Supporting Information S1

The benchmark dataset $S_{\xi=7}(K)$ used to train and test the model for predicting the possibility of carbonylation at Lys site. It contains 300 positive samples and 1,949 negative samples. None of the sequences included has $\geq 30\%$ pairwise sequence identity with any other in a same subset.

(1) List of the 300 peptide samples in the positive subset $S_{\xi=7}^{+}(K)$

SAPAKGRKKSKMGFQ
ENSLAEDKDGRRKGK
VTTGIRDKKGVTVKT
QCRLKLLKLERIKDY
NIGYGSNKKTKHMLP
VTGGAASKLSKIRVV
QTTWRKYKLKTDLKR
RKLLMLKKEKLEKVQ
QKKWKIMKKAALLIQ
GIGRLILKEEMKARS
DKLTTADKLLGELQE
KLQVLPQKASERLQF
GITNLGYKVLIQPSN
LAEKVFAKVLREEDS
SERSGTPKKRKAPPP
AAKEMEEKISNLKEH
ADAGRDKKEKVFEKH
SSAINILKEKKKREK
GPRLGDAKLKEKFKD
VSGEHKQKGKVKRKL
QQEVVRMKLRLQHSI
DLQEQMNKNKMMPVL
GQQDTIKKALNYSTA
SKVTYLGKELLRYVS
DEKMKAMKRSRTWGQ
KRQLQTPKEKAQALE
KEPVGEGKGTKAFKQ
EADRTSSKKTKTQEI
QRVLDILKKSSHAVE
HKKPIDLKAIGKLPI
VKQTVQSKILTTGEN
KRLGRRFKKLDLDKS
IPVQLVFKNKIKLYW
SQEIKDTKECVQNKE
TRFLEEIKNQDKLNK
CNHIREVKSLLEELS
DCSQDLLKKREHHIE
KTLFSCWKSENEKLL
SPPAACEKGKEQHSQ
AKEHKEPKQKDGAKK
YRIELLRKVREQVLK
RKFVEHVKSQGHKDK
KTLNDSLKKVENKVS
RQICVTQKLLPAKRS
TEKEMIQKLDKLELQ
KQIEKELKQMELIKD
YDASVRNKQQELHLE
KHLKTLSKTVFSETW

```
RSIIGFVKLRSNKVK
NKVKRSQKLVYMGKL
YTKKDTKKNAKKSSD
CNKLLAAKNINLRLF
SGYEEALKIFQKIKD
PIKQVEEKPQRAQNI
EPTERPLKGIQSRLR
EPAPKKQKLSASVKK
DKLRISTKKLEEYET
LQIIPGAKHGNIQVG
RKLVERRKQVKQLMK
FKLGNKVKSEVNKLY
IEEKDLEKLDLVIKY
ECREEILKFLCIFLE
GLFAGPCKVINAKDV
MASLRRPKRIIIRGH
KVPADLLKRAFVRMS
GNGGVVFKVSHKPSG
IPEQILGKVSIAVIK
DSRSLAQKIVATYRL
NIPPLTMKRIRERFI
ISGLGGEKDRWTEAA
RNIYQGVKKKQYDIL
LKLESKLKADKLYLQ
EVQEVKDKAQKIVDE
KVLMPKLKAKPIRTA
LLNRLNRKMEIKPLQ
HFRSLRTKLLLMSRN
VIQEIVDKSGVVRVR
GRNQEKLKKFMIRIG
KKRIKKRKGEAMALN
TKEKAGPKGSKVSEE
ERIFTGNKFTKDTTK
KILKKINKAIVSKKN
VANAVFVKNASEIEV
CIVILTGKDPLGETF
TEKLAGVKQEVIKES
EREAELKKLQEARER
KKYLEDVKLQARGQL
AECVKQAKGVRQQAV
SISQKKVKIELDKSA
PLAGKKAKKPAKAAA
AAKARVTKPKTAKPK
FQQATKIKKLRAGKL
SVRGPRIKHVCRRAA
RFQLPLDKGNGKKHK
MASQRQAKIQRYKQK
VFDKVSEKLGDKRLF
QGKENLWKENLRKEE
SGKLYKTKSNKELHG
DRKTIPIKYPLKEIV
ASFQAERKFNAAAAR
GSLLGSMKLRKSESK
ESPHFYRKGTTPPRS
KKKYEKEKEKNKILR
IQYVVRLKKENVRLA
```

```
TANIQDLKEEYSRKK
ERLKRLQKSADLYKD
DILQALQKLSEDKNL
KHLQQKLKEKDRELE
GEDDERSKRIWLDYP
PDLRVVQKYPLLKEP
TNDFTKIKGWRGKFH
TLAPVVAKLGNSGAS
AKEADQLKQDLQEAR
AKKPAALKPKVPKKE
PGTTKTTKSSAVPPG
IDIYLLEKSRVIFQQ
INRALDAKLSRQFFI
NKLMTNLKSTAPHFV
HKLEGNLKLNRESME
LEVKSDSKEDENLVI
KLSLEESKKKGYDLR
KETYEKQKGHYLAGK
SQQMQSGKEYRKDYE
LLSVKQEKEIQMKMI
IISFAVQKETQFHTE
GEKYLQSKEDLRLML
LMLIELKKKQEAGFA
VTGREGAKDIDISSP
IPKMKMPKFSMPSLK
KLKGPKLKMPEMHFK
EMHFKTPKISMPDVN
NLKGPKMKGDVDVSL
EKAKSPVKEEVKSPE
APKKEAPKPKVEEKK
KARLERFKLRRMKAN
LQSQEETKERRHSHT
ADAKKKKKKSWVYEL
ISLKKAKKPPLPPSR
IEIKSKNKVRPGSLF
LAAMDATKLEYERAS
KEEKKKKKIKTIKSE
GLGVMPPKAGQTITV
GLEAAQIKELEELRQ
ELKAWQEKFFQKEQA
RKVEEVTKVCEGRRK
TLVPAGIKSLGLAVS
REVSQSRKPSASERL
RGGGSPEKPPSRRRP
KKFLPVLKEILDRDP
IHVIQQGKTGNTEKF
KIIKSSSKVHSFGKR
ERMKRHQKALVRERK
EMKKQTLKSLIIKLS
PPPIKSFKEMKFPAA
QKATTGLKPVDSGCV
IHKDLIKKPTISTAV
CRGIIGGKSQRVNGL
QLDRLLPKLARGSLS
SIVRIIGKMLPLEPC
KEKELENKLEELKKQ
```

```
KDLQEQNKKNEERMF
EGKDEDAKKSLRTME
EMSAKRAKKDVLHSS
LLLRMLEKRQMDRAE
LESMIPIKMVNFPQK
SIEKTSAKDLAPTSK
VLAKPTPKAETTTKG
MPRVRKPKTTPTPRK
KEERVSLKVLAKNFG
KNPKEEKKPKEKKKK
SRTVRLLKELRRALG
ARVTEVMKALKCNVR
LLLAPSSKKGKARLS
DMLRLANKDSIGFFI
GDRIQVRKVPPLKIP
TLADIKAKAQLVKAQ
KSIRAFAKGFLAEEK
QMYEEQIKNLEKENY
AVLLKTKKKGQKKSG
HGVCSVLKGSEFMFE
AMPGAEYKAKKAKGD
RLLDEATKRSNRDSL
TEHPYKSKKAVWHKL
LLSVRMGKEEEKLMI
MVTSLFCKLGVLVRH
HFLPLMEKLKKKAAT
EDGEKEEKAKEDKGK
VIPVLQTKTRTNVPT
PEPKKIKKAASPSPQ
TQNSYRAKEPILALR
DCENLLKKLLVLNPI
LLLAYKIKYPENFFL
ERKSPEDKKSMILSV
MYLMLDNKRKEVVHK
WSQRGKGKVFPKLRK
RGKAAILKAQVAARG
AYPRLVVKLMPNGLR
VPAAYVKKLDPAQSA
KKFDDFQKDLKANES
MAASRRAKLNESHRL
VEHWKELKQLAAARG
TELIRQEKLEQLAAR
LKQSGQKKSYTLRVR
KYDAIKFKINQLSEL
DELKSSVKTKDKKII
GRIRLPRKATKTKKN
RRKKRAPKADISKSL
ALMEVKMKEMKGSIR
SKKIPEEKVPVPVQK
EKPKAPPKGPEISEK
PIKGVPKKTPSPIEA
VVPRSEGKVHTLTLR
PKIKVDVKFKDTVIL
LKVTEITKDSVSITW
LDGGSKIKNYIVEKR
ESSVLRAKEVTWYKD
```

VQEEISQKALRSEEI
IGAARTTKKRIPNTK
GRNVKSRKLKEIFNK
PRPPSLKKKQTTKKP
GAVKGLCKLFCLTLH
LSSLGGFKLAHGLLE
HGLLEELKTVLSSHK
RLIGNESKGEHVPGF
KAADALGKLISIHIK
IGSINLKKNEPPLTC
FEALKYPKFSKAIVI
LTKQDKLKSPFKFSD
RKMEVGFKARGQPKS
FAKNIPVKSKNIRVV
SKKPDKEKPIKEKDK
EKLKEILKERELKIY
KKLALSAKKASTLWR
VNEQEPCKFLLDAVF
VFKTRILKIIDEGLK
LQLSKKLKTVLDQAR
KKKLTSKKPPGELLV
GLTVGFDKKDMMDIK
QQLNGVQKQIIICGR
DDVALVSKALQAEEM
TGDFESKKNELPDGL
AARRFLKLQQDQQE
SKAKPSYKQKRQRNR
ETRESETKLKELPGV
ANRLFGDKSLTFNET
LKLEDTPKINSRFFG
FFGEGTKKMGLAFES
EKLIKDLKSKEVPEA
EKLTALTKKYRITEN
KFLDMLIKKLKSFDY
SYNETKIKFDKYKAE
IDLLKIKKITAIITQ
VSEHEATKCQSFRDH
YLAPNNLKPVVAEFY
YAVAVVKKDSGFQMN
MNQLRGKKSCHTGLG
LFSSPHGKDLLFKDS
LSHHERLKCDEWSVN
EGCAPGSKKDSSLCK
VPQNTGGKNPDPWAK
LCLDGTRKPVEEYAN
AVVTRKDKEACVHKI
YLGEEYVKAVGNLRK
TVKVPMMKRLGMFNI
RSVQLTEKRMDKVGK
SNTKVDKKVEPKSCD
LHTLFGDKLCTVATL
NECFLQHKDDNPNLP
LFFAKRYKAAFTECC
AEFAEVSKLVTDLTK
LLVRYTKKVPQVSTP
PVSDRVTKCCTESLV

CFAEEGKKLVAASQA
SEVAHRFKDLGEENF
KAACLLPKLDELRDE
SSAKQRLKCASLQKF
DSISSKLKECCEKPL
VLLLRLAKTYETTLE
CCKHPEAKRMPCAED
SEKERQIKKQTALVE
KHKPKATKEQLKAVM
DNEETFLKKYLYEIA
WLNGKTFKCKVNSGA
VATALAHKYHXXXXX
VHLTDAEKAAVNGLX
GMYELLLKVRPEQLV
EVERTRNKFLFLKAD
LTAFTNLKIKHPTYC
DICTLPDKEKQIKKQ
AADKTNIKNCWGKIG
KVADALAKAADHVED
KAGQYTDKGLRKCCE
MSAKEKGKFEDMAKA
DDSPDLPKLKPDPNT
KGACLLPKIETMREK
VGTRCCTKPESERMP
EKTPVSEKVTKCCTE
LVELLKHKPKATEEQ
XXXXXMVKVGVNGFG
TVDGPSGKLWRDGRG

(2) List of the 1,949 peptide samples in the negative subset $S_{\bar{\xi}=7}^{-}(\boldsymbol{K})$

SIAVQMMKRIHSLLE
SPHEQEIKFFAKILL
FGEQLIQKSEPLDAV
MASFSFLKDNSTDVC
HSRALLVKTTLNISF
IKVLSSSKVLSEEIS
ETWRLWQKFLDDYSR
RINEYVGKAATRLSI
EKQKSEAKDRKVLEI
EYRLDWAKWKAKIQS
PGEHVFVKGYEKGQP
GILDRSSKSQSSASL
ARLEADEKKQLCVLQ
QNLQDAAKDMKKFEA
FVILRDEKWGGNKTY
RQHPDTLKYSTLMDS
PTTLIPAKAPEIIDV
QIQALEDKKEKEMSA
THFVARLKSWRGNEP
WYNLLSYKYLKKQSR
AQPAAPAKVPSPGQQ
SPQPESFKTSRSSKQ
AARVTGMKKWPRTPK

ECLVSAQKVLEGSEL
VTSNEQVKGYGTHLM
NRGWFCEKVVILCPF
RILAAGGKSNHLHLW
NEMASTEKLTDVARY
GSKDREPKPKRKREP
QWQFLVQKSAEKSQK
HAPQFPEKEFNITVQ
KMALEVYKLSLEIEQ
RNTRILRKCIEKVAK
KAVMTSIKQLSSEEL
IRMSARQKQIMEIEE
HLSSPPTKFFVSTPS
LENSSGTKSAFVTVR
EELRDYVKARLKVFY
IRVCALNKVGLGEAT
VINYIVEKQDTRKDT
KCSKTSFKVENLTEG
FSQPIINKVKPQLLK
YRVEAAWKLSQWDLV
KARKLFYKAIVRGKE
NAEMHKNKLQYFMEQ
PRLPKLFKQEEQREL
ALEVDFEKMKKERDQ
YSAVNPEKDIHSGLI
SSVNSLRKAAHEALQ
RNTILWVKENKVPCL
HTLVESLKLSITDQQ
ALDEFATKLIQNNHY
EVLMVWYKQIEQVLI
NWCSQMDKGMLHFGS
EKAPATPKTEEKKDS
AIPLLAAKANTKNTS
ESIHNSPKSCPTPEV
KEVYQRNKSNCTIEP
KREPKEPKEPRKAKE
VSFIENSKFFEQYEV
ALLADSEKPSHKSFA
KLVVTGLKEGAFYKF
FESHQNYKDVQDPSV
PLQQFCSKKLFSIPE
REPAGDGKSIRTFKE
RYLYTLEKVCQPLYN
NDVLEHVKHFVINLI
QALEDIIKKFASVPE
KGPALSRKRKKEVDA
VFLTLYEKVFEDTYT
TKVPEVSKKIVPQKP
TEFQNHEKQESQDLR
QEWFLGAKAAAKESS
YGALAPFKPSEPGAN
TTILQHWKKCDSWWA
CRGLLLPKVVEDRGV
GDYVDRGKQSLETIC
NVLQYCRKSGLQTDY
PKPPSKVKSSSKESS

GRPLVASKKYRISFK
MQLPYRAKKFSLYCT
DLSGADIKAICTEAG
SKADAPVKWFKDGKE
DGFVDQKKKLXXXXX
DVQFYQLKLPMVTVA
LVNPAGEKAVFVNVR
KEGLLQMKLPEPVKL
RVNAWQAKANNNKQW
ITGQLGVKPQTGNSI
GCFWVIIKGCSPFLD
AISDVETKETTYSSG
SKPPGNLKECSPWMS
LRWTRVNKDYVVYDT
PVAAPEPKKETESEA
LGTNGTVKYSISAGD
EETARREKQQLLDVQ
HQSADRAKSEMATMK
EPPPSTVKTYHYLVD
ENYRPLAKTRQQNIS
LTMEINPKVPVNLLR
DGPLSPGKMEDISPV
EMDTDQLKLYEEPLS
LYRSVYEKNKMKIHI
TEFLALMKKVSKSPL
DNAPQFLKSKYFTPV
KPEELVSKELSTWKE
EKKNEVHKVEMFLGE
RGPQVDVKGPFVEAE
EMLTPEEKALLYEAI
ITLGPPSKPKGPIRF
IEKSKSLKQTDQPKA
RNYLPALKVEYNTSA
KRGAASVKTSGSPRS
LCNQCRGKKAKNPKE
EVCRAGSKHSRPIPL
QKELVLSKPCKFEEK
SRTRDVVKSALGFIK
STDWSGVKKPIYLSK
TGASGSFKLNKKASS
KKAKKPAKAAAASKK
GSELELAKMTMLLLY
YYSALSGKSISDGHS
KGSLEEEKRRAADAL
GSLEQIMKDRWMNVG
GETTHTDKVPGGEDK
PLLPSLLKCLKMDTD
SLESISTKVIVTQTT
TSKMDLEKPNYIVPD
QAHPVMRKCLQSLCD
FHLVRDVKQGNLPPG
IMEAPLLKHKDDIED
TGPPQVGKTGSYLQF
EEDWEINKDSAVEMA
RVLQALMKRFYLPGT
LSFINGYKNIYAEKV

PLQEAESKVSMALEE
VNRETDTKRPDARLA
EHSEIHTKLYFLQWL
ETADQFLKRSLEMRE
RGHEELRKLESTLDG
REHPFLVKGGEDLRQ
TLLENGNKQHQLGVW
ERSAEESKPRSLQEL
KYQALLSKMRAIDLQ
NFSPNQTKFTVQRTD
LVFCERGKLQPPRKA
AFSASARKVVNRDSC
EKPPPVNKQENAGTL
LLETGQEKMAGDQKI
SLLIGVFKHGYERYN
DITLTNDKPATAIGV
EVSEIHIKVKPTTKS
CFPEGLVKSCSETLL
TAFVQEPKVGETAPG
PFSSADVKPKVEDLD
DLLSLSGKTLCVTAG
SIEEPEGKLKGPKFK
INQVNTIKNEAEVIN
YQPLRRSKRRWVITT
TAAVQLLKPEGVLVY
NTSLLSQKKKLEADV
ESTPLANKPGKQSGH
SCQLSREKANVKWYR
DIAEQLLKANPPGKD
DLTTILTKLKAKTDN
RVMGFPEKSDIFDVD
EDTATSSKRRPRTRA
YSIQDWAKRMKALVE
NSFISIPKMEVKSYT
AGAQALEKLEAAESL
MTTQLICKLLRSREA
LRLLEPVKRAQEMDA
XXXMAKRKEENFSSP
IEIGLEGKGFEPTLE
RSVDSSAKREKPVVR
MESICIMKGMKPERK
EISESVGKNQFTSES
APVPRGRKGKKTKNQ
QSPAVLDKADGQKPV
KLNMMLSKGELLSTL
YTTQSTIKPXXXXXX
HTQAGWRKEGNLGTH
AGTERWMKVVTLKPT
KEAYEHTKAYGYTLG
WYRLRLLKPQPNIIP
EEFVTSYKALKSRIS
WNLRLDVKKNPVQDK
TSLAIAQKLMELKLG
KVEDLPLKLTIYSEA
VTACNPYKPPGPPST
RGQVDYSKIMNGNLS

GVIYTVAKEELKKFE
TWSKDGQKLPPGKDY
LANSAARKKKLLEAQ
HPVSGVRKEQGGGCH
IFTDSEIKDEAFLEY
GEVQLTAKDFKTHAN
SPTSPFYKACDTVFK
SHDNIRVKWFKNDQR
HMYKMPEKILPKILG
KEKQEANKQKVPNPS
ISFDDSSKTQSHSDA
GAGEEEAKGGKRPKE
GEEEVPAKKTKTIVS
VFLDSGRKTRSARRR
NMVLPDEKGAGALPF
SSEKLFVKYDLILTS
TNLTDDLKALYKVAG
FNEGPCSKILIQCKA
VVPGASIKYDAVKEE
MLQVTQAKKSQAIAS
KLLKELHKESKTRDD
IEAHEQGKDIDLNKV
EVREAAAKTFEQLHS
PIPPPDAKELELMFG
TVWRQADKHNIPRIC
PRIQQGAKTSLQEEM
QSLLAPGKFSPAGVE
IDFRLFAKLNAFCVI
TSTSDILKVPKPEPI
ELYSPVQKANPGTLA
GATCNNNKLSLSNAI
EELPEISKTKEAATT
QDPACSDKAPGMEGT
QTLAGTQKFSIRPSP
GKGSTFAKASFVASG
IIPHISTKTIDSWMS
ITSEERTKHDRQFDN
CIDDALRKNDFKPLK
NRFEDCQKEEETKQQ
HESTSVAKDKSSTAS
AAHTQSFKQPSCTHI
IQNVHASKRILFSIV
ETTTAAPKMTKETAT
QHTLNYWKEQSLNVS
YKADTVAKVQGVEFS
DFEEQFLKEKRFHDQ
IVALIAQKGNFSKTS
LKIEAENKYDAIKFK
QNEYNAVKEREFHNQ
IETLRLAKNYIWALS
GAWEQQIKQLEAELS
WAEMCSIKAVTAIEK
PVDNHRYKWNGRWWE
ILLNLLEKIPGKNAI
FAQVFLSKFTMVKNK
FLAGCQAKVEQAVET

RIEEALQKAAEKGTH
LRDGQTYKFRVLAVN
TLSWFPPKDDGGSKI
DATRLTDKLLELCNK
ENQTEINKPKAKRCL
NQTVTEFKQLISKAI
PEAKTPAKEEARSPA
SSSIFCLKMHKEMIF
YRCWRAHKEYLAILK
KKEEKEPKKEIKKLP
HTSVTLLKCTCTISM
QLWHHWCKKDKELYH
VSSPVIGKLNISETA
DYHKTKAKINIPADM
GHTFLLEKPPVPPKP
QVTMPGIKVGGSGVN
GVSVADQKGKSTVAS
FLAQADCKLLECRNV
ALHPNAFKRLGASLA
SSCIDMGKELLARSH
NASKEERKRWQATLD
DSGTYTCKVSNVAGG
TQILEELKKTKMDNL
TVEARLIKVEKPLYG
TSTTQHNKVSTSPKI
MSLDFYTKVLKSCRS
PSPVGALKGSDVILQ
DAMTHLIKISRIIRT
GENMTSGKLTVAGGA
ENRFGVSKPLESAPI
GTRPGDVKEEAPQEM
ARPAFNLKKPSKYCN
HIFFDAAKNLVWKER
VRYENNLKPVLKHVK
KYEGAKVKVLEERQR
FISDGVRKSLKVVPE
GIPPPTLKWEKDGQP
QELEQENKLFKDDME
QDNLGLQKQLDLQRE
SEAGVEVKKEVGVSI
AAAEENSKMEQSNLE
LAMAVIRKKHGMXXX
ELLETLQKFKAVDKE
WYQETHEKQQDLNIM
RSVLESLKRYRFGKD
LSSHHSSKRDWSKSD
VQVITHVKEKLHFMD
ECPDAKLKGPKFKMP
DVETQSGKTVIRLPS
IDYNPFAKGFRDDGL
LNELTQLKQLVDAQK
QLSDLNYKVEGEKLK
WQEKFFQKEQALSTL
VVSITCLKHKSGGHA
SASLAHMKAKFRETL
EEKQTSWKEIDNDFT

TDEDTDFKEFLKKLA
SCTLLSEKDSESSSP
EKASQLEKTGNDASK
FSFPASPKVFGGRLY
CTKPITCKDELAPPT
KVFERASKDTFQLEA
MSPDQADKLPQLQGE
EVLDGVEKMKSMSQL
LLEFHYYKCLVLGLV
DLSAAPAKQFNLEII
GFVLRGAKAETPIEE
ESFFSNCKSLPEAPL
QSLDRETKERFVLMI
LRYRQAAKPPDLNPE
HLNEVSLKLPLSDVA
LRIVVPLKDTRVKEQ
SAATQALKRNGAQIA
DSASIMVKAINIAGE
NTFSFAEKNFEVNYQ
ASLENCMKLSQMAVQ
IQLPVVSKQHCKIEI
RQSWSVCKQVFPGLN
MDWGAQQKKANRSSE
SGILDREKEERVSLK
LSLQSEHKILHDQHC
PYIAHSQKMQDLFSP
LTSEETAKMMVKIEA
RPEEEEPKVEPKKLE
DMMATLAKSQVTTVK
EGKEEWEKGKDKEVR
AAVSEDGKSDDELLS
QKELAEQKSVETAKR
AYEDLSQKYKAAQEK
NEEIDISKAAQTTIE
KGDKGRYKIVLQNKH
PFLVMIIKNLKKYFT
SHVILGDKVTKNSSG
TQEWDNSKSILGVQC
TPHLRILKVGKGDSG
HLLQYFLKFVPAGYE
TSNGLVTKALEHAFQ
RADQRKAKLGDSHDL
RGGSGFLKSGGEMLK
PDKEKPIKEKDKGQW
LLFGQRAKTIKNTVC
GEELVSLKEKSKSPK
EDDCLAFKVHQYFNV
FHLEFGEKSNAKIKD
YFTTQSWKTAQQHLR
EILQDLQKRLESSEA
AETEIVCKVVESSSI
LDLKINLKASYIIVP
VQAASSSKERGGVSL
RAEVHLRKSAFDMFN
FLCIFLEKMGQKIAP
PKKTRNLKKITREQR

```
IQIVYLHKFLMSLLN
GLDTPCFKTSVNDSQ
KREVELEKNTKEEED
SDVVWETKPKKKARW
DLSSQLLKVLGTRKL
GQSQDNEKELAALFQ
LILKEEMKARSSSYA
VANLEEGKSYFFRVF
VFFQSCAKAVMWQKM
LSSDDYTKKKKKGKK
TARFLYGKKVEGTAF
DGAKKARKPREASGT
FTVQLSNKLIMDIRR
EKEQLRSKLEEMYEE
QDLCSIGKEHVFSLE
VEEGLTYKFHAAWSS
LSSHDSIKMEIVDHA
GFIDDAVKKLNELSF
SSQDVESKRSDKTDF
QDLEDLIKAKEKEVE
DPMAGLLKFADDLGM
PPGQLQVKAQPQARM
EYDVTMPKVESEIKV
FTEPAIAKNPYDPPG
GHRKRASKSPRKTAS
ESTIGHLKDVVGRIA
QAEFADLKKLQDLTL
LPYQNMDKTKTDYTR
QKEIYQVKQQRLELA
KNDFKPLKTLLQIDI
EFDVNLSKANVDISA
PYMVKMLKXXXXXXX
RLYRKANKSSKLVSA
STGTLVSKREVELEK
TRVVHAAKAALAAFN
ADEMDRMKKEYTDCV
YYWGMESKIPVEDNK
DSFTSAPKQTPDSGK
FHKSVEIKPSDKHRL
ETQIIPGKGQEKPLQ
RENFDKEKGKTPKYN
TQQCVQCKESVGSGL
KVASLLGKKGASATK
IFENLANKADRDQYE
LQDQLLQKEQENAKL
CSALLSVKEPATITE
EGEEEGDKNGDGGCG
GRSEERDKEELEDLK
EEGLRSQKSDERQGV
VKKLEASKIVKAGDS
ERYLENYKCYRKMEE
EEKAKEDKGKQKLRQ
XXXMDDDKPFQPKNI
SLPGSITKAGDFLEA
ENLAGPGKPSKSTEP
SSHRSNFKDPLQVAN
```

HLCRIIEKKHVSLNK
GDRVNYIKRSLQSLD
DEVPPTIKLRLSVRG
AALDNLRKVVPCYSK
LSGVPLYKGDVPTQD
IIPTFQDKSLSFPQH
IGLHLVEKEIDIEKQ
GTDYQLSKEYTLDVY
KRRPQTPKEEAQALE
NDSQEIFKLAKDVLI
EEPWLEIKHLHESLL
VNTATCDKNPSASKN
TVVERLVKVTLYGSQ
TQDTTPFKITTLKTT
VVMAFSMKTMQNAKN
VTKNVRHKLTSRNER
AIITQGCKSLSSEMY
AFAICDQKSVHGTFS
VSTSQAGKLPTRITV
LKAMIMCKGCGAFCH
ILSLNQRKEDLLVDL
VVYVAPAKSLVGQVA
VSISSAVKSAPEENS
LPYFIRNKLCKVIVD
DELALFDKSINEFWN
YKDGKRLKNYNMPWG
LDDISQIKAQVKELS
IFDDWARKVQQRNLG
KAARGEEKDKETKNA
AESDTDVKLSIFCEQ
DRTRVPNKFRYPFYY
GKYTCQIKNDAGMQE
PVSASTPKIGAISSL
ERPDVDEKRSDLLKL
QIPKRNGKENGTENL
QKPVLDLKLSGVLTV
FVCSFLQKNRKHPSS
VKSMKEPKTEAXXXX
MELLEAEKERNNKLS
LEADIRGKPIPDVVW
YGEALGIKYPVQVPY
PTMSIDLKYLGVQLP
DFIQHIMKSSPLMFV
ARQSFFLKTYLDHEI
GPIIVEEKLEAPDID
IVTRTTIKDQEDLKW
THTELVIKEADRGDS
QAAFRGMKTRRHLKS
VTMQSVVKPFSIFGQ
SESSAPAKTEGRSLE
NSYAGQLKSARFEEA
PKFGFGAKSPKADIK
LDATPVTKGVASAVV
EDSDSPTKSSSTEPS
GYSISHVKRVLDAEP
RSPGGPTKLIEIISD

RQSNRATKEIQLHQM
IALLAARKNKILEMK
KSQEVEDKTKEYELL
SSNQEEPKAQGKVEE
YYRAYVSKKEFLSLK
WQAFLDLKKRIDDFS
ELWIRDTKYVQQPCV
QEMEAFGKKAANRSW
LVGIYGRKGLPPSHT
ETSQRISKLQEDTSA
LRKKMNLKPIMRMNG
KGPELDVKGPKLDAD
SKAMSVAKQVFNSLT
KKLINFMKMLSQKLK
KKLTRCLKPFENSRL
PDRDERLKKEKQERE
LTQLSLLKDTLSAYI
TETTEPVKASEAPSP
VFETLIPKPMMQHYI
SDEAEDMKEAMNRMI
ISNVSAEKATLTWTP
PGPLPPGKYMAGLNL
PYLEFRYKRRVYAQN
KLLESQIKQLEHGWE
LEKEKAAKLEILQQQ
EAQVRERKLQEEMAL
ISQTPLLKESGFRVD
ARPGLPEKEENEKSE
KKEAPPAKVPEVQKG
TRWVPVNKSAIPERR
EKYLTILKAVKVLQA
LNVIVLDKPGPPTGP
GLWLPLTKNVVGIDL
KEALKVQKNQKEILL
ESDSVVAKNAFVTPG
EASDVETKLREFEDL
LEVVDVTKSTVTLAW
STDIVPAKKWDNAAI
QKTTLVVKDCIRTDG
DASVPSMKVHAPGLN
KMVVAELKSELAQTK
LLRDLEEKVDKTELL
KVQINVVKLIAALKN
EDITTIFKTLHGNIW
VEANQAAKEDTVVLK
AELCGVLKAVLQRKT
QEQLNELKTEIEALK
ASDVSTAKSPFSAVG
GANTGIGKETAKELA
LESMMYGKKVLSQEP
HLEITIPKDLKLKEA
RVGAGPMKRENGLHL
VVFKVSHKPSGLVMA
IEDEQREKSVSHQTV
EEFCVFYKMMSLRRD
NILEGLEKLKHLDFL

YREDSNDKISEVSSD
LTNADEAKAQQELMK
PSEPTITKEDKTRAM
MSDQECVKRNREEEI
ELEELRDKNESLTMR
ELQLLNDKVETEQAE
HPGYALPKAQTTVTC
NLGSGTYKLTTMKKV
VKSPEAVKSPKRVKS
KLKMEIKKSRRHPLG
KAKLAVEKGRKARKL
LHQSRLTKLQSVQAL
FVVPDAPKAPEVTTV
SQDSEWSKREQIPLQ
SSSFPLPKQYLDVSS
CLLAILLKTTNRASD
LFKKLNFKRLGAGML
FGPAGTGKTESVKAL
ASTTLDIKDSISAVS
REAERAAKASSSAHE
VPLTQVAKEALREAA
LKKMDSIKGSFIDLE
ACPFLREKRGKNITT
LEASKVAKQGESIQL
WFLAAGRKWSCEAIE
GKAMHTPKPAVGEEK
QEMMSQIKYKKNALE
ELCRRLYKLHFQLLL
KASMASLKKTISQIK
AFDADMLKVFRNAEK
AVESIREKLKAKPLL
EIQQLLDKHKEYFQG
ADLTNPTKVTHLYNH
NQQQNHPKYSGQKNT
VHGSGPSKDSFGELS
MSVEEELKKDHAEMQ
CIPNEYRKGSTVKSS
LYGPIAFKDFILSKG
ECSLSLNKKSRSSTP
IEVVFVTKFLYSILA
FDAAMNEKMDLLNDA
YPNFEETKSLPWILL
LTAQASLKTKPIHTI
CVDCYRMKRKNCQQG
PIFVQWLKEKEEVIR
IITEAFPKPKDIPQA
IINITMTKKLDVEEP
KKFEALIKKDLQAKE
CSIQQQRKIAQYLKG
EWEESEVKLKKSKIK
WTLAAPFKEQHHHRG
AAASGGAKELKNIHV
RMASLCLKPPKTRRV
RCNKLLQKVQKNKEL
DKMSFVEKVNKMTAK
RYYHTLIKKDLNMVV

RPRSELLKAPSIVAD
ELVVSGNKRLKELSY
EQDSLNAKLEGVSDF
VTVENASKPDFTKNS
EEPGTEGKQLGQAVN
QTFYLLQKITKDNNA
IKDTEAQKVSTQKTS
DLLAYIHKQTEEGGN
VTRFEISKLTEHQEY
NGHVNFEKFLELAKQ
LTEKRMDKVGKYPKE
KTKNAVEKFTDAKRS
RPDLIDMKSVKHRSN
VRMLYYLKAEILGEA
VGTALPGKFKIVLDS
PAGPPPRKTSDAVVQ
EWSRCTEKMIKVRQY
DAEANQLKEELEKVT
EVYLVSGKNVGYAKY
LVSDVDYKNYLHQWT
KEKFEKTKGQPYAST
LSSTAQLKVLEADPY
KSQAIATKKIISDYH
TMEQALNKSNKLNMD
EIISDVLKIPVQLVF
NHLETCLKTQQIKSH
LIQELQHKLNQKKEE
QKNKPISKKEMIQRS
EPFSAMSKSTVSTSH
SDMMENSKEEGTSSS
VAHPLNEKMWSAVFG
NMVEELEKALGMKLP
LKQMELIKDQYQKKN
DSDSEESKKPRYRHR
YTCDCSQKEGVYRVA
PISVVWLKDRGEVIR
GLVYRGVKVMPFSTA
SSYFHREKSNIFSPQ
IEQFSTVKGVELEPT
QELIALAKNRFSLKD
ELDLWHSKLNELDSE
REQRRHMKRWEYFIA
DTRTHFLKTKAAVIS
PNSELCCKPLCLMLA
KPEKSLAKVEKELTP
PKIMSPEKSHRVISP
GKRHYEEKQKEHLDT
VFCAETEKELSKQWK
YELTEEEKFYRNALT
RRQLLLQKLLQRRRK
DLLLETRKPDQVFQS
ALHRLIAKSRELYEK
WRSSETEKHGQADSE
ELIAKMGKSXXXXXX
EKEVRAQKLADDLST
SPTEGKRKIIIXXXX

EYNTVVQKPRQILCQ
CQITGTPKIRVSWYL
MRYSAEWKSTVSRLG
PVAEVAPKAQEGEST
ATLMTQCKEGAQKEG
KKTNEILKGIKLLKL
FCDQQLRKEEENASV
AILTALLKGTNASAF
ARIHIPFKGRPTPEI
QTERGRNKDKAPEEL
VCNPHSGKEYFRVPD
EAMTITWKPPLYDGG
TSTGHKNKEFVARIK
HSEAPFAKSDNQPST
LNEDREVKNAEILRM
LHAVWEHKLGSQVSL
EVLHYSQKAQEKLLV
PLNSAAHKESPPTVD
LWLFEGEKIVLQVPA
PYHSHCMKVTAQNSD
VMDWINDKEAIVTSE
AFSTLIDKGTHVAFL
DSDINANKEELLRVL
HSLNLLLKSIGATLT
CDKRLWTKIDLSRCK
LLISGGRKGFTYVFD
QPLHRDGKLSLDQAL
DFDDLERKYWKNLTF
FGDTEEAKKQINDYV
TKEVEMMKAAYRKEL
VSSSMEFKYDFNSSM
CPPSIDLKEFMEVEE
NFSSVLYKEDVSPGT
APTIEWYKDDKELQT
RNLPSDYKYAQDRAS
PGQDNLVKQLRDAFD
ELEGQIAKLEEQWLS
TPDLLRAKRGQKLQS
ITNYIVQKRDTTTTV
STSDRAEKFCGQKKR
NALDFVTKSYNETKI
EHLEHEDKQNRLRAL
QASLHVWKIVVSNTP
QGEQLIEKSEPLDAA
FVQAFVNKDPLASTK
RSEVSQVKREKENIQ
KPLAASPKPAGLKES
RQRNRKNKMSLALKN
STVEVERKQVHILQQ
AVDPTLLKTFEALKF
MEKEHLNKVQTANEV
QHAELRDKTTMLGKA
KGEVETLKAKIEGMT
PASGADSKDWLAGPH
DYIPYVFKLLKENLC
ARQLLREKHKASIVI

QALELPQKAEALKLF
VLTEDALKSLENQTY
HTILPGSKTREPSNA
ESPSPAPKPRKVELS
PDSTVEFKAILKGTP
ENHTPEMKHFVGLEK
NYFAIRNKKGTELLL
SAVGEHEKVINNQKE
EEEYEPEKMEAEDYV
LGREVGAKFWELLLV
YKRAMEIKEAETSLL
QPGPDYVKQRFQEGV
EGLQPAAKAGTEDEL
SLSQGRIKTFSQSEK
VPEPILRKIETFLNH
YSDDEYIKKIQKRLE
RETQGIEKLVLINKS
RPEGLREKNSVIRSL
LFLPFSYKNKYGMVA
ETTSASLKKRHADSL
ELQQLDSKFLEQVHQ
QVPTPVVKQTKLKQR
CFESSETKKSVEQKL
QLQLRHLKAIGKYEN
FADIPTGKTSPYILR
SRSFREEKDRSNKAE
LQINELQKEIEILRQ
YSPSPPPKRRTASPP
KVVSVTRKPEEDGAR
CLSAFLDKYELQIPS
LLSIKWNKLEDVAQL
VTDVIEMKDRLVSPD
VQSLEGEKLSPKSDI
DCIEAEEKVHVIEKK
CVLRLGNKEKTSTAK
LLSLLQAKVASLTLH
KAALKTQKTSERKYE
VREEILQKEITKLLE
RLFIISTKAGSLGIN
ESGPEQSKQKKVAPR
EGFFRSNKVFVESSE
LCLAVPDKTVRWCAV
VSELATVKRPQERSY
GVKQEVIKESKVEEP
NPNSSGRKKLTDSPG
TPQILHAKKSYDLQS
NVTFSFVKYRDTDTN
GFVYIAGKCGLVPVL
RASLKRLKPHVPLGR
EKKLLMFKRVQEAED
DSCVVAWKPPASDGG
FCIQEIEKSGGSKEA
FTSLEVRKVAETKEK
FTGDEDIKIPENPLE
EELAQFLKVEQKMNN
KKEAGNKKSTPVSIL

ITCPADPKKTLGIKL
DVDDLIFKLAYYEIR
ATYLLDGKVVNVAGG
IDHPVHLKFNVSLGK
GSMNVSVKLKTCTLD
KKLTKMDKVLGSSIH
SWMPTLEKICEEVIV
PEHLRTTKVNKQISD
ENAAGVGKPSHPSEP
MLEQTLQKVTELENR
SLEMPLRKQEELEST
KAEHVKQKGHYVGVP
XXXXXMPKSGFTKPI
KLEYSSFKTNQEKEF
IELFSTEKKDKNDSE
KMPQDARKILSNQEM
LLDIILAKNPSGLTQ
ERQVSHAKEKLTALT
RWRKLVSKTQLEMNL
NYIDEREKSISELSD
PIEECRYKLYQEMFA
QDRIPSGKQHQETSA
ASDRLTMKNDHISAH
LQMISACKGETGAMV
VPSDEKFKCLQDQNT
EDDDGSLKNVHMAVI
KELNLLQKLYGLYDT
TVSIMLTKYAAYNTF
SKAGVPSKSSGSAAL
STFDYSLKDRAPFTV
LFQQQGSKQKLHDCL
DEEDEEAKAWLQAHP
EVINRIQKLRKKCNL
IKVDVKFKDTVILKA
LPLMESGKILPGVRI
IRTALFEKVLDKIVH
AGDGKSIKAFKESAK
VKPQAKKKAGDQEEE
IMIDRLEKASSNQLV
ANDLAAQKTKLWSES
KAITCGEKEKQDLIK
PISITWLKDDQILDE
ATTTEEAKGRNPRAG
QIIQIEDKTTIINNT
ASQELDVKRREAIYN
AVDELNQKKEGFRST
EYEEDRGKGFFPQTI
SRSPRRRKSRSPSPR
RLEERARKLQETNLR
VMYNAESKGTSAQTW
IFNQLEVKNPDSKMM
ELLSLVYKGIAPGDE
WDSPMRVKLSIWKPY
GISDLLGKSGITEDA
SEEADSLKHQLDVVI
NDEFEIEKPKAYIPD

AGPTQGVKAYPDATM
RCFVLQPKEKIVISV
RISSSLGKKPSLTSE
VLQDLEDKIRALKEN
CCPDMITKFLGVIVK
GVAQANIKNGKHLLM
DGTSANSKMEGLESK
EASEGSIKLPKMKLP
QYIFEFDKNDRLSSV
EQSFYQLKVVASGGT
QENTLTVKNKEGEIH
VLSQVYSKVHNGSEI
RHSDGTEKNKLPSQP
IKEVEQSKATSQELS
LQTLLDTKFVHFIDA
PIGAAASKEYSLEKN
FKSVLLNKTIEESLI
RPGQQDGKEGSQPPP
NPVEARSKYDVPGPP
AYLGPTVKPDDKAKT
GTSQISQKIIPLLEL
VSHSPPPKQRSSPVT
FGGRNIKKQCCKMRK
TVTIVDSKRSDTGTY
EIIADGLKYRIQEFK
QTNYRAVKEQLKQWE
VSGPADQKTGTATVL
DDEPEEGKKRTGKQN
DVIEVTGKXXXXXXX
RGWTFHAKGLWLYLA
LILMIFSKLLKHSVI
DTGERPLKKSDKSNK
XMDVKERKPYRSLTR
GEVTEAGKAYVPPRV
QNIAAFLKENHDIDD
FANFSGGKITKDVGY
LFNTRYAKAIGISEP
KDTTFRDKKSSLMEE
EPGVPTTKTPAATKP
EFGRLHCKFVVQTSP
YSFVIDGKKTTCKLK
IISQFTEKIEKMQEL
VTGYPQPKVNWYLNG
ANVSEGEKKTEESSQ
AKRAELEKLKSQVTS
HARDVVAKLSEDRVS
CCELAFCKIINSYCV
DGEGRPAKSQEEGKR
GYEKKRSKLIGAYLP
SAVAPQRKMPLDDTK
YRQNQELKVRVELLH
KSPKADIKSPSLDVT
IDSVECNKLLAAKNI
RIKYIEFKKSTVILQ
KRLQILLKGYGEYPT
DAIVTWYKGPTELTE

LPYMIRSKLKLLLQG
XXXXXMKKEGSSGSF
YSYETIGKTTKTPED
PRTEISDKITSELVS
KKDKKDSKTDNKKSV
NDVQSELKYKAEHVK
VTHEQIRKETEKTFV
QGLNFTSKTAEGHPK
RHHHLVRKTDAVYDP
XXMRLLAKIICLMLW
SLLVAGLKKEVQALI
LKVLGTRKLMHVAYE
HDPIMGNKYAANRKS
VAGSDTTKSKVTIKD
RELRERMKPGFEVKP
LTDLQVIKNETDARW
IACNFESKYQCISKS
SELSDQYKQEKLILL
ERREVTGKWVRVNKT
KTSEDIIKSKNLRSI
TENAIMLKRFNRYPL
PIRFDEIKADSVILS
QENEENSKNISTMES
LPSEDLLKELQKQLE
TILDDPGKKVKKPAP
GESTTDWKQIRSIIM
TTVGLLSKDSGKKLG
QLQTTFTKALQDSNR
DDTYLQLKKDLEYLD
QEKPSRLKQRFLLRR
NYFHEILKTIEEEVK
PTDFEWLKQSRFYFK
VTDYRDMKSDSVAKL
SEFVRFSKTENKITL
SYRTYSEKEGPEKKK
VLYDYVNKYHWEHTG
VIRTTHFKVTTISAG
SLSPQRSKLRIMSPM
PHDIPESKDCVLTIS
KLAAQMQKLNQKRAE
VNCKVLDKPGPPAGP
EESLIYRKKYILSKD
KAHVTDPKADPRALL
VAHIIDPKAISKDHL
KHLASHIKYREEYEK
EFQREELKAKQVSLT
FPNGVERKDAGFYVV
QQISFTEKGSDEKKP
RVAAVNVKGMSEYSE
EKQAELDKVQAKFDA
KEYTIVVKVLDTPGP
RKLALTVKGKELSEE
VTEVTSEKCVLSWFP
PDLDLNLKGPSLKGD
DLGPERRKSTTELTI
GVEIISIKPSAGHLS

PTMRDDPKLVWFEHA
PYWLDEPKNLILAPG
LENAASLKLFAGSLA
ISWLREPKSKPQLID
REQQLVEKSGELLAL
DPVQEPNKSMFEKTK
SVSKSGRKIKGRGTI
SEHMDDSKQESTTSK
FEGNTNTKGHVKNFF
QGVLDTGKKLSDDNT
LFSKPASKEDLKQTS
WFPASHVKLLGPSSE
MHISSQIKKTTDLTT
HIMDQHWKFCVECEE
EEVPPPPKVPALPKK
RQLENSEKYEITYED
HEENISSKMKGLNGK
EEVQLTSKQKEMLQA
GKVLVREKRVDLIQD
SPGRCTPKSSEPAPA
FEEPPFEKEANIIVD
KEVWNREKLTLQKSL
GLALIQNKKEDVSSI
GILEHQYKESTTYIS
AQSQQLQKHFSSLEA
LQVANFLKDSYTALT
SLPPLPFKRGGGKDG
KAENDRLKVAPGPSS
SRATPPTKAVDPIDA
HDAHLNGKVIGTLKN
DKTSDGPKLFLTEED
RGTSLVVKGLKENVE
VNWYFENKLVPSDEK
EGELESLKENLPFTM
ETISTANKANAERLK
ISLPNIRKPDGYDYY
DSAHGFLKVPPRMDA
QAWSSLGKRADQRKA
NPDDGARKPIPRDKV
LEELSGAKISDLHRE
EEHMEQQKENMEHLK
SLLCNFTKSMEEDPQ
LTENQFLKQLQLRHL
IDYYIVEKRETSHLA
HKRRPKVKRSKKRRK
LSPAVRQKETVKNDL
GHQNLHQKIQYVVRL
SAVSQERKLQNSMLK
KRLEQQRKDYAAPIP
NDWFSNIKVNLKECF
TAFYDSLKLVRAEQI
RQPPDRFKFTSVTDS
KMRVLQEKKQKMDKL
LEAIGAPKPAYPYLE
QENLNSMKKRNVNFS
HPVEGHQKLDSSPDP

```
IENQLAIKSKALDEL
HYLQENLKSMEEIQG
RAPEHSNKSKDRDNE
VVLRQMSKHKTELKN
MDAQTIAKLEKAFSE
YSHGENHKLVSEHVQ
EDDVTDPKRPLSFGV
SNMPQSLKVGSFIIE
TKGGWQQKSKGPKKT
YTITLENKLGSATAS
VKKENRGKGGAVSKL
YRTYRAAKLQRENYI
YDYESYEKTTRTSDV
NLRKQYEKSSRFMKV
KPGEMEKKMKILRES
GPFTVVVKESCDGMG
FEELVEMKSNYYSFI
VATSRSEKEQLQALN
LSFGVPRKAILYPHE
EWLRNIEKVEVEAVN
AIAAQVWKLSTCDLC
DIHGPDAKLKGPKLK
ELHRTLSKVNAEMVQ
KGDSPELKPRAVHKQ
NFSAAETKSDKKDCA
VPMTEETKSITLFPD
NTARLEIKSTDFATS
DFQAANVKSSAHLRV
YDEITDLKQLTVVME
LNTLSEAKVETICHR
FRTWQTRKQFLLYRK
NRLAALPKGRPDRPR
EIVATASKVDIAPHV
ALTILRRKDESETEW
LPLHDSLKLLQEQLT
TNRGHTVKAVCESFH
VVLPPIMKQPPMFSP
CGAQAGIKISSVHKR
DPPPFMFKPVKEEDD
IIHDTQFKAQNLEEG
PTLLSLFKHECSRVI
LKALRVRKKTGGEKM
QQKELLNKRNMEVAM
IFPSNYVKPKDQESF
EASKEIAKLTCVVES
PTVTWLRKGQVLSTS
RVAELQRKPTRESEY
CTAQLNIKERLIPPS
DSSLLKTKRKHKKKH
VSRAPITKVGLKGVE
LTRSPRVKIDTRADS
PFEVTWLKNNREIRS
XXXXXMSKLSFRARA
LAMETIDKAKEVGAT
LSMSTHFKMLKKLVE
EIYRWPIKESLKAMI
```

ECSSSQLKASVCKPF
EGILMHIKPENHCSE
HPDNQQLKNEQEESI
LPARQLWKWFGKPTQ
INFLENLKILEISED
TVLLSMSKGGKRKDS
PPATPPPKTRHSPTP
QMYEFRVKAVNEIGE
LAIESCLKGLLQAYC
QPEDDDVKGQWDRLV
QYRLLSLKYHPDKGG
PVRPSVWKKHGVVNT
TPTYITAKESQQMQS
QGLASLRKSANKLKP
WAGYLITKVRHDLAS
KRMAQDLKDIIEHLN
SLSLSNVKGHGRADE
PTPGITRKEEAPENV
KSGRSKEKRKSKHKL
KAREFGKKVQEEXXX
VIQCLRSKEVFGFYN
AKSPAEAKSPEKEEA
EVTIMEEKERAYTLE
NAKGMVEKWLQQVEQ
LQEQIVQKNQEIKNM
KKEEVEQKKNEYNFK
KVKSEVNKLYKLLEI
KLQRKKKKRKGSDEE
SEEDLSGKHDTVTVS
DITMWLSKRLPQFLQ
RKSKHKLKKRRGHLS
DHLRNGSKPKRKNYA
QCDCWLSKSEGDGQT
ETYPRPTKEDVYVVG
SIRENGAKSQPDPFC
QVQITQEKMRKEAEK
IGVLAQSKFQAVRKK
ELHTEPMKLALSIEA
RRAPSPVKPASLERV
QVIDMSVKSLTLKVS
PDRTVAEKRGRKITL
KKEPDDAKAKEPSKP
KASLFINKEDVYEYY
IDIHYGVKSNSLAFI
MAGLGVEKLEKLMPE
LSCSQYHKMYRTVKA
PGFQGRLKYNAFCEL
LEVLSGQKLPCEQGR
NWLSNAIKEFGFSGL
SAMQSLIKYSGSFAR
QMFLNDYKEVPFDAL
RKESITEKKPKKGLV
EKQAQLEKELSAALQ
HALAIAEKSQEFLEA
SNDSTHFKALHSNQA
LFEVALPKTAHHFVI

SPREESNKRQRTYXX
LLPTPEEKKPPPKRI
KLKEQLTKVKECYSP
DMNIRGPKVDVNAPD
HDKTAVEKDEITLKC
KGYVVEVKEAAADEW
EEEPAQGKGDSIEEI
SLHGTVMKINQERAE
EPWTTDAKIQNKNCP
PGIDATTKLNVGAPD
RARGHAVKNSSHVDR
QVFFNMVKETTSNCF
SEECTSRKPGQTVNN
MFCLEQEKRLPFSKR
GSEQDGWKAEPGEGT
RCICGLYKDEGLMIQ
SLQFEEDKENSSKEI
QSILEQDKLTPESTR
VFVGRAQKNQGDRSN
LTNVSRAKAEDALSE
FRVRAVNKAGESEPS
PPVTSPLKPNTPPSA
KELDLCEKVSLQYTG
VTPPAGLKATSYTIT
TETQNALKEEEERRK
SLEEIMDKAYDKFDV
LYCSMNSKHLNGTIT
CNNVSSDKIGDKGLS
PKTAKPKKAAPKKKX
NIAIEFNKPQYFSIM
LINGAEEKILEDFRK
ELEALNGKHQQISGR
SASPAVTKDQYSATS
STETIICKDEYEAPT
PDVNVNLKGPKIKGD
MHIDTVLKEIAKTVL
SFQAANAKSAANLKV
DLASLPEKAAANEEG
FTLPLYNKTAYVGEN
AHREIEDKPVATSNI
AYFVPMSKIADFLKA
SQYTYALKEKAKYQC
ARLDQMRKFRRQHEQ
DLEYLDLKMTGRDLL
HMVALLDKDQAVAVQ
KEIDMLLKEYLLSGD
QAMGHRHKAGGAGRS
SMLEERIKRSAKRPS
KKSLLEEKERFQREV
ASSMTEMKFASMSAQ
ETSKVTLKEQQHQLE
DRLITLSKKKAGHDQ
CYYHVISKVPRIRTP
PQKTAFLKDRLNAIQ
KPSVGEEKDIIIFVG
ESLMQSLKMNIFRLQ

STAELLVKAETAPPN
EKLLRESKELADLAR
DTVVIASKPYAFDRV
LPTKASSKAELEIAN
LYICDYHKNLIQSVR
VEHPQRSKKAVWHKL
EACIELEKIVGELKK
TRHLDEMKSAMRKGH
FLLNEIEKEEKEKLW
CEVTGTIKAVDRLTA
STQTESNKKMEENEK
LPESGATKVEDMLTL
ALCSWTAKKDNHLNF
SGLDRMAKHLSQRKT
RLQICDIKPRDQGEY
NNCDYEQKERFLLFC
KQEIMTLKEEQNKMQ
KIQFLKMKIAEKQRQ
TVAQYPFKVPGPPGT
ERSDTNEKRSVDSSA
KNLWVSQKRQQASFP
NRDSQSEKDDGSDRD
SSRESVSKAGMPVSA
EDLSEELKKENPFLR
LAVEIRQKSFGPKHP
RQENMVLKLANQKNK
KSEAGGLKQEIMTLK
EATENTLKEASVESE
EGLKDDEKLEVSELA
THRDSGFKNVGDITN
NCERELHKLEGNLKL
LSNLEEAKKKKEDAL
NAKCIHTKFQVFKHQ
KEEAQLQKCLGQVEH
EITVSWYKGDTKLRP
TPEIVLAKNNALTMS
ANVEGDLKGPALDIK
LLLPVLIKEQLEPPE
NSAGLLIKAYNRAGS
DLVVDVGKPLTMVVP
LIEPSVFKEAATPEE
GNPFQANKHGERPVD
AELRIELKKANNKVC
EIIHDRIKREYGLET
HVYYSRSKTTHIIAT
DSEVDNQKRGKRHYE
DSGNYILKAKNVAGE
GTRNYYRKIGYRLQG
SLLESVQKLDLEATI
NENSNRQKLEHVKVN
VGASQECKKEADEQL
LQEERTCKVCMDKEV
PSLDLSCKQLASGLL
MEVHEGDKYRMHSDR
RQYCFYQKLQWATKI
YVEPHELKEVSQTGD

IQNERLYKEDYHKTK
TVHDFPGKKLVLSSE
GPPEGPLKVTGVTAE
AELQEMWKRLGHELE
YFHISPIKLHLSVSL
DAELRNTKWEMAAQL
PQVSGELKGPGVDVN
RRTTQSTKHGEVDAW
SDNSEYSKSEKGKNK
PKVEGEMKVPDVEIK
ENGVGEPKSSTVSVK
VRQDSPTKRPALCIQ
STTKRSPKPPNKKKT
DFLAEIFKKIGSKEN
SVAWFHEKTKIVSGA
LDYELCQKHEMTISA
TPHSHREKPGIFYQQ
ENLVLQAKIHMPEKT
VVERRDIKRKTWVLA
TTASAHLKAGCPMLA
NEARINAKAKRRLRK
CCLTFLVKVKSKALL
LRQCMLPKSAALSEK
AYEEQKKKVADHPNR
MLRDLQEKINQQENS
QVFNQIVKSRPHDNI
PQCVAIDKLKKCKSS
SDDYDRIKCHVDEQS
CVGSIALKAPPRFVK
QFMGQAFKSIHEKVS
LPTRALQKIQELEGQ
GHLNPSEKLQKLNEV
LEQARQEKDKLPLEK
PAYDGGSKVVGYIIE
ARHSGHHKKRRKEKK
FDTDTGLKQALETVN
IPQIILAKANAINMS
FEDSGNPKSPDKAQA
SNINVEAKFNTVVSQ
QQQIRPIKSVIPSPT
QLKASLGKVGVKEEL
AVYAVLDKVKTKKST
IQVASNAKVAASRKP
SIEHQEIKHTEFLKE
LYCLSDEKLIQKISX
IPATIQDKLIIRFTV
PQQGTQMKTKEFTSP
LKEEYSRKKETISTA
LSVSESQKGEERDRK
SVLRECDKLRHDGYR
WEESQTQKIQNLEKL
KSTRSSVKLQGTSKI
GTDCLEAKEKVHVIG
VFQREQTKQVEELSK
SSEPWEYKSSGNGRA
PCGRWLAKNEDDGSI

AGAKSRPKKREGVHL
SEPSEFYKAADPIDP
GNQTTALKHAKDVKD
RGPGRPRKRTHAPSA
VTDEPLLKMLLTKAD
GMDNQGHKNLKTSQE
LPSPCKSKSLGDLTS
KPLDLRPKKTRAMRR
STFVDQCKTVLKLAS
FRICAINKAGVGEHA
NEIHTNHKTLYDLKT
NLQDDLEKQERSLQK
KEMTNHQKKRAAEMM
IRTEQESKGPMTRRL
KVEAPPAKVSKKIPE
EKARDREKAQELLKF
SGGLIPGKSLVFATM
MVVLVQEKENDYYCS
LRSAPAYKMQFKNNV
KKRLERIKSSRKXXX
IVCPFFIKTGMFEGC
SNCNGEEKKLDDSCS
WLQETDHKSIEDQLQ
CLHVSAQKMLDGASF
RKYDKRRKKRAPKAD
EIATSQLKSDLTSRD
LVQWVVQKVAWFLQI
TFAFNQMKPKSIPYS
AFQDHPEKSEQFEEL
CASSLPIKKIAGSNY
ETRRYALKNVTILVT
LDENIFPKQISLRQA
RKVEEIVKDHLMYLD
LGPEEVRKSALTLVM
IISEMESKLVEQQLA
PPLVPAIKEDAGGMI
DGYEFFAKRQLVTLF
PALQMAFKLGLSYTP
MHSNSQDKEVTILAE
IDGLSGEKIRWTQQS
GIIVRWCKVDDDFTA
ENLEFHIKVSYFEIY
EKQIEQQKKIQMSNL
KRRGPKKKKMTKARL
MYLNILRKKITVNSS
SDTVHQMKPYLKKSV
QKPTKAPKKPTSTKK
KDKPLIYKGEGKWLR
QDLRATAKVPSPPDE
FDSRGDVKVIMATNR
LLGQAEDKAEVDMDT
PELVQAVKNTAMFSK
TSTLRHCKYSVTKLI
LDERAFEKTLTPIIQ
DFLDFMTKINGLEVQ
LRWVRAIKTPVSDLR

PEMTTTAKDKTTERD
AASPLEPKELEELRD
HTISVRVKAAPYWLD
VENPAGSKTVSVKVL
EKKAQLSKYKSLQQT
ILDITDVKVEDSGSY
EKSPLPQKRITNIIE
QAKIQEAKEKVQINV
KLKVEAVKIKKTLKN
KSALGFIKVAVTVMD
IKVKSEEKIKVVEKS
RTWAEAAKTVLEKYP
TVLHGNRKDNELIRV
VKQAEKAKDSVYKRS
TNGEHTSKATLELSP
DFACANNKICDELEE
LHSNPEPKTFTFDHV
TVARTMMKVMKLTTG
RGPKVDIKAPDVEGQ
ISLGEACKKVFLDCC
LGYHVERKERNGILW
PGPCPSVKVKEVSRD
KVTNQVSKQKYKEDY
FTCPHCAKVCKTAAA
SRIDALEKAGQKLID
NVDISAPKIEGEMQV
NALNVSNKLYREGWD
ALIRTLGKQLNVNRR
VTLVEAIKQVVKHLP
ISVEQDGKVKTLMEF
HQNTMEDKFIVCLNK
NLNSLCRKYHSAELE
KAAALACKKLKSLGL
PSVSTHIKVKSEEKI
DEEEVSDKGSDSEEE
VGPDGLMKLNGSALN
SSEHDVLKEAVLKKL
STYHIPTKAEASTSY
VTMDSAPKPFTDVSI
LVSDNNLKDYFERGR
GLRPILTKLYKEKGV
LQLWQRYKDYSKQCA
LGIENQHKISVPRAP
DCQRIELKQGYTFVI
TDTVSDVKYKEDLTW
EPPSFVTKPGSKDVL
EQWLDDCKRTFGAKE
QEVIGKLKGLCPSVA
ENGSLEIKMIRKEDQ
MEEIENLKKQHDLLK
LQYMTMAKLQKEFGP
FRRLQFNKHGVLRVE
NAFLEFSKEVDAQSS
RDHLAEAKEKLSILE
QLPLSRIKFTGKECE
WRESSLLKNLWVSQK

DLKTQAGKDDKGKRK
KNNGQFVKASASLKG
LQGPSDTKDLHSTKW
KKEPVTRKAGLLFGM
DSAEALLKKHEALMS
QTYGSTLKVTXXXXX
SFDFLAEKLQFYQRQ
SAPAPQCKALCRKPE
ATGLDEEKRARCDGK
KPENLSSKTRGSIPT
SCASNSVKFGVEMLG
AGPPTPCKPSSPGAD
DSRVSSAKLNNFSRL
ISPTNPTKLQITENG
QLLLEEIKSSKDQLK
LKNVSEEKMAAIVNQ
QKNSTLPKAIVSSRD
KQLEKMRKFLDEQAI
LFNWYNRKNCILADE
NDDISGYKENYTRWL
ITNYIIEKKEVGKDV
FKIFRETKENEIQDL
QRVRENQKNFSLLQY
SRLSTIWKGFINMQS
DPLVTVMKVEKAPQE
PELPLQTKDAADGEA
KRCKLLGKSSRKRQD
VSFSAGVKVRFLLGG
LENMQKLKLIRMCPY
LKKRKGRKPRWTKVV
VETICHRKEGGSSLI
KKETTVKKAVVVPAR
VFEYFNAKSKHSWRK
PSVENRPKIAAFLPA
CLYQTNSKLYESLTY
QKLVYMGKLKEKESF
AALRPLVKPKIVKKR
ALRAEVSKLEQQCQK
KQKLYHKKIFRTAML
IISDSQIKASEFLGY
FDTSAAYKSILRKKV
DELTTEIKELKETLE
EILMSLQKVLSGLGG
KHGFSVLKGQAELQM
NRVNHIVKKIETENE
TYVAQFLKHYPDIHN
YGVNGVGKILKKINK
IFRLQTEKDLNPQKT
SESELESKESQKDEK
SPSPAPEKKEKTPEL
EAPLPSPKLEGELQA
ASPDEGWKRCNAAAQ
KLKDDTEKREQKDEE
DEIRTSLKNLNQFLP
YGAVTYRKSKRGPKQ
KVLASYQKTRSAVIV

NTMMELRKCCNHPYL
KFRLEEQKVRRQEQE
YGVFMLHKKTVNQKT
FAPDPENKGTEVCMV
XXXXXXMKKGSQQKI
VPTQPLLKAHPKFSK
RISDAENKHILELSN
KQVTQESKEKKVSEN
TFYSQREKPGIFYQQ
ENYQTLPKSTRHPSG
DCEEAFLKNPEGKPR
RVFCSYQKKGFDIKW
TLTETRFKTGTTLKY
KQGDELAKLSSSFKA
GFNTSQGKLLRTILF
PIETERSKCDITGLL
HPTEEALKISVASEP
EIVECILKRGQKATL
PEDKKEEKEKPKKEV
GSDQERTKKKRRGDR
PENHNNAKEPDSEQQ
SELSQLYKEAQAELE
VANVVGQKTASIEIV
LSITVRTKKQELSEA
EFEQIMLKAMQEFTL
QEAMLFFKKSEQEKK
FTLQYLNKLSMKPEP
PRKAKEPKKAKEHKE
APECMALKADVVAKK
TSTPLSGKESVFFAE
EIMEHRFKTYQQFRR
YLFIQTDKTIYTPGS
LEAVRQEKQKETAKC
SEFSSFDKNDSRGQE
SLLTFIDKAMHGELQ
FREGQILKVFGDKDA
KERISAAKVAEPPQR
IKASDITKSSCKLTW
EVQVLDDKNVRRFR
FHKLNIPKLDFSSQA
DAGLYTCKVSNDAGS
QTEETLTKVSATPGP
EDSEQELKGYACRHC
LNINPDRKEHNCREE
NFILLFYKVSTSSVV
LLDVNLAKFLAQDVP
FSTALYGKLLKLPTC
LDPESQRKRTVQNVL
KNSNGIPKLKYIPLK
SSAPSSPKLQHGSTE
KLEYIYAKQQALEAQ
DQLKLQEKDIQMWKV
RRVAQAQKALSDVAY
KESARDPKPEASRAS
DRMTEESKVEAELHA
FIRTRLEKEGIVFKS

YSSVSQRKYYSTKHF
KALLLEAKQKVPPVL
EANPLTWKQLLCDLV
IAEELSYKTAHEDIS
LQNTDAHKRAFHEIY
CKATLSEKLGGAEVA
KEETNIVKLLEKQYQ
CNSVHLCKQLVESGA
LITPEDRKYGTTNLH
LLDYQDHKGADSHAG
TECHARFKTPEDAQA
ATVNIQLKGTNEYVP
KNLDHVNKILKAKKL
ANQGCVRKVYLKDTA
LQQATPPKNYQPAAH
LCSQLLCKDPAERLG
QEEVIEVKVPAVHTK
KTKRISAKKRDVQQQ
CLTENVLKEIAKAGR
EDEETDQKCSLSSSE
LISLQEAKDICEEQV
LRNLVHSKHKMIAMG
IVCSSEIKRQVVETM
EPPRFVSKLNSLTVV
NCRCEISKSLCCRNL
STGGEHKKSDRKEEP
AETLNQSKERNRFSL
SEPSELIKVLGEKPE
VAGGAISKPLTDQTV
AFKNIASKIANELKL
PVLQEEIKESQDPQH
RSLVEQHKRERKELE
VPGPDAQKTETPSVS
QSISENEKLQKLKER
DLKGTLGKLLDERLV
GKTGNTEKFGRFRQC
KWIQLLEKIEEALKV
VTAEPEPKIPDVAPI
VVRGWLHKQDSSGMR
TEVELFHKDIEAFKE
LTKQIQEKQGQLSEL
VPTFTIPKLYQLQVP
MGLKGPLKTPIAAGH
VVYCEELKCWCRAIV
SSEEPTEKEPPGQLQ
HLRTRVQKVQVKEEP
GRIETDTKVLEILYE
LACQENCKKTLTLIE
IGTQTNLKCRRGIEN
NSKVKHKKKHYKPSS
EPIRNSVKERESERN
SVTRSTFKFTRLTEG
TEDLEDAKKELAIRL
PDSSPSRKPPPISKK
IPMAPSCKLLFEVHN
VIIENVSKSTSTSEP

GVNLIRSKVEDPAVI
PICRGIIKGTVRTFL
CTTHLYIKAPAKFVK
LVPGNIFKSSGLTDG
RAENLRFKEEGVRDV
EIEVSQLKEQSVLEF
IENIDFNKSGSSTAS
LEDEAARKAKQGKEN
LVAEFPEKEAQLSLV
EISLNHNKIEQIIAQ
QSLKDLEKEFVDFWE
PLGLIMLKTTSEELA
GTAPTTLKEPAPTTP
VSVNVSSKEFLQTES
LSMGTRRKVLVLGSG
LKEEIKKKKGIKEEV
EFWDVIQKHCLEAHQ
NVLLTCCKDNVCRLW
TTASNYNKSESQESL
AKRELESKLQRLQAQ
TIAEQESKLNEAHKS
LKQQAKEKQAQLAQT
QLSLLPNKDLFIWDL
ATQLEQMKEDYSYIM
HQPVSFYKLKLTNNT
VRQVEALKAENSHLR
WLQDVDDKYDRMKTC
ALGMHESKYLILPVH
GAVVEFVKELQDIEV
CEIPVEVKPKARLLR
RLGSPSGKTGTCRIS
QDLLHTTKHQDVLLS
YVHTNGPKKKKVTLH
LHEFDSLKDPEVNEF
LEAENQIKKCDIRNK
KQLTDNLKEKALSFM
KEIKKLPKDAKKSST
TKMDANDKKIQSLEK
LAHWSPAKLLLQMDS
LEELQDLKGVWSELS
LEEHMFHKLIEAGET
AVAVARQKDSGGSGR
EHLCGCWKQCPSCGA
SELHESWKYNMSFIN
GFYVVCAKNRFGIDQ
SHNVQQLKSRILMTA
KEAYEKSKGKHVGFR
KLRLMEQKFQQVDEW
IFHGVMGKDEREGNS
NMHSARFKNGKQIRK
DKVDGLQKQTESSSD
HLKQHCNKVLLRVRM
NSDQHYSKESNYFYS
TGFALNLKPLAPAGI
DGNHLYLKIKNAMPE
QGLPGPEKLPGSLRK

PDFERVNKRLCPTTS
NKLPLPSKSVDWTHF
KFTMVKNKALRKGFP
RQQSSSSKGDSPELK
KEFSDLEKKFKDKSQ
IGSLDVEKCKRATEI
ISLNDVSKAEGILLL
VNLHRIEKDVQRCDR
AQTADPIKVAEVPQP
FPLWESTKVVPLSKF
AGFEDEAKTSARDAA
CVSIEGYKTPYTDVN
AFEHFAQKGPTWKEP
LSPDSFIKDNYGLNQ
SGEQVDGKGQAGSVS
LEEIRQAKAQRKQLH
VVEVFLEKFTTEWDN
RANKKKQKVFQHNEL
LETGDRVKMKTLSAY
AEREQVIKLGTTNYD
DYCLDIIKKFEVSEE
DMDEAIEKGEAEQSE
SSLSDDRKWIIFDGP
YCTSPDFKMGGASDV
FGVPSEPKNARVTKV
AMEESPRKKKKRKQE
WELYIPPKQNKSVLV
DSTADDSKDSVAQGT
VIQFELGKQKNIMPL
GGLRRAVKVQCTWSE
ENYIEELKKQGFLLR
SFLLIASKSVNMKKE
TFRWTSLKHSRSEPQ
IQESSPTKNNIAVSS
KEAVYEKKQAVHKEK
FLAPVFTKDEQKHRR
NQKRAELKLVVDRLQ
HFQTHLAKLLETQEQ
QQRRYLEKRSKHSKK
GRTTDLSKGLQSPRG
RWIREIQKVTKLDRD
KDKEEHGKIKREELD
CSGILTVKEPPSFLV
QKQFRLFKTAALVIQ
SNANMLRKGPAEDMS
KSEWMDLKKEYLALQ
TICSAYLKINSKGEG
DFPKVRAKEQTFMFP
IYTYFVYKTRNMDMK
KFAECLEKKVDMSKV
SIKAFAEKLSVKLKE
IANITALKDQLLAAK
VMVKAYLKKYDPEVQ
EKSLKRIKDTNKDIS
NDLSTENKLKVLKHD
KMVAWFEKSKDIIQS

KYRAALGKLPQQAND
FHCNILMKDDKDPQC
PDHMVLGKRLKGAFK
NLSVTSFKKIVYREY
VVENSEGKQEAEFTL
DTCVCTVKLKEPPRF
VQAENLVKEASEIPL
GVNSLEVKFRTRSEN
ADLGVIYKIKLRHDN
IQEVEYVKEEAKMVY
FGNAIEGKSNTVASL
SICQARAKELEDSLQ
ISQLTEEKNDLRNMV
IPTIQWIKGDQELSN
GLSLEDSKKLTASPS
LLDSPLNKEVVKVSA
GSTFPDVKPDLEDVD
RESVNLGKSEGFKAG
VVDSGLHKEERYDLK
DTVRSSFKPGLEARL
PWMRLFFKIKPLVKS
SSHEAEEKLSEHKAL
KLKVKDNKKNRTKKK
GGGGGMAKASAAELK
TQTPVSNKKPVFDLK
DFLIQTFKVFQKNQV
NGCMNRFKNALNSCD
NIELVDLKPLPTARR
RVQEAEDKWRGAQAL
SDLHVNAKGPQVSGE
LYTQLEAKKAAIKPL
FLENEDTKFSLSDDQ
CLIHPRIKVGNEYVT
VEVPTVTKRERKIPE
SSSEDSEKVIAGLYQ
PVSMDICKQEEKLGF
MACTWSVKEVASYQP
AQNRARAKEFHKKTL
GQPDEELKPKKKVFE
GEVIALDKIFNKVEK
TLEERLDKACEPGVD
CGRWLSLKKEDGRLE
RDAEELEKWIQEKLQ
INLAGVGKPSLPSEP
GLGKVGSKGREAPLM
RVTYSPRKEKALKIY
ERFDVGPKEKPLEEP
NEVINSPKGKKRKVE
RYLNEEYKKKMSYMI
AVYTCEAKNDYGVAT
QSSGIGSKAGVPSKS
QMIEKRSKLGESQTL
SDPTPDGKLFATQYC
RFTMEVPKKLTEWFK
AGAGETLKQECDSLG
DAHEDSFKSADESGQ

NVFNNLAKGFSEGTA
EPHIQPTKSSSRQNI
EPQLNRLKEKAQQLW
MPIVDKLKEALKPGR
LDVEIQDKPDKPTGP
IVSALWEKWLSLLEA
PGLELMEKEVEPEGS
EILINNFKYIFSHLV
NIQNWVYKLSTSSDI
TTMDTNGKPLYILTY
SDAGIAGKVERNIPE
EEAGGSVKEEAGPEV
RLGVSLSKGLHHKAV
ELALARMKDLGIQCR
ELLLGHEKAPAFRDY
EVQISTEKQQKNSTL
KHRLRELKLQVKAFA
RMIVRCRKQDFPLVK
LGSGTYFKKQFSDPV
EWIVESIKAGRLLSY
EHHPRVAKYAMENHS
EEEDEEEKEEDEEET
KTVPLTAKEERGMGA
GDPIPNVKWTKGKWR
HLTEETLKVTAIPEP
LEAMLRNKELKLQES
ELTQEALKVSAVPQP
RHLIEDLKFRQKVNL
EIAPEEEKPVPVAEE
LPEPPVLKVLSSLKP
SKGIAFVKFARSSQA
NTMDVFVKEPSFDWK
VTVIGIPKPKIQWFF
YIAETDEKKINRIRQ
EQAEIQIKEESKTAV
LQTGACYKSLPSRNG
WRADINSKKVLREAG
EKLSILEKEDETEVQ
TKIIEEEKKTLELKD
LLAPHRQKWVERIRY
RWKPMSVKHTEKLER
KNTGELKKSKVLSHH
ARCLRVGKLNIDRLY
VRSSGYGKPSSPLKM
SAIIIQRKWRAILPA
TVIQSAYKSLIQHLE
RLAEETEKLRAENVF
ETREDRTKRLFRHYT
PMIPQSGKQHLLDFF
HERSRRRKHKRRPKV
EFPLPDEKTKKRIFQ
TCASAVQKTTFRVTR
ELDQDLAKAYLTVLA
EQNHLYLKTKAAVVT
GLLGASFKSTGSSVP
RLSSWCLKKPGHVAE

IISSLTPKHLQAGMQ
EFSDDIVKIIQAAIN
YLDSTFTKRDPRGEY
IVSNEAGKASCTTHL
NDNAPIFKEDPFISE
ENIQTLLKSKEQEVN
PVLIPEPKIRAVSPH
DRVQAIVKGTKQRLV
STGPITAKDEVELPR
HRDADETKEWIEEKN
LEPDLAEKNPPLFAQ
ENIHSRMKTVKLPKK
VPGFCLPKKGVTSIL
VKAGEVLKINADIAG
PGVQGDVKGPQVALK
TKHLETSKQLAAAFQ
PTSYQAPKHFHIPAG
SVHSREAKESSAAQA
ERNIEAGKDDGLTDA
CSISQELKLKSFTYE
SPQRLCSKHMPQLHS
VPVPVQKKEAPPAKV
ALKMSPYKDILETHL
VSRYGELKSVPTTQC
QPSLRAAKEHAMPKD
AKDLEQHKIKQEEIL
RTLTYTAKPVGGIGG
LKKLEEEKAEAAEAA
MDAENMRKELAELES
SPLKEDAKAPEKEIP
NVSSRGGKKFLPVLK
EEEVTVVKGQPLYLS
QEERAFLKVDPELVV
DSLAAHVKNLLQCES
KEVMAAYKGKRRSEA
GLVAQLQKTVKELQT
RLHTRQMKRTKCADI
NGFSAMAKTVGLPTA
QRLLEAQKRLEKLTT
KAKAQLCKIEKSKSL
LPCLSLRKLPPRSEI
QKLAELEKILESNKE
ADSLYVEKIDVGEAE
XXXXXXMKWVTFISL
VFRRDAHKSEVAHRF
DLGEENFKALVLIAF
CPFEDHVKLVNEVTE
ESAENCDKSLHTLFG
EMADCCAKQEPERNE
DNEETFLKKYLYEIA
PELLFFAKRYKAAFT
ECCQAADKAACLLPK
DELRDEGKASSAKQR
LKCASLQKFGERAFK
RLSQRFPKAEFAEVS
KLVTDLTKVHTECCH

DDRADLAKYICENQD
NQDSISSKLKECCEK
CEKPLLEKSHCIAEV
AADFVESKDVCKNYA
VLLLRLAKTYETTLE
DPHECYAKVFDEFKP
EEPQNLIKQNCELFE
FEQLGEYKFQNALLV
ALLVRYTKKVPQVST
EVSRNLGKVGSKCCK
QLCVLHEKTPVSDRV
VDETYVPKEFNAETF
DICTLSEKERQIKKQ
TALVELVKHKPKATK
KATKEQLKAVMDDFA
DFAAFVEKCCKADDK
TCFAEEGKKLVAASQ
NECFLQHKDDNPNLP
XXXXXMGKGDPKKPR
LDSEAFQKRKQMQEI
PQLSEYCKMPLKMWH
IILDTLEKCLAGQPK
KAQAAEMKAVHEAEG
KVSTEHNKECLINIS
MHLIQTGKGEAIRIR
QELISLAKSRYSCRD
VVFVQTDKPVYKPGQ
ALKLQKGKQYLMWGL
TVPETGRKDTVVKVL
ALNDHFVKLISWYDN
AEFAEITKLATDLTK
SIGLGYHKDLQTRAT
NCLAPLAKVIHDHFG
ALVGHLLKGYRHPSP
GRVHESVKRSNCYMV
EIIQDISKDVVDESN
VTLGCLVKGYFPEPV
VFANMQMKLSQSRPI
ADGNFDPKPINTMNF
VVGYYCLKVPLVVFK
NRTWQDYKRGFGSLN
AEGSPQYKKWYFELI
GIISIPLKLPSLNKE
KQLRKIVKHEIMSVK
DGFAESTKRKAAVWP
LYPMLFNKLKNTISK
TRHMEAVKTQFLEQA
VVAASEAKEDSPAQT
ATAPEDDKAPAEPPE
NNLHLQEKSDDPAVK
EENGMAWKEILNLLY
QHSPGLDKENVELSP
TALAELVKHKPKATE
TVRKQVTKSYKMADE
EFWANLEKETDWLRN
DVAVWFSKRLPTFVN

RFAGPFSKQTQIPDY
CAIYTYGKPVPGHVK
DLVPGSFKSQLQEGP
AAVRQRMKCSSMQRF
GELLITNKGSELAFA
WLTAYVVKVFSLAAN
ARYEREMKTYIPPKG
LSGFIPLKPTVKKLE
SVSRVELKPGDNLNV
RVTAILDKLITMTIN
NMVIADVKMLSGFIP
KCCAADDKEACFAVE
ISCADAEKEIMRLGT
ANHVLFTKEEHMRPF
SVRFVCQKGFVLRGS
VSREINLKDYEGQQK
QMLFYICKKLTSHQM
FPNNYWDKFVKRKVI
GGLVEGAKNIRVTKI
SSPCSSDKQNVXXXX
LEEPDLEKVVTYLAG
RYNMPLEKQQPAFAL
QYDSTHGKFHGTVKA
KEYDPDGKGIISRKE
NLLQMTEKFFHAIIS
DLLVHCGKLCALVYK
PSFSSETKSACKEED
LMVCNHEKVGLQIRT
VERSYAVKSGKWYFE
YAVYSLSKSYIYLDT
MVLSAADKTNIKNCW
LLKETAFKFKALKKV
GMRDIPMKYSCQRRA
AVINSLEKAFWNWVE
GMGVPEIKYGDSVCF
FAYKFLRKILKYVDS
NSLFDFQKNNKDSNS
RGNFLEIKNKFLARR
SLSHSPGKXXXXXXX
NAFERQNKAEGLGMV
AADFVEDKEVCKNYA
SVLGMPDKVEDMCPD
TVKAENGKLVINGKA
LFIVSISKTLAANEP
LQMISASKGEMSPMV
PPVEPNTKVFPAVFL
DQTEQWEKFGLEKRQ
EIPLQILKTKALSML
ERAATVTKKTYEIWH
FQVELPRKAFITNFS
EISCWTVKIXXXXXX
TDQHGLAKFSIDTSS
QAELLPSKSVSLIIL
NTLCDEFKADEKKFW
PTLSQQRKSQNEQDT
YINVDCAKLKRLLKE

SLLNGNSKYMVLVPS
EAMQAYPKQDRPLPS
TNVEMILKFFDMFLK
LLALLLLKDFDSVPP
VTVQDFLKKQVLTSE
MSFSAGVKNWVDHMM
GVELMEAKSQLFLKY
ALLDERTKKSNRDSL
ESHQHLSKESMAGNK
SPEALVGKSLYVSVT
MAPSLTFKEKVTSLK
PMTVICQKDLSWSML
KVTSLKFKEKPTDLE
EILKSLDKEAIREDN
VTYPGLSKRRLKPEA
LQLSGPDKVTISLLS
VAFNFFRKFYNKSED
VIPELNGKLTGMAFR
VDESNINKKLFLDSL
TVYVDAVKDSGRDYV
RWPSHFQKEAFVKPG
VREDMETKCFICGIG