLVVTVFD'TTRLFEIAWSEADLLTLCLYENLVLYL
LILILFTTEDSLDKLIAWMTWLSSRATGATNAASISGCDLLREVQRNLTRTMA

>P03701

MRNVCIAVAVFAALAVTVTPARAEGGHGTFTVGYFQVKPGTLPSSLGGDTGVSHLKGINV
KYRYELTDSVGMASLGFAASKKSSTVMTGEDTFHYESLRGRYVSVMAGPVLQISKQVSA
YAMAGVAHSRWSGSTMDYRKTEITPGYMKETTARDESAMRHTSVAWSAGIQINPAASV
VDIAYEGSGSGDWRTDGFIVGVGYKF