**Supporting Information S2.** The benchmark dataset used to train and test the model for predicting the possibility of carbonylation at Pro site. It contains 126 positive samples and 792 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 126 peptide samples in the positive subset

GKWELYI**P**PKQNKSV

PLPRKRP**P**VTQAAGA

QRPARRG**P**PPLARAV

ASKLQVL**P**QKASERL

GYKVLIQ**P**SNRRAIH

KVSAAPG**P**ADQKTET

PASKDTR**P**KEKRLVN

GRKSTEF**P**RKIREQE

KQRAKKK**P**NPVEEEP

YYYWKKT**P**EAASSRA

QRNRETA**P**RTIFQRV

KSLQAKF**P**SNLKVSI

NSRTMLL**P**LLRRFPE

EPQKQVQ**P**QVQPQAH

IAKQSLR**P**FCTVCNR

ERKQNLR**P**MPKKYHS

SPVQKAN**P**GTLAAEI

PTLPKRV**P**RAGEARN

VVRLSCL**P**AFKDLIA

KERKMAV**P**MPSKRRS

LKKAGRI**P**EQILGKV

ASLKRLK**P**HVPLGRN

STKEKAG**P**KGSKVSE

VGIPDKI**P**FKRPCTY

YQDLDGL**P**CIVILTG

KTVMGAA**P**ELKVRLE

AGRLFLI**P**ATIQDKL

AKARVTK**P**KTAKPKK

PNSRFSA**P**SCGSSEK

VFKLLEC**P**HLNVRKA

CSGKKVE**P**SALACRS

KEKQPTM**P**ILKNEIK

RLKNEKE**P**MVLKLKD

STIPLSH**P**KIPRCQE

VAKAVEC**P**PPRTTAR

GSETPDG**P**LSPGKME

KQATDVK**P**KAAKEKT

KKEEKEK**P**KKEVAKK

EVKKETP**P**KEVKKEV

KKPAALK**P**KVPKKEE

KESLKVG**P**LDSVTYL

LFNRIGD**P**SVDEDAQ

LGGKLKG**P**DVKLPDM

KIPKMKM**P**KFSMPSL

VDLHLKG**P**KVKGDAD

NMPKISM**P**EIDLNLK

KKSKIKM**P**KFNFSKP

PKGRPDR**P**RDLELTD

DSKKEEA**P**KKEAPKP

KEEKAKK**P**EEKPKTE

ADGRGML**P**KRAKAPG

VVRPEEI**P**PVPENRF

RKRVSLE**P**HQGPGTP

KKATSCF**P**RPMTPRD

FPENKKQ**P**YYYPPFD

STLERVL**P**RVRFTRH

VVERRRR**P**RDSPTLS

YKPKRGR**P**KSKEMPQ

IPREKKP**P**RPPKKKY

KIHRNRY**P**VCTKPYA

ERSPGGR**P**IHKRKRQ

KGSDEKK**P**FKEKGKT

RALGVQT**P**SIRKSLV

ELIEELH**P**IIKEALE

NPKEEKK**P**KEKKKKG

MDPSAPQ**P**RAETSGK

ILDLCAA**P**GGKTTHI

KHTFSVL**P**GDCQQRL

RVPRRAS**P**LRTSRSR

KKKGRPD**P**YAYIPLN

LLKTHFL**P**LMEKLKK

KIKKAAS**P**SPQSVRR

KKAKSPT**P**SPSPPRN

KKLLVLN**P**IKRGSLE

RRRAPPP**P**KRAPSTT

LPGPEKL**P**GSLRKGI

PDMLKNA**P**MLKSMKN

LMEKEVE**P**EGSKRTD

KLTRCLK**P**FENSRLR

SYRTVEK**P**PKFTEKG

EEIRSHR**P**TLDALRE

KVSRTVS**P**KSFAKAF

KSTLAID**P**ATSKEIP

KEPFKAA**P**HTIHPPC

KVPIKQV**P**GGVKQLE

PAKVPEA**P**KKIVPEK

PEEKVPV**P**IPKKLKP

PEVKPKV**P**VPAPVPE

RPIKPPG**P**PINPKLK

PFVVPDA**P**KAPEVTT

TQEGRQQ**P**VPAPQQK

ARKALCD**P**LEEVREA

LRLLKPQ**P**NIIPTVK

KKPLKKK**P**TPVLLPQ

VIRSLKR**P**PPADAPT

PRKRTHA**P**SALSPPR

KQDKLKS**P**FKFSDSA

GNGIISS**P**LTGKRQD

IREATAA**P**HGKRKRK

LLVNEQE**P**CKFLLDA

KTLGIKL**P**FLVMIIK

ESKKNEL**P**DGLNKKR

LKLYRLL**P**MRAAKRP

ASQRCRD**P**PSNPVAA

RRAHTAN**P**LHRCRCG

AVETEPE**P**ELRQQTE

VYDAYLA**P**NNLKPVV

GLLYCDL**P**EPRKPLE

CLDNTRK**P**VDEYKDC

EFQLFSS**P**HGKDLLF

FFSEGCA**P**GSKKDSS

QNTGGKN**P**DPWAKNL

KTPKYFK**P**GMPFDLM

RNGQWSE**P**PKCLHPC

EFLRACT**P**GASLRKG

SHFRHGI**P**FFVKVRL

ENVGTPC**P**VSAFPSS

LSHKDDS**P**DLPKLKP

DKGACLL**P**KIETMRE

ARLSQKF**P**KAEFVEV

KVPQVST**P**TLVEVSR

GTRCCTK**P**ESERMPC

CVLHEKT**P**VSEKVTK

ESLVNRR**P**CFSALTP

VELLKHK**P**KATEEQL

KRVIISA**P**SADAPMF

**(2)** List of the 792 peptide samples in the negative subset

LCGMLSL**P**WIYSHSD

SGGSYVY**P**PVQEFPL

YIVPDYM**P**VVYDKLP

TTAGATQ**P**AAPTPTA

QMANNFT**P**PSATPQG

KGSETYS**P**NTAYGVD

STNMGIL**P**SVTMYNF

GKSTDFA**P**IKEDFGQ

HSWQEQA**P**NQPPGPT

REKLKAK**P**LLLQLPI

SSFIKRQ**P**LVIARLA

VLGEYYT**P**ALKCEVN

LRRVAGT**P**CGVFGQR

ALEHEYE**P**YSRLKSD

VIINQLQ**P**FAEICND

KSDDPVV**P**SLFKQKT

SDFDFSA**P**KINLEIE

VGGTIKS**P**KNNEKNF

TGRQIFQ**P**LHALRNA

YLSTEDV**P**LARMLVP

PEKEAKS**P**VKEEAKS

SHGEKEL**P**QWLREDE

SPTASQA**P**NQPQAPT

LEKQFVK**P**DGENRAR

RFMEIES**P**HINENYI

RSLWRLE**P**LRISWSG

EEYIALY**P**YSSVEPG

DALKSRE**P**QAKPQLD

AIIALFT**P**TTDPEAE

SSALGEK**P**ITFYRQA

PIPKKLK**P**PPPKVPE

SLSSAGR**P**GPSEGGD

RLCNDSL**P**PGTVKLX

AELEIAN**P**PELQKHL

NIVGIGK**P**SKVSECY

RDIPAML**P**AARLPTT

VMTLKWT**P**NQLMNGS

KPSKSTE**P**ILIKDPI

AELLANM**P**DPTQDEV

IKLNYAL**P**PPLHQTE

VQGLGVM**P**PKAGQTI

ISQNESR**P**HFLIELF

PLVDHTS**P**SSDSLLF

SSTVNGV**P**SRSPRLV

TPEDQDL**P**PCPEDIA

RCISSAL**P**RRRPPVS

SSISDHT**P**WLMPPTI

SLSHSAC**P**TPNPLSR

MVVCWGH**P**DSDGGSE

SESSFGK**P**XXXXXXX

RQFDNLK**P**SGGYITG

TVLDAND**P**PIFTLNI

SWSRISN**P**SAFSIVP

SSIALGD**P**HIPTSPE

FNQSLKS**P**SRLGCPG

PPLLASM**P**AQLPPRD

IAQDPKQ**P**PDPPVDV

ECHISGY**P**SPTVTWY

KASVVTL**P**VYLNFTR

NTSSDCR**P**SEESELL

WVKENKV**P**CLECNYK

SWLKDGL**P**LKESEFV

EEVDCMD**P**TCSGRGV

AESPEVL**P**HIEKELS

SCRSEGT**P**AWYMHGE

VAFQLHL**P**LSQRACS

DRFCLSS**P**TEALKMG

QVRVVAQ**P**SLPAVPQ

WIRDALR**P**PLQNINS

HKESFLA**P**VFTKDEQ

AAVQKAI**P**MYKIATK

WGMESKI**P**VEDNKRF

AQNPVDA**P**GRPEVTD

SEDSVQV**P**RNLVGKV

YPDIGWN**P**CGGERYL

IGSKAGV**P**SKSSGSA

SLEPHQG**P**GTPESKK

LWGCHSA**P**HGPGLIY

EIGDGFL**P**VCSLGPG

GAETLTF**P**CDRWLAT

LKASVCK**P**FLFLLKK

ESGKIQE**P**FSAMSKS

GTQVSQR**P**GAQALKV

KEQQKDS**P**VFCRFFH

QSRPTGV**P**TPTSVLS

EMFGVWE**P**LLEPLEI

YTEVANI**P**VARPGRR

SRLDLDL**P**GCQAEPP

YLNKKGN**P**KKFAGLA

AIESLVK**P**QAKKKAG

ENIDGIL**P**LHDAVAN

GKLYIEA**P**TFDLQGS

VQIKKME**P**DEQLRND

KRKKKKR**P**EDTAASA

SERLQRA**P**LKSVGPD

DPRQLGV**P**VIARDIE

ELLSTTH**P**ANKASLT

SVTVTTI**P**ASQAMNP

LSAHTIL**P**GSKTREP

ALKISVA**P**GLADQKT

DRGEIID**P**FVEVEII

HKDLIKK**P**TISTAVG

DRETILD**P**NLQATLX

DGAVPAA**P**ASADAAR

VLAERKS**P**EISERIV

SPHPPLQ**P**LTGSAGQ

FLPASTA**P**MQGKRKS

LNNSSSS**P**QRSVDQR

KNALENY**P**NFRSVVD

ERLALAG**P**QLRPEMS

LKSQTQH**P**WEKLLNL

NANQNAS**P**NVPGKRG

SLEMGIL**P**REIRKLV

EATPTAA**P**PTLPPTT

EKHTHHA**P**LSPRTFH

QYRKVLM**P**KLKAKPI

ALSELDV**P**FKVKAGQ

DAEVYNV**P**LDSQSDD

DKLTLKI**P**WKNLYGE

LSPTLNT**P**APVAMPA

TRSLGEE**P**VGGLGSL

EEQVNSL**P**GSITKAG

KVVIGFV**P**LAEIMGY

TYIELMR**P**VSELIRS

PDVEAHG**P**EWNLKMP

YFPNRKF**P**SSSVAFK

EGSRLSL**P**RLIDMSA

PRLTPVR**P**AAASPIV

PQPGLVV**P**VPTVRPL

LLVTAMG**P**PGGGRND

KIEKELK**P**YGSSAIN

ENLELTN**P**QEFGSSW

AVDKRFL**P**DDHGKNA

PSAFSIV**P**RRAAKSS

RLLSLSK**P**ILPQQDN

IKRQCVW**P**FIVMMDD

ANLHPAK**P**KDFSAFI

TQITTES**P**EKTLFSS

SDCCGQK**P**TGPGGPL

ACVTACE**P**PKTCQDG

TMATRTS**P**RLAAQKL

ALTALTT**P**IQTAALQ

KLQSILK**P**MMLRRLK

RKKMNLK**P**IMRMNGN

EKAFEGS**P**ARELDVP

GSSGAGV**P**GGAAAAS

SKQKKVA**P**RPSIPVK

LPGSVLS**P**PPPPPLP

GDSGGLI**P**GKSLVFA

NKEEGAR**P**GTLLGTF

RRTEGQY**P**SICPRLE

HEAIILK**P**EAIEMYI

HPFASFH**P**GLNPLER

PDGLKLY**P**TLVIRGT

PQRPASE**P**HVVPKAV

EIPKKKV**P**EERKPVP

RLFNDSS**P**VVLEESW

GDRSGIF**P**SNYVKPK

QAAIIQI**P**KPASNRT

EKRDYRE**P**FFRKKDR

FTNTSLH**P**QRMKALA

PQFSSLQ**P**PCFPPVQ

ACRCSIF**P**DLSFVTF

LMDYADL**P**YQIEDIF

FIDDINM**P**VINEWGD

GTLVLCL**P**QIKIISA

ISTSEAA**P**YAPPSGL

TPSWLGL**P**NNAERVL

DLEDEST**P**IVKLGDA

EDVVASL**P**LCHAALR

RTHTGER**P**FQCHSCG

GNLHLLS**P**GNSARLT

EAIIDVR**P**ASTRFLP

EMSDSNS**P**PIPYSQK

DSVGVMS**P**PLSRSSV

PRSVSRS**P**IRMSPAR

NMVNGMG**P**MGTEGLF

TNTSVLG**P**MGGGLTL

TQGSEKC**P**QKTTRRD

EESQHQQ**P**DDSNNIA

FAGTEID**P**ENEELML

KVTIKDK**P**AVAPATK

NHDESLL**P**ESLESMM

QHLPSVS**P**SVSDAFL

LNRDLST**P**GLEKDSG

CEELRAA**P**RQGPAPA

PSKEEPS**P**VKAEVAE

NSSRQER**P**PVKPFIL

GTEEKKK**P**SDFKKKV

GGESSAS**P**GEPQRTL

QADRKIL**P**FTSMRHQ

RARTDEV**P**AGGSRSE

RFEMEGL**P**VDESDSC

LLSQLGQ**P**SIFDTQK

FDAEDGE**P**QTCKSMK

AASLAEL**P**LPTPPPA

ISDTTLL**P**DLSQISP

NRFTMEV**P**KKLTEWF

TPESSSL**P**VALQTPT

ITESQSS**P**PPVIDLI

PLLPSSE**P**PPAPPLD

EMVSGFG**P**IYNYKDT

VQVLKYE**P**FLDCALS

LGRASVV**P**LPYERLL

KKTKGDN**P**NSHPEPA

ARNLFSV**P**GAPDKPT

AKVMELL**P**THAFSTL

HAVAWGT**P**GFLTVML

TTSQDEA**P**REELAVQ

SYQEAQL**P**ALPESVP

IVEREMA**P**DFELDAE

ATLEEGN**P**TDEVPST

LLKPEAD**P**RPCERAP

PKKVAKS**P**AKAKAVK

SRRRPRA**P**KEKAQPL

IPMLYVV**P**RPGKAAF

PMNPIPA**P**LPPDIPS

GNDSDEE**P**PAAAPTP

DKVVIQD**P**YRLPGPP

FAKSDNQ**P**STEKAME

IAKEKRT**P**APEPEPC

SAKITSL**P**PPPPTLF

LHLQRIN**P**TTVKMKS

LDIKAKA**P**KVKMPDV

ESGDAPR**P**PPAATPP

VDLHRFH**P**EPYGLED

RRMQQAA**P**TDPDLPP

KAVSSAD**P**RAPGESP

TIELEWE**P**PAFNGGG

XXMNTSI**P**YQQNPYN

GALMTLG**P**SVVRYHL

VTFKANR**P**FLVFIRE

VKVAQRS**P**VDSGTIL

PMKPLLL**P**QPEVLSP

NLVLKRV**P**SPVLIKK

MDDLISS**P**GKNKSGR

PRESACL**P**EKLKEKE

RAWGCAG**P**CGRAVFL

NWRQPEL**P**EVIAMLG

MDRDCRM**P**MGLSTGI

TVENASK**P**DFTKNSQ

RTDLYFM**P**LAGSKLA

DRGLQLV**P**VWDIILS

NALNSCD**P**WVQAKCY

TNMGIII**P**DFARSGV

SKTIYTA**P**LDMLQVT

RTPPRRM**P**PPPRHRR

XXXXXMH**P**PKTPSGA

YYRQSEV**P**DSVYQHL

EKDDSVR**P**NMTLKAM

ECLGAAV**P**ARLRKVA

RSLQKDF**P**IQILSID

VDSDVAK**P**NQACIST

SLKKAKK**P**PLPPSRT

EKFRIEL**P**DLGRFYK

AARISTS**P**IRSVRSP

AFPDGHL**P**EEALKVS

ETYTEED**P**EGAMSVV

ALDDAEE**P**ESPPPPP

ECTTIEH**P**FCMYDAD

GAAGSMS**P**SRVEANV

EDVQVLI**P**FKGRPPP

EKVSEVL**P**IQKNDEE

LKFSSMA**P**DLDRLNE

HPQRPPS**P**HPPPHPS

NRLCFLE**P**TSNAQNV

TDEILLN**P**NNKWFKP

GKLSNWE**P**KDDAMSE

GSGTDME**P**SLYRCSN

FKSDGII**P**FAHAIWH

DMLKRVE**P**LRNELQK

FGYIKII**P**FYILNAT

GESKQHI**P**EKKNEVH

ATLSVLE**P**ATIVEKP

LQGGNKS**P**VTLTAYI

DIIQVTV**P**DPSPTSE

ALGNQST**P**APPTGEV

FTNFMRS**P**ACDIFNP

PDVSMVD**P**EALAIEQ

TSLFSNK**P**FLKLGAV

LLCGSDL**P**LHKMAIQ

CTVLTLE**P**NSQVQQR

YNGKLTG**P**AAVELKR

GDASIAA**P**FTSKLSS

IRDWLVK**P**IRDQHVK

SQVFIDH**P**VHLKFNV

AEDVSGM**P**ALCSPIS

NHARLRT**P**PPPLSHA

LSHTLGD**P**IKIRAWQ

FQELSQT**P**GHTEELA

LGRPPSA**P**ADKDGSK

GFLSYLG**P**FNQIFRN

ESQGEVQ**P**NFSTSSE

LMLTAGL**P**ELTSVKD

RPGPHSV**P**SQAPRLE

ILDEFYG**P**EKSLQPI

KIINFKT**P**EDARWIL

DREDQEA**P**DSDAEVD

KQNGIPH**P**RGSTAIF

AGQVAVL**P**EVQVTQN

LAELLRG**P**CGSFDVR

KAKKRRL**P**RGFPPSA

LQRAFPA**P**AACQCHC

VTRQLYE**P**LVMQLIH

EWFLRTF**P**DPTSWYS

NPKAITA**P**QMFGRLD

KLNQHDS**P**RIKALEK

MMLCSEC**P**TVFVDAE

ILKKYGI**P**FSRITQE

EKTAAPS**P**SLLYKST

AECWRSR**P**EERPTFE

TLTYTAK**P**VGGIGGG

VTSSQVQ**P**VTIQQQV

CQTSTVF**P**LKKKVSI

RTVKDST**P**SSLDSDP

QVTDAND**P**PAFHPQS

LADTDLS**P**MDGISTC

EADASGL**P**HKLGEED

KKIIKEQ**P**QPSGKQE

DCFSEEA**P**DATKHKL

SGQRSAS**P**SVPGPTK

KTVKLKS**P**VLSNTTT

QKSLSLS**P**GKXXXXX

FVASGMG**P**SASSHGS

LGCTAAT**P**PSKDPRQ

QWIRNLV**P**EFGVSSS

MRKFLSD**P**QVHTVLV

DRSLESN**P**EQLQAMR

VASSEAK**P**AATIRIV

DLQVRAC**P**FLPVKRK

ESVHGTA**P**DIAGKDM

XXXXXMV**P**GEENQLV

FLRMSKV**P**TDLAVEE

DFLKRRL**P**KLSKSTA

QECKTKH**P**HHLSLKE

VLKFSCK**P**GFTIVGP

RLEVELY**P**SEVVEIQ

TTKSSAV**P**PGLPVYL

QLYDDSF**P**MEIRQYL

KKDTKKY**P**ESTDTES

SHGDVIR**P**LRKQVEL

ESGLMGE**P**QPQGPPS

GHETEEF**P**DEEQLRR

GDSRGPS**P**SYLGGPR

ESRYQTL**P**GRGLSGS

DNANSRL**P**EDTTSVL

KAAYLSD**P**RAPPCEY

IGFLDNW**P**LLEQWFS

LFQKKID**P**VTMDPEK

SVAQTGG**P**PEADGLV

EMEVCKL**P**RLSLNGV

SVQSSHF**P**SGSYSVR

ITEVWGI**P**SPIDTVF

SAQDMHV**P**VPKQLAH

SKDIPLS**P**PAQKNDP

YRGVKVM**P**FSTACNT

YQEKLYR**P**PPVLDAL

PAGPTES**P**ASKGVTA

WVRCNKM**P**VKDTTYR

EKECAPT**P**APVTRAK

RHAVKGD**P**QSPGRHW

TASLLAI**P**HTPESSS

VEAFFYE**P**YQGAIQG

FMAVKNW**P**WMRLFFK

ENYVQET**P**LVLSRCS

MTISWHE**P**LSDGGSP

SAEKNQR**P**RKQTSAP

FAASSTV**P**HCEQSCR

VATPQVS**P**NTVKRAG

EEDLLDD**P**SLEGMCG

ASYESES**P**ENQTEKE

WQMWTLL**P**GFCTRPT

RQLDRLL**P**KLARGSL

GKKVRKM**P**VSYLGSK

ICHEISE**P**RLQNYDE

IVDVLDV**P**GPVGTPF

KKSRFWE**P**NKVSSNS

QRFCIGH**P**TQEMTFG

RARSLKF**P**GLISGCT

LIKEQLE**P**PEIDMKN

HSDSGYE**P**SFGKSSR

GLAWARA**P**QPKAWVS

ASSVETK**P**GASKVAT

LIRICSK**P**VVLPKGP

AKKEDKT**P**IKKEEKP

GKASIKD**P**HFLNFNK

MTMMSHI**P**SVMKAHG

ISELTGT**P**FDVESDS

DYALLTL**P**AEFSSQL

GATAPPA**P**AEPTPPP

EIYGKWS**P**LILNPSY

EGKSLRF**P**LALEEKQ

NRFGVSK**P**LESAPII

PQLSVCN**P**HSGKEYF

SNRKGET**P**LKVANSP

YGIWKLI**P**NGQYEFR

GGAEEER**P**GTPELAP

APLNVVP**P**EVPSEEL

ELVQAMF**P**KLNNQER

GTELSLN**P**TCAVFIT

KVFEEAK**P**NSELCCK

VTSAISL**P**NIRKPDG

CDEDHQT**P**RDGETSH

VNYCECN**P**CFNGGSC

AFYQQTL**P**NSHLTEE

VYIESRR**P**NTPYFIC

YLAQQIH**P**VVARICE

VEDKSDP**P**EGSEEAA

XXXMTAE**P**MSESKLN

KSVHLQK**P**KPDKKKF

ADSVSLL**P**AAPEGSR

GCEGSSK**P**FFYNRQD

RRREQRD**P**TVHDDVL

LDGEEEE**P**EEETTLG

SGIRAGY**P**LSERQQV

MKRESQN**P**DQKDETN

NKVGASD**P**SDSSDPQ

MKVVTLK**P**TVLEHTV

ITPKSDV**P**IQAPHFK

ENLKIDR**P**EDAGEKE

ISDHSET**P**NMELSCR

SCSENQE**P**GYCTVSN

SFIKQNN**P**KFSAVQD

PEDDGGS**P**ITNYVIE

TLVSGSL**P**FDGQNLK

LARYSYN**P**FEGPNEN

PASTRFL**P**QGTRIAA

LHEIYFL**P**DHPELKK

LNDRMTF**P**EELDMST

GRIFVFE**P**PPGVKAN

QSKLYHL**P**PPTVGPH

IKNSNGI**P**KLKYIPL

FLEVDEY**P**EHIKNLV

QETAEGI**P**PGSQDSS

XXXXMSE**P**GGGGGED

VYCLALV**P**ANTLPKT

RCFLEAD**P**YIDIDQN

ASLEVEV**P**AKIHLPK

KKVPAPV**P**KKEKVPP

QVFKPFL**P**SIIALCM

YSVYGTE**P**YIRLGVL

MSPARMS**P**ARMSPGR

PVPEEKE**P**ASSPWAS

KKTFVGE**P**QRLGSET

KDMVKMF**P**LLVEKLR

LGAMSAA**P**SQPNSQI

TDRLVIT**P**LTDRCYI

SLRQNET**P**QPAAQRS

LSKDLLK**P**IQDVNSL

TASPSSR**P**VASPGAI

KSQVRDY**P**KHNGQIS

GKDVLRL**P**PSSITTD

GLETYLG**P**LQVAYRE

REYQEYN**P**YEVLNLD

RVELLHN**P**AFCSLAT

SVVSITG**P**LIRILGD

PEQNQDM**P**PHLPTAS

RLTPIGD**P**TMVVEWL

TCRDEYA**P**PKAELDA

GDEDIKI**P**ENPLEPL

GEIRSVR**P**LDREKVS

LSPPPTE**P**TDGEQAG

KKELKVG**P**YANTTRY

GKDELSK**P**SSDAESR

ATLVDQS**P**ESLKRKS

KLKVLDK**P**GPPASVK

CTDSMGV**P**RALGETW

ANGDSDG**P**SYMSSLS

LGRVFIH**P**ESPSTGH

SLKTAHS**P**NVFLQQG

DKIQQMN**P**PEFEMIE

ALLKDQQ**P**GTFLLRF

AVGTSAA**P**VPSDNHX

EVKISHF**P**AADLGFS

CVDGCYC**P**NGLIFED

EGVLLTL**P**RQSGGSG

ESMENSC**P**QVTEVTA

SSSGAVD**P**SVQRSFT

XXMMAQF**P**TAMNGGP

EEVRVRS**P**TRSPSVK

CLEENER**P**LEDQLEN

SSSSGIS**P**DNRDFYQ

ATSTSSE**P**LSSNQPA

TTCNCKK**P**DDDTRKG

MLTGESV**P**QMKEPIE

VPGGSVK**P**KIMSPEK

ILTTLQE**P**HDKHLLD

TNVEDVC**P**NIPSLEK

RLEYLEG**P**TVTSSYR

QSPTKAK**P**KVEDEAP

PPPSTQA**P**SVNGVCT

HGDACFL**P**EECPCTW

EVSIGSA**P**LAKQQSY

NILVELD**P**DGCPWLV

DDTERRS**P**TPERTRP

GSFGYKK**P**PPATGTA

XXXXXMA**P**VHGDDSL

APQQEEL**P**LSSDMVE

KKLRSDE**P**LIDPKPV

AFTRTVS**P**GSVSPIH

QLSELAD**P**LKDELNL

TPNPLER**P**IKMGWLK

FKELFQT**P**GHTEELV

EASKEYS**P**CVAPCGR

KSFATGR**P**KKTKKRS

HPYVTHI**P**SPTLPGS

DAWLQTD**P**EIQPSPF

SEIKISN**P**TEFQNHE

HNIENAS**P**ATVSRMG

SLESSSF**P**LPKQYLD

TGQLGVK**P**QTGNSIP

CSPLMAD**P**LLAGDAF

FVYKDIK**P**GRKRKKH

TSANEVS**P**VSSSGVT

FQRSRSE**P**SPDAPES

ADSDGWT**P**LHCAASC

VEENNAS**P**HFEPDLH

NQIGCLT**P**ELSELKK

ETQAEDH**P**PRLYGCS

FKISFGT**P**APGFSSM

GVCTDGE**P**SQPLLGQ

LSPDSHY**P**LEEEKTD

SPISQRL**P**PKVESLE

AAEFISK**P**QNLEILE

NKCLIKN**P**AERADLK

ELVAAVW**P**YRRLALL

RQTVAQM**P**PQLVELP

NPALTLW**P**MFLQGKD

QDIAKDF**P**RGEESLE

VTLKWAK**P**EYTGGFK

QHVFSAG**P**VSDEVLP

LYPPGRS**P**LHHAQQL

HLSSKWL**P**SSPSGVA

QLAPDPG**P**AGHTLFQ

GLCAALE**P**ILSAKTK

PADVTGL**P**MPKIEWS

QKFRFNH**P**AEAAVLR

SAMEMNE**P**QAPAPSL

FHKVKAL**P**LVSWSRR

AFKRAND**P**VNYLESK

VRATRAS**P**DPRAYVP

KVIKMEE**P**LPAKVTE

QQVTAEI**P**VDLNTRE

GKVIGEF**P**GVVHCLD

MIKKQDQ**P**TFDNSGN

SASDGYC**P**REHMLPC

LKHSPED**P**EKYSCFA

LIKKHLL**P**LWNDGCI

SSSLGKK**P**SLTSESS

GILKRRY**P**NSLPALI

SKRQLQT**P**KEKAQAL

VYVDYPG**P**ASLTQIY

AEEDLTD**P**VRCNSLE

NSAGRSA**P**RESRPVI

FQEFLKR**P**DHKQDFV

QAESQVE**P**SYSEQAD

LTAKQKL**P**SYILENN

RSTEGIV**P**QYDGSSS

LSLSLAQ**P**LRSERSG

EMKEIER**P**FETYKEN

NATSAFG**P**NLRYIVK

VDQKIEH**P**LQPQPGL

HRRTSAI**P**RAFTRER

GDFKVTT**P**DTSTTQH

VISRTGV**P**QPTQAQS

RGKKAKN**P**KEEKKPK

YTLKKDV**P**DGVKELS

GNNDVKQ**P**EGTGMID

VESEIKV**P**DVELKSA

KLGVQVR**P**ENWCSQM

SWLKDDE**P**LYIGNRM

PAYDGGS**P**IIGYLVE

LAKILQN**P**ITHSLQV

KLSSQAL**P**SFGYIKI

PPARQLG**P**RSPRVGR

SKVWYNC**P**XXXXXXX

GKDFERK**P**LLEMNDP

GQAVTLL**P**FFTSLTG

KLAHGQE**P**WTTDAKI

PGGKEEA**P**EPPDGGD

QMHSTSD**P**SHRSSSP

LKNKGIQ**P**LLDAVTM

SNIDIRI**P**TGQLTMI

CDEHFET**P**SGSYWDH

IDKNTKI**P**CKSPPPE

NAWERRA**P**LAPKHIK

TDWEKTK**P**VTGNLRP

CYHFGLS**P**DLPICKE

EALVIRE**P**ITVPERP

GDMGDTT**P**AEPPTPE

FVRPDLG**P**KMYNAYG

DQKYRQH**P**SNFQFKK

QMSFSQS**P**FLSIAKS

FSALRAL**P**DMEVVGL

TSPEGVA**P**TSGHDRR

GSHVNQG**P**LHLGGIY

GRITDVI**P**SEAINEL

VAEYEVT**P**DDKRKAC

LVRDPLS**P**AVRQKET

GNICVYC**P**GGPDSDF

RLLERYI**P**KHQKCLT

PLTTALN**P**PVTATEE

FLTYHMW**P**LTFLSPS

GHGIMEE**P**ELTLIST

KLDSEAW**P**GVLDSER

LRISLTH**P**TIPFNLM

ENSIEGT**P**SKCYCRL

NESKKTA**P**RQEAIPD

INPSTSL**P**LIGSPPV

QHRTFIK**P**KELEGLP

RVRVVGK**P**DPECEWY

LQALDLL**P**LLIQTVE

WAQQIHS**P**LTCEQLT

ITVITLG**P**PSKPKGP

AWDSGQG**P**RLPSSVA

STIGLNQ**P**STPTHAA

GEYDVTM**P**KVESEIK

FKTSIFS**P**MKKEVKT

DEREFSG**P**STPTGTL

GDEGADE**P**RGAGRKA

STSKNTS**P**MRNSFAH

LRPHVFM**P**EVTPDMD

VDLSLGS**P**KLKGDIK

LLVGAKS**P**NLPEHIL

ILDGLYG**P**IAFKDFI

ARHRCQT**P**HLLAAXX

AAAARTE**P**YCAICTL

LFHILVT**P**LFLLSTV

AQILILE**P**PYFVKQL

PPPGQLD**P**EDQDSEE

CLVKDYF**P**EPVTVSW

TPELSAE**P**TPKALEN

TAALAVA**P**GPRFLVT

GRPAEPQ**P**EPEEPAE

DLRTDAI**P**IRAAKAA

AGIAAIG**P**AKELEAE

EEITDNY**P**QAIVYPF

APPGLQP**P**AEDEARA

LVIPPQP**P**TTGPPRK

SRGTDGQ**P**LLLPYKP

KVDALHC**P**PAAVVTV

TMKMRIL**P**DQFSPLN

QVTTVKV**P**MMKRLGM

KKACMFE**P**LPIKLSK

RQQNTRD**P**DTMVPLT

EEALKVS**P**VSIPAEQ

IWWANKD**P**LPPLDPQ

ELKLSES**P**EPGRSPE

MMGLDLQ**P**DLCSLLI

TRLFVKE**P**AAFLKRL

FDMHKRS**P**LLRAPRF

ALGRTYH**P**KCFVCSL

QESARAV**P**EKSTAPL

VVLCNLK**P**QKMRGVE

DASSSWG**P**QMSASVH

AGVSSGA**P**PGRNSFY

PSASERL**P**DEKVELF

TVSQFTL**P**KSVSDGI

TNSKATT**P**KPQKPTK

KESKVEE**P**RKRETVS

ALSSVGS**P**KAKEALN

EGQTQRQ**P**GSPMYRR

KKPKTVK**P**KKVAKSP

AAVRGSE**P**ISVTWMK

QFPFADT**P**DTYKSEA

LAQADKL**P**MTDKRVA

XXXXXMD**P**PRPALLA

PEVRSTE**P**VAPPEQL

EDEAPPR**P**PLPELYS

EKPVDPP**P**FMFKPVK

IQTDVVL**P**SCKKKAP

ALSLFTV**P**EDCPIGQ

SQPRDGW**P**MMLRIPE

EGGGRSL**P**RSSLEHG

TLSISVQ**P**GEGNKAA

TKSNDTL**P**SSHREIP

ASAEGRE**P**SPQCGGS

SGGEWKG**P**QVSSALN

YTIIQQM**P**RGNHFTI

PGQHLLA**P**EDVISCC

SLCVNEV**P**SFYVPMV

KGVTSIL**P**VLREGVL

EQQAQFT**P**LADPPDI

IPLLLPT**P**EEKKPPP

PPTLSRT**P**EVQSRVP

TDQKDSK**P**PEKATED

EAMGLGT**P**SEAIEIT

IKTLSDV**P**APRRLAE

TLLIELL**P**WALLINE

RIEEGER**P**DRAKGYD

WARVTKD**P**IHPYTKF

KLLVNFD**P**KILEVVR

ESKQATT**P**ASRKSSK

LVKMKFL**P**VDDNFSL

NIETSNF**P**KLLLALF

HLDFLKQ**P**LATQKDL

LNKLSMK**P**EPLFRSV

KDKIHTT**P**DTPEIRQ

DAAHGYS**P**RAIDMSN

KTKEKAQ**P**LEDLAGL

ECQITGT**P**KIRVSWY

FGTLEVM**P**QNESETQ

ESEKSDF**P**PTPVAEV

FVNKADF**P**KVRAKEQ

SGQLVVI**P**SDHLPHF

ATGRPWL**P**LHTLSVS

ADLGVSG**P**KVDIDVP

WKQVHNY**P**MFNLLMD

PPCLQQN**P**GTMQGVY

GSAGPVH**P**LVDPLTA

AHLFSWI**P**LSASITP

LKWNFTT**P**RDEYIEQ

VKKTDSQ**P**TSVMLLD

EGCANRL**P**NGHVNFE

RGGKKFL**P**VLKEILD

STCSSTF**P**EELSPPS

RSKLEDD**P**LYTSYSS

SVVSQRF**P**QNSIGAV

SLLALGS**P**MYSIITP

ASKGEMS**P**MVVETLK

HCLSFTI**P**RLPSSEE

IDKTCLS**P**TPTLEQH

TIFQERD**P**ANIKWGD

NTDGLFF**P**VMSFSAG

KADIGCT**P**GSGKNYA

MTSTSLS**P**QGISTLE

KTVGVWS**P**SPPTCER

PRERSMW**P**LRKCPWK

VDYASKI**P**AQGGTNI

NRLCFLE**P**TSEAKKI

ETERCLQ**P**LQMMALH

AVRMKIR**P**LVGQXXX

TLVVKGD**P**RDNRQPA

LLRRIEH**P**TAGNIEA

HVLLTLK**P**YYAKPYE

DEEVLTD**P**KIQALLL

LISDTKA**P**KRQEMES

SSVTFEY**P**SNAVEDV

TSSWLVT**P**KSLGNVN

KVYDRYI**P**PDLCVCN

CWYSQWT**P**AAPQCKA

SNGYKPA**P**LDLSDVK

LGVAEGS**P**QYKKWYF

IALTKLQ**P**LLNKDSP

DTWVEHW**P**EAEECQD

GYLAATY**P**AVGQTSP

GGEFEMM**P**LGVNKSP

LSAEAAC**P**MFNFSIR

YLCSAVL**P**LLTRCAP

MDVSLHL**P**SRSSPTV

TIEELKE**P**EKNGIST

AIDRFNA**P**GAQQFCF

GGFQVLL**P**LVDQYFT

ILLTDGE**P**TVGETNP

KMLSGFI**P**LKPTVKK

EELTYSV**P**YGREKDV

SLKFKEK**P**TDLETRS

DLGQGLM**P**VLEAWKA

SRLANRA**P**EPPPQQV

RTKKEGI**P**DARQATR

GCKEICC**P**VPDPKSV

FSHIDVS**P**GSAQVKA

IILLILC**P**EIIQDIS

KAHIQFK**P**TLSQQRK

MGNCGHK**P**ESPTEVE

EKASKEA**P**VKGLLQT

LKNLLFN**P**SKPFSRG

RGGKGIS**P**PSFSSYL

SRTQWEK**P**QVKESKR

NSALDWW**P**KIDAVYC

GVQREDV**P**AADLSDQ

YRDRSFS**P**GSYERET

KVNSGAF**P**APIEKSI

AIMRYGM**P**PQDAFTT

KEVKSTA**P**EATVECA

SEGNIDS**P**VSRFMDR

TLGEKAK**P**ALDDLGQ

GAAFSQK**P**QSNALIG

SNQAELL**P**SKSVSLI

LRNDKDK**P**LPPLLAR

EKSLSHS**P**GKXXXXX

GYYCLKV**P**LVVFKRE

FYQEVAN**P**LLSSVTF

ALLESSL**P**DTDPVVS

EKCLAGQ**P**KDTMRLD

GTSKAGN**P**IFYYVAR

YSRIKYD**P**DLFRMAL

TQKTVDG**P**SGKLWRD

EAERSGI**P**IVTSPYQ

TNLKIKH**P**TYCYEMN

YAAFNNR**P**PSTWLTA

ICLAQNS**P**STFHYVL