

>ileS

MKLKETLNLGKTAFFPMRAGLPNKEPLWQAAWDEAQIYQKRQKLNKGKPSFHLHDGPPYANGNIHVGHALNKIS
KDIIVRAKSMMSGYNAPYVPGWDTHGLPIEQVLAKKGKIRKEMDLAEYLEMCRDYALSQVDKQRQDFKRLGVSA
DWDNPYVTLDPAYEADQVRVFGAMADKGYIYRGAKPVYWSWSSESALAEAEIEYHDIDSTSLYYANKVKDGKG
ILDTSYIVVWTTTTPTVTASRGLTVGPDMDYVLVQPASTSKKYLVAEGLLDSLAPKFAWTDFEILSHHKGSE
LEYIVTEHPWDSVDDELVLGDHVTLDSTGIVHTAPGFGEDDYNVGTKYKLEVAVTVDERGMMMENAGPDFQ
GQFYDKVLPTVIEKLGDLALLAKEVINHSYPFDWRTKKPIIWRAPQWQFASVSDFRQEILDEIEKTTFHPSWGK
TRLYNMIRDRGDWVISRQRAWGVPLPIFYAEDGTAIMTKEVTDHVNLFKEKESI IWWKKEAKELLPEGFSHP
GSPNGEFTKETDIMDVWFDSSGSSWNGVMNAREQLSYPADLYLEGSDQYRGWFNSSLITSVAVNGHAPYKAVLS
QGFVLDGKGEKMSKSKGNIISPNDVAKQYGAEILRLWVTSVDTDNDVRVSMIDILGQVSETYRKIRNTLRFLIA
NTSDFNPNTNRVAYEELAPVDKYLTIKFNKLVEVFNNSYENYDFMAIYKAVVNFVTIDLSAFYLDFAKDVVYI
EAANSLARRRMQTVFYDILVKITKLLTPILPHTAEIISYLEFEFEFEFVQLTEMPVAEIEFEGEEDILSAWDAF
MTLRTQAQKALEEARNEKVIKGSLEAHLTIYASDEVKTLLTALDSDIALLMIVSQLTIIADFAADAPTNAISFEN
VAFTVDHAVGEVCERSRRVDPTRMRSYGAFVCDASAKII EENFPEAVAQGFESN*

>pbuX

MQTETQHSQSASAILGLQHLLSMYAGSILVPIMIAGALGYSAKELTYLISTDIFMCGLATFLQLQFNKYFGVG
LPVVLGCAFAQSVAPLSIIIGAKQGGSGAMFGALIASGLFVILIAGVFSKIARFFPAIVTGSVITTIGLSLIPVAM
GNMGNNTPKPTGQSLILAFLTIFIIILAIQKFATGFIKSIAILIGLIAAGTLVAALMGLVDTSAVSSAPVHVPT
PFYFGAPKFEITSIVMMCIIAIVSMVESTGVYLALSDITDETLDNRLRNGYRAEGLAVLLGGIFNTFPYTGF
SQNVGLVRLSGIKTRRPIYYTAAFLVVIIGLLPKFGALAQMI PPSVLGGAMLVLFGMVALQGMQMLNRVDFQGN
EHNFI IAAVSISAGVGFGNGTNLFASLPSTANMFLTNGIIVATVTAVLLNLI FNGNTKEN*

>ftsY

MGLFDKLFKGKTKKEQQDEKDQVENELGVDSSFPPEESQAEEQNIFEKPANSQHIEEELSLSADPASIEEVIIQQ
QELTQSENLETVIEGDHLNDNKDIQGENSDLERQEQVSDHADHQLDNLVENAEENQEQLSTEKTYISIVDEYYQ
RKAAFEGLQNGDFEPLSQNDITPSPKPQLEESDQEKYQRSCLKKTRNGFAARLNAFFANFRSVDEDFEELEE
MLILSDVGVSVAATELTELRLQEAKLENAKKPDDLKRVII EKLVDIYEKDGTFNEAINLQEGTLVMLFVGNGV
GKTTSIGKLAYRYKAEGKKVMLVAADTFRAGAVAQLAEWGRVRDVPVVMGAEKADPASVVF DGVEKAVAQGV
ILLIDTAGRLQNKENLMAELEKMGRIIKRVI PDAPHETLLALDASTGQNALSQAKEFSKILPLTGLILTKIDG
TAKGGVLAIRQELEIPVKFIFGGEKIDDIGEFDSDFMKGLLADII*

>group_415

MRKFLMSCFAALLLLFAGVVSQADADQYLRVGMEAAAYAPFNWTQDDNSNGAVPIEGTNQYANGYDVQVAKKVAK
SLNKKLLVVKTSWTGLIPALTSKIDMIAAGMSPTEERKKEIAFSDSYTSEPVIIVKADSKYAKAKSLDDFA
GAKITAQQGVWHVNLI PQINGVKAQTPMGDFSQMRQALSSGVIDGYISERPEAMTAENANSFAKMMVVLKKAFT
VNESDAAIAVGMRKDDPRIVQVNTVLADLSANDRLDLMDKMVTLQPKKKAENGAQPSFLDQMWSIVTKNWKQ
FLRGTLTLLISTIGTIVGLIIGLLIGIYRTAPKSKHKVLAFFQKLFGWFLNVYIEVFRGTPMIVQSMVIYYG
TAQAFGISIDRTLAAIFIVSINTGAYMTEIVRGGIFAVDKGQFEAATALGFTHGQTMRKIVLPQVVRNILPAT
GNEFVINIKDTSVLNVISVVELYFSGNTVATQNYQYFQTFSVIAVIYFILTFTVTRILRYVERRIDDDNYTTT
VNELP*

>group_704

MKVEIIEERNKKVLKKTKKDGSISYTIKGAYLGIDKKTGKQVTTTTISAPTLKALDRKYQDVRREFEENDSTRK
ETAIEGTIQDLAEIWFDSYKTWVSSDNTRNKVRGYLDTYIIPRFGDYKIDAIESTEVEQQWVDSLAEQARSEVA
KGKGKAEKGKASDYGAVVHKLIDILDFGMVHFNLKFNSAKQVRIPPKPRAKNKRVRLHEKDLAKWLNLYLDTL
PNTRANRRFKIICNTLLASALRINELLALEIDDLHIDTNEIEVNKTLMWRSANKKLGTGKGVCKNSAKTESG
NRRITVPESIMQDLLNIHNEMNEYFKKHNLPESKLIFPTIYGNMYTDRNERATLIKRLKSLGLPDTGFHLFRH
THASLMLNSGADWKELQERMGHKSISTTMDIYAELDPNRKNEAVDILLEKLSDIKNSN*

>group_2949

VISSTPTNGLSDENGNLVGTPTVDDWGTDEETREVKVPVKVTNTLPDGTEEVVEVEVPVTNVRTDGDGIID
SEDPDDDNDGVEDGSDDTPKAFTGFTATAKPVTVPEGQAVPADTKVNTLPLTGDSSEAFFTSAAALAISSVGL
LGIAGKKNKNGSD*

>group_241

MLEIKIAVNDPQIMAQDGSSILRSLQNKSLPLVDLIVRESIQNSLDATLEESDTTKVDIKVDKFDTLTCSNHF
QGIETELKEKYLNNNYFLSIRDSNTSGLSGKVIGESIAELNDSNFFKLVFALGKNQEKAGSGGSGWGLGKTSYF
RIGNGIVIIYYTRVKSETGYEERLIASLIEDNKEEDRLLKNSDRGIAWWGKINNEDVILPITNSDEISEILSIF
KIRQYVGTETGTTVII PYLSNEFIHRESTKNQIEKSCFWLNDLSEEIKIAVQRWYFPRLNNDTYRLHFNQSKL
ICSINDSLEINYFPEVTFKLLQNIYNSALLGNTIDENTINVYPIRLPRKGAVDVNKPIGHIAFMEVSKEDAKM
LPPDNLQSPLAYLGNVDEKQNTHEYKSKFLCYCRKPGMIIEYDFTDNWLPNGNIQKENHLLFSVFPVNSDLELS
QTTYKTYPNVENYLRSIENADHAKWEDEHGYTFVKRTKSYIIDTVKKHFDNDENSTVGAATNRLSKILGNALL
PPTGFGKRSSRVVTGNSNINRNSSSKASISVEDVSFVDFKTVKVLVKGRIIKNRKTKIKILVETQEGKIDKDR
WEGDLNTIDFPFIITS AKFINDNFKLQEDYIVEGNCILINELPTDKNIEVEILINKSTNQYSPVISITS*

>group_4349

MTKDELIARLRLSLGEQLNRDVSLTGTKEELALRVAELKEELDDTDETAGQDTPLSRENVLTGHENEVGSQPD
TVILDTSELVTVVALVKLHTDALHATRDEPVAFVLPGTAFRVSAGVAAEMTERGWPECNNGSAVADSITCSML
PLPAPMKRYAGTWERQPPPLHPVSSQVR*

>group_4355

LQAAKQELTNLIASAKTLSASGKYDDATTTALAAATQKAQTALDQTNASVDSLGTGANRDLQTAINQLAAKLPA
DKKTSLLNQLQSVKAALGTDLGNQTDPTSTGKTFTAALDDLVAQAQAGTQTDQLQATLAKVLDAVLAKLAEGI
KAATPAEVGNAKDAATGKTWYADIADTLTSGQASADASDKLAHLQALQSLKTKVAAAVEAAKTVGKDDTTGTS
DKGGGQGTTPAPAPGDTGKDKGDEGSQPSSGGNIPTNPATTTSTSTDDTTDRNGQLTSGKGALPKTGETTERPA
FGFLGVIVVSLMGVLGLKRKQREE*

>group_4438

VPVPVVTTTRTEVIAHTTKYIADDTLAYGQRQTVVAGVDGEKTYTTTTDGVEDAGVVTKAMVAEEIRVGTKPTVE
TVTIAKPADVEEADATLAKGTRVLKTAGSTGSTTTTTTYSMDATTGVVTANTPTVETVAASPDVYRVGTKEDV
IEIVEEIVEKTKSFGKVVINNPDLLKGISRQVRLGKNGIIEVYKIVKVNGLEISREIIESEKEIEKPVDEIVE
IGTKEPVENQNQKVIKANESKMNSTDVKKVENKGELPKTGEIKESVFGMFSVIMTGLFALVTFLGKRERK*

>group_1680

MKNITKIITLTSVVALTSLAPLDIFADQVDSTTPSTGVVVPGDIA PSTPVPBGDTTTPSIDPSTPVPBGDTTTPS
IDPSTPVPBGDATTPSIDPSTPGSGDATTPSIDPSTPGSGDTTAPSSDPNTPKPEDTTSPSTSEPGDTSVPVTP
GIIDKVDPNTPGSVTIKPIPVNPGTSVIGTQDGNLIVLDNEGKQSIVSAESLGGRVNNNGTVTIKDAEGKEKTL
PHTGQETTFIVQLIGFFSILISGVIILIKKKFGNS*

>group_181

MKKTIFLAGLATVGLLQGTQVKALDATNVTTTTDPDPQPSAPTNPVGVSEQDVAAAQTA YNTATTNAATAQAAYTT
QANAVSTTTTAVNTAKAELTSQQSQVSALEQTASTVSTNVAQTSEQLAAAETAAATAQADATQAQDTYNASVS
EVTAKTAELQAAANQTTTQTETTTTVTTTTPATTTNDANQPYGKTAPVSYSGDKTEVITLNSSQVTDLTTKGEFT
YTPDYKAVSHFAVDYINEIRALNGIEGRVYYDDTTVEIAIKRAEEMLAAKTLSHFTNLASQYPQYAGDNAAMV
SDSFTYTGLKTLSGVYSDKEYAYKLVLWFSDHTNVFADDTNNGALNYGHRNALLEHNGGLAVASAHQPLDPT
VPLTQTNRQQLGFSDMPFTVNHQITQWVEIMPGVMEEQTIGDGRVSLDFTYVANADGELVQYLNNGKEVVF LPR
TVFEYVYTNTQRQVVDAAAVQALNDYKKSAAARLLAEKTAMETTQSISAAKQVSVLALQEKLKQLQSEQATINA
DLSAAKAQVAATTEEFNSLLNKLESEAAQLKALEAANKAAQALRSQLFQTYKDVLAARQATILAKEYASILAE
NKIPVAVVDSTGKIIDYTAQDKPVDTPVDAPVDTPVDAPVDTPVDAPVETPVDTPVDTPVNAPVGFYNNKNT
SGSKQKIQFVSSKIIDNKFLQNYAGKIELPSTGEDNSVSYAILALSGFSGTLLFIKKRKVQ*

>group_159

MLNFKIHTATCIGNSSNCIYPNEVLVSDRDSFIKAISFDHVCGSFNGSYRSKDNFIKSDCIPMDCDNDHSDDP
DDWVTPFDVALAFPGVCFFASYSRNMKVGNKSARPRFHVYFPIEEIKDAGEYSSYKERLYKEFSYFDDNAL
DAARFIYGISNPEVELYDGDLTVIDYLGRRKFE DLPIILGSQIQEGSRNSTLSHFAGIILKRYGKSEKAKKAFL
EESEKCNPLDREELSLIWKSAISFYENISKQKGYIPPDEYKKVSWEKPLPFTGEKMPDFPIEALPKALRN
YIAVGKSTQTPVDMAAVGVLATVSACMKNLYKVEGKADWHEPTNIYSVIIAEPSEKRSAVISLVIKPVDEYIKK
YNQIHKVEFEMSKVIKQRLNKKNSLLSQSKKKGEDKTASEFNDEIRSVVEELVNFTE SKPLKVYVDDTTTEK
LTESLAENNNAIISISSEGGIFDVISGTYSSKVNIDVFLKAYSGENISVDRIMRNSIYVENACLSILLSVQPV
VIGELMRNKKFRHRGLTARFLYTTPQSFVGKRTLESECISKDVYREYKELIDNILMEEKTGNTQIIKLSEKAK
ELLKEYFDWVEQKLVGFTMYSDWLGLVGNLT LRIAGILARSSVIKKDVG DALLEEDSPIVIDEEVFSNAVKI
GKYFLVHAVNAYGDMGVRSDFKAALMVLEKLKEKELVNITRREVMRLCRWVGSAAEEAQSILDNLEDYGYIRLS
EIDPAEKMRNGRPKNVVYSINPSVLSE*

>group_610

MFRRSKNNSYDTSQTKQRF SIKKFKFGAASVLIGISFLGGFTQGQFNISTDTVFAAEVISGSAATLNSALVKN
VSGGKAYIDIYDVKNKIDPLNLIVLNPSNYSANYIYKGGSTFTSVNQLQTPGTATITYNILDENG NAYTKS
DGQVDIVSLVTTVYDTTEL RNNINKVIENANDPKWSDDSRKDVLDKIEVIKNDIDNNPKTQSDIDNKIVEVNE
LEKLLVLPVPDKDKYDPTGGETTVPQGT PVS DKEITDLVKIPDGSQGVPTVIGDRPNTDQPGDY PVTVEVTYP
DGTKDTVIVTVHVTPTPDKDKYDPTGGETTVPQGT PVS DKEITDLVKIPDGSQGVPTVIGDRPNTDQPGDY PVT
TVEVTYPDGTKDTVIVTVHVTPTLPTHK CSTSSKKLPGLQNSSKVS KLPATGDEISPFFTSIALTIIASAGLM
LNKKKDN*

>group_0

MRYFDKSQTKQRF TIKKFKFGAASVLIGIAFLSLMGETASAEVISSTSI VADTTSNATNSPDSTALTPTSGQS
TTPSTSTPVESNTSSSSVPI SATLVNPTAVEVPPTAPTANALYT DSTLLSGIASAGDTITVTFQNGITATSIV
QPDGSWSLSIPEGMTLNTGDVITIFSTNQGLKSSLT SVTVIASQSELETPVASPSSGTNSLDLVNNSISPL
STSSSTA EVQPTALSSNTTTTVQPIITSSDISTTQPTASNNVT LKPFNSASVSLSALT TATAGDSSLTVDGKNV
IESAAFNGTNNNTVNGKV VQWGTGNPLEINNP AVATPLAGVRIYAQWVEKSGAVSPIYTTVTMADGTYHIVMQ
DFVLPDGTIVQFDADPNFP EGEKFRIWADTPAGLQLYYSWESGQLGPQSAVMDTSFNASNAIGPNQLNNFNFI
YVSQTQSNVMHNLGDATTSIPAPGETGFINGTVFWNNAVNFGAQTMGASVTNSPGSVDT PATNITVVGSYLSD
YALQQIYTKAATAIGTSNIRGVGWT DANEQALQNWIKQQIAIEGTS LWIGETVTTTTDANGKYNLQFN GTYGK
EWNDRGFNGLLLDSALNSNTVLPDAIAISEGLPIGSKYSDFLNVHSSSPQIGTWYGDASAGTNAMTTDAAPKH

INLDWTVYVSLLDVGAYGFSSPYGNKFLFDTASTSWTNTLKVDAVADQYITGANFGLFLDQLYFDIVNFNTTI
TPASPGNTVTTAVQGLPSPLDGNKYQIVWFDKNGAEVATSPVQSPSTTGTLP SAD FVPTTAQNGDVYTAKI
FAINPTDDTTTRAVYPLAVDSFVVSVDATTNQPDYIDTTVGAGTSVTIAAPLNTNGTAVPMGTTFTASDPATL
PSWIVVDSSGSLTVSPGVTVSPGIVAVPVTVNYPDGSSEVIDVAITVVDVTAPTITPIANVTTPENKVMTPIT
VQVNDSTATTTVTGLPTGVIYNPTTGQITGTPTIPGTYPITVTSTDTAGNTSIRFTTITVTPDITTVSTAQTVN
PVTSEDTKVTGIGVPGNTITVTFPNGTTGTAVVQPDGTWGVVPVPTGVDLVGGEVLPVTSTTPAGDVSPTTNVT
VKDVTPPPAPPVNSVTSEDSVITGTGEPGDTITVKYPDGSIGMTVVKPDGTWTITIPATVDLVGGEVLPVTET
DPAGNVSPVSSVTVKDTPPAAPTVPNPVTSEDTAVTGIGVPGNTITVTFPNGTTGTAVVNP DGTWSVPVPSGV
DLIGGEVLPVTSTDPSGNVSPPTTNVTVQDVTPPAAPPVNPVTSEDTVVTGTGVP GDTIIVKFPNGTTGAAIVQ
PDGTWTVTIPPSVNLVGGEVLPVTEIDLAGNVSPATSV MVKDTTPPAAPTVPNPVTSEDTTVTGIGVPGNTITV
TFPDGSTGTAVVQPDGTWSVPVPAGVDLVGGEVLPVTSTDPAGNVSPPTTNVTVKDVTPPPAPPVNPVTSEDTV
VTGTGVP GDTITVKFPDGSTGTAVVQPDGTWSVPI PAGIDLVGGEVLTATEADPAGNVSPASTVTVTDDTTAPN
VPTVPNPVTSEDAAVTGVTGTPGNTITVTFPDGSTGTAVVQPDGTWSVPVPAGVDLVGGEVLPVTSTDPAGNVSP
TTNVTVKDVTPPPAPPVNPVTSEDTVVTGTGVP GDTITVKFPDGSTGTAVVQPDGTWSVVI PSGLVLDKGGEV
PVVEIDPAGNVSPASSVTVTDDTTAPLAPPINTVSSTSTTVTGVEPGNTVTVTFPDGSTSTAVVQPDGTWSAP
IPAGVTLSGGEIITATQTD PAGNVSPASSTTVSDVTPPTPAVNPVTSEDTVVTGTGTPGDTLTVTFPNGSTAT
TTVQPDGTWTVTIPAGVDLVGGEVLPVTETDPAGNSASVTVTVIDKTPPATPTVDPITSEDSVVTGIGVPGNT
ITVTFPNGITGTAVVNP DGTWSVPVPASEDLVGGEILPVTSTDPSGNVSPPTTNVTVKDVTPPPAPPVNPVTSE
DTVVTGTGVP GDTITVKFPPTGTTGTAVVQSDGTWTVTIPTGVDLVGGEVLPVTETDLAGNVSGATTVTVTDDTT
APDAPTVPNPITSEDSVVTGIGVPGNTITVKFPDGTTGTTAVNP DGTWSIPVPAGVDLVGGEVLPVTSTDPSGN
TSPTTNVTVRDVTPPPAPAVNPVTSKDPVVTGKGVP GDTITVKFPDGSTATSVVKS DGTWSVSI PVNVILVGG
EILPVTETDPSGNVSPETT VVVQDVTSADAPTVPNPITSEDSVVTGKGVP GNTITVTFPNGTTGTTVVQPNGTW
TVSVVPV GIDLVGGEVLPVTSTDPSGNISSTTNVTVKDVTPLAPPVNPVTSNDNVTITGTGVP GDTITVKFPDG
TIGRGDVQPDGTWTVSIPV GIDLVGGEVLPVTETDPSGNVSEATTVTVTDDTTAPDAPTVPNPITSEDSVVTGKG
VPGNTITVTFPNGTTGTTVVQPNGTWTVSVVPV GIDLVGGEVLPVTSTDPAGNTSPTTNVTVKDVTPPPAPPV
PLTSTDSSVVTGTGVP GDTITVKFPNRVTVTTIVQLDGTWTVTIPVSVDLIGGEVLPVIETDPSGNVSPATSVT
VTDTTAPVVDVNPVKLGDEVTGTSEP GTTITVTLPNGTTVTTVTDHNGNWI VPSPLQPGETVTAVSTDSSG
NTSQPSSETVPLDTTAPVVDVNPVKLGDEVTGTSEP GTIITVILPTGTTVTTVTDHNGNWTVPVSP LQPGET
VTAVSTDSSGNTSQPS SATVPLDTTAPLV DVPVKPGDKVVTGTSEP GTIITVILPTGTTVTTVTDHNGNWT
VPVSP LQPGETVTAVSTDSSGNTSQPSSETVPTMSVENTSATIGQVKTS GNNDKVN NVNVS QLPKTGEKTNIL
ALLESLSIATAGLLL MNKKKNK*

>group_28

MNTNNTKGHG YFRKKKWKGLASGIALAGAVAFSSGAVLADEVTT PASEMETPAVVAQPESMNDNNAYATQAN
TSTGTQTVVIDNSTVEAAAQTATEIGVAVTKEATVDQGTDTNTTDLTASKAEIKADQDQ QIKKIEAATQTQAD
NNAA YNEAQA AIDANNDFAEDATAKHEKDRVTVTDDNSTTTDGSATKNKQAAETAKKVLTTNQTAVQTYIGE
KAVYDATVAQATTLN HAVEAATLELKNKGVTVTTTAKVVSTVAEVKALRKQNEEAIAKANQQVTLNTAILAAY
DKTKQASDATNTDADKKSDELKSGV TITSSSTEVS SAAEEAESIASQNQSAYDQAKQKQADWQKKYDELKAKT
GTEGYTKEVVLQAVDLAAANPQARHGSSAAGATVVGTRDIASSSGRSGYARILDSTGVIKYANVGAGWTT EID
YTNLKGTLVTTADGHKHDISRIHRTFRLVNNGATGLNDVYIPNDPTEGFIVARNNGTDSYSDFMNFVTDNYY
YTANGKEVAFTATEKTPLALTYSSLNNNQIGREGARATGNGSKMVEINGSTVSVHSDQFAYSNHYNRQEETGV
NWDMDTSPYQYKGAAIGVFTQGEAFTTEFVQWDGPADSGGQTYWF AINTKVVSPPVQVPASATIVKTRVKPVS
VEPVSAQLVKAVNPEKPRLELTKLADTRNQVSVSYHDYKLSYKPTVTKT VTDTDKKNTDGATVVKNTPQVYT
LNHDNIYGNIKVGD TITIVDPLEAGAVPLVEDNAADAKTKGWTVSYDDTKATYTYTATYQGGKLAAPT IKWTP
IYDKGFFDNTYKVFKNDYEAFSNTVNTYPEPPKPVKSITDNSGADINGAKTFDRNVNFHVTTDYSPTYTTVTA
SSEAILKGF AIMDDVQDGAFTVNEEAITATDIEGKDVKDLFTMYHVLSDDARTQVIQDILDQANLLPVGEFY L
WVAKDPASFYANYVKQAKNVTIDL PARLLVPEGEVVKNDFFQVDFGNSYQSNLVTV EVPDVAPEKHALDQSDD
TII LDDQTVQIGDYIRYLLDGVTVPVKHDTLWQYDGD KLDVVDHRYTGNWKGI IKGTEYTAKENLVLT YDVT
LEDGTLIKAGDTIKAGSKYAFTFEFDQDTNSDFIKKIVTVTWDEKKGEWAYS IDEEFLRSLGVEGTFDADFY L
EVERIAAGEVENTFVNIDNGKEMTAKVTTHTPEPPQPTPPAPNTPAK PANMLPSTGEASSRLAILGGILLSGF
GLAGVRKRKED*

>group_1597

MKKKQLLSALALSTLVLSQAGFVSADEVTPIDPTAPSTE VVVPPTDTPAVTTDDTTSSDSGTTLPVDPTTPSTD
PTTDNSQGKDTAGDDV VAPDTGTQGKDNAGENLDPQPSPEPQGKDTAGDNVDPNTGN TTVPTTDGGAQVTPD
KTVPTNNPNISADTAHQAGASQVGTSTSTVTGQVQNV SADAPVYNTGAAIVSTQNGQLVLSDGSTVAPATIG
AVTNADKTITV TQADGTKATLPETGEAHSLLALLGSSLLGVVAWFARKRKMN*

>group_4515

MIPLTPIGGV TIAKENKLPTTGEDENPFFNIAALTIIASVGLLSILKKKED*

>group_4520

MFRRSKNNSYDTSQTKQRFSIKKFKFGAASVLIGISFLGGFTQGGQFNISTDTVFAAEVISGSAVTLNTNM TKN
VQNGRAYIDLYDVKN GKIDPLQLITLNSPDLKTQYVIRQGGNYFTQPS ELTTVGAASINYTVLETDGSPHMKP

DGQVDIVNVSLTIYNSSALRDKIDEVKKKAEDPKWDEESRDKVLISLDDIKTDIDNNPKTQSDIDNKITEVTN
LEKILVPRIPDADKNDPTGKDQQVNVGETPKAEDSIGNLPDLPGKTTVAFETPVDTATPGDKPAKVVVVTPDG
SKDTPVDVTVKVVDPRTDADKNDPTGKDQQVNGKGNKLPATGENATPFFNVAALTIISSVGLLSVSKKKED*

>group_4525

MQDKTLLSEESDIPLLTGQLLKGTGLGVGALLGGVSVTQTAFAEQVSSTLPVTSETGSTLAGTDSAVVVASTSS
STSDTSSQSSSLSVSNSQSLSLSTSGSTSLSISQSDSVLQSSSLSQSTSASASTSSANTSTSLAPKSAMTMST
ATTLAATGTVNNNLVTVTSSYIKDTSYNDGYIYPQDGESLEYTKFTIDNAATAGDQFTMTYSPYTSPSDFDM
ANWTPTDITDSSGEVIATGTWNATTKTVTYTFTTYVDKYQNIVASLDAYSIDRAKVPNDANLTLTYSLGGET
TSKLEYIRYQSPIVSNNTSSIQSVFNELDSVNHTVEQIFYVNPMDYAAYSTYFNVYGYQIDPSTGAMVQQGSTV
INSTTSIQVYKAGTGSTLPDSMDVLDYSTLTKVTPSITYGTDSARIFLGNISDVYVIRVVSTYDPNSANPIVQ
SGTMSSYDYYGYYSIVETNNYVVFSTSDTSGGTGTAVTYSIGDYVWLDTNKDGQLQTTGENPLANVLVTLTYADG
TSKSVYTDATGHYRFDGLADKTSYTTITFTPTGYLFTTSTNVGTNDLVDSDGKTVTVTINGSNMNTVDAGFVPD
TSTSTSLSLSTSVSKSKSTSLSTSTSGSLSTSTSSSLSASTSLSTSISKSTSTSASTSSSLSLSTSSSSSLSA
STSLSTSVSKSTSTSLSNSTASLSASTSTSQSTSTSLSTSAASASTSLSTSVSKSTSTSLSNSTASASTSTSI
NSLSLSTSLSASTSASTSGSLSLSNSISGSTSSSTSLSTSTSAASKSTSSSLSASTSGRSASTSTSTSLSSN
STSLSVSSSLSSSNLSQSDSQSISSSVASNSASTSVVNSASASVSQSISSGSQSTSNSASLSTSTSTSLSTS
ASQASALSQSTSTSRLLSQSVSTSVSESDSTSTSTSLSVSASLSKETSQSVSTSTSVSDSASSASQSTSI
SQSLSTSNSQSESQSTSLSLSTSVSSSSSLSQSESTLQSMSISSSTLGLSLSDSSSQSESASSSQSLSDSESQ
ASTSLSLSLSESISSSESLSQSEENSQSEIVSTSTLGLSLSDSNSQSESVSSSQSTSDSESQSISSSLSLSLSE
SISSSESLSQSEAVSQSESVSTSTLGLSLSDSSSQSESVSSSQSTSDSDSQSLSESQSLSLSGSQSGSESLSQ
QSESVSLSLSDSKSASVSESVSLSESLSTSAANSTGVADSASVSDSESRSTSISESISKSQSESDAQSDSESI
SISQSFSLSVTISESESTSFSESLSDSRSLSESTSTSESTSVSDSLSESVNAESLSTSIQDSVSTSVSESI
VSDSGSLSDSVSQSDSRSLSESTSTSDSTSIISDLSLESLMASESMSISEHDSVYTSISESISVSDSGSLSDSV
SLSDSHSLSDSTSVSDSTSNVSLNSRKEIDSIASQFILASAIASDSTSVSDSVSTSDSLSQSDSHSLSES
TSTSDSTSVSDSLASLMVSESLSTNIQESVSTSVSDSISVSESVSISVSILOQSESQSLSDSVYQIESLSMSS
SLSDSISSESEMSTSQNQSLSVSVSDSTSVSVSLLESLSLQNASQSLSSSGSTSTSLVISDSISQAISDMS
QSDKDSVLLSISISDSTSVSASMSDSSSLSQDNSEYVNSVSGGSLSESIISTSQLTSVSGSDSLSGSHASD
SLSQSIHSHISASDSVSQSLSQSISTSLSASSASASTSDSVMVSLSVSGSTSLSTSLKSTSGSISQSDSQ
ISTLVSQSVSVSVSESLVSSMSISNSLSLSSIIISNSLSTSQSSSNKTSQSLSNSQSIQSESDSLSTSLQSA
AIVGSETAKMYQIGNFVWEDKNANGIQDAGELGIAGVTVTLTKPDNTVVSTVTDNNGYEFDHLDIGKTYTIS
FETPSGYIPTVADVDPTKDEIDSDGTTVTITINGSDDLTLDSGFVKETYISIGDFWEDINKNGIQDIGEPGIP
GVTVSLTKPDGTTVSKATDVNGFYQFSLHVDGETYTISFETPQGYSPITSNVDSRDAIDSDGSTVSVTINGAN
ITSIDSGFVKETYITIGDVTWEDSNKDGIQDAGEPGIPGVTVTLTNPDGTTVTTTTDANGHYEFTDLPNGDYTV
TFETPDGYTPPTSNNTGDDTKDSDGQVVKVTVAGSDNPTIDSGFVKEVHTIGDVTWEDTNKDGIDGTGELGIPG
VTVTLTNPDGTTVTTTTDANGYEFDTLDPNGDYTVTFETPAGYVPTSSNIGDDSLSDGTRVTVTVDGSNNT
IDSGFVKVEQPTPGSNSSSESLSQSTTQSSSQSSSAKPVASQTAAQLPHTGQAENNGLYGSAALAILVALGLA
GKKRNEND*

>group_743

MKVKGLLKGAAICAFASLTIVACNGGGGSASKGAVEIEYFSQKAEMQSTLKDIIKDFEKANPDIKVKLTSPVDA
GKVLKTRMANNEDPDVINVYPQNMDFQEWAKDGOFLDLNKSGLLENLKDGAAKTYAINDKIYSLPLTANAYGI
YYNKDKFKELGIEVPKTFDEFKALVDTINKDGKAAPFALSINDPWSLNGYHQLAWATNAGGYDGAEEKLIRSG
KDAIKADDVTVKVNQALELLHGNQKQGYEGAKYTDTVAKFAKGEALMMPQGTWAAPVINQQKPDFEVGMFAF
PGQKEGEELTVGAADLALSISKDSKHPKEAKKFLDYMSKKAIQKYVDVDSGPTSVKDVDTEGKFKETEGVTK
YAFSDKHIVWLQKEWQSEEEFWNITVKNIKNPSATELAKGLNKFFDSMKK*

>group_4356

LTSGQASADASKLAHLQALQSLKTKVAAAVEAAKTVGKDDTTGTSDKGGGQGTAPAPAGDTGKDKGDEGSQP
SSGNIPTNPATTTSTSTDDTTDRNGQLTSGKGALPKTGETTERPAFGFLGVIVVILMGVLGLKRKQREE*

>group_14

MSNQFIQQSQTKKGHGFRRKKKWKAGLASGIALGAVVTFSSAYADDIANTVDTQPEVTVVSDSNATNLADAQ
ADNSQEHIDLTNQSGEQTGQMTDKVTS DGLDQSVSDAQ SAGVEVDQNGTVTHSSYADAQADLAKQKEAIDQAA
ETQKEVDNAKATAEEAAGAAGVEVSTEASKSAKTYTSVEDAKADAKQVDNLSTVTEAQKQVNEQLPKAVEAA
TDSGVKVNKAGAKTSDAKKALENLANQVANLKAQDTQNTISNSMAQALEEAKKAGTVSTTGSKSYDDLK
ALADAQAQVAKLTFEKTQASIDEAIKKSADNAKASGVTVSTGDAKTYTSAADAKADSQKTASLDKAAETQK
AADKVISDLTKEAQNAGLDLTGKTSQTYTSVDEAKKDIKQEQALKDAATAKANAETTNSSNKKAAADAQAAA
NKAINDAIAKATASGTTVTKGTGTIETSVEEAQKIAKEQSDKVNDVAAKNEQIQNENTAKKNAYDAKKGEVDKK
NKAADFDRAGLNYTGDIATDKATADKFNSDEIERVKSQFEKANDNNVLGATDGWVQHVGSNGYTQSGTSQFKKA
QNVNGVEIVANGSLEPGKATVYVKGNVNQSHVIANMWNVGNVAPQGDNLTSNIGSKWAEAYDYNADGTTTQIWK
GKILKWYRIPNAITMLDGSQHDAYVLFHTDTSGLSSPGDEVVFWNQDGAINAVDGYRYGVNDKSDGIRTVIRV
DSPDNNDNYLWISLLGDFDIGQFLEADGVTVLGVGGGFATNSGVGGHRVAVDENLGLTYGKNTTPTSALTGFN
SAPDGVAlIAQYSNQYSTVIRNTAGNGVAVARTDFGAQAKVSIVTPRHVNVENVETPTYTELLPVSVKDVHI

TPLQVTFPKAVTLHGTVNVAETHAIKLATEVAPVTVETEVTTTPALSATVHDI TMKAAVHPVEVTQKPTNEKAV
VNNDNVNDINGKTVAKGSTVKWLLKNGNLLAGRKELTAVVMNDPFPAGFEVDSEATAEANADYYTLKTTEGDSY
QLVGTAKLLSLLNAHRDQDVEVPGFVFGKPLNDNAHYDNAFETIYSTAKGYYKVVSNI PSIDTPDNPKPTKT
VTDKAGSDIDGKTVFDKDVTFHLITDYSTYVNTVVDVNTMTKAACIYDDTDDATVNPDVAKATVTDNSGNDIT
AKGTFYDIEAGQPIPAEIQIFLDGAHLTPNGEFIVWIPNDLKDYENYVLTGNNVTVNL PVEVIAQPGKEAEN
TFVQVEFGNGYQSNLVFVKVPDVSPEKHAIGKDGTVLDGQEVGLNQVFNYKLDGVTVPKNHDTLWQYDGGDKL
DIEH DRYTGNWKG VITGTEYTAKEDMTLT YD VITEDGT VIKAGDTIAAGTNYHFI FEFNQDTNDDFINKIVKV
TWNAQAGEWAYSINEDFLRSLGLEGTFDADFWIEVERIKDGEVENTFINTVNGRELI AKVTHTTPAAPVPSTP
QVPAKPTAETPVAKASVLP TTGEKTSALDVLASITGLLMIVGA AFGLRKKEKH N*

>group_47

MNKAKVVTAATVAASFGVVTQVHADDDLATVSNNTEVAQNQATDNVTKADV DVTKANLDAANQAVSSQE QV
VDDAAS TVDSAQEAYNQAQQATS DAQN LADQATPENIANAKTDVNSAQNQVEQADNTEKVAESTAQQAQEAVN
NQSEVVDS AQSDVNTAQS AVDAAQKDVNDKQAILDGTGQAEI IANRDKAQADVNSAQQAQVSQKEADLTKAQEA
DANRQIAIDNAKTAVDKANQNVSTSKSDLD AKTATATQTQATEDAAQSAVDAAQKDVNDKQAILDGTGQKAIL
DEADA AKVDKANKETALTEAQKGLKEAQEADANRQATIDNAKAAVD TANQNATSTKS DLD AKTATAQQTEQAL
KDAETAYKTAENDYKAINKITLTQAYVEALKKYNDYSLSSSEERQQVLKTLVEESEALAKVNSYKSNPNDDNST
KYQINDLPEDVIKELSQFAADLVNQIRKAFGTPQVSVTSSSVKFADLVTDGYVSDGWNVT KAINEGQVGHDAK
AVNNAARYFGLPTSPSDDEATGGQYYEDWASSAVYDGITFATLTKQRVYEAVTDFMFNADEWRHARDIVGGGDV
NSHVYFAIDFSKPSSSENGAFAGVHFISVAEDQLTTKNNNFDTNAIANPKSAENV TATYNAAKDSYNQATVANT
AAQSAKSTAQTAYDNAISQLSSASSVLTQAQSVAIQTPTAQANL TSAETALNKA KDR LVKAQKAVEDLNADIK
AKQEALTS AKAVLETAQSTLAQAQSKNSTAQQA KADAQAAYDNAISQLSSASSALTQAQSVAIQTPTAQANLA
SAQSTL TEAKENLVSAQKAVEDLNADIKVKKANLASAKQVLATRQAALATKQNV LKTEQNRLANLQNSLATAQ
DNVSR AQANVVA KANL DKAKHYLTSLQTAPQLLADAKKQEETARTHLLDALD TLETELLKLKDLQVKQAAAQ
AVYD TTSKAYQTILDAQE KQRLQDEYNAIVAQ GKTPVPIVDETGKIVGYQVENIQYSALATEVNNTKATALS V
KTANDTLQKSSEISEATLPETGESSSALVILLGFLMTLWGLVGVHRNKVK*

>group_1647

MKKVKALALFSTAILATASTGVYADEVISDGTTSPSTEQIEPTPAPSSSEVAPGGESNTPTPSSSSSTEAP EG
SKTVPGA AEDESSKRGSETTETETPTQ PATEPEKSTETPSEQPGEVEVPTTDGGKATVVPDTSVPTNPNITAE
TAQNAGASQV GTTSTVTGQVVRDVTRSNPVTLYNGASLVDIRDGLVTLSTGEKVAPEVVG VKANADGTYTAKT
IQGDTVTL PHTGEKNSTVLSVVGALILSVLGFGLKKRKNQEI*

>group_4620

MVVNSIKASNGQDVSSLLEMHHVLSKDALDEKLQAILSNSGLSPIGEFYMWVAKDPSSFFTAYVQKGLDITYN
LSFKVNQTFTEGEIVNGVAQIDFGNGYLG NLVNDLPKPEVHKDVL DKQDGKSINNGTVKLGDEV TYKLEGWV
VPTGRSYDLFEYKFVDQLQNT HDLYLKDSVVAKV DITLSDGT VVKKGADLA EYTETIYNKETGRYELAFKKEF
LEKVARSS EFGADAFIVVKRIKAGDVYNDYTLTVNGNPNVKS NRVTHTPEDPKPVP TPQKLTPKTPALPHTGEA
SVAPLTAIGAIILSVLGLVGFKKRKEN*

>group_4637

VLIQDRSGRYIVSASELGGR LNDDGTVTIKDSEGKEKTL PHTGEEKGFLSII GGTILSFVAFLFKKKK*

>group_4701

MAILDK E VAKL KELQAVQAE AQNQYSIVFEAFKAVQEAKKQAE LTEIYENI IAGGGEAIPAVDETGKITGYVD
GSKKITNETNFNL TSTD KSPANQENKLVS SLTNDLPNTGESSIAPFTAIGAIILSVLGLTSLKKHRT H*

>group_4723

MVAEEI WVGTKPTVETVTIAKPADVEEADATLAKGTRVLKTAGSTGSTTTTTTYSMDATTGVVTANTPTVETV
AASPDVYRVGTKEDVIEIVEEIVEKTKSFGKVVINNPDL LKGISRQVRLGKNGIIEV TYKIVKVNGL EISREI
ISEKEIEKPVDEIVEIGTKEPVENQNQKVIKANESKMNSTDVKKVENKGELPKTGEIKESVFGMF SVIMTGLF
ALVTFLGKRERK*

>group_20

MTTNNTKGHGYFRKKKWKAGLVSGIALAGAVAFSSGAVLADEV TSPVTESEATTVVVTQPESTNDNNAYAAQAN
TSTGTQPVAIDNSTVEAAVQTATETGVAVTKEATVNQGTDTNSTDL DASKADIKSDQEKQVKEIKAATQTQAD
NNAAYSEAQGAIDANND FVEDAKGHEKDTT VTVTTD GSTDTDGSAAKNKQATEVAKKTLKGNQESVKTYELN
KATYDATVAEAKTLNQAVESAANALDKGVTVT TTEKVVT SVAEVEALRKQNEASIASAEKQITLNTAILAAY
DKTKKASDADTADADSKVTDLKEKGVTVNVT TTVVSSAEAEATLRKQNEAAATSAGNQVASNNAVLAAYDKTK
KASDATNADADTKAKELQKGKVSVT TAEKEVSSAEAEAESIRKQNADAVALA EKQVTS DNAILEAYNKTKKASD
TTNADAEAKVNELKGAGVVVT TSSKEVSSAEAEAEKITSQNK SAYDQAKQKQADWQKKYDELQSKTGTEGYTKQ
VVLQALDLSASN PQATHGSSAAGSQIVGTRDIGSTTGSSGYGRILDSTGVFKYANVGKGT TTEIDYTNLKGLT
VTTTDGQKH DISRIHRTFQLVNTGATGLNDVYI PNDPTEGFIVARNNGTGSYS DYMNFLVTD SYYYTVNGKEV
VYSASQETPIALTYSSLNNNP IREGARATGDGSSMVEINGSTVNVHGDQFAYS DYNRQEETGLNWDTTDSA
YQYKGAAIGVFTQGSSFTTEFVQWDGPADPAGQTYWFAINTRV VAPV VETPT SARLTKTTVKPAEVKPI SASL
IKTTVKPSEVEPVNVSIVKTTVD PVSPDPVSAELVKAKNPEQPTLELTKLADTKNQQVSVAYHDYKLSYKPTV
TKTVTDTDKKNTDGATVPKNT PQVYTLNHDNIYANIKVGDTITIVDPLEVGATPVVDENEADAKAKGWTVSYD

EAKETYTYTAVYQGKKLEAPTIKWVPIYDKAFYDNTYKVFKNDEAFSNTVTTYTPPTPDPVKSITDNSGADI
NGAKTFDSNINFHLTTDYSPYTKVTASALAIAGKFAILDDVQDGAFTVNEDGITATATDGTDVKDLFTMYHVL
SDEARTEAIQITILDNAKLSPVGEFYLWVAKDPASFYVNYVKQAKNVTIDLPARLLVPEGEVVENDFYQIDFGN
SYQSNLVTVEVPDVNPEKHALDQKDDNIILDDQTVQIGEYIRYLLDGVTVPVKHDNLWQYDGKDKLDVVHDRV
TGNWKGIIKGTETAKEDLVLTVDVLTLEDGTLIKAGDIIKAGSKYAFTFEFDQDTSNDFIKKIVIVTWDEKKG
EWAYSIDEEFLRSLGVEGTFDADFYLEVERIAAGDVENTFVNIVNGKEMTAKVTTHTPPEPPQPTPPAPNTPVK
QANVLPSTGEASSALTVLGGFLISSLGLVGIRKRKEN*

>group_1653

MKKKQQLLSALALSTLALAQAAGFVSADEVTPVDPSAPSTEVTVPVEPSTPAVTTDTTVPVDVTTTEQTTVPVDPT
TPSTEVTPPSDEQKDNAGDNVDVPEPEPQPKDSAGDNLEVTDPSQGEQPAPSGEQKDNAGENLDVPSEQPK
TTEQANQQGASQVGTVSTSTGQVVQNVTPDAPVQTNNTGVSIVSTQNGQVVLSDGSTVAPEAIGAVTNTDKTIT
VTKADGTKATLPETGEAANFLAVVGTGLLGLVGFLFKKKLV*

>group_4777

VVKVTVDGSDNPTIDSGFVKVEQSTPGSNSSSESLSQSTTQSSSQSSSAKPVASQTAAQLPHTGQAENNGLYG
SAALAILVALGLAGKKRNEND*

>group_1835

MVKKQVAIIGMGVSGLAVLLALSQLPDKLLEKIDITCFDDPKHFGRGIPFQEDSSTAWINSPIDAISYDYHDM
NDFQKWMEQKGLDTDQSYVPRSLYGRYMTERRAHDLLQKLKTSVIHEKVTQLNYEPDSQKWNIGTSQKTIPTRF
DEVHLTCGELPVLDPYHLQGNPNYIADPYPLKNLPKQPGKKDRIAIIGTGLAAIDTLKWLKNSQADLLAFSP
SMTFPTVRILKTETIDWQFLTDTNKQKLFEENSFNCKSLEDLFLSELQALGFQNWEECTCRQFLAEGIPGISLS
LAFFPAQLFLLQQLASHLVDWLTDFWPQMTLSDRQYKKNYKAIINLRNPMPEEAGRLLIEATAQGRLLQIEA
VTDIETDDHGFVLKREVGEVLSVATVINATGYHLKESNVHQARTLIQQVIRDGLVQIDPEGGLSILPQTGQVI
SPKYGILATLYAHGSLVNGVIYQNNSTIKIQQMAERAIGNVIKKPTI*

>group_2281

MKAARLLAYNKEKIDLQIEEIAKPIPSRTEVLVEVKYAGVNPVDNMISRGEVKIIVPYKLPITAGNEFSGKVI
AKGTEVTGFDIGDRVYARMPNSKTGAFAEFLTIDQKELAHIPYLSYEEAASVPLTALTAYQALDLLNVKSGE
SLFISGGSGGFGAMAIPLAKARGLQVITTSARNAERVLALGADRYIDYQKEDFSQILTSVDHVIDTLGGAEL
EKQFSILKKGGHLVSLKGLPNKAFKAFGLPKWKQFLLGLVGKKYDRLAEKNHQTYHFLYVTSNGEQLADISR
IFQENKVATSIDSIFTDQVNQALKKVDKGHSQGKTLLKIH*

>group_2504

MISEQKRRFLPYLLIGIVMFYFSHWLIKLYAIAPKSSEVTIFDTTAFNWMYDHWSSKPLLDQFTMFSLYGG
IVVFFIVMLFYFRVTDNGKYRHGEEYGSARYATIKELATFRDKEPDNDMIFSQNARMGLFNKRLPFNKQLNKN
TLTVGLPGDGKTFTFVKPNLMQMNSSFVITDPKGLLVRETGKMLEDNGYKIKVFDLVNLSNSNQFNVFHYMHS
ELDIDRVSEAI IAGTKRSDNNGEDFWNQAEILLMRALIGYLYFDGKVLNKEYEPNIAQVADLLRNIRRDDEDIP
SPVERMFEELEKELNGNYACRQWELFNSNFEGKTMTSVLSMISSRFVFDHDAVRNLI SRDSMEMEKWQTEKT
AVFIAIPETDKSYNFIATTMFTMMFRQLPLTADEILQGKHKTCKPGDLLHIRLILDEFANFGRFPNFTETLSS
VRSREISIDII IQAVSQLKALYKAEWETIFNNCATLIYLGNTDKETMSYFSMRSGKQTLNVKNQSQTRGSQGS
SSQSIQTIQRDLLTPDEIARIGVDEALIFISKQNVFKDKKTNVLQHPRAKELSNSPDDGNWYNYIRSMGSGIDD
WDANVNHDRKIIASASEVEDTQLPFQEALAI*

>group_2707

MYHFIPAWYQKDRKWYDDTKETHVDQFLQFDDAVNQIRIFKDNQEELEVYVLNYAPHLRFTLHEKDILLEVTV
WSLFDLQGRQTQASSLNWRHLNWPKGVSFLYSPFTVTALLDDILFARIQFTTDGRVQFQILLYDNDSSALKE
SLVFDDRGLSSIQYFEQQLRRQDYLDDEGIWRLREHLNQDRGFVIEVNIIEKLPLFKQATYANWETLIQEIL
ISKFSNIPANDVLVISADDQHNHFFKEVPAQKVYSFFQDRLPITGGFADNMSFLEGADLLVSDSQKNTEALKS
LDSVETLPAAIQVSPFDTRFELGLSIRRKELEIYFFADQLPDDDIKSSLKIIFKEMAKNPLITLTVGTYQKFG
PRLQWLESLLIGDYSKHIPVEKVSFEDLLPKNTEEEEPRAKVLVLKAENDVISQFKTSRLVIDLGMQPDLYH
QIAAISSGIPQINQVQSDYVTHQQNGYIIISQLSDLPGKIDYFFKGLANWNKSLVYSIEKIADNTGSALISKWQ
SHLKGQDNNE*

>group_2786

MTVNDITLSKEVIQERIARRVAQELENNSLVNLGIGLPTKVAHYIPKDVTTITLQSENGFIGLTGLKDKEIDPT
IVNAGGQPIGIAPGGAFFDSSTSFGIIRGGHVAATVLGALQVDKDGNIANYLIPGKMVPGMGGAMDLLVGAKK
VIVAMEHTNKGKAKILNQCTLPLTAKGVVDLIVTELAVMAFRPEGLCVLEIHPDYTFDYVQSVTEVALIDKTQ
EELLA*

>group_2838

MASLKTIDNKIDYGIITPVFCLILIGLLSIYVATYHDYPQNLSKVMLQQVLWIAFGSLLAFILMFFSTKSLW
KLTPFLYLLGIGLMILPLIFFSPNLVAATGAKNWVTIGSVTIFQPSEFMKISYILALARMTVWYKGGKDRTHF
QDDWKLLGLYLLLTGPVLILLGLQKDLGTAMVFLAILCGVILISGISWWIILPIVLGTLLLILAFFCVFLSPQ
GKTLLYKMGMDAYQMNRI SAWLT PFD FSEGIAYQQTQSMISIGSGGFLGKGFNLHDL PVPVRESDMI FTVIAE
NFGFLGAIFLLTLYLTLYRMLKVTFQFNNLFYTYISTGFVMMILFHIFENIGAAIGILPLTGIPLPFISQGG
SSLISNLIGVGLILSMHYQHILELENESRQQLRRSYKYD*

>group_2886

MSFKDYNFKEYIQKALDAIRFQEPTEVQKKLIPIVNSGRDLVGESKTGSGKTHTFLLPIFEKLEKENSQDVQAV
ITAPSRELATQIFEAARQIASYSEEEIRLANYVGGTDKLRQIEKLQVSQPHIVVGTGPRIYDLVQSGHLAIHK
AATFVVDEADMTMDMGFLDTVDKIASTLPKSVQILVFSATIPQKLQPFLLKKYLSNPVMEQIKTETVISDTIDN
WLLSTKGRNNNEQILMTLKTMQPYLAMIFVNTKERADELHQFLISNGLKVAKIHGGVPPRERKRIMNQVKKLD
FEYIVATDLAARGIDIEGVSHVINDAIPQDLSFFVHRVGRGTGRNGLPGTAITLYQPSDDSDIKELEKMGIAFV
PKVLKDGEFQDSYDRDRRQNRREKSYQKLDTEMIGLVKKKKKKKVPAYKKKIQWVVDEKRRKERRAESRAKGRA
ARKAKKQSF*

>group_2903

MNTVEHLIEKFVPENYNIFLDINRDTKTFSGNVAINGEALDHIVSFHQKDLISIQSILLDNEPVIFHEDHKSET
VHVELPVTGLMTLVIEFSGKITDNMTGIYPSYYTVEGVSKVISTQFESHFAREAFPCIDEPEAKATFDLSIK
FDQKEDEIVLSNMPEINVELRKETGLWTFDTPPKMSSYLLAFALGEVHGITAKSKNGTEVGVFATKAHPSTSL
TFSLDIAVRIIDFYEDYFGVSYPIPQSYHVALPDFSAGAMENWGLITYREIYLLVDSNSTASSRQQVALVVAH
ELAHQWFGNLVTMKWDDDLWLNESFANMMEYVSVDAIEPSWNIFEDFQTGGAPLALKRDATDGVQSVHVAVNH
PDEINTLFDPAIVYAKGSRLMHMLRRWLGDDAFAKGLKTYFERHQYGNTIGRDLWKALSDASGKDVGSFMDSW
LEQPGYPLLTIKVENDQLIISQEQQFFIGDHKEHDRIWSVPLNASWTGIPDVLTEKTMVIPNFNQLLSENNGPL
LLNTENTAHYITHYHGQLWESLKSSVSEELSEITKLQIIQERRLLAESGHIRYADLIQLMSLFQSEKSYLVTSA
IQQIISGLERFIPENSVEEKNFKTLVSRVFSDDYHRLGFNKQEGESDEDEMIRQIVLNQLIAVDYSDAIGKAN
SLFKDYSSNNLADIPAAIRRFVLVNQIKHFETDALVDQYFDLYINTTDNNIRLDLANALSKTQNMMLTKRIIAS
LKDKSIIKPQDLMSWYNVLLSQEFTQNPIWEWARHDWEWIIQAALGGDMSFDKFVILPANHFKTEQRLSEYKAF
FEPKLDDMAIRRNKIGIGINEIKARVALIEKEKKSRTFNSSKPLMMRRHLV*

>group_3151

MKKSRLAKTSFISLLSVAALSQISVENNHLSFFNVKSVNADEVVEVDATDQGALTDTEVVENATLTSESTVVK
IVSTTVVETPSASFDLTNQNTASIVSEVGSVEVLEAKTSEALSNSVALAATDPIAPDDQTEDVADAKLRVDIG
KVLDDLEAMIASGDIAAAKESYELANTFQGASLEFYGYKDEIDQIPNVKIASVRLEYEARNAAIYQAIYGTPL
PTGPVVTTTRTEVIAHTTKYIADDTLAYGQRQTVVAGVDGEKTYTTTDTGVEDAGVVTKAMVAEEIRVGTKPTVE
TVTIAKPADVEEADATLAKGTRVLKTAGSTGSTTTTTTYSMDATTGVVTANTPTVETVAASPDVYRVGTKEDEV
IEIVEEIVEKTKSFGKVVINNPDLLKGISRQVRLGKNGIIEVYKIVKVNGLAISREIISEKEIEKPVDEIVE
IGTKPEVENQNKVIKANESKMNSTDVKKVENKGELPKTGEIKESVFGMFSVIMTGLFALVTFLGKRERK*

>group_3154

VVEVEVPVTVNRDTDGDGIIIDSEDPDDNDGVEDGSDETPKAFTDLTATATPVTVPPEGQAVPADTKVVETNKS
DAVISSTPTNGLSVDENGNLVGTPTVDDWGTDEETREVKVPVKVTNTLPDGTTEEVEVEVPVTVNRDTDGDGI
IDSENPPDDNDGVEDGSDETPKAFTDLTATSTPVTVPPEGQAVPADTKVNTLPLTGDSSEAFFTSAAALAISSV
GLLGIAGKKNKNGSD*

>group_3204

MFKTKKEIFSIRKTALGVGSVLLGVMLTTQVASADEVSLMTPSVDKSLTTTSPVLESTSSQLAATTAPTITDT
TSNVTATSPVLAAATTAPSVTTTPIASPIRYVSDPNQPVGYRATQVQGTGDSIIITTQTGALDANGNPIVTVR
IEPTETVIVLGTKSTSQVTSTQAATTTYSIDVTKPVGTDVVIQAVDGTQTTTTTYKIETATTAPVSPSVLSEG
YKWIDQPFYHVDTTQTLPSDRISIDQLFVPMPLVTPDYNTTESTVREYAQYAENMYDYITVTNPDGSTTRQQ
VIRPVTSEMLDPTNTRLRLTLGLTDDNAFYSLRFDASQQDLWNTSAQDYGLEIVPEDLSTSTDDFLRYHSSNI
INDALYADIKADYLAQLAYDQLVSLGLTSLSDQQSAMDMMTSQFESLTLRYNNYKDSDAIVVDYSNTTMSATQ
QADFEAKLAALPVEVQRAISELTIYDQIIPGMGETTLGLANSADQTIALKYEANNLNLVSTVLHEMTHIIDFK
SGLYSETTDRNTDGLSLSTMVAFSDTQEFLDVYHTYFDRPDVWSYYRDNSEEAFAEGLSQYIMHRLFGTPYSTY
IANPYTGDAYNPGDGSYSPFAETEFYFASLYNRLFYEYPRTAQVVPYLVTTTTTAPVNGQVIYIGAMPEETTTT
TPYTTVYVGDTSFAYDPTGQTDREVQAGVDGTETIRTTYSLSNNQLVATQTVISSTPVQNQIITKGTQPTVVD
TSVPMITIVYQEVTDGSLGDWQVKVLDAGQDGLIRSTTTYSVDPVTGIVTPSTTEATITAMRPMIVQYQVGSEK
VTAIPYQTRYVIDTSLAAGTQVIVQAGVNGSSTESVQSYNFIQDGSNSHFDIVYSSPVVAAQDQVIAVGGQ
DQVTDQAVAKTIIYQEVTDGSLGDWQVKVLDAGQDGLVTRTTTSYSVDPVTGIVTPSTTEATITAMKPMIVQYQ
VGSPQTSTIPFQTEYVADPNLAYGSEKVVIVGQDGVVTRSVASYDFIQDGVNSRFANITYSEVNNPQVVNEVI
AKGTQTVISDKMIARTVTTYQEVTDGSLGDWQVKVQGTGQDGLIRTTTTYSLSVSGIVTPSTSSQSTITAMEPM
TVQYQAGSPKVTPIAYRTRYVADSSIPVGTQAVIQPGKDGSTIYVDSYQFIQSGATSHFENIVYSNPVVDA
QDLVIGVGVVNQTHDEIIARTVNYQEITDGLSLGDWQVNVLDAGQDGLIRTTTTSYTVDSQTGQVSVSTTRVQNL
EMLPMTVQYQIGSSKVTAIPYQTRYVIDNSLAAGTQVIVQEGVNGSSTESVQSYNFIQDGSNSRFDIVYASP
VVVAAQDQVIAVGGQDQVSDQIVAKTIVYQEVSDGSLGDWQVKVLDGQDGLVTRTTTSYSVDPVTGIVTPSTT
EATINAMKPMIVQYQVGKSKLSAIPFLTEYITDDSLAVGLEKVIQEGVVGTQIETVQSFNFIQDGANSHFENI
VYSSPTIVVAAVDQVIARGTKVDEVVAVPEVVTCPKPTSEVISPEKGQTAPTITVEAIKAPAQKKAKVEAVT
TPKESLPTTGDDQNLLVTLMSLLLSLGLGLKKKEDE*

>group_3320

MSITITCQARIMTTFSPKLAIRLNKKYVAEIGNKETIALDIPNNMTLLSYQFLDYPNIKVSNGDHIIKRN
YITLIRWAYILINFFFIYFFKKTLPSTGFINTFLGFSIIPIALLPSYRFIKE*

>group_3370

MRNCKYKLRKLSIGLVSVGTMLMTTTLVGEENTTTTPTTEVTTENQNATSSISNKGDLSNSKTDEVVEETPAP
PGITSTVTTELTELTETTQKNEKEIAKTINEGGDSVISQQTEDASTTGQTPPQVQTHSRIRSIRSAEAGQAMEAETME
VEKVSVDKEKTEVTVSDEKKLIKNRDDKQRKIFDVKRDRVVDPDGTTLNVTLTVTPKEIDKGAEVIVLLDTSK
KMTDDDFNTAKENIKKLVTTLTGKDSRNDNKNPNYNNRNTVRLIDFYRKIGEPINLSGLDANKVEEKLNEVRK
KAQEDYNGWGVNLQGAIHRRARQAFNSVNEKPSGKRQHIVLFTQGEATQSYEIKEKKDLKKLTVDEPVVTSNPL
LPWPLYFDTTTTTRKANLVEEGKKIQDYLLKKIGIKRYDGLLDLSIAQNGNKILNFGSSSLVGFNNPLNFITMSDLDT
KNLSEESFDYETKIGEGYHHRSYDRESGQVPLDSIVKKAVKAKIKKKLKEAEENSYDIFGLSKWSTSLSKVF
GFDITLEEKMVNKFVDYLFYKRNYYVYNHNLSAQAEAKMAREEGINFYVFDVTDPERSSKMDKAWPTGAYGERL
KQKVEEARELSKKRNEKFDKYLKEMAEGKSFFKEVENAEKFKDILTELKVTETFTDKVSVIKDDQLKGEKDVK
HTPASSNFWLSTSESIIWTTISKEKLKDAFEQRTPLELTYKLVKVEKDKFKSASPTRAKRSLETPESENKPI TEK
AISGKVITYKINGQDGTGKELGDVYVYTSKEKVRKPQEESKVPEAPQLIPLTPLKPSNPVPAPELPDVEKPEI
ENESIVPDLPEMEQPDQDSPEISGQSQIVDIVEDTLPGVSGQHSSSEETEISEDTRPESDNEIIVGGQSELVD
IVEDTQSGMSGQHSSSEETEISEDTRPESDNEIIVGGQSELVDIVEDTQPSLSGHQSESQETVTVEDTQPNQS
NILIGGQSEIVDIIEDTQAGMTGQYSSTDQLTIVEDTLPEQMEETDEIKADSQVMDIPKVNNTNNDKGAKASV
AFDVEESKVVSTQDIKPSTYVKGDNLPLQTGDDDKVNAFFTLAALSVIGAAGLRQNKRREREKERN*

>group_3415

MKNVFRKSKIVETERKSRVKLHKSGKSWVKTTLAFSGLLQLFKGLRVDRELSVSDEDIHLSHHYFIKGIIGIG
GVIGAGALSHTVAADEMVSDLPASETSATIATLDSAIITSQLDSSQSVSSEQRITSEDNSQTSVVEQSQFAL
GKNFSFERTSVSPTELSSESILINTSLAVETNQSDRQTLSESVVYNSISSNNSGSLDNLVSSSESIISSLDSVT
VIESNTLVSDSTLLVSATSETLNNPETTSLTPKNAFKTTQPIILLATGTVDNSLITVTSQSYLESQDQGIID
PFVMEYLKHTVNFSISNTATAGDQFAYTFGTYVAPDDLNIKQYVPDPITDASGQIIATAAYNQTTKTITYTFT
DYVDKYTDVTASMTVVNLIDRGTVPNSTTVTASSTFASAPAATKSIKVQYSNATKINGTSDLSGMLNYRDKIN
NTIEEIFYVNPLDYISYNYSSLDIYGYDVNTGISGSDYINSSTIIIEVYKVPSTVALPTSMRITDYSLYTKIT
PTISYGTDMATISNLSTFNGETYLVRVVAKYDPASTDPIKISGVMTSYDYGYSGYTATVVNGLKSTSGSSTGT
GTLVITYKIGNYVWQDNDKDLQGTSSSTEKPFANVKVTLTFADGTTKSVFTDAIGYYQFDGLLDGQTYTVSFET
PAGWLPTVSNVGTNDAVDSGDSSITVSINGADNMTLDTGFYQGPPLSASTSTSTSKSTSDSAYSDSVSKSTST
SVVNDSVSKSTSDSVKSDSISKSTSTSI VSESVSRSTSDSVKSDSISKSTSDSVVSDSISKSTSDSVISDSIS
KSISLSQESDSVSRSTSNSQLSDLSQSTSLFQESDSVSRSTSNSQLSDLSQSTSLSQKSDSVSRSTSLSQE
SDSISKSTSLSQESDSVSRSTASQESDLSQSTSLSQESDSISQSTSESAYSDSVSQSISMSEESDSISQST
SESAYSDSVSQSI SLSEESGSISQSTSESAQSDSVSRISVTEESDSISQSNSTSAQSELSQSTSMSQESDS
LSQSNSTVAQSELSQSTSMSQESDLSQSTSDSAESDLSQSTSLSEESNLSQSTSQAEESDSISRSTSMS
AESDLSQSTSDSAESDLSQSTSMSLESDSISRSTSESIASDNSYTSQQSDLSQSISLIQESDLSRSL
ASEESDLSQSTSLVEENDSTSRSTASVESDLSQSVSQIEESISLSQSTASVESDLSQSTSAEESDST
SRSTSVSVESDLSQSTSI VEEESDSQSRASASVESDLSQSTSI VEEESDSQSRSTSVSIESDFLSQSTSVKQ
GSESLSDSKLSQESDSVSQSTSVKQESSESLSDSRSLSQESDSVSQSTSVKQESSESLSDSRSLSQESDSVSQS
ASVKQESSESLSDSRSLSQESDSVSQSTSEKQESSESLSDSKLSQESDSVSQSTSEKQESSEISDSVSLAKESD
SVSQSLSNELEYESSSESLSQESSISSELSITVHSESVSQESSELSQESSESISESLYRVNSISVNSNDSS
SLSTSI SQSQSTIGSEAPKVYAIGDFVWEDTNKDGIQDAGEPGIPGVTVTTLKNPDGSTKTTTTTDSNGYYEFT
DLNDGDTYTVTFETPKGYKPTVANIGDDSKSDSGASVTVTINGANDITLDSGFVRDLHTIGDTVWEDTNENGI
HDASEPGIPDVIVTLTNPDGTTVTSTDANGHYEFTDLPDGDYTVTFETPKGYNPPTPNTGDDTKDSDGQVVK
VTVAGSDNPTIDSGFVKKAYTIGDYVWEDTNKNGIQDAGEPGIPGVTVTTLTYPIGSTVTTTTDANGYYEFKNL
NDGETYIVTFETPKGYAPTIAN TGDDAKSDSGSSVSVTIKGADDITLDSGFVKVHTIGDTVWEDSNKDGVDQDS
GEPGIPGVTVTTLTNPDGSTETTTTDEKGHYEFTDLPDGDYTVTFETPDGYTPPTPNTGDDTKDSDGQVVKVTV
EGNDNPTIDSGFVKETYTIGDYVWEDTSKDGIQDAGEPGIPGVNVTITYPDGSTKTTSTDSNGYYEFTELNDG
DTYTVTFETPDGYKPTVSNNDSDSKSDGATVTVTINGADDLTLD SGFVKDLHTIGDTVWEDTDKDGTDQDAGE
PGIPGVTVILTNPDGSTETTTTDENGHYEFTDLPNGDYTVTFETPNGYTPPTTSNTGDDTKDSDGQVVKVTVDG
SDNPTIDSGFVKVHTIGDTVWEDTNKDGIQDAGEPGIPGVTVTTLTNPDGTTVTTTT DANGHYEFTDLPNGDY
TVTFETPDGYTPPTTSNTGDDTKDSDGQVVKVTVDGSDNPTIDSGFVKVEQPTPGSNSSSELSQSTTQSSSQS
SSAKPVASQTAAQLPHTGQAENNGLYGSAALAILAALGLAGKKRNEND*

>group_3432

METKEIYSFRKFKTGTHSALIGKFGIALTTTSIAIMTAGGVVHANTITTTSTTSQVTPPTTSVASAEIPTSVPAET
SQSTAECTINSAQDTLRTNVSNAAKAGVDVTTGTTTIDIILNNSNVVDKINEVLTDLSHQDKAVADAKAQEVN
QKAYTDAKTTTRDTAVSKGQSDLSHAEQGVDDQITIAKKDSISVTTDSKDLTPKYIDTKGLTGSAALTAMEKNI
ALYNQAVNDGVGIMDASSVQMKKQIADYLTALTNYQKGVSSNTGLQWQNSVVLEAGSGASRQTGTENVVDFSD
GTIKAAGMYATQGGQNLQNTDANFDNIFKIEGTGI IWKNTTNGDVKLTFSEINSPYNTGTYVAIWGDDKGGI
AWSVFALYYGGASGGAGETGSNTGSGISGRILNYVNSYKATVETTRGVSVVTFNDIDNQQTVKMSGLDNAKVT
TGENINQGTGNDFVAGSGDVSQSSAGVLGTNGVKWTFTSADRRLFSFTHSTAGTKNTSIVGGIFGSASNPQKP
VIPKLTAHKAIVTGPKGPSIPVNQKVTVHYHKVSVTPTPENPVTPQKNITYKPIITTTTPQTPVSVQANVGLPTTG
DKADGSIVQMVGALMVSFVGFSALKDRKKEK*

>group_3433

MRETKGHGFFRKSAYGLVCGIALAGVFVAFNNGGNVSADEVTA PATQ PQT VVTTLTEPTNNQAQSVTSTALDTA
VSTAKEAGVTVNTTSTVSHTDVSSALSDLANQTKAVNDATAKAEANTQAIKDATAQNAQIDAQNAQAEATRVAE
ANKAGQLAVDQKNAQGGQALVDKQNAQAQADATNAQLKADYEAKLAEIKTVEDYDYNKAVVERNALAKQQADAT
NAQAKAEYDRLLSIYNKALADRVNLVANDITFEYGYGKSEVLDSYGQKTTDVVTVNNDGTFVLTEPQSDTDGTF
GNTKVTGKFIYTVAFDPATNQAVIKITGISLNTWQLDLTRQTTAVNRNISASYKDLAGNVLYTRHFDGLSSIG
VQTINKSFITLDKEIILSNGQTSEEVMFLKTVPSWIYEAPSSLYAKLSYNTNALPPKPQEPTLVTVTPEALKPV
PQTPPVVYVVKPVVKTFTEPEVYTPIKPTVKPHVAVPDKLVVTVSVHPVVVPVANPAKEVVDQAGKSINGLSVL
PNSDLNYVTKQDFSQYKGMTASNDKVAKNFIYIDDYDEKSLDGKSMVNSIKASNGQDVSSLLEMHHVLSKDV
LDEKLQAILHNSGLSPVGEFYMWVAKDPSSFFTAYVQKGLDITYNLSFKVNQTFTEGEIVNGVAQIDFGNGYL
GNLVVNDLPKPEVHKDVLDDKQDGKSINNGTVKLGDEVYKLEGWVPTGRSYDLFEYKFVDQLQNTDHLFLKD
SVVAKVDITLSDGTVITKGSDLAQYTETVYNKETGLYELAFKKDFLTKVVRSSSEFGADAFVVKRVKAGDVYN
DYTLVYVNGNLVKSNRVVTHTPEDPKPVPQKLTTPKTPALPHTGEASVAPLTAIGAILSVLGLVGFKKRKEN*

>group_3633

MKKKQEMQYYLRKSAYGLAAVSVAVLAVGSPVSAQEKAASTEATPKVGPKVPKPKPDRELLKKLLEEAKAKEET
AKKAKEKLIKPAALTDEESLVITEKAEELSDKVLGEALSKEEAEPSEENKTDEEESDGVESGEIEKTESEES
TESEGVKEESEEEKEDDPSESETEVENVEAINLSEAEGNDSSKPETSEEVTAEEDRQETDRLAEVKTEESAKE
GDEDADKKDEVEEKAKKAAELSRVKAELAKLEALNARRLMKKIVESGKTVEGILSFMKESLPQLEAARASEQ
AKAPEVTQSPDHLPSSEKKAVHNPVQAAKRSESLQKAENATTSTNLQNTQSPVQEAQRTQAQLPSTGEDYQAY
LVAAAAMALIASSGMVAYGSYRKKKQK*

>group_3639

MSKEVLERRVEVSANLDDYVYDYEEKKHANFLKSGDSNDVAKQNAVQDLKTNLTNDNLTLVDTTYDPDTGVAAI
AVKDSQTGETYISYAGTNLDADGLKDVGSDLISGLNDPIHLKKLGNQATAFYDQIQATNGTNITVVSNGHSYGD
FLATRVAIERQVPFKFGFQGAPQSVSKLTAQEIIVNQIYREGGIPPKELLDVADSARVESDRVSKLMENYKGY
AVTFSTTGDLTNGLWRQSDREINFDGSKINGPSTWGLGFLSGTLGTFTPDYKYVGHVIAIDVNTESHSYNYRV
DQSVMAYTKQIISENIGVDFNQDGNQEFFLTQDYLTTQSLPMTAQSAKEIKLDDKKAMTVLSNNLRNVKSQI
ESLIDLTSQAKRSNDDVLSRLSSRRQSLKDAIMRELEAVSLIEAVRQIDDAFNELGNLNDISTVANYDSYSF
TRKFDWFGDSGNIEWYQDGSKWYGPISRLSDSSWEASCLKSLISDNQLTCTPGYGYRYNTNFCAYPAKTDV
AGEGVELINSFEKMIETTSGLDNRSHFDDGIPQATKEILQVVLNDLTLTKGCTQYILISVTELIHRLSDSD
SLATDIKNLDSLRVPEVRVTISDDYQTFLEESKVFDDKTVLVAFDDQVDQKSEDLDKMSKAYGEYLSTVEHQ
IESVNRYLGLKLAENFSQLSHDMPKQLQYKKEEFHLFSEDNDQTPMDYKGIGDHITVSSTISKLAGNGAGTGI
EQVDTKLTQAATTINTVYSYLGNFQKQPFRLGMEEAFYGVSGLDEIVKSQALAVSSVLKAMHQRFEFQSTLSYH
SGLAVEALSCLKLDGTLTLLTSLNSVSKMLDECFGD*

>group_3645

MFKKTNKYAIRKLNNGANSVVIGAIGMMALCGGFARTVSAEQVTTNTTQTQPIITVVSPTNQLSPAIALAATQ
PAVVVSNPDAEIVTTTTTAPNPAPAKTEDKMTVTEPADQVNQVVEGGTHTETSTRRGDTYSYNAQGSTTKT
IVQTGVDGTTTVTYRVVYDNGKEIARDVISVTQNPTEKPEINNYTPDIIVDEHEVFNHLYPGDTPYTGGIHK
PMFVPLSVPFESGVALKLDDIVWTDGITKEFAQNMVFTFNTWDEFISKQSVPGFELFLNLLAQEKYRNAFT
DTSHTTLPNRYWDPNTGEQITDPDSIQEIVNRSTNGNAPDAEQVDLTNYSKDKIALLNEKFYTAPVYYPEG
TLITNPDGSPVFDEFNPTFYFANTPMLKEDGTPLTVRDLVQSFQYYLNIREYWKAHEDIKRDVIVRFNPNL
KSSEKVLVQQGQDGRIEYAHFDYAKYIDGPAQYFFGPGPMVAMSIEPATSTPISTSTTDINGITYVTDPMTG
EKYIDANSYLLSEINKNVAQKNFSYSSMQVGDYKKTLPYVGTQVNNPNYDPSQPESSTNLRFIIEYSIQE
MPLDVIITQFVQATPDIYEYGVDAFTSIATVLPRTDLALTDTELEVFPFNTVEVRYNPNLKPGESRIVQVGSNG
TTEVTVTYDSATDTYPTDIIETVISTPVSEIIIEVGTGVTGQAVEHSLEVIDYAIEIVYDSNLTSGTVQIDQKG
VNGYKDVLTSTTETLNGQVIGTAKTTTNQKDLNTIIRIGTKIEPTPEKNVVVTHEESVQPPQIEIERYNDQLTA
GTSNIIQKQDGFVTKIIISQPTVNGIPQGDPTVQLVTDQAVNTIIELGTRPIEVIKEVPVEVIKEVPVEVIVE
KEVIKEVPVEVIKEVPVEVIKEITVEVIKEVPVNPVVDKIVEVKEVIEVTEKEVPVEVIKEVPVEVIVEKEV
PVETSITVETIKEKIAYPTEVRKNPDLSAGEIHVQKGFGEASVTTKTTVVTKNGQSTSTKTEDISTIIDPV
AEVIEVGTKVDIPVIPPPIEHETPTDPVNPTEPTVPLTPPIEPIVPSVPDSPTPAPTANATALASYQVAPVKK
EGVLPVTGSHDQTALLSLTSLSSSLALGFALLQTPKKKED*

>group_3657

MVQIVLEFFLGNISDVYVIRVVSTYDPNSANPIVQSGTMSSYDYGYYSIVETNNYVVFSTSDTSGGTGTAVTY
SIGDYVWLDTNKDGLOTGTGENPLANVLVTLTYADGTSKSVYTDATGHYRFDGLADKTSYITITFTPTGYLFTT
SNVGTNDLVSDSGKTVTVTINGSNMNTVDAGFVPDSTSTSLSLSTSVSKSKSTSLSTSTSGSLSTSTSSSL
ASTSLSTSVSKSKSTSLSTSTSGSLSTSTSSSLASTSLSTSIKSTSTSASTSSSLSLSTSSSSSLASTSL
STSVSKSTSTSLNSTASLSASTSTSQSTSTSLSTASASTSLSTSVSKSTSTSLNSTASASTSTSIKSTSL
LSTSLASTSASTSGSLSLSNSISGSTSSSTSLSTSTASKSTSSSLASTSGSRASASTSTSTSLSSNSTSL
SVSSSLSSSNLSQSDSQSISSSSVASNSASTSVVNSASASVSQSIGSQSTSNASLSTSTSTSLSTASQS
ASLSQSTSTSRLLSQSVSTSVSESDSTSTSTSLSVASLSKETQSQSVSTSTSVSDSASSASQSTSIKLSQSL
STSNQSESQSTSLSMSTSVSSSSSLSQSESTLQSMSISSSTLGLSLSDSSSQSESASSSQSLSDSESQSASTS

LSLSLSESISSSESLSQSEENSQSESVSTSTLGSLSDSSSQSESVSSSQSTSDSESQSISSSLSLSLSESISS
SESLSQSEAVSQSESVSTSTLGSLSDSSSQSESVSSSRSTSDSDSQSLSESQSLSLSGSQSGSESLSQSQSES
VSLSLSDSKSASVSESVSLSESLSTSAANSTGVADSASVSDSESRSTSISESISKSQSESDAQSDSESISSQ
SFSLSVTISESESTSLSESLSDSRSLSESTSTSDSTSVSDSLSESVNASESMSTSIQDSVSTSVSESI SVSDS
GSLSDSVSQSDSRSLSESTSTSESTSIISDSLSESLMTSESMSTSEHDSVSTSVSESI SVSDSGSLSDSVSQSD
SRSLSESTSTSDSTSVSDSLSESLMTSESMSTSEHDSVSTSISESISVSDSGSLSDSVSLSDSHSLSDSTSVS
DSTNSVSLNSRKEIDSISASQFILASAIASDSTSVSDSVSTSDSLSQSDSHSLSESTSTSDSTSVSDFLSA
SLMVSESLSTNIQESVSTSVSNSISVSDSGSISVSIQSESQSLSDSAYQIESLSMSSSLSGSISSESLSSTS
QNQSLSVSVSDSTSVSASLLESLSFSQNASQSLSSSESTSTSLVISDSISQAI SDMSQSDKDSVLISTSISD
STSDSVMSDSSSLSQANSEYVNSVSGSGSLSESI STSQTLSVSGSISLSGSHASDLSQSLSLSISASDS
VSQSLSQSISASLSASSASASTSDSVMVSLSKSGSTSLSTSLSKSTSGSISQSDSQSISTLVSQSASVSVSE
SLVSSMSISDSLSSIIISNLSLTSQSSSNKTSQSLSNSQSI SQSESDSLSTSLSQSAAIVGSETAKMYQIGN
FWWEDKNANGIQDAGELGIAGVTVTTLTKPDNTVVSTVTDNNGYEFDHLIDGKTYTISFETPSGYIPTVADVD
PTKDEIDSDGTTVTITINGSDDLTLDSGFVKETYSIGDFWEDINKNGIQDIGEPGIPGVTVTLTNPDGTTVS
KTTDVNGFYQFSLHVDGETYTISFETPPQGYSPITISNVDSRDAIDSDGSTVSVTINGANITSIDSGFVKETTYI
GDTVWEDTNKDG IQDAGEPGIPGVTVTLTNPDGTTVTTTTDANGHYEFTDLPNGDYTVTFETPAGYVPTTSNT
GDDSLSDSGTRVTVTVDGSNNSTIDSGFIKETYYIIGDTVWEDTNKDG IQDAGEPGIPGVTVTLTNPDGTTVTT
TTDANGHYEFTDLPNGDYTVTFETPNGYTPPTTSNTGDDTKDSGQVVKVTVDGSDNPTIDSGFVKETTYIGDT
VWEDTNKDG IQDAGEPGIPGVTVTLTNPDGTTVTTTTDANGHYEFTDLPVGDYTVDFETPAGYVPTSSNTGDD
SLSDSGTRVTVTVDGSNNSTIDSGFVKETTYIGDKWEDTNKDG IQDTGELGIPGVTVTLTNPDGTTVTTTTD
ANGYEFETDLPNGDYTVTFETPAGYVPTSSNIGDDSLSDSGTRVTVTVDGSNNSTIDSGFVKVEQPTPGSNSS
SESLSQSTTQSSSQSSSAKTVVSQTAAQLPQTGQAENNGLYGSAALAILAALGLAGKRVEK*

>group_3674

MSKPMTKKKKAISIQSVKPI LGFTFGALLSTVFTPSVFAEEVASSLGHATSGLLSVSPKELTSLETTTYL
MASESPSNTLTSDTISSDNGGTASNPNETVTTETTSEAI PFDTEVIQNPDLPIGEIKVVQEGVAGEVTVTKTT
TTITQNGVVSQSTTTESRVPVKPKPINKIIEVGTKEISTSPSSSDVITVSPSPSSTSSSENQQGSLTPAPKSRQN
SQEKKGSQTKKSKDDAKEKEGDKKELPPTGSQESGIFSLFSALISTALGLFLLKSNKND*

>group_3695

MKKEKLLSLITISGIALLGSTSVFASDVTDTLIDNQPVVTTSPSPDPVSDTTDTSTPI PPSDPVSGTTDTST
PI PPSDPVSGTTDSSVPI PNTDDTANNDSSSTATSDT NIEKPLDKPNENIPTDPSKPVPIQEEPVKPVTKPII
DAPI TTETGAQIVGTQDGKVLVQTETGTQVKEAKEVGGEVQKDGTVVIKKADGKIEVLPHTGDSKKVFTVLGI
ILILGAFWVGFKENIKKFLTMFSKKEKIRVGKAMINFMFAFFFLDNI *

>group_3850

MEIKQKHGKHALRKAVTAAVLAGTAFSSLGGFAGNLTSVAADQIVKGRVFRPNLVDTQDYEKIIISAFASFRAK
EGSVADLRKVLKDASLDTLYIILLSGLDPADLTYN GESNIIQT LAKMLNTFENDKEENEDFYHRRGINFITLLI
QEITDRIVQNEQRGKEIIKKDSELI LMNDKVKTLES DKENIENKLQDKSQELAEKEEKLKETEKELSESRHDF
ADVNEELEKVKETVKEKEASVKDLTEKLETSKKNHAELAEFEKTKGEYEKELSENDKKLSEVAKENS DLVSE
NKLKENLGTAEIITNDLQKKVMNAEKAMKELEELKAEKEALEAEKAKLAESEKANEKLTEERDAAKKEAEK
VPELEEQVEKLVEEITA AKKEAEDLQAKAEALEKDFEAVKAEKEKLEAEIAKMKEDHQKEVDALNALLADKET
MLKNLQDQLDKAKEEAMKNEQMSQEEKAKLQAE LDKAKQELADKIKDMPNKVAPQAE GKANAGQAAPNQNN
QAQANQAKNGNNLPSTGDKPVNPLLVASGLSLMIGAGAFVYAGKRKKG*

>group_3863

MLKKQLLVLTCTVSTLGLVGTTAFAEDVVPVDPTTPSTE VITPSTPIDTGLPSDTTDPSTPVEPETPTNPTNPS
TPIDPGTPTDTTTTPTDDTTTPTGTSTSPSDSESTNTNPGTKPTDEVKPVTPSTPEVPSTVTPGVVDNVDKETGN
ITIKPIPI SPDKTVVGTQNGNVLIQDRSGRYLVSA SELGGRLNDDGTVTIKDSEGKEKTLPHTGEEKGFLSII
GGTILSFVAFLFKKKNNFKLTKKQ*

>group_3895

MTHMNKNGRYKQRF SIRQYKFGAASVLLGTIFALGMTGTTAQAIYNDYGGYYGAPEIYSRGFNGLATFPGV
GSNPTYSNVDGYPI SNVNPNLNNGSYGYQPSTYSTASYDSNYYNNYNLLQGPPGIPGPPGTPGIHGMNGASG
LAGPAGPQGPSGAPGAPGPVGPKEPGQVGPSPGAPGSPGAPGKPGEPGPAGPQGEAGPKGEKGDGPKKGEKGD
RGKKGEKGDGRKKGEKGDGRGPEGRKGDRGARGRDGQKGAKGDSGRDGRDGGQGPGEKPGRDGHNGAPGRDGLN
GEKGDGPKGNGLNGKPGAKGADGQDGRDGRDGNPGKPGKNGAKGAPGSDGAKGRDGDQDGAKGRDGDGTNGVAG
KNGRDPAKGRDGDGTGKGQDGKHGRNGQDGRHGQDGDVDRNGLDGRRGRDGRDGRDGDVDRDGLSPIIKT
MTHSDGSHTI HFLNPDGTRSEITLRDGKNGQDGKD GAPGRDGDGKD GAPGRDGRDGDGMPGRDGRDGDHDK
DGMPGRDGMNGKDGQAAAGNTAGKANASDMKPKAMAAPAAMTNQNAHANNNGPAKAQLPSTGDKANPFFTAAA
LAVMASAGMVAVSRKRKED*

>tolA

MTKTCNHHFLVNQEKGEKHVFRKSKKYRTLCSVALGTMVTAIVAWGGTVAHADEVTPKVDTTIQR TENPATNL
PEAQPTPVSEQTESLVSTGQSNGGIAVTVPHDVTVQAVEEAKAEGVSTVEDSPMDLGNTT SASETSQQISKAE
ADAQKQVEAINEVTESYKADKATYESNKARIEQENNELSQAYEGANQTGKETNAWVDTKVNDLKNQYADADVA

VKEQVVSNGT SVLDYTNYGKAVETIQSTNEQAVADYLTKKTKADEIVAKNQAIQKENEAGLTKAKADNEAI
ERRNQAGQAAVDAENRAGQAAVDQANQEKQQLVSDRAAEIEAITKRNQEKEAAARKENEVIDAYNTKEMERYQ
RDLAEISKGEEGYISEALAQALNLNNGEPQAQHGAI TRNPNQIIISTGDAMLGGYSRILDSTGFFVYDSFKTGE
TLSFNYQNLQNFARFDGKKISRVTYDITNLVSPAGTDAVKLVVPNDPTEGFIA YRNDGNGDWRTDKMEFRVVAK
YFLEDGTQVTF SKEKPGVFTHSSLNHN DIGLEYVKDTSGKFVAINGSTVQVTNEGLARS LGFNRASDLNLPEE
WDTTSSRYAYKGAIVSTVTS GNTYTVTFGQGDM PQNVGLSYWFALNTLPVARTVTPYSPKPHVAVELEPVPEP
ITVTPDVFTPKTFTPEKPVFTFPKPLEEVVQPSLTLTKVNL PVKPIPKELPTPPQVPTVHYHAYRLTTTPEIM
KEVVNSDQANLHEKTVAKDSTVIYPLTVDALSPNRAQTSLIFEDYLPAGYLFDKETTQKENGNYVLSF DATK
NFVTLTAKENLLQE VNKDLTKVYQLNAPKLYGSVQNDGATYSNSYKLLL NKGTTNAYTVTSNVVTVRTPGDGE
TTTLITPDKNNENADSVLINDTVVALGTTNHYRLTWDL DQYKGDRSAKETIARGFFVDDY PEEVLDVVENG
AVTILDGQKVSGITVKTYASLNDAPKDLQYKLARAKITPTGAFQVFM PDDNQAFYDQYVKTGTSLALLTKMTV
KDSLYGQTKTYTNKAYQVDFGNGYETKEVTNTLVSP EPKKQNLNKDKVDINGK PMLVGSQNYTTL SWDL DQYR
GIKADNSQIAQGFYFVDDY PEEALLPDEAAIQFITS DGKTVSGIMVKS YAQLSEAPNTLQAALSKQKI QPKGA
FQVFIPEDPQAFYESYVT KGENITIVTPMTVLETMLNSGKSYENVAYQVDFGQAYETNTVTNFVPKVT PHKSN
TNQEGISIDGKT VLPNTVNYKIVLDYSQYKDMVVT DGV LAKGFYMVDDY PEEALTLN PDDIQVL DKGDN RV S
GISVSTYASLSEAPKVVDAMAKRQFIPKGAIQVLSSDDPKTFY EIIYVKTGQILVVTLPMTVKNELTKTGGQY
KNTAYQIDFGLADVTETV VNNVPKLD PQKDVIDLSHKDES LDGKEVALHQTFNYRLVGAMIPSNRATDLFEY
GFEDNYDEKHDEYNGVYRSYLMTDVTLKDG SVLKEGTEVTKYTLQQVD TENGLVSISFDKSFLETISD DS AFQ
ADVYLQMKRIAIGQVENTYLHTVNDYVISSNTV VTHTPKPEEPS PNQPTPPQPP IETIEPPVPSSVLPNTGEQ
ESLWGLIGAGILLGTAYGLKKKEER*

>group_3961

MKIKIKNNMRNNIKKNHRQPRYALRKMSVGFTSCVIGWCIFSSSTTVHASVLPDTSISTSTTSINNTTQTSP I
TENVVDTSIIITKPIDSNKLIESTATTIETTSTVETITVALENNTAVEPTNTTVENTTAVKPSTATLETTTAIE
PITTSVENTTAVEPITTSVENTTAVEPITTSVENTTAVEPITTSVETISSTDGIATNSTSTSTSQLQPKMLMM
ISPMSVSPTAPSIYISNMTELLNAINNYS DVNIFIKGTTYLIDKAISIASGKNISFNVDPSSTEPVIFLRDVS
KGYLSGSLVTV EQGASITFNSDPGQMIIDGGSTDNPSSGIVSETGLLLTNKGETIINGAEFRNQNLFYGSYIA
PIYTS GPGSKLVINSGSIHDNKITANSERIIYSSTAILAENESQVFLNGGEIYNNTIDKSLINSDSVGVINIS
DGSYFEMNGGQIHHNSARSGVTIGDSRFYAYYIKPETWQTEGTIKNNPTGLLTIDQIFVPGK RMPKIAKAVIN
NGTINNMMAYANGGGLNVWSASEVTINGGKIENNETNGYGGGISVIDNYLNYLGGYTRKIANV KRSEWEQYLG
AKLTINNATINGNKSTDGRGYGGGMFIASNLVTINKADISNMMAVNGASIAMTDKAYTLKKNVSTDQSIGNS
SFSNSSDSNLYIYEKDG LVFTPINPGTGTSNSNYDLTANWATNNILD PAYRMLGGGLINYIDKLGNTVNAEPK
YSNSVTPPTPAHYIDPIDGYATKSSLDYGKTLASITMNGNYSVYYYR ALNRYYP SYSSGSNFFNGGTLIIGTP
DNIKNEIDVNATLNFKTIYDSFLAQDPNNILKLT VNYETPMGNVILEEHDIPSLEGLVQIIIGNLSETNDLSKL
KFQFSGINTDETITLTADNFIKSETISSTERVWNSDTFMFDEKTYTTLNHFSILLDTVIKSVETITSLPEKI
IYDSTKPIGTEIVEPGVDGLTQTNSTILKINGIESDVIMTSQ NVIKEPIAQIRIIGTG VVGQNV DVTFEV IPI
PTQPTIIEDETIPLGEEVIVETGSEGQIKVTTTTPTLNGIPNGEPIVTREIVTQAKAKVIKRG TGISSSNIDT
IKARITKAISDKKHEVNANETIILPNTGFKHNHKQTYGLLSILLSLILFVFKSKRRTK*

>group_3392

MSASTSASTSASTSTSVSASTSTSTSASTSASMSASRSASTSASTSTSVSASASASTSASTSASTSASTSAST
SASTSASESASTSASESASTSASTSASTSGSESASLSASTSASTSASESASTSASTSASTSASTSASESASTSASTS
ASTSASESASLSASKASLPNTGETDDSSSFITSL LALLLALGLISK RDEK*

>group_4004

MFKKLNHQTVFETEAKSRVKLHKSGKSWVKTTLASIGLIHLCKGMQDKTLLSEESDIPLLTGQLLKGT LGVGA
LLGGVSVTQTAFAEQVSSLTLPVTSETGSTLAGTDSAVV VASTSSSTSDTSSQSSSLSVSNSQSLSLSTSGSTS
LSISQSDSVLQSSSLSQSTSASASTSSANTSTSLAPKSAMTMSTATTLAATGTVNNNLVTVTSSYIKDTSYND
GYIYPQAGESLEY YTKFTIDNAATAGDQFTMTYSPYTSPSDFD MANWTPTDITDSSGEVIATGTWNATTKTVT
FTFTTYVDKYQNIVASLDAYSYIDRAKVPNDANLTLTYSLGGETTSKLEYIRYQSPIVSNTSSIQSVFNELDS
VNHTVEQIFYVNPMDYAAYSTYFNVGYQIDPSTGAMVQQGSTVINSTTSIQVYKAGTGSTLPDSMDVLDYST
LTKVTPSITYGTDSARIFLGNISDVYVIRVVSTYDPNSANPIVQSGTMSSYDYGYYSIVETNNYVVFTSDTS
GGTGTA VTYSIGDYVWLD TNKDGLQT TGETPLANVLVTLTYADGTSKSVYTDATGHYRFDGLADKTSYTTFTT
PPTGYLFTT SNVGTNDLVDSDGKTVTVTINGSNMNTVDAGFVPDTSTSTSLSLSTS SVKSMSTSLSTSGSL
STSTSSSLSGSTSLSTS SVKSMSTSLSTS SGSLSTSTSSSLSGSTSLSTSISKSTSTSA STSSSLSTSTSS
SSLSGSTSLSTS SVKSTSTSLSNSTASLSASTSTSQSTSTSLSTSASASTSLSTSVKSTSTSLSNSTASAST
STSI NSLSLSTSLSASTSARTSGSLSLSNSISGSTSSSTSLSTSTSA SKSTSSSLASTSGSRSA STSTST
LSSSNSTSLSVSSSLSSSNLSQSDSQSISSSVSASNSASTSVVNSASASVSQSISGSQSTSNSASLSTSTST
SLSTSASQSASLSQSTSTSRLLSQSVSTSVSES DLTSTSTSLSVSVSQSKETSQSVSTSTSVSESASSASQS
TSILNSQSLSTSN SQESQSTSLSLSTSVSSSSSLRQSEENSQSES VSTSTLGS LSDSSSQSESVSNSQSTSD
SESQSSSSSLSLSLSEISGSESLSQSEAVSQSESVSTSTLGS LSDSSSQSESVSSSQSDSESQSDSSSL
DSVSSSISLSTSVFESESTSVSVSTLGS LSDSSSQSESASSSQSKDSESQSASSSLSLSLSEISSSESLSQ
SEENSQSESI STSTLGSQSDSRSQSESASSSQSTSDSESQSSSSSLSLSLSEISSSESLSQSEAVSQSESVF

TSTLGSLSDDSSSQSESVSSSQSTSDSDSQSFSESQSLSFSGSQSGSESFSSQSQSESVSLSLSDSKSASASESV
SLSESLSTSVANSTGVADSASVSDSESRSTSISESISKSQSESDAQSDSESISSISQSFHSHVTVSESESTSLSE
ESLSDSRSLSEFTTSTSDSTSISSDLSESLMTSESMSTSEHDSVSTSVSESISSVSDSGSLRDSVSLSDSRSLSE
STSSSDSTSISSDLSESVNSESSESMSTSEHDSVSTSVSESISSVSDSGSLSDSVSQSDSRVSESTSSSDSTSISS
NSLSESVNSESSESMSTSENNSVSTSVSESISSVSDSGFLSDSVSQSDSRSLSDSTSVSDSTSTSVSLNSRKEID
SISASQFILASAIASDSTSVSDSVSTSDFLSQSDSHSLSESTSTSDSTSVSDSLASLMVSESLSTNIQESVS
TSVSDSISSVSDSVSISSVILQSESQSLSDSVYQIESLSMSSSLSDSISSSESLSTSQNQSLSVSVSDSTSVSV
SLLESLSLSQNASQSLSSSGSTSTSLVISDSFSQAISDMSQSDKDSVLLSISSISDSTSVSASMSDSSSLSQD
NSEYVSNVSGSGSLSESISSSQLTSTSVSGSDSLSGSHASDLSQSLSHSISASDSVSVSQSLSQSISASLSASS
SSSATTSDSVVSLSVSGSTSLSTSLSKSTSGSISQSDSQSISTLVSQSASVSVSESLVSSMSISDLSLSSI
ISNSVSTSQSSSIKTSQSLSNSQSISSQSESDSLSTSLSQSAAIVGSETAKTYQIGNFVWEDKNANGIQDAGET
GVAGVTVTTLTKPDNTTVVSTVSDSNGYYGFDHLIDGQTYTISFETPSGYIPTVADVDPTKDEIDSDGTTVTVTI
NGSDNLTLDGSGFVKETYSIGDFVWEDINKNGIQDIGEPGIPGVTVSLTKPDGTTVSKATDVNGFYQFSLHVDG
ETYTISFETPQGYSPTISNVDSRDAIDSDGSTVSVTINGANITSIDSGFVKETTYTIGDTVWEDSNKDGIQDAG
EPGIPGVTVTTLTNPDGTTVTTTTDANGHYEFTDLPNGDYTVIFETPNGYTPTRSNTGDDTKDSDGQVVKVTV
GSDNPTIDSGFVKVHTIGDTVWEDTNKDGIQDAGEPGIPGVTVTTLTNPDGTTVTTTTDANGHYEFTDLPNGD
YTVTFETPDGYTPPTSNTGDDTKDSDGQVVKVTVDGSDNPTIDSGFVKVEQPTPGSNSSSESLSQSTTQSSSQ
SSSAKPVASQTAAQLPHTGQAENNGLYGSAALAILAALGLAGKKRNEND*

>group_4009

MSKKQIGSDYVQTDKSRVKLHKSGKSWVKTTLTSMGLLHLFRGAKTEERVLDTDLDLEAGFSSTHWFKGAAA
LGALVGGAAVTNTAYAAETTSTLPVTSEVATTGTTQVDSTVVASASVSDTTSASASTSASTSQSVSEIQSNSF
SLSTSISESLATSSSISSQSTATLNSQALAPKAGQTATDTSKAATTSGLAINEAITSSSTITITDTGNYDSNGSL
IPSTIGDGTVLPTSAENIAIDINMTFSDAATAGDQFTIQFSDTVKTDDLAAPGDVYTPIDITDSSGEVIATGT
YNATTKLVYTFFTNYVDKYENVKAVLHLEKYINRDLIQYDTNAVPTVYTVGNNSVTEYFVKYGLPMTLNNAQ
IESSFSYYDTTNYKVEQIAYVNQDNLNLYNGYIDIFAGANTSGIINTTDTIVEVYRVPAGMQLPDSMEIQDYS
QLEQVPVTTIDDRGTNNIGFYLPNPYQYDTSTYVIRVVSSYDPNSTNPIEQFVRLTDGYDTTYVEYSNFIKEAE
DTSGGNGTQSYSIGNYVWVDTNEDGIQNESNTGLAGVLVTLTFPDGSTKGVRTDANGYYEFTGLTDGETYTVS
FETPSGYLPTKVGGGALDDTDSNGLSTTVTINGANDSTLDTGFIIDSTSDSTSTSVSESTSLSDSTSTSTSISS
DSTSESVSTSTSVSDSTSTSDSTSTSESTSTSESTSLSESTSVSDSTSLSESVSTSDSVSVSDSTSVSESTSV
SESTSLSDSVSISSDSTSVSDSASISDSTSVSDSANSASDSTSTSESTSVSESTSVSDSTSLSESVSTSDSV
SVSDSTSVSESTSVSESTSVSDSTSLSESVSTSDSVSVSDSISSVSDSTSVSDSTSLSDSTSVSDSTASLNSG
SDSDSASVMSSESLSTSTSVSDSTSLSESVSTSDSVLVSDSISSVSDSTSVSESTSLSDSVSISSDSTSLSEST
VSDSTSVSDSTSTSDSISSVSESSSLSTSISESTSIIGSETSTVYTIGDTVWEDTNKDGIQDAGEPGIPGVTVT
LTNPDGTTVTTTTDANGHYEFTDIPNGDYTVTFETPNGYTPPTSNTGDDTKDSDGQVVKVTVAGSDNPTIDSG
FVKVHTIGDTVWEDTNKDGIQDAGEPGIPGVTVTTLTNPDGTTVTTTTDANGHYEFTDLPNGDYTVTFETPDG
YTPTTSNTGDDTKDSDGQVVKVTVDGSDNPTIDSGFVKVHTIGDTVWEDTNKDGIQDAGEPGIPGVTVTTLT
PDGTTVTTTTDANGHYEFTDLPNGDYTVTFETPNGYTPPTSNTGDDTKDSDGQVVKVTVDGSDNPTIDSGFVK
VEQSTPGSNSSSESLSQSTTQSSSQSSSAKPVASQTAAQLPHTGQAENNGLYGSAALAILVALGLAGKKRNEN
D*

>group_4104

MNQAKIITGLAVATLSTSVGIVHAEDVTPVTEASTTNETVTVPEVTQSQVDSAKVIADQATSDVNAQQNVVND
AQQAQTNQAQSNVVSATTAVKDATATVEQATPEVVTKAQAEVTEATTAVKNAETNLITAQSDVTKVQEAQVANQT
QVVAENQTSVDQAQSDVTKAQAEAVGSAKQSIDTTKANANLAEAEKLVAEKTNTVATAETSLSEAKQVDAKLAE
EIIQAQATVTKELAVKDTQDLLNQVASEMSKEQVTTSLQNQSYYNQRDGAAGYGNNTFASTGCVPTSLAM
VFTELARRGVTPTEVANYLYNNTNYNKFYSGTSANGIVSATRAFGFVPTHLDSONAIAEALQAGHYVVGAVQ
NNKFSPWGPQYSHEVVMRGYSNGNTYVYDPYNRANIGWYPVANLWVERSGDKDDNALGVPFFKITTTQKMASIE
TQKAQLTSAVNTAKSQLDQAKQVLSNLQATSLQTPDAQSKLDQAKVELALAQDNVYKAEAVKLASKELAVKE
ANLKNAQADLLAKQNILKEAQATLAESQLLLLASLQDNLKEAQASVADAKTSLDTAKTNLAQKQANLLSLQAN
KVLTEAQAKLVTAKTDLANKMAILDKEVAKLKEAQAVQAEQANQYSIVFEAFKAVLEAKKQAELEIYNHIIA
GGGEAIPVVDETGKITGYVDGSQKAVGNEVTALTSNGQAPLESFVNKENHDLTKTSQVLPHTGEAGVSLLSV
LGAGLISTLGLISLKKRRTH*

>group_4120

MNSNIKGHYFRKSKAYGLVCGIALAGAFANSGNVSADEVTAQATQSQTVVTTLTTEPTNNQAQPVTSALD
AVDSAKDAGVVVNTTQPVSHTDVPSAQADLANQTQAVKDATAKAEANTQAIKDATAQNAQIDAQNAQEAAT
EINKAGQLAVDQKNAQAQALVDKQNAQAQAEATNAQLKADYEAELAEIKTVEDYNKAVDERNALAKQQA
TNAQLKADYQTKLDAYNKALADKINLIANDVSFQGYGKSEDLNALGAVTTDTVTIDGAGNFTLKEPQSDPQGI
IGYITTKGKLYSSVYDPTTGKAKITIDSITLDTWQLDLTRPSTSGITNASAEYIGLDGTSFLKQAYTGLES
APITIGKTSAGIQTFTLASGETTPEFMFLKTNPFWQYFAPSSLFALKTYNTAPLPNKPEPTLVTVTPPEAL
VPQTPPTPVYVPTPLKSFTPEVYTPIKPTVKPHVAVPDKLVVTVSVHPVMVPVANPSKDVVDQAGKSIN
GLSVLPNSDLNYVAQQDFSQYKGMTASNDKVAKNFIIYIDDYDEKSLDGKSMVVNSIKASNGQDVSSLEMHVLSK

VLDEKLQAILNNSGLSPVGEFYMWVAKDPSSFFTAYVQKGLDITYNLSFKVNQTFTEGEIVNGVAQIDFGNGY
LGNLVVNDLPKPEVHKDVLDKQDGKSINNGTVKLGDEVITYKLEGWVPTGRSYDLFEYKFVDQLQNTHDLFLK
DSVVAKV DITLSDGT VITKGSDLAQYTETVYNKETGLYELAFKKSFLT KVVR SSEFGADAF LVVKRVKAGDVY
NDYTLYVNGNLVKSNRVVTH TPEAPKPVTPQKLTPKTPVLPHTGEASMSIVSVIGASII SLLGLSRLRRKVEK
★