**Supporting Information S3.** The benchmark dataset used to train and test the model for predicting the possibility of carbonylation at Arg site. It contains 136 positive samples and 847 negative samples. None of the sequences included has pairwise sequence identity with any other in a same subset. See the main text for further explanation.

**(1)** List of the 136 peptide samples in the positive subset

ASKLSKI**R**VVRKSIA

IIIQSYY**R**MHVQQKK

HCKAFKI**R**KHYLHLR

IHLELRN**R**TPAAVRE

LEELDRE**R**CFLLNEI

YEMHKMT**R**DPLNLAL

RPPVSGL**R**RRKPRAT

KTQRSPV**R**IPFMQRP

KQFFKDW**R**DKDQSDG

KSEYKSY**R**LRAQLYK

KSRDEGP**R**LGDAKLK

SLRKLPP**R**SEISSEK

RKSTEFP**R**KIREQEP

GSVTGTK**R**LRCMPAP

ELQDMAA**R**QQQQIEN

RKLANAP**R**PLKKRSS

ASSRAHR**R**HRRQAVF

RALLTAP**R**SLRRSSC

TMLLPLL**R**RFPEQVR

SIDSLKQ**R**LNVAVKE

EADEEKE**R**ILAQLRE

VADLEKQ**R**DCSQDLL

EEEEEIL**R**LHQQKEQ

IELLRKV**R**EQVLKCP

LLAEVER**R**IRRGHAR

PDGENRA**R**FLPGKDL

WERKQNL**R**PMPKKYH

RVLSGRW**R**LPLRALP

VSWGKLK**R**RVKGWAP

TLPKRVP**R**AGEARNL

EGGHFKL**R**VTISSNN

TLVEKEM**R**VTLAKLL

GDSRLKM**R**AELGEYI

REQEPTE**R**PLKGIQS

ALAKHKE**R**KMAVPMP

KDLGLHP**R**AVSTSFG

CLEVVLC**R**VEGMTEL

IKGLTYL**R**EKHKIMH

MQKLNQK**R**AELKLVV

RIPQLWK**R**VSWDSST

GYKRKVM**R**MLKRQVS

ALVPPWT**R**VYRVQRD

KCSEETF**R**FELGGGV

GAGSGLG**R**LLALQFA

SQDKMEI**R**SCLPKLL

VDKSGVV**R**VRVEGDN

KERLQEV**R**DKLGLQY

YVGKAAT**R**LSILSLL

RSVFTVM**R**RQLLEQA

QRGLTEK**R**EELLSVP

RKQLIPI**R**HTKTREK

LLEAEKE**R**NNKLSLQ

STEPLPG**R**KQVRDTL

RQLAIET**R**KTLAAQT

LSHPKIP**R**CQERLLR

RDRPGDP**R**RKRSSDG

FKDTDMK**R**LSMEIEK

LKRSLEM**R**ERVLGPD

DRGHEEL**R**KLESTLD

SGILRQL**R**QTVEATN

TAEDLRT**R**KSKLQEL

KAAAADG**R**GMLPKRA

PPSAASS**R**YSMRNRI

SRSYTRS**R**SLASSHS

HIAQKLF**R**DGHFNPE

ETSLNKA**R**EHSLLRS

HVTKTRR**R**CKTVRVD

RLSSVFL**R**VRTNVGV

GQDGSGR**R**GPQKRGV

PRVFPPR**R**PHTPAER

APYVTLR**R**GLNAESS

TSQGKLL**R**TILFGVK

IYKPKRG**R**PKSKEMP

ENELTFL**R**SQIAAIV

IANVFQN**R**DYQCDTV

LMEFITS**R**KRGPLWN

VVRPKDR**R**VGTPSQG

RRSVRMK**R**PSSVKSL

PKTMPRV**R**KPKTTPT

HFHTVIA**R**RAGMAAS

RIKNPKA**R**VTEVMKA

LFTQELA**R**RLKGSGV

WDATEDL**R**ISRTDSF

ELIRLTQ**R**LRFHKVK

LPLVSWS**R**RLEFQGE

FLQVETR**R**AGERLGW

LSYRSLR**R**RVRRLRR

VGLGLPN**R**LRFFRQS

GLAARLQ**R**QFVVRAW

SLRKGIP**R**TKSVGED

ETRQRLL**R**TVKKEVK

AIVRQRE**R**ESHAAAL

ASPWKSA**R**LMVHTVA

LGKRADQ**R**KAKLGDS

IKSMAAS**R**RAKLNES

QRMKALA**R**ALPLQPR

SASKASV**R**SISVEDK

EEELLNV**R**RELMSVF

AQLNIKE**R**LIPPSFT

SDPIKAC**R**PIKPPGP

HHVVSGL**R**ENSEYFF

VVESSVL**R**AKEVTWY

LKRTKKP**R**PPSLKKK

KESLLQD**R**RQEEDQR

QEERLEQ**R**LKREEVE

PEKENAV**R**DNKVYCK

EKLHLIT**R**NLQEVLG

HGKQPSL**R**AAKEHAM

NCILDDK**R**PHVKKAT

NRRTLLV**R**PISKQDP

QADNPLG**R**SVLRKDI

GPPTGAG**R**GPSQLQA

KVLKNNI**R**FMNHMKH

SRLSRHR**R**AVHGPPE

LAPLNDT**R**VVHAAKA

YCDLPEP**R**KPLEKAV

CALSHHE**R**LKCDEWS

VKVPMMK**R**LGMFNIQ

YQLEGNK**R**ITCRNGQ

HKSEVAH**R**FKDLGEE

LPKLDEL**R**DEGKASS

SLQKFGE**R**AFKAWAV

EKTPVSD**R**VTKCCTE

CTLSEKE**R**QIKKQTA

PLPKEPP**R**KDPPPKD

XXXXXXE**R**IDGGITG

RGGEGNA**R**THGTPDL

VLGNDGH**R**EVVKEVV

RGADYSL**R**AVRMKIR

QQARLHE**R**GAAEMVL

LPKIETM**R**EKVLTSS

CTKPESE**R**MPCTEDY

YLSLILN**R**LCVLHEK

TESLVNR**R**PCFSALT

VGVNGFG**R**IGRLVTR

AITIFQE**R**DPANIKW

**(2)** List of the 847 peptide samples in the negative subset

VTPHTQC**R**LKLLKLE

VLSFSET**R**AHQVVQQ

ELSKANS**R**FATTFYQ

EEAAGSD**R**ASDLDSV

GSNGVDH**R**SYRIIRG

HSGASQH**R**IARPSRQ

DSLLGKN**R**KALAKGL

IVDEVHE**R**DVNTDFL

SDVEAHF**R**AAHENTK

ELFVKGV**R**EVYDYYC

HASCYGI**R**PELVNEG

ASPATVS**R**MGMVYIS

RVQRLGE**R**VVDSGRS

RSLSRFE**R**SARFDIF

PNTYDIH**R**LEKILKA

EKPGGKE**R**GSSASHP

VTEASPW**R**ATNPFLN

KKKKKKG**R**MKKEDNI

GEPKDVI**R**KDVRAIL

PHQLLLQ**R**LQEEKDR

LSLSHFC**R**SPFLCFG

TTEAPVK**R**ASLLGDM

QEECGYL**R**RHHQEEV

VSNTLES**R**LDLIAQQ

KKKMTKA**R**LERFKLR

VYTTKSP**R**ERAILGL

PASPISQ**R**LPPKVES

GAQMEAA**R**PLVQENP

TKDNEHK**R**SLTKTPA

LHVTDAG**R**KHIAIAW

INYVLES**R**LIGTEKF

SSASLIW**R**SEAEASV

KFTTEEK**R**FVESRDV

EGEQYLF**R**IRAQNEK

PPIYTAA**R**EQTPFRH

IRGNEAG**R**FRLDITL

EWDKDVA**R**KLGSTTV

TDDSGTY**R**AVCTNYK

ESKGPMT**R**RLLLHEV

LDLKETL**R**LRILSED

DELLRHH**R**DEQKPAT

ALNMLTW**R**AEQEGGM

PSSGLKK**R**PISRLQT

CVIQSYW**R**MRQDRVR

DVWIDVQ**R**RWVYLEG

WLLAHCG**R**PQTECRH

AACVMPS**R**LKALGTL

HKQTKKR**R**ILIPSDI

SIQDWAK**R**MKALVEQ

QENCILL**R**LDNELGG

SQFNLSG**R**NPQKQAR

KSMSIVL**R**ENHQKPE

LPEEIAL**R**ASDGPQL

EIEKLEF**R**VRELEQA

EKQVELY**R**NGQRLLE

AKKPAAT**R**KSSKNPK

EGQEYSF**R**VRAVNKA

SISGCKM**R**LSYLSSR

RFRTLMM**R**RRFLSLK

TYDSYYS**R**SRSRSRS

EKETLLQ**R**LTEVEQE

HWGQQQI**R**PIKSVIP

AVPPAKG**R**TVLEEKV

DLSTIQE**R**MEELKGQ

SECGFCL**R**QFPRSLL

TPSEAHS**R**IFESVTL

RIQVFLA**R**YSYNPFE

KREKEVA**R**KLEFDGL

PLKTDIP**R**HPMPFAA

CYIFPGG**R**GDSALFA

RDMEKVE**R**QAVPQAN

LDLDSEL**R**KGIVVRA

SNQSVLH**R**WERKQNL

ECSKTSF**R**VANLEEG

MSTIEPH**R**QVAWKRA

YKDKREL**R**SGKKYKI

LEAQVLP**R**FQPRVLQ

SPQPPLS**R**AEAIKQV

ELEAHLC**R**MMKHSMD

EKKLEDA**R**ARCEGQE

TYSSAEE**R**MQSEQIR

GLELSID**R**FRLRKKT

SEIRITH**R**SHHFIPR

CSKDELE**R**ALHYLKN

HFIPTIG**R**LRKRAGK

KQLPQVP**R**PLQLFSA

CEPLSNK**R**NSNSVTN

HVHEASS**R**SHAIFTI

KNPASSK**R**RLKTSLG

GPKLEAL**R**AEVSKLE

WICTKCV**R**CKSCGST

NCVSGQE**R**GRLGVLA

KLNETME**R**QRTEIAR

NYSESLY**R**QAMEEAK

LAMLAKE**R**LQEVRDK

EEKQEEE**R**SKVDDLR

SEPGANM**R**HIRKPVI

HVEAVYS**R**RCVSFIL

QNQVGIG**R**PAELKEA

IHYLMLP**R**VREELID

YGLITHQ**R**IHTGEKP

TGLETPD**R**KGKGTTE

KVVPPSM**R**RIRAQKG

RRELLEL**R**RKGREEK

PLGLILN**R**FSADTNI

GNLSILS**R**QESSFFT

IQGKDVL**R**LPPSSIT

LNTSIDL**R**TNVLNDA

EKTINQQ**R**IHARIGQ

KAKEQET**R**ISEEITK

VLIPFKG**R**PPPTVTW

MALEETL**R**QYQAAKS

PGFKGEG**R**EVDVNLP

RDLLLQT**R**LVNHVSS

DNKKKKQ**R**TSSKKKA

WDDIITN**R**CFFLSKI

RKTRSAR**R**RTTQIIN

YIRISPT**R**AYNRPTL

LEEELPK**R**VRSRLSL

QLKSFPA**R**LRQYASY

STGLLPG**R**GPGTSAP

VKAGDSS**R**LECKIAG

ATEISTE**R**DLGQCDR

CREVGLQ**R**RSVQLFC

EFKQKLC**R**AEQAFQA

KVNGCLG**R**NCKLPIT

HQAAAAT**R**EASSESD

GAQMLAM**R**GEQLGVV

TEGMDDF**R**YACQSPE

IKKANMK**R**ENKAYSF

FSCQRRT**R**FISLGEA

NSIEASL**R**CSSVVPR

ALTDEES**R**KNWEEFG

EEKPEEE**R**SAEESKP

SQELDVK**R**REAIYND

EKKSKKI**R**DKTSKKK

DVFLDSG**R**KTRSARR

KIDHELH**R**LQALLKH

LCPEGVH**R**FQWIRNL

QDQKVNI**R**VAVLPCS

GATTSTI**R**SSCLVGC

APMTPAA**R**PEDYEAA

KNLVQKE**R**ELEEQEK

QFGGIDQ**R**KIFTFAE

QERDRKF**R**LEEQKVR

VVLPLDE**R**AFEKTLT

ALAVEGK**R**IGKVLDH

EFLELKY**R**LLSLLVL

AKWFKDG**R**ELSADSK

RVRPLSK**R**ETKEGGR

IDVNVIA**R**PSAPKEL

EAYITAT**R**QYVAQAN

AQSEAAG**R**TEPTGPK

CYFHSGP**R**GTHDLWD

VNLVSVG**R**VLWQWTK

ALHTEIT**R**NKDSTLE

LQVNDHV**R**KVTDQIS

VTLQSAY**R**GMKVRKR

IVTEEHL**R**RAIGNIE

CPGMMLW**R**YPEPRVL

EDVHIKG**R**IIDTVAQ

ELVPPSF**R**GGKQIDS

KLQTAIH**R**VTLAQTA

VHSPDIY**R**VVVEGER

DIANVMQ**R**LQDEQEI

KFVGDAT**R**LTDKLLE

GLTNPSL**R**SPEESTE

KGNEIVL**R**QDIMPTT

RSRSGKG**R**GLPKKGG

SGAGCSS**R**DPGPPEP

LEKIKAK**R**SDSGKYC

KFSSKYL**R**TEHGSEM

SKSIEKK**R**GRPPTFP

RKVFSPI**R**SEPRSPS

WYNVENW**R**TAVTSPD

HRLIAKS**R**ELYEKTE

GLLFYSL**R**TLTTMAP

ESDSFCL**R**NIMEAPL

VKEVAAW**R**ERYEDSQ

KLGKVND**R**WQHLLDL

LGYWLEK**R**EVNSTHW

ERDYLLE**R**RDLAVDF

IIDEGLK**R**FIIEITF

AEAAHNQ**R**AGAPGIQ

TSVPTTP**R**ATPILQP

PLDQTLP**R**QGPGQSL

IKALADE**R**EAVQKKT

LINGGSE**R**KSCFWKL

VADLGNI**R**IRFIRKN

AKISDLH**R**ESVDHLT

EDVTFNK**R**KMKELED

KKQQQQC**R**PSISISS

PKVLQDY**R**KLKNTAE

IKKLEPS**R**IVKQDEF

GKPKVLA**R**TKGSMLV

AWIPVPA**R**REAEKVP

SLDWLIS**R**LERLEAS

RLKAQEH**R**AEMEEKG

DEAKTSA**R**DAALAYH

PTTRSSA**R**RSQAGVS

LYYYNFT**R**TVISSGG

DTQTIWA**R**TGYLGFL

NELAYFK**R**ENQELMI

ETQIDET**R**MGYKPVA

ALFPKGS**R**SVVPRDT

EVFTYSV**R**GLYENHK

PRARPRE**R**DPGRRPH

NAALDNL**R**KVVPCYS

ESTSQIE**R**PLSQEPA

YDGPKNV**R**SDISDQE

PFTPVPP**R**GEPVTVY

MKKFEES**R**AELEKVL

GTIDGEV**R**LTGELDR

SETKDPL**R**VQPHLKK

LMVHVDK**R**ITLAAFK

FFAKLNC**R**LYRKANK

QKVAEQE**R**TAQQLRA

SEILSNI**R**SCRLQCF

PLPTVNE**R**DTENHTS

EQEQLYL**R**SGVVSSA

AVVDVSD**R**AVPPSFT

LEYCEVQ**R**DPYSNYG

EDISKLN**R**AAKVVQE

QGETNKN**R**TKGGWQQ

EFITSRK**R**GPLWNHE

DMSMLEE**R**IKRSAKR

AKQFRAV**R**TTEGEEE

VKERESE**R**NQCNFKP

DKINKTY**R**SQLSSEE

EGVKYQF**R**AMAINAA

GPPASNL**R**KQKSLTN

DVLLGAS**R**TLSLALD

EEVMHTP**R**TPSCSAD

PGLPVEV**R**GSNGAFY

PGSTVLY**R**IFTVNHK

AQGALPE**R**DASRGGL

PVWAAKQ**R**VLCALNH

LEKRQMD**R**AEHKGEL

ITESDAQ**R**TMYPGSC

ETPSYAK**R**RRLAGPS

EDEVQFL**R**TDDEVVL

AVVLTSY**R**STAERKL

RSEGEKV**R**VGDDLIL

LHNSLLQ**R**KSKLQSL

GWWVRNT**R**ILRKCIE

LLANAPE**R**VVERRRR

RNGPAQT**R**VSSSSSH

SVQLTEK**R**MDKVGKY

EDTWKLG**R**NWGQSVE

SWIWVQL**R**TGLARDG

HCFHMCD**R**MVIYFFI

STPSLTS**R**KIHGLSH

QSVAKMF**R**AVEEGLT

LSVKDAV**R**VDSGNYI

KIYNIED**R**TRVPNKF

DADLEAL**R**RKLANAP

NFLASKH**R**QPPEYNP

CEPNCYS**R**VINIDGQ

PPEPVDK**R**AKAPKAR

NKPIPAL**R**VVEEKKK

SKLLKNW**R**DLDARIT

IRNGVTK**R**GETLSWL

YTMESTD**R**NQTFSKE

CNGQNTL**R**NIVHLSK

PLPIDLY**R**YDDVSGK

LRGLLRS**R**GAWEQQI

EDELQLP**R**LPELFET

FKHIKDT**R**YMSSYFK

QRAAAMA**R**TKVEIAN

QNSSVLD**R**LPQPAES

TSKSSDN**R**ETPRNHS

KDTLMIS**R**TPEVTCV

QLKNLKI**R**TNRIQRF

HQWIQET**R**TYLLDGS

QALLSKM**R**AIDLQIK

KKTWAKV**R**TAKFFIL

MSFKDAE**R**GDDTSCE

VEEIMYL**R**QIYKQQL

TTAATVQ**R**PGPGQTG

VGLWSMC**R**TNRGTVG

XXXXMSA**R**TPLPTVN

PHPSPEA**R**APLASAS

DTPPSAL**R**GSQSVSL

HCSLDAL**R**EFFSTIV

STSYMHQ**R**SPGGPTK

LLIQKYY**R**AYSIGRE

ESESDNH**R**SSSDFFE

ARFDIFS**R**YESMKAA

TNYILEK**R**ETMSKRW

KAPSESA**R**GPPGPTD

RKKKSKD**R**PSKLEKK

KVELSQL**R**VAKVTGG

VTSQFTT**R**DDILRDW

GHDEDNT**R**ISSAGCL

SGSFVSS**R**ARREKKS

HLFYKIH**R**NRYPVCT

GLCIFFV**R**CRNDVAI

SLEKNLD**R**MKLDWVN

MTCLSID**R**FIAVVHP

AETVPEE**R**ARGDPDS

NEHCPSK**R**DPSKLYK

LQQQEQL**R**ALQGRQA

RLPTLPK**R**VPRAGEA

PDSMFGD**R**GEIIDPF

EGQEYYE**R**EEGYDEG

EVNAAWQ**R**LKGLALQ

TFQTLEP**R**ASDAKNQ

FLALILQ**R**TFLQASY

EHQMYEF**R**VKAVNEI

QVWETWP**R**AQELCPE

IVNLLKM**R**QGEEVFQ

QIHMIAQ**R**RDAMAHR

SLSEEIQ**R**RLEATGM

VYTTAEQ**R**PNAYIPE

RRVRETQ**R**HISTVKY

AFYEKMG**R**MLGSAFP

PGAKVTI**R**ASLGPLQ

TKLLKAL**R**VRKKTGG

ITLMDVT**R**NSVSLSW

LVAVLLS**R**TWHVRRQ

YLTVTLQ**R**PTKELHG

AAMFQSE**R**KNPAPQC

QDHFQVA**R**QNHFPQR

SPRRRKS**R**SPSPRRR

SSEEKTQ**R**MSLMRHH

KHKGTEK**R**ESPSPAP

FLDVFLP**R**VTELALT

HFFFPKS**R**IVRSLMP

AVQKTTF**R**VTRLHEG

GYTEVQF**R**PYRTDDF

DPVEHGQ**R**HTATKRE

HEYEPYS**R**LKSDIKD

TKFQKLH**R**DMEEATL

PSAMPQS**R**STESAHA

VWEPLIE**R**VEGKRQW

PNVDYQF**R**VCARGDG

SWIADCA**R**HHCSSTP

VLESTTG**R**ENSDQHY

ELSLNKQ**R**KLNKSES

AWYMHGE**R**AHYSHTM

SATLKNF**R**YHISLKT

YIKSELD**R**ELQDRYV

LGLDEHN**R**VKVYRFX

LDTYNVK**R**EAAEQWL

KAWKKHC**R**GEGKISK

VFKHGYE**R**YNAMRAD

QLSEEQG**R**NLELQVL

IKVLEKK**R**AMFECEV

GTPVFAD**R**LPEKMKT

DPSAPQP**R**AETSGKD

KSLLDWL**R**QQADYSL

KAAQQHW**R**HELASAQ

LNKHTAL**R**GEMSGRL

TCMKFYA**R**VCRSGSG

PKKVEMQ**R**SLPGSLL

AKCQVTP**R**RNVLQKR

FVELGVT**R**VGHRMNI

IMEEAVT**R**KFVHEDS

EPIQFTK**R**IQNIVVS

LVIQKYW**R**RVLAQRK

LEAKQKA**R**KKEERQS

LLEEALL**R**LLPAELP

YQTMLKG**R**CPSVSAP

LTKEKVS**R**LDRIVAE

KELADLA**R**LHPTSCA

FNELEAL**R**AESVATK

YLVLQQL**R**CNGVLEG

QMVTPPP**R**LVVGTYD

VPYFEQC**R**RDACRCG

EQEIVQK**R**TFTKWIN

KNQGLFH**R**VLREEIS

AACCGSC**R**NVSCLFT

SLAKFAG**R**KLKDCGE

CSSLSPQ**R**LCSKHMP

DAAAAGD**R**RPAPDTW

LDINVVI**R**QVYHLME

DDQEIQK**R**LAEQQDL

RYIRSAL**R**QEKALSA

LEKAAAQ**R**KAKLDEN

SGEAPGI**R**KEMKDVT

REEAEKQ**R**VASENLR

VLQKCSH**R**LQELEKN

MEFAVNG**R**PLEPSQA

AREHSLL**R**SDRPAYV

SNYQSTT**R**VKPFICT

AVASAAS**R**PDSTHMV

LKYQSLL**R**MFLSYCI

NSQVTFP**R**IQVTSLS

DALTDPS**R**KTQKCLQ

EVLFLSS**R**SKVFERA

DTIAHLE**R**TRENMEQ

AEQKSLL**R**SVASRGE

TSAEQDN**R**FSNKQKK

PANAGNM**R**YSHSTGK

MVAVLVS**R**TVGPTQR

NHHMEVL**R**ELCEELP

LVNKDST**R**RDSGAYT

AMVFMDV**R**MPKFKIY

LVSHVTL**R**LLKPECV

EKFGLEK**R**QGALELI

EGCEYEY**R**VYAENAA

YACGVED**R**KSRARLF

QPDVTGI**R**LLSLGAG

EELHKEI**R**LIGNESK

SFLKLTT**R**AQLGAKS

LRVIEME**R**EGKGRIS

QIYGQKG**R**TDEILLN

IYNTKSS**R**ERAALSL

VELLFNT**R**YAKAIGI

YGGSYGG**R**GRGRRTG

EDTQYQF**R**VYAVNKI

NSQPVAV**R**GGGGKQV

TPSPPPR**R**RSPSPRR

LEFDGGS**R**GKGEHFP

YLMGGTY**R**CTYTRKR

MAVGMGE**R**PEVLHLT

EGGKGEK**R**KRETDDE

ITMWLSK**R**LPQFLQV

NIVWKYQ**R**YHFIMAY

LTASPSS**R**PVASPGA

XXXXMAV**R**QALGRGL

QAMPYNC**R**NNLAFPA

LFGKMIA**R**AGRAGNL

LIVCFIK**R**SRGGKYP

AHEKRKE**R**LQMLQTN

ASLQPKG**R**RDPPLLP

LQRQHYL**R**QRHSAVI

ELKALAQ**R**GIGYHHS

QDEWEIP**R**QSLRLVR

AALESLN**R**AVLAAMD

QFLEEEG**R**TLEDVAR

KDDIKST**R**KLKESLI

YVRVKTE**R**KNFLAVQ

VNSTHWS**R**VNKSLLN

VDTKYQI**R**IQIQEKL

AGLSTDN**R**NVCLWDT

PRRASPL**R**TSRSRPH

MPSKFSC**R**QLREAGQ

GENQLLE**R**RRLYHCA

TTIDDGS**R**CFFTKSK

AAKVFGL**R**SRKLKLF

ERLKEPT**R**QALQQRL

VIFYDFL**R**GLEASWI

DEEASYL**R**FRNSIWK

PKNDGGS**R**IKGYIVE

LQVITTL**R**TAAKEME

SGEGDLS**R**HLGLLGL

NDQLEEQ**R**QEQALQR

QEGSLSA**R**WPVAGQI

RQILRLL**R**LRCTEND

DPVDSTL**R**DFCGRCI

PVYKTSA**R**RGTLSTT

KTIGSPK**R**IQSPLNN

TNSVTAV**R**IQPLEDI

LIPVIGP**R**KNIKKQQ

ILVNTNL**R**ALINKHT

WGTATVG**R**PRGPPKA

WLEAQEE**R**LKTLQKP

GETPEAC**R**QARSYLE

LRQCLAP**R**LPHSVTC

ADEIYEL**R**VTGRTQD

NAEKYYG**R**KSPVGRD

LELESQK**R**LYEKNLT

KSEAKLS**R**KQVDSEH

SSLSSAG**R**PGPSEGG

KVFAKVL**R**EEDSKDA

FIAKAFK**R**IKDSEFE

QHHRLES**R**YSSSSGG

LAVGKFT**R**TSGETTH

VKMCVGS**R**RMVDVMD

QRPSGED**R**WQEKDQD

GSKKVTI**R**SPVQIRN

TSENTLK**R**VSSLAGF

QPGTFLL**R**FSESSRE

TDFWKTL**R**YLSLLYP

AEPGEGT**R**PATVGDS

PLSPPQH**R**YSEGERT

GCFLFIS**R**SKPLKTL

VFLHQKP**R**LPKLFKQ

HSYSVLE**R**FVEECFQ

TNDLTAK**R**LLHVKGR

KSLKAKF**R**KSDTNEW

LNELEIL**R**NSAVSQE

EWKFKLF**R**VRSFEKT

KASAAMV**R**LRLYDIL

EYNAVKE**R**EFHNQYR

MVTMGFA**R**DEINDAL

INIDQFV**R**KYRAALG

DISALTL**R**ISELESQ

SCFKDGV**R**QIPIESI

PWAKYVV**R**EGDNVNY

LSFLDAY**R**NYAQHKR

SESEDEQ**R**PRTRSAK

IGSDSQG**R**ATAANNK

EEPGSSS**R**ESVSKAG

GPIFTTA**R**QLVHALA

GQSPFQQ**R**KKKIKRE

DALEEQQ**R**CISELKA

YKTLHDT**R**THFLKTK

HSSEHAG**R**NGRNAAD

AAIGATP**R**AKGKAKH

SKRKVLG**R**RDSDDDH

INKEAQK**R**WTRREQA

SFMSLQL**R**ELVIKSL

EHKTMAS**R**AKVMADW

VKQKELS**R**IEEALQK

QYSFNSQ**R**FAELLST

SLHYALA**R**KGGAGGT

AHTKPLN**R**RSVLEKN

SLDAANA**R**LMSALTQ

SGLDIFE**R**INTSAFE

AVDSPAG**R**QQLLQRG

LVFHSIT**R**SHAENLE

AEYTVVA**R**NKYGEDS

NAQILSQ**R**IEKAKCL

EDKGEPK**R**RGYLQVN

YGQIESV**R**FRSLIPA

ENQQCLL**R**EETWLAS

HQVESFI**R**KKLESLL

MKFKQKP**R**EEQLEAD

ASAADTS**R**LSLWGIQ

LPKVTLS**R**DGVPLKA

ETLGVLA**R**AVGEPMR

EPRRLSS**R**RSVLTSP

SVVWLKD**R**GEVIRES

ILTLAWD**R**VDIAKNH

VYNENSN**R**QKLEHVK

FKIAGDM**R**ATCPAFN

PSHGLAD**R**VINCREV

AQQEELE**R**RKRLEQQ

MEDKIYD**R**QVTKQSL

EEQPTSQ**R**DRLSQVL

TNYVIEL**R**DKTSIRW

QKKSAFT**R**TVSPGSV

EVLNEVN**R**RSLKKLA

QTAGVID**R**WELLQAQ

RSKVDDL**R**GTPMSVG

REGITTK**R**EQVQITQ

EASTSAL**R**ESCQAEH

PANGFLV**R**YLRRKLV

NNTDEMA**R**LIRSVMQ

ASERKLN**R**RGSMSYL

YSASTPV**R**KPRPGQQ

EMDISAD**R**VKVEGEL

WDLSKKL**R**IPAISRL

SLALKNI**R**CRRGIHK

ALQAALA**R**VVPSYMQ

GMYSEPL**R**QFRDSSV

MDGAAWD**R**RNGKLME

GSYLQFL**R**ILFRMLI

LKHFIGQ**R**EEFETAR

LVSKDLI**R**KAGVGSV

GFKSHLI**R**LIGNLCY

AHEDISQ**R**VAAENQD

EIEEDKA**R**RILELSG

SMLQHLL**R**RLVFDVP

ISPPSPD**R**PPHSQTS

RSRGKEQ**R**KLARQRS

EEAEAAA**R**ALARFAQ

TLTIQKA**R**VTEKAVT

EFTQDLF**R**FLQLLCE

GAHKATL**R**IGQDGIS

ANLKMEL**R**DALCAII

ENSLETL**R**FSISNLS

MFKWLVA**R**INRALDA

HKFLTIP**R**LEELYNT

RYNERKG**R**SELIVVE

SVDVKSV**R**MGSIQRK

KGGSKLL**R**MKLSDFN

TGTPSDP**R**RRLRSYE

PEVFKYN**R**FLNPDGS

ESVKASE**R**PLPPGKI

AYPICEW**R**YDACASP

PETIDTA**R**LHYRNSW

RSALFAL**R**YNILTLM

EPLKELF**R**QQEAVRG

GDEEQKL**R**FAFSIYD

YLMLDNK**R**KEVVHKI

RLDPVSG**R**LSTISSL

KPQDYIP**R**APTFLMN

NSHHECI**R**KLLQSKC

TSSLSPQ**R**SKLRIMS

QGLKIHQ**R**FHFASAL

SSKPIKI**R**GPSHCAG

HANNVTI**R**ESMQNDV

PGHFVGP**R**GPHPSQF

GLDPTQF**R**VHHYHKD

PSTNTES**R**KDVITIS

IDAIGTK**R**YDSNSGG

AESNSNM**R**CTCRIIE

VEVAATE**R**TLLGFFL

YDYVHSV**R**GKVAPTT

SRALSHD**R**QLISQDA

GLPNNAE**R**VLLTTQG

FDFPDLN**R**FLTRTLQ

GGKPVQN**R**ELQGYES

LSLQVVN**R**EIREENK

KAFVTLE**R**FDQLYGS

EAVKYSE**R**SLTKCIG

FGKKAAN**R**SWQNVYC

TLRMRSH**R**VPCGQNT

QVSETLK**R**FAGKVTT

EAVNVCG**R**ATAVVEV

EKALRSM**R**KAQVSPQ

EKVLSPL**R**SPPLIGS

LDLVSHP**R**ATQTHVE

LISQINK**R**YKTKDDI

SSEKVAE**R**KETEFFS

TAKPINP**R**PSLPPNS

EAKAELE**R**KLSEKDE

ITKVGEE**R**KGPLPKS

LVEETEE**R**LVERISF

QIYSTGL**R**KGNLHNV

KHFPNID**R**EKAMSRP

SMSKGGK**R**KDSVWGS

WADPTGT**R**KLTWTYA

AQIADGF**R**IRVDLAS

LRDWNLI**R**DAATLIL

AEAADNQ**R**EEAADNQ

VYQIKVH**R**KYTGEDF

ARLETLV**R**KAGEQQE

DASPGDK**R**LAAYLML

LSSYATA**R**LMMNNEY

KLPVDFE**R**SFQEEKP

GPLPYPE**R**QKRARSM

DSAKELQ**R**SVEFDRE

TGTLVSK**R**EVELEKN

GLFNKNP**R**HSSSENN

FNVPDEE**R**CSFATVN

ERERERE**R**EAERAAK

QQLNKDT**R**SLGEEPV

FSPLISA**R**FGMGSMP

KTKTDYT**R**IKSLSIN

SSPSSGQ**R**SASPSVP

QHTTMAT**R**SPALPPE

VQQVTFS**R**GTDGQPL

KKKGGLF**R**RGSLLGS

GNFTDAE**R**RKCEEEI

FLNPDPL**R**ADGISDL

WFETQEK**R**LKQQHRI

QGYIIEK**R**RHDKPDF

SLWSNTV**R**CYLIYTD

FDFQAMG**R**IFVGLCQ

TPGGTIS**R**ASRDLLK

DTDMDSL**R**EARREDM

KKLRLKF**R**DVKCIPD

EPSFEAT**R**SRNRYSV

SSGKSVA**R**TSLQAED

VMLADSP**R**EESNKRQ

LSCLTTS**R**SMTDKLA

KNYVIDK**R**ESTRKAY

VWGGIWP**R**HRFSLQV

GRREAGS**R**AHPLLRL

YIQLEVP**R**EATISYI

DSVSEAT**R**DQDGVLR

NQLAEKV**R**LRLRYEE

ILKPMML**R**RLKDDVE

VSPFPKK**R**GRPKRQM

ATDQEKS**R**KLHELTV

MVVLSLP**R**IQSQRYQ

XXXXXXM**R**GAMELEP

FREEEQL**R**QEREEQQ

QNLPVQD**R**NLMLPDG

DEFTTLA**R**DLYAFYP

DMAAHEE**R**VNEVNQF

TAEDIQE**R**RQQVLDR

KMAEESE**R**RAKEADQ

ALHQKTL**R**VDVCTTD

LAAERHA**R**LNDSYRL

RLDDEIG**R**KVYMRDR

IKIQAMW**R**RYRAKKY

ESLHQNS**R**ATLLYGK

QSKALTE**R**LKLNSKR

PSLEGDL**R**GPDVKLE

TFYVTKS**R**DALTETA

AVSSADP**R**APGESPC

PISDGGS**R**VIGYHLE

PHQIGGI**R**FLYDNLV

GNMELEK**R**RQALMEQ

EEQQESA**R**AVPEKST

PDLEKVV**R**YLAGCGL

EMEPLVP**R**VKEVKSA

YNINTYA**R**FCKCCRE

EKIDLEV**R**SLETCMY

LAKMFLN**R**INYWHLE

TILDMVE**R**ECGTIIE

KLAYYEI**R**YQFYKRD

VPTVTWL**R**KGQVLST

DMNKVKL**R**AIERSPG

LLDSPYA**R**VHIAEDQ

GKMPLEE**R**IQKIKEI

KLKPVEL**R**ELLNPVV

LQVVACT**R**GGMSRLF

YLQHLCY**R**NDKVKTD

DEVEVYS**R**ANEQEPC

EDFSQLL**R**NFGQWLQ

LLFRGAN**R**DVRNYNS

TKSEKLV**R**LHQEYQR

GPLASHM**R**WFIWLMA

QSVQETV**R**VRKVDVS

SLPCLSL**R**KLPPRSE

KEVEVAK**R**NLAQQKI

LSPGNSA**R**LTPLRYD

TVGDSSA**R**PARRVLL

QEEITEY**R**RQLQART

MPWKTCW**R**YLTSGGF

IVVEEKV**R**VPEEPRV

PSFSRQL**R**DVQETVG

KKFGGPG**R**MKQSCVL

TSLQKLF**R**EVRIMKI

PPVETIQ**R**TTARDPI

SQLLTSV**R**RMVLTLT

CKLAALV**R**HRVSLFG

TLFSNVQ**R**VQGVSER

EVQYEIF**R**SLMYWMT

IELPDLG**R**FYKIRVW

AAEELLQ**R**CISSALP

ELSHDNI**R**VKWFKND

VPDALND**R**HPDVRKC

NNAENLF**R**AFLGELK

VIAMLGF**R**LDAVKSN

KNMLIHG**R**AMLLVGR

AQAGAKS**R**PKKREGV

DFEYKLD**R**EFLKGCK

LDALVMD**R**VAFVKLL

VSIEAPK**R**EPQPIKE

PIGYQCS**R**VYWSTTD

SANKENF**R**ILVESVK

HQLTEED**R**ERLKEEL

EAAQLSE**R**TAQEQRT

ECIQMKA**R**LTQELQQ

AGPGLLP**R**KPPGLAA

NFPVHLS**R**IRRLRAE

KVDQFVT**R**FLLRETV

HYPFENK**R**EFEDAFP

NNLHSGI**R**KLEDYLT

RNLWQSS**R**PLICLLD

IQSAPNV**R**FQWFKAG

ISSPPAC**R**SPSPVAN

GYPGGKA**R**IIHKESD

DQRLHTT**R**SVSMQDE

APPVTSA**R**FPSWVTF

DKCNDII**R**NKDDTAA

IEVPISG**R**PKPTITW

TEINAKL**R**NIIVLDS

VRQEQHK**R**LEDLRRR

ITDASAL**R**LVDKAQV

SILLDFT**R**GSRGRKN

FHHCEQC**R**QYMDFTS

EYEMPLY**R**KYWPNHI

IRPRPPS**R**SRAAQSP

RAARTLY**R**IELLRKV

VGGSGLP**R**GRRPPRW

MHHDPFL**R**SMAYWIL

LCTVATL**R**ETYGEMA

CAKQEPE**R**NECFLQH

DDNPNLP**R**LVRPEVD

KYLYEIA**R**RHPYFYA

ELLFFAK**R**YKAAFTE

KASSAKQ**R**LKCASLQ

FKAWAVA**R**LSQRFPK

LLECADD**R**ADLAKYI

MFLYEYA**R**RHPDYSV

YSVVLLL**R**LAKTYET

FQNALLV**R**YTKKVPQ

PTLVEVS**R**NLGKVGS

CKHPEAK**R**MPCAEDY

CTESLVN**R**RPCFSAL

PIPQCPP**R**LDVQTIQ

WNPIEGE**R**YLSFLRF

QFLEQAQ**R**ELAELLD

LNFTMND**R**HTGPAWN

CDCELQH**R**VEAIVAF

FCQEMKF**R**NKMVEYL

QEESQAA**R**IIRNTTA

LVIEEQL**R**RAAYLNM

DDIKCLT**R**DLDQASM

DMKADVT**R**LPATIAR

YLHLLTL**R**GSVLRVE

MKCSSMQ**R**FGERAFK

GKNIYVE**R**QRPALGE

SKPEGTP**R**GPQVYTM

KLTGMAF**R**VPTPNVS

SLGKVGT**R**CCTKPES

GPLSLAS**R**NRSQDSG

LKNVNNM**R**IFGEAAE

CQKGFVL**R**GSGVIHC

VVVKLLI**R**RPECFGP

CAKQEPE**R**NECFLQH

TCKKFDL**R**VTIKPAP

KATLKNL**R**KQYRNVK

LVEVMMA**R**RDDLSFC

TENAIGS**R**KFLSLGL

YHKDLQT**R**ATFMEVL

ANIHKEQ**R**KFCLAAE

CSGSLVE**R**RPCFSAL

YDDKAIE**R**LLDRNQD

EHMPNDS**R**LKCLLDP

VLVFEQV**R**LLMKSYS

TEILKWL**R**EILICRN

SFEVDPT**R**LEPSESL

NMGQWKI**R**AFYEHAP

FSKKCSG**R**WKTMSAK

EIMEACM**R**DIPTCKW

VNGILVS**R**GVKDMEL

TLGSTVG**R**LQEQLGP

GLAESGA**R**YTEMPHV

PLPPLLA**R**VGGNIEV

TYPGLSK**R**RLKPEAI

STLQTIY**R**LSKGRSL

YPVGTVL**R**YICRPGY

GGITGNM**R**QEAIDRF

EKETDWL**R**NEMNKDL

ANQDFTN**R**INKLKNS

KYLPGLL**R**SGFYDLL

GVVWVPF**R**GADYSLR

RIPPVAV**R**LQMSERN

LFLHGYQ**R**FWIETEA

LAFDVGL**R**EHATGYL

VCALGNS**R**VASALCS

ENAQTGG**R**KRGMSRR

VSLHLPS**R**SSPTVFR

LHKTIVK**R**RMSHVSG

YELLLKV**R**PEQLVKH

PMKYSCQ**R**RARLITQ

VLVNSLH**R**IITNSAL

DTTVKKT**R**RRGQKAE

EKADMED**R**EKEDKIK

SCSRSVS**R**EINLKDY

WLISKLD**R**LESSSGI

PAVNGTL**R**TSSYEEH

WVQAVVS**R**FDEQLPI

VPLVVFK**R**EKEIARK

DLSWSML**R**GCKEICC

KGDPKKP**R**GKMSSYT

NPDGSPA**R**RVPVVTQ

IPITSST**R**NIRLYPD

MLQQLLR**R**LVFDVPQ

IQQQLEQ**R**ISASGAE

TSFSNYY**R**TWNVVVH

EAGGAGT**R**ARSLWRV

GQPKDTM**R**LDETMLV

DDIAILA**R**YMLMLSF

FQNAILV**R**YTQKAPQ

AHTVVTS**R**VVNRADT

GDLYGAE**R**IAELLGL

QTLWLWA**R**PRAGGAW