

Supporting Information

Multi-dimensional deep learning drives efficient discovery of novel neuroprotective peptides from walnut protein isolate

Like Lin¹, Cong Li^{1,*}, Li Zhang¹, Yujiao Zhang¹, Lu Gao¹, Tingting Li¹, Lihua Jin¹, Yehua Shen^{1,*}, Difeng Ren²

¹Key Laboratory of Synthetic and Natural Functional Molecule of Ministry of Education, College of Chemistry and Materials Science, National Demonstration Center for Experimental Chemistry Education, Northwest University, Xi'an, Shaanxi 710127, China

²Beijing Key Laboratory of Food Processing and Safety in Forestry, Department of Food Science and Engineering, College of Biological Sciences and Biotechnology, Beijing Forestry University, 100083 Beijing, People's Republic of China

*Corresponding author. Tel.: +86 13571961479, E-mail address: licong@nwu.edu.cn (C. Li); *Corresponding author. Tel.: +86 15191818556, E-mail address: yhshen@nwu.edu.cn (Y. Shen).

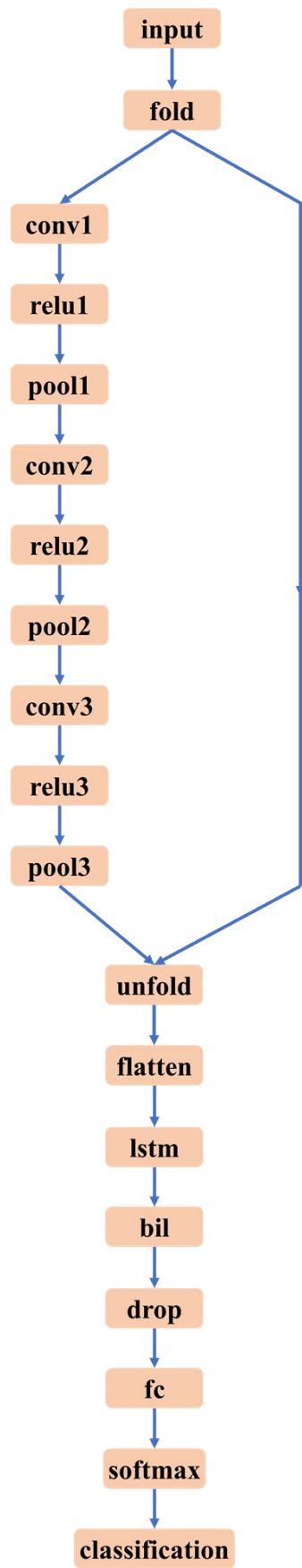


Fig. S1 The detailed framework structure of MiCNN-LSTM.

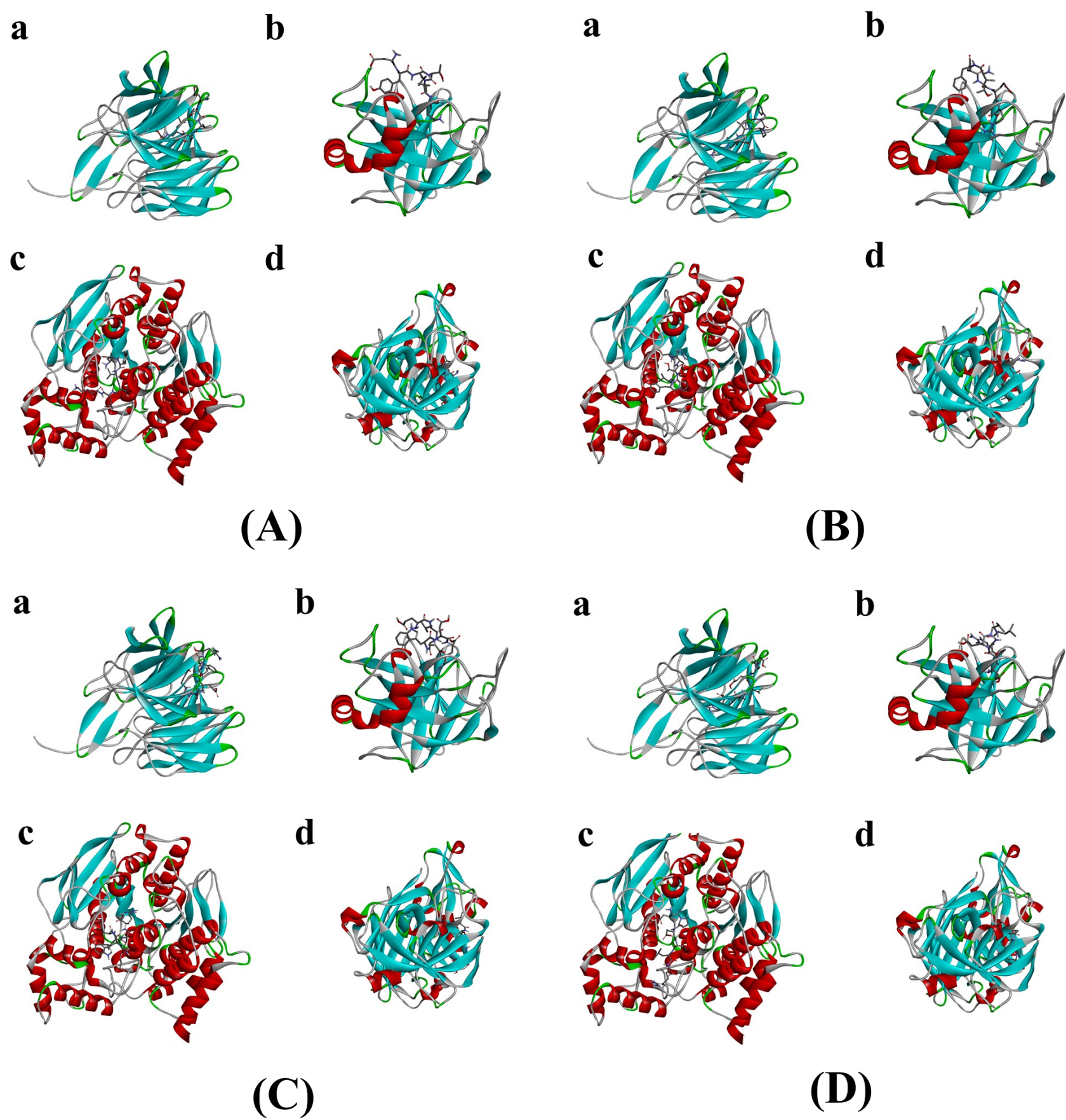


Fig. S2 The 3D maps for the optimal location of EYVTLK (A), VFPTER (B), ELEWER (C) and EPEVLR (D) combined with (a) Keap1, (b) p38 α , (c) AChE and (d) BACE1.

Table S1. Binding sites between peptides and Keap1, p38 α , AChE and BACE1 protein.

Peptide Sequence	Docking proteins	Binding sites			
		Hydrogen bond	π -alkyl	π -cation	π - π
EYVTLK	Keap1	Val467, Asn469, Val608, Val418, Val465, Asp422, Leu457	Val420, Ala466, Val514, Val467	/	/
	p38 α	Gly219, Gly216, Ala190	His57, Phe174	/	Tyr99, Trp215
	AChE	Gly123, Gly118, Ile275, Asp276, Trp84, Trp279	Tyr70, Trp279, Tyr121, Tyr334	His440	Phe330
	BACE1	Thr72, Thr231, Thr232, Thr235, Gly34	Ile110	Asp32	Tyr71

VFPTER	Keap1	Val369, Gly367, Val606, Ile416, Val604, Leu365, Ile559, Leu557, Val608, Thr560, Val467, Ala510	Val420, Val467, Val561	/	/
	p38 α	Cys191, Trp215, Ser195, Gly216, Val227, Asp189, Gln192, Trp228, His440, Ser122, Asn85, Tyr334	Phe174	Asp189	Tyr99, Phe174
	AChE		Phe331, Ile287	Glu199, Trp84	/
	BACE1	Thr232, Phe108, Lys107, Asp32, Ile110	Tyr71, Ile226, Tyr198	Tyr71	/

ELEWER	Keap1	Val467, Val561, Arg326, Asp422, Gly423	Val514, Val467, Val369	/	/
	p38 α	Cys191, Ser195, Gln192, Gly193, Trp99, Gly216, Ser96	/	Tyr99	Phe174, Trp215
	AChE	Tyr121, Tyr334, Asp72, Arg289, Ser286, Gln74, Tyr70, Ile287	/	Phe330, Trp84	Trp279
	BACE1	Thr72, Lys224, Asp288, Thr329, Arg235, Ser328	/	Asp288	Tyr71

EPEVLR	Keap1	Ile416, Ile559, Ile365, Val604, Gly367, Val561, Val418, Val467, Leu557, Val420, Leu557, Val420	Val514, Val561, Cys368, Ala607	/	/
	p38 α	Ser195, His57, Cys58, Tyr59, Tyr99, Ser214, Gly216	/	/	/
	AChE	Ser286, Trp279, Ile282, Asp72, Asn85, Ser81, Tyr121	His440, Phe331, Ile287, Trp279	/	Phe330
	BACE1	Gly230, Gly34, Thr232, Arg235, Thr72, Thr329, Gly13	Phe108	Asp32, Asp228	Tyr71

Table S2 Matlab Software License

License	labels	Uses and options
12345678	MATLAB(Individual)	Academic-Total Headcount

Table S3. Neuroprotective peptide sequences and their activities in the dataset

Sequences	Neurodegenerative activities			
	Anti-oxidation	Anti-inflammation	Anticholinergic loss	Anti-A β aggregation
WY	1	0	1	0
VAPTAVK	1	0	0	0
SQK	1	0	0	0
PPW	1	0	0	0
QQRQQQGL	1	0	0	0
IISVSIR	1	0	0	1
HYVEK	1	0	0	1
GGHRYYA AFFARR	1	1	1	0
VLLALVLLR	1	1	0	0
VYTE	1	0	0	0
LRYL	1	1	0	0
RPNYTDA	1	0	0	0
EMVELPLR	1	0	0	0
VAGTWYS	1	0	1	0
ADNF	0	1	0	0
VSLPEW	1	0	1	0
QCHKP	1	0	0	0
TKGQ	1	0	0	0
RILDWYKK	1	1	0	0
MQIFVKTLTG	1	1	0	0
MMLQK	1	0	0	0
GFLHQQK	0	0	0	1
KLPGF	0	0	1	0
LDDDGRL	1	0	0	0
EVYVVAENQQGKSKA	0	0	0	1
RYKF	1	0	0	0
FH	1	0	0	0
RYNAMNDYT	1	1	0	0
YS	1	0	0	0
RVNQGEMFVVPQYY	1	0	0	1
TW	1	0	0	0
IM	1	0	1	0
WQ	1	0	0	0
GNLDFVQPPR	1	0	0	0
PKPQQFFGLM	0	0	1	0
LPF	0	1	0	0
GNLLGPLGKGLTH	1	1	0	0
HNLDQTESDV	1	0	0	0

IDLGTTY	1	0	0	0
LAGNPDDEF RPQ	1	0	0	0
CVPIVTIFYGVCY	0	0	0	0
TY	1	0	0	0
PSYV	1	0	0	0
IGAF	1	0	1	0
FQRISGAQPSLRIQS	1	0	0	1
IMSP	1	0	1	0
CNTATC	0	0	1	0
NYDNSAGKW	1	0	0	1
VPY	0	1	0	0
RPKPQQFFGLM	0	0	1	0
QWFCT	1	1	0	0
RIIGL	0	0	0	1
EW	1	0	0	0
SSEDIKE	0	1	0	0
GPE	0	1	0	0
MGPAMMRTMPG	0	1	0	0
IVEF	1	0	1	0
LAGNPHQQQQN	1	0	0	0
GR	1	0	1	0
KQLAIAN	1	0	0	0
TINSLTLPILSF	1	0	0	1
FYY	1	1	0	0
YPYFDPL	0	0	0	1
SLAFVDDVLN	0	0	0	1
AH	1	0	0	0
ADIYTEEAGR	1	1	0	0
EKHIMEKIQ	0	1	0	0
EPDPNAFYGLM	0	1	0	0
PETMQQQQQQ	0	1	0	0
KWFCT	1	1	0	0
SGVYKVAYDWQH	0	0	0	1
EQEEEEESTGRMK	1	0	0	0
KLVFFAE	0	0	0	1
VHH	1	0	0	0
FS	1	0	1	0
SW	1	0	0	0
IA	0	1	0	0
VNAVIH	1	0	0	0
KDHCH	1	0	0	0
GLLGPLGLTGGA	1	1	0	0

YVLLPSPK	1	0	1	0
SEAGVTE	1	0	0	0
HIVVNPSPF	1	0	0	0
AYLH	1	0	0	0
GRFREF	1	1	0	0
MEIFVKTLTG	1	1	0	0
LVNPHDHQN	1	0	0	0
MMLDF	0	1	0	0
RYYA AFFARR	0	0	0	1
IL	1	0	1	0
IW	1	0	1	0
WDQWCIQ	1	0	0	1
LPC	1	0	0	0
YLGAK	1	0	0	0
HQKLVFFAEDV	0	0	0	1
VEGNLQVLRPR	1	0	1	0
YLSQGRSPYGGEE	1	0	0	1
TMKLLLVTL	0	1	0	0
WSVWEQELED	1	0	0	0
MW	1	0	1	0
TGENHR	0	0	0	1
VAENQQGKSKA	1	0	0	0
EM	1	0	0	0
PSLGVFCTPNAP	1	1	0	0
MPVGTPNLVPL	1	1	0	0
PGKLVYA	0	0	0	1
AHTVEKLPFFD	0	0	0	1
GPR	0	1	0	0
YG	1	0	0	0
VPP	0	1	0	0
LDAVNR	0	1	0	0
RGAVLH	1	0	0	0
EDR	1	1	0	0
FY	1	0	1	0
MFSGSFR	1	0	0	1
SGGY	1	0	0	0
LK	1	0	1	0
YPWF	1	1	0	0
NLPR	0	0	1	0
NFRVLTQ	0	0	1	0
IQW	1	1	0	0
MDPGQQ	1	0	0	0

AIW	1	0	0	0
VETGLPAGPPV	1	1	0	0
WVSPLAGRT	1	1	0	0
TPCDIFTNSRGKR	0	0	1	0
DS	0	0	0	0
LPFFD	0	0	0	1
IGPW	1	0	0	0
ILY	1	0	1	0
LLGPLGRDGAH	1	1	0	0
RWALNEDQMATEKL	1	1	0	0
LPFFD	0	0	0	1
GRQVIAVAK	1	0	0	0
MKEQ	1	0	1	0
FRFK	1	0	0	0
LPYFD	0	0	0	1
LPF	1	1	0	0
ISYNY	1	1	0	0
HVRETALV	1	0	0	0
STITCHER	0	0	0	1
IATP	1	0	1	0
SMAKEGV	1	0	0	1
SRADIYNPRAGRF	1	0	0	1
SQK	1	0	0	0
FPVGGDDGAPE	1	1	0	0
GGVLVN	0	0	0	0
FVAIGGW	1	0	0	0
EINQ	1	0	1	0
YPF	1	0	0	0
TGAGAIR	1	0	0	1
LCFFD	0	0	0	1
EISEVKMDAEFR	0	1	0	1
APDTKTQ	0	0	1	0
MEHFPGP	0	0	0	1
GR	0	1	0	0
LQQQ	1	0	1	0
AH	0	0	0	1
EL	0	1	0	0
FFGLM	0	0	1	0
PAY	1	0	1	0
RGAVLH	1	1	0	0
EDR	1	0	0	0
YS	1	0	0	0

WDQW	1	0	0	1
PTT	1	0	0	0
RPRTLHHRNR	0	0	0	1
FSEY	1	0	0	0
KE	0	0	1	1
VW	1	0	0	0
YT	1	0	0	0
LLRW	1	0	0	0
SRHEEGIWPF	1	0	0	1
VVLIGDSGVGK	1	0	0	0
GIGTSIGVTK	1	0	0	0
CGHKAKGPRK	0	0	1	1
DAMGO	0	0	0	0
LAEQ	1	0	1	0
GSH	1	0	0	0
GHC	1	0	0	0
KVPIRVF	1	0	0	0
SPGGRTGGY	1	0	0	1
RPKPQQFFGLM	0	0	1	0
LN	0	0	0	1
VYY	1	0	0	0
TWY	1	0	1	0
GTEDELDK	0	1	0	1
CHC	1	0	0	0
VGDGGLFEKKL	0	1	0	0
HVRETALV	1	0	0	0
TWLPLPR	1	0	0	1
YSVH	1	0	0	0
SEGFEB	1	0	0	0
KVPPLLY	1	0	1	0
ECH	1	0	0	0
NHAV	1	0	0	0
LPFD	0	0	0	1
ADAF	1	0	0	0
IRW	1	1	0	0
WS	1	0	0	0
QQMMQYLPRM	1	0	0	1
KE	1	1	0	0
DNYDNSAGKWWVT	1	0	0	1
YT	1	0	0	0
KAEFHRWSSYMVHW	0	0	1	1
K				

SDIKHFPF	0	1	0	0
TMMM	1	0	0	0
NHRMLTRRY	0	0	1	0
QQLQGRQF	1	0	0	1
TSQLLSDQ	1	0	0	0
QEPQESQQ	0	1	0	0
FYDWPK	1	0	1	0
GPR	0	0	0	1
NEGLTW	0	0	0	1
GFLHQB	0	0	0	1
VR	1	0	1	0
PW	1	0	0	0
SPHLHTSSPWER	0	0	0	1
VGW	1	0	0	0
LLPHH	1	0	0	0
DAPAPPSQLEHIRAA	0	1	0	0
APTLW	1	1	0	0
RKWHFLW	0	0	0	1
VSAFLA	1	0	0	0
YW	0	0	0	1
LIVAGD	0	0	0	1
RER	0	0	0	1
IR	1	0	1	0
rpDT	1	0	0	0
SGGY	1	0	0	0
AELQVVDHLGQTV	1	0	0	0
WYPGK	1	0	1	0
QCQQAVQSAV	0	1	0	0
LN	1	0	0	0
LQEQ	1	0	1	0
GVSLQFFL	0	1	0	0
GGW	1	0	0	0
HKQLPFYEED	0	1	0	1
RLCIPYIIPC	0	1	0	0
QEPV	0	1	0	0
TRTGDPFF	1	0	0	0
NAPVSIPQ	0	0	0	1
CR	1	0	0	0
VNPHDHQN	1	0	0	0
TGNYKALHPHNGC	0	0	1	1
YVLLPSPK	1	0	1	0
VQIVYK	0	0	0	0

WIR	0	0	1	0
MWCAAFCSSRGKV	0	0	1	0
QIESW	1	0	0	0
KCHQP	1	0	0	0
CART	1	0	0	1
MWCDAFCSSRGKV	0	0	1	0
DMHDDFFVGLM	1	0	0	1
FDLPVDSEHK	1	0	0	0
YP	0	1	0	0
FR	1	0	1	0
YECG	1	0	0	0
IVQQ	1	0	1	0
IATPE	1	0	1	0
SCNTATC	0	0	1	0
CPQVKPGIKLEC	0	0	1	0
CGHKAKGPRK	0	0	1	1
NQFDHYSKQDRCSDL	0	0	1	1
QAGVQ	1	0	1	0
DADLE	1	0	0	0
ADNP	0	1	0	0
MWCDAFCSSRWKV	0	0	1	0
WTWYGYWVW	0	0	1	0
GSNKGAIIGLM	0	0	1	0
RPKP	0	0	1	0
MASVPTK	1	0	0	0
PRARIY	1	0	0	0
CHQP	1	0	0	0
FRFK	1	0	0	0
HNLDTQTESDV	1	0	1	0
IRW	0	1	0	0
IVD	0	0	0	1
VGIP	1	0	1	0
GHC	1	0	0	0
VPRKLM	1	1	0	0
LVAVSINDVNHL	1	0	0	1
YG	1	0	0	0
QQQQQGGSQSQ	0	1	0	0
AGGRESKAN	0	0	0	1
PPKNW	1	0	0	1
SCNTA	0	0	1	0
NW	1	0	0	0
LF	1	0	1	0

IF	0	0	0	1
DIFAM	1	1	0	0
QQQQQGGSSQSQSQKG	0	1	0	0
QHSQITKV	0	0	0	1
LQ	1	0	1	0
RNRRWCIDQQA	0	1	0	0
FLAGNPQGGRQS	1	0	0	1
DAEFRH	0	0	0	1
DAFCSSRGKV	0	0	1	0
DVEFRH	0	0	0	1
SQELHRLQTYPR	0	1	0	1
QWFH	1	0	0	0
GEGFFEM	1	0	0	0
IWCDIFTNSRGKR	0	0	1	0
ME	1	0	1	0
TWYS	1	0	1	0
PIKTLPM	1	1	0	1
RPKPQQFFG	0	0	1	0
YK	1	0	0	0
NEVFLDTQ	0	0	1	1
VH	0	1	0	0
TWLPLPR	1	0	0	0
AGIVAMR	1	0	0	1
QQRYWQSLQQHQQHR	1	0	0	1
LPC	1	0	0	0
DTMH	1	0	0	0
LLPF	1	0	0	0
RLTAQEPTIRFE	1	0	0	1
LLCIALRKK	0	1	0	0
ARSRLAELRGVPRGL	0	0	0	0
NFHPQ	1	0	0	0
YTIWCDIFTNSRGKR	0	0	1	0
GPR	1	0	0	1
AGNDGFEYVTLK	1	0	0	0
NAPVSIPQ	0	0	0	1
VSRRR	0	1	0	0
RLSFDRVGSMEKA	1	1	0	0
AEDG	0	0	1	1
VNAE	1	0	1	0
MWCDAFCSSDGKV	0	0	1	0
KED	0	0	0	1
AEDG	1	1	0	1

LLCIALRKK	0	0	0	0
LDAPGHR	0	1	0	0
NLNQCQRYM	1	0	0	0
RMLRTTRY	0	0	1	0
GFSGLDGAKGD	1	0	0	0
AAGAQ	1	0	1	0
RGAVIH	1	0	0	0
YQIGL	1	0	0	0
DW	1	1	0	0
QWFM	0	1	0	0
GSH	1	0	0	0
LRKLRKRL	0	0	1	0
KVPPLLY	1	0	0	1
SVDGKEDLIW	0	0	0	1
YPYDIPL	0	0	0	1
NHAV	1	0	0	0
YA	1	0	0	0
FF	1	0	1	0
GGVVIA	0	0	0	1
PRRTRMMNGGR	0	1	0	0
SVRVPY	1	0	0	1
QSHYRHISPAQVC	0	0	0	0
SLAFVDDVLN	1	0	0	0
RPPGFSPFR	0	1	0	0
EHP	1	1	0	0
AEFHRWSSYMVHWK	0	0	1	0
WEKPPVSH	1	1	0	0
SCNT	0	0	1	0
INW	1	0	0	0
WSREEQEREE	1	1	0	0
IGFLIIWV	1	1	0	0
CGR	0	0	0	1
VF	1	0	1	0
VRPDEEEGEQEHRGR	1	0	0	1
GTEDELDK	0	0	0	1
LVFFARK	0	0	0	1
VL	1	0	1	0
LGRLSQELHRLQTY	0	1	0	1
LETVNQ	0	0	1	1
VGDI	1	0	0	0
RYYAAFFARR	0	0	0	1
DTEFRH	0	0	0	1

LLANSLLA	1	1	0	0
YAEERTPIL	1	0	0	0
DYVTVKGFSPLRKA	1	0	0	0
IMEF	1	0	1	0
QDHCH	1	0	0	0
RGAVLH	1	0	0	0
SDLKHFPF	0	1	0	0
FQPSF	1	0	1	0
RGAVIH	1	0	0	0
RVGSMEKART	1	1	0	0
ENCPRG	0	0	1	0
AYPF	1	0	0	0
TWCDGFCSSRGKR	0	0	1	0
YEVHHQ	0	0	0	1
VNAVLH	1	0	0	0
MKP	1	1	0	1
VNAVIH	1	0	0	0
CCNQSSQPKTTTNC	0	0	1	0
KEGV	1	0	0	1
YPYFDPL	0	0	0	1
YVLLPSPK	1	0	0	1
NFDVLTEQ	0	0	1	0
MWCDATCSSRGKV	0	0	1	0
AADGPMKGILGY	0	1	0	0
VEGNLQVLRPR	1	0	0	0
ATHVLAAAERN	1	0	0	1
QCHQP	1	0	0	0
MQIFVK	1	0	0	0
SHNLPILR	1	0	0	0
DAAGRLQE	1	0	0	0
KHNRGDEF	1	1	0	0
VEDELVAVV	1	0	0	0
CA	1	0	0	0
WVSPLAGRT	1	1	0	0
MDFMKGLSMAKE	1	0	0	1
YGS	1	0	0	0
CIK	0	0	1	0
LVYIERGQGLI	1	0	0	1
MMLDF	0	1	0	0
KCHKPK	1	0	0	0
KELEEK	1	0	0	0
RDPEER	1	0	0	0

IPP	0	1	0	0
NSHNLPIRL	1	0	0	0
SALLRSIPA	0	0	0	0
NAPVSIPQ	0	0	0	1
DFNKF	0	0	0	1
GDMNP	1	0	0	0
TWYFITPYIPDK	1	0	0	0
SSEDIKE	0	1	0	0
QAW	0	1	0	0
DIFTNSRGKR	0	0	1	0
EV	0	1	0	0
NFALAIRAG	1	0	0	1
GLATAGGDAGLW	1	1	0	0
VATLSVGPV	1	1	0	0
GPW	1	0	0	0
CHEASQC	0	0	0	1
WPK	1	0	0	0
GPETAFLR	0	1	0	0
TLKIVW	0	0	0	0
LKQELEDLLEKQE	1	0	0	0
DHG	1	0	0	0
QAW	0	1	0	0
FY	1	1	1	0
RYKF	1	0	0	0
SCNTAT	0	0	1	0
YPYDIPL	0	0	0	1
CA	1	0	0	0
CHC	1	0	0	0
LDAVNR	0	1	0	0
DGFCSSRGKR	0	0	1	0
EVSGPGLSPN	1	1	0	0
GPE	0	1	0	0
EM	1	0	0	0
VFRGFDVQILSEVF	1	0	0	1
PAYCS	1	0	1	0
GGLEPINFQ	1	0	0	0
QCQCAVGEGGL	0	1	0	0
DYVTVKGFSPRL	1	0	0	0
GFAGDDAPRA	1	0	0	0
FLV	0	1	0	0
VNLAG	0	0	1	0
SGFDAE	1	1	1	0

RADARADA	0	0	0	1
HKQLPFFED	0	0	1	1
AK	1	0	1	0
YSVH	1	0	0	0
IGFLIIWV	1	1	0	0
CVGSY	1	0	1	0
TPAIS	1	0	0	0
TGENHR	0	0	0	1
RYYA AFFARR	0	0	0	1
YK	1	0	0	0
LDYE	1	0	1	1
AL	0	1	0	0
LPFFD	0	0	0	1
GVYY	1	1	0	0
EYVEETTGD EYVSLK	1	0	0	1
APW	1	0	0	0
VHHANEN	1	0	0	0
VSGGPI	1	0	1	0
DLSGLLLLYSSLGDA	1	0	0	0
CR	1	0	0	0
LLPHHADADY	1	0	0	0
NW	1	0	0	0
VMVGM	1	0	1	0
VNAVLH	1	0	0	0
NFGAIL	0	0	0	1
LAYLQYTD FETR	1	0	0	0
TQGDIGF	1	0	1	0
YRQQLEPINF	1	0	0	0
VHV	1	0	0	0
LNA	1	0	1	0
ERLDIVLPG	1	0	0	1
WCPFSRSF	1	0	0	0
HNLDQTESDV	1	0	0	0
SALLRSIPA	0	0	0	1
EC	0	1	0	0
IEEEEEIEAER	1	0	0	1
AEFHRWSSY MVHWK	0	0	1	0
LRLRLRL	0	0	1	0
IQW	0	1	0	0
YA	1	0	0	0
LVRL	1	1	0	0
GTWY	1	0	1	0

EQRPR	0	0	0	1
-------	---	---	---	---

* 1 indicates that the peptide has the activity, while 0 indicates that it does not have the activity.