Supplementary Materials

The Dark Side of Alzheimer's Disease: Unstructured Biology of Proteins from the Amyloid Cascade Signaling Pathway

Kundlik Gadhave[1], Bhuvaneshwari R. Gehi[1], Prateek Kumar[1], Bin Xue[2], Vladimir N. Uversky[2,3,*], and Rajanish Giri[1*]

Affiliations

1. School of Basic Sciences, Indian Institute of Technology Mandi, Mandi, India
2. Department of Cell Biology, Microbiology and Molecular Biology, School of Natural Sciences and Mathematics, College of Arts and Sciences, University of South Florida, Tampa, FL, 33620, USA
3. Department of Molecular Medicine and Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL, 33612, USA
4. Laboratory of New Methods in Biology, Institute for Biological Instrumentation, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia

*correspondence to: vuversky@health.usf.edu or rajanishgiri@iitmandi.ac.in

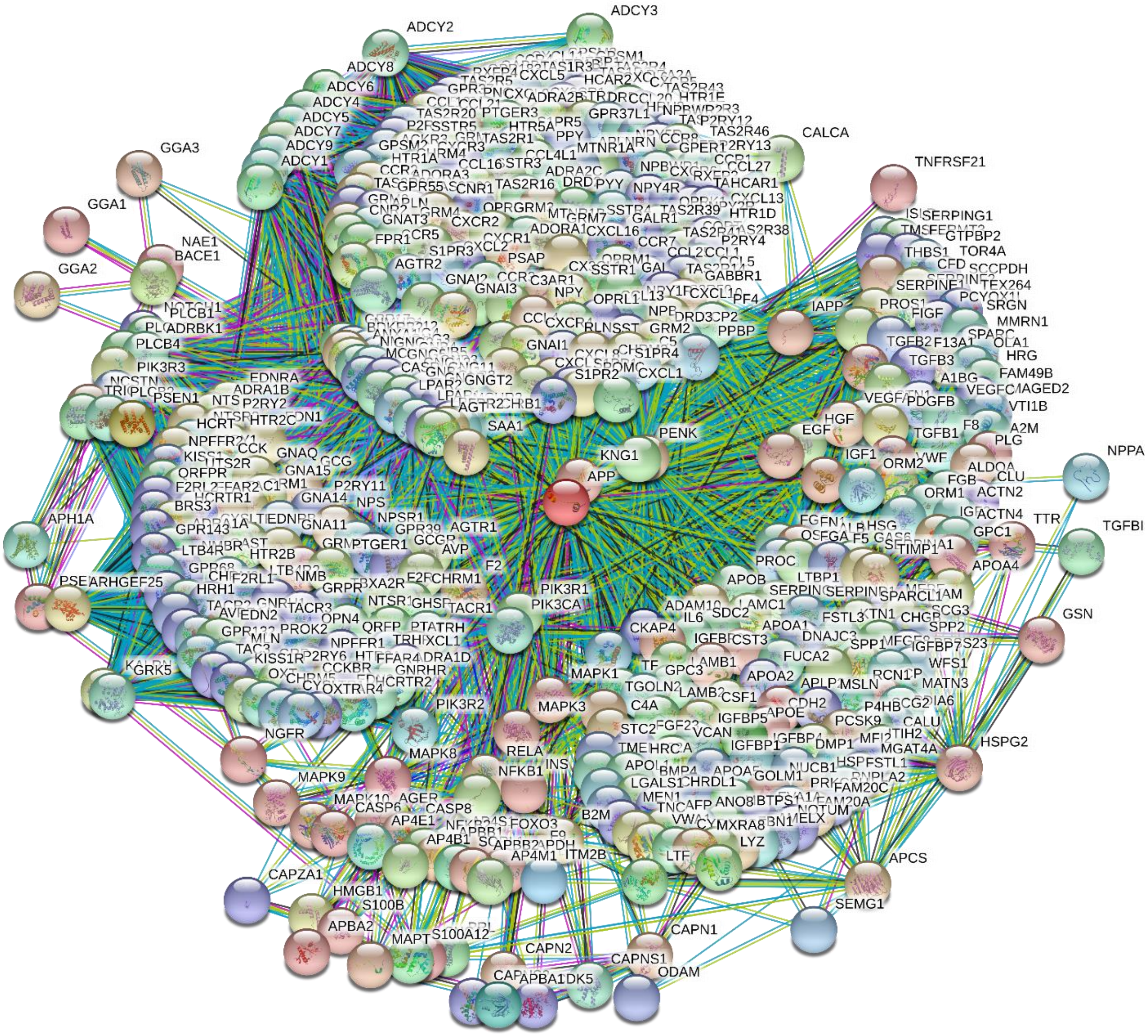

ADCY2
ADCY3
ADCY8
ADCY6
ADCY4
ADCY5
ADCY7
ADCY9
ADCY1
GGA3
GGA1
GGA2
NAE1
BACE1
CALCA
TNFRSF21
NPPA
TGFBI
TTR
GSN
HSPG2
APCS
SEMG1
APH1A
ARHGEF25
GRK5
NGFR
MAPK9
CAPZA1
HMGB1
S100B
APBA2
MAPT
S100A12
CAPN2
CAPN1
CAPNS1
ODAM
APP
KNG1
PENK
SAA1
PIK3R1
PIK3CA
PIK3R2
MAPK1
MAPK3
MAPK8
RELA
NFKB1
INS
CASP6
CASP8
FOXO3
B2M
LYZ
LTF
ITM2B
AP4M1
AP4E1
AP4B1
IAPP
PPBP
VEGFC
PDGFB
TGFB1
TGFB2
TGFB3
HGF
EGF
IGF1
PLG
APOA4
GPC1
ACTN4
ACTN2
CLU
ALDOA
HRG
MMRN1
SRGN
SPARC
A1BG
F8
A2M
VTI1B
FAM49B
TIMP1
APOB
PROC
CKAP4
IL6
SDC2
TF
GPC3
C4A
STC2
VCAN
IGFBP5
APOE
APOA2
FUCA2
MSLN
MATN3
WFS1
CALU
MGAT4A
PCSK9
DMP1
GOLM1
NOTUM
FAM20A
FAM20C
MXRA8
LGALS1
CSF1
TGOLN2
AGTR1
AVP
F2
NPS
P2RY11
GNA14
GNA11
GNA15
GNAQ
CCK
PLCB4
PIK3R3
PSEN1
NCSTN
EDNRA
ADRA1B
HCRT
KISS1
GPR143
HRH1
OPN4
TACR1
TACR3
KISS1R
NMB
F2RL1
CXCL1
S1PR2
S1PR4
GNAI1
GNAI3
C3AR1
NPY
PSAP
AGTR2
CXCR2
GRM4
CR5
FPR1
GNAT3
SSTR1
SSTR4
OPRL1
GRM2
DRD3
CXCL13
CXCL16
CCR7
CCL27
GABBR1
HTR1D
P2RY4
P2RY13
TAS2R46
TAS2R43
TAS2R38
TAS2R39
TAS2R16
TAS2R20
TAS2R5
TAS2R1
NPY4R
MTNR1A
PYY
PPY
ADRA2B
ADRA2C
HTR5A
HTR1A
HTR1F
HCAR2
CXCL5
GPR37L1
GPR55
GPSM2
ADORA3
ADORA1
CCL16
CNR1
OPRM1
GRM7
PTGER3
PTGER1
GPR39
GCGR
HTR2B
HTR2C
EDN2
NTSR1
PROK2
TRH
GNRHR
CCKBR
OXTR
CHRM1
CHRM5
EDNRB
TBXA2R
ADRA1D
FFAR4
NPFFR1
QRFP
BRS3
LTB4R
F2RL3
UTS2R
GNGT2
LPAR2
BDKRB1
SERPING1
SERPINE1
THBS1
GTPBP2
TOR4A
SCCPDH
TEX264
PCYOX1
F13A1
OLA1
PROS1
FIGF
ORM1
ORM2
FGB
VWF
SERPINC1
SPARCL1
LTBP1
FSTL3
CHGB
SCG3
SPP1
SPP2
IGFBP7
RCN1
P4HB
ADAM10
LAMC1
APOA1
DNAJC3
LAMB1
LAMB2
CST3
IGFBP1
NUCB1
FSTL1
PNPLA2
MELX
ANO8
BMP4
CHRDL1
MEN1
TNC
AFP
APOA5
CAPN2

Figure S1A

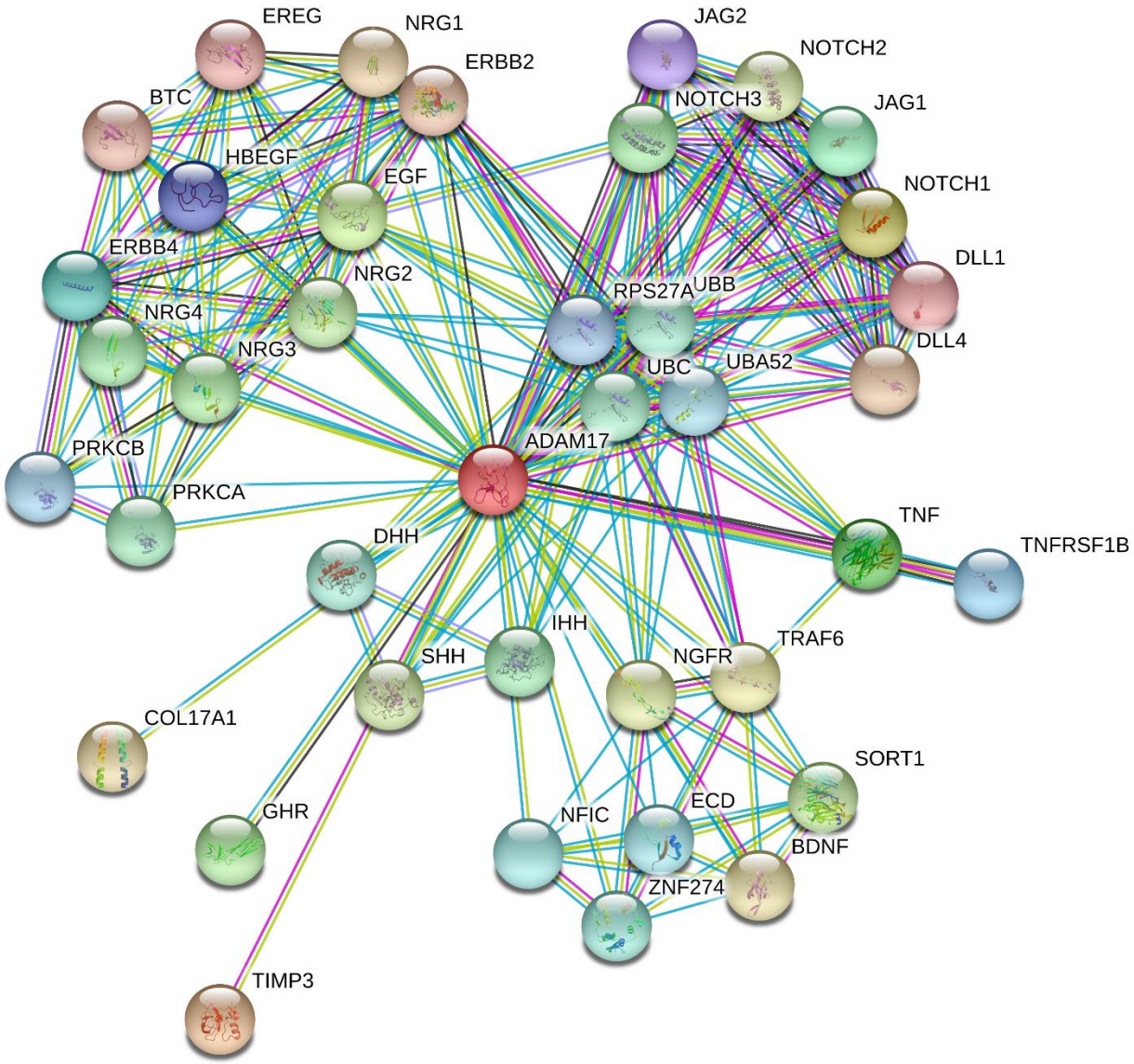

EREG
NRG1
ERBB2
BTC
HBEGF
EGF
ERBB4
NRG2
NRG4
NRG3
PRKCB
PRKCA
JAG2
NOTCH2
NOTCH3
JAG1
NOTCH1
DLL1
DLL4
RPS27A
UBB
UBC
UBA52
ADAM17
TNF
TNFRSF1B
DHH
IHH
SHH
NGFR
TRAF6
COL17A1
SORT1
GHR
NFIC
ECD
BDNF
ZNF274
TIMP3

Figure S1B

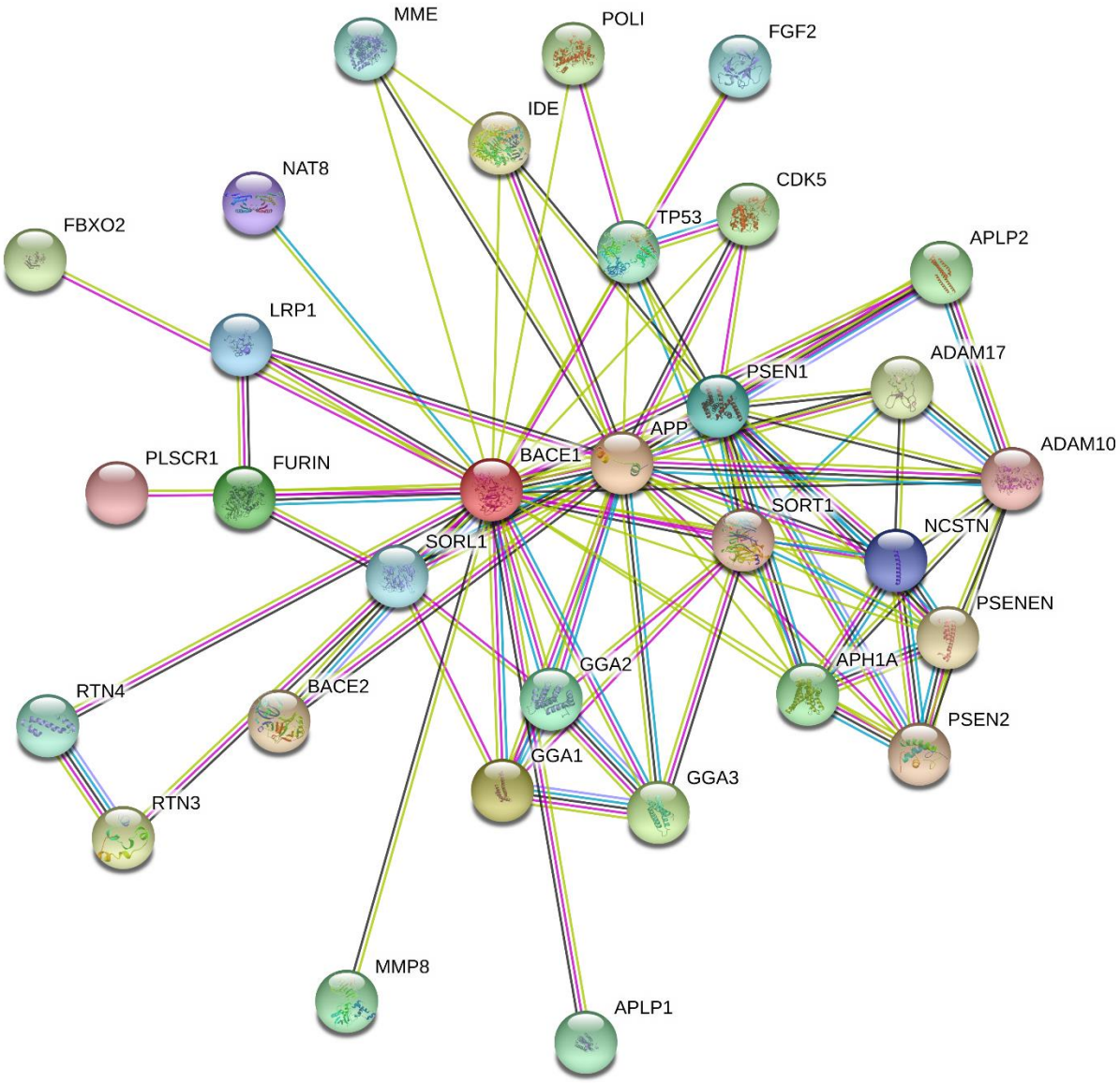

MME
POLI
FGF2
IDE
NAT8
CDK5
FBXO2
TP53
APLP2
LRP1
ADAM17
PSEN1
APP
ADAM10
PLSCR1
FURIN
BACE1
SORT1
NCSTN
SORL1
PSENEN
APH1A
GGA2
RTN4
BACE2
PSEN2
GGA1
GGA3
RTN3
MMP8
APLP1

Figure S1C

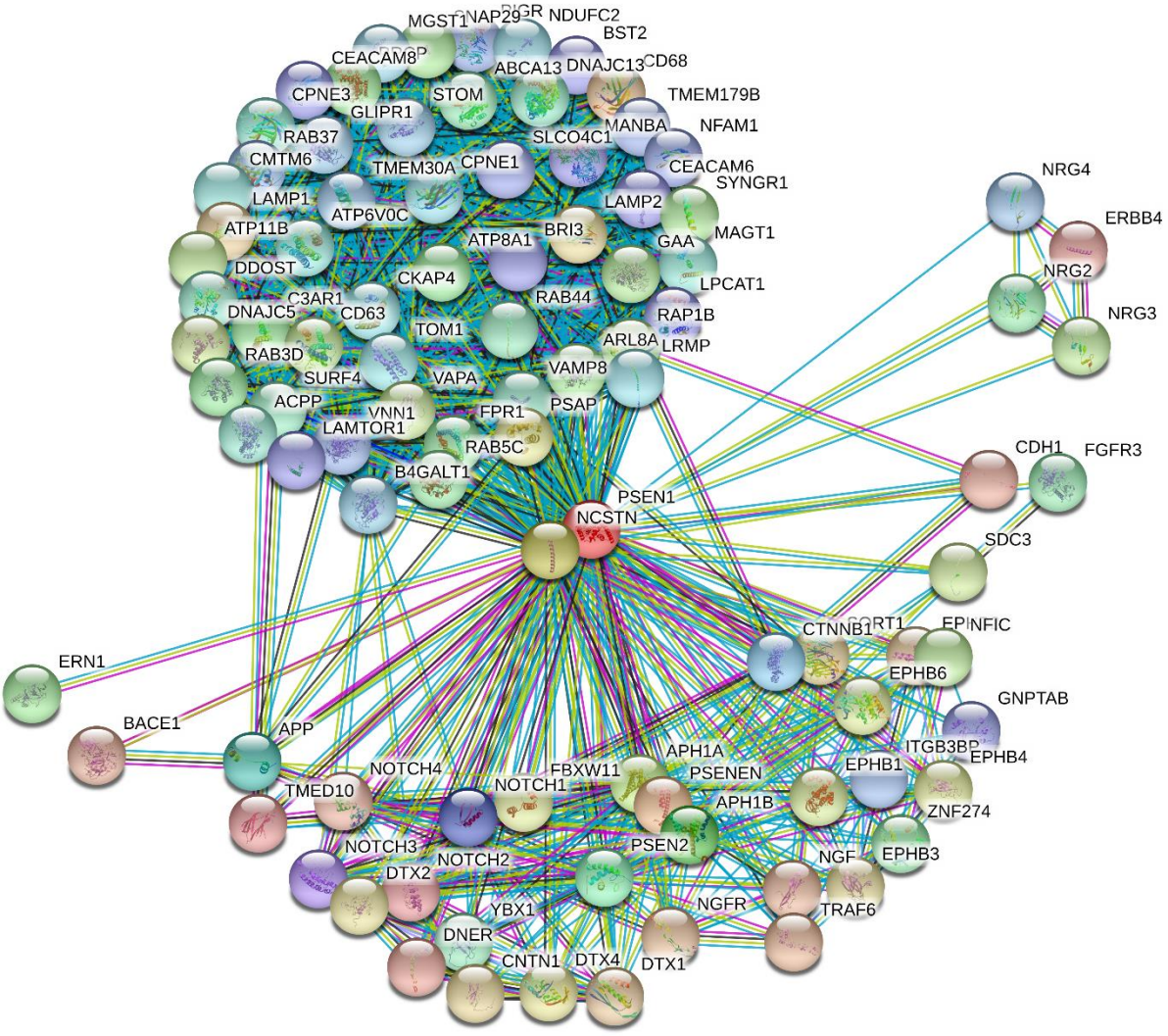

MGST1
NDUFC2
BST2
CEACAM8
ABCA13
DNAJC13
CD68
CPNE3
STOM
TMEM179B
GLIPR1
MANBA
NFAM1
RAB37
SLCO4C1
CMTM6
TMEM30A
CPNE1
CEACAM6
SYNGR1
LAMP1
LAMP2
ATP6V0C
ATP11B
ATP8A1
BRI3
GAA
MAGT1
DDOST
CKAP4
LPCAT1
C3AR1
RAB44
DNAJC5
CD63
RAP1B
TOM1
ARL8A
LRMP
RAB3D
VAMP8
SURF4
VAPA
ACPP
VNN1
FPR1
PSAP
LAMTOR1
RAB5C
B4GALT1
PSEN1
NCSTN
NRG4
ERBB4
NRG2
NRG3
CDH1
FGFR3
SDC3
CTNNB1
EPHB6
GNPTAB
ERN1
BACE1
APP
APH1A
PSENEN
EPHB1
EPHB4
NOTCH4
FBXW11
NOTCH1
TMED10
APH1B
ZNF274
NOTCH3
PSEN2
NOTCH2
NGF
EPHB3
DTX2
YBX1
NGFR
TRAF6
DNER
CNTN1
DTX4
DTX1

Figure S1D

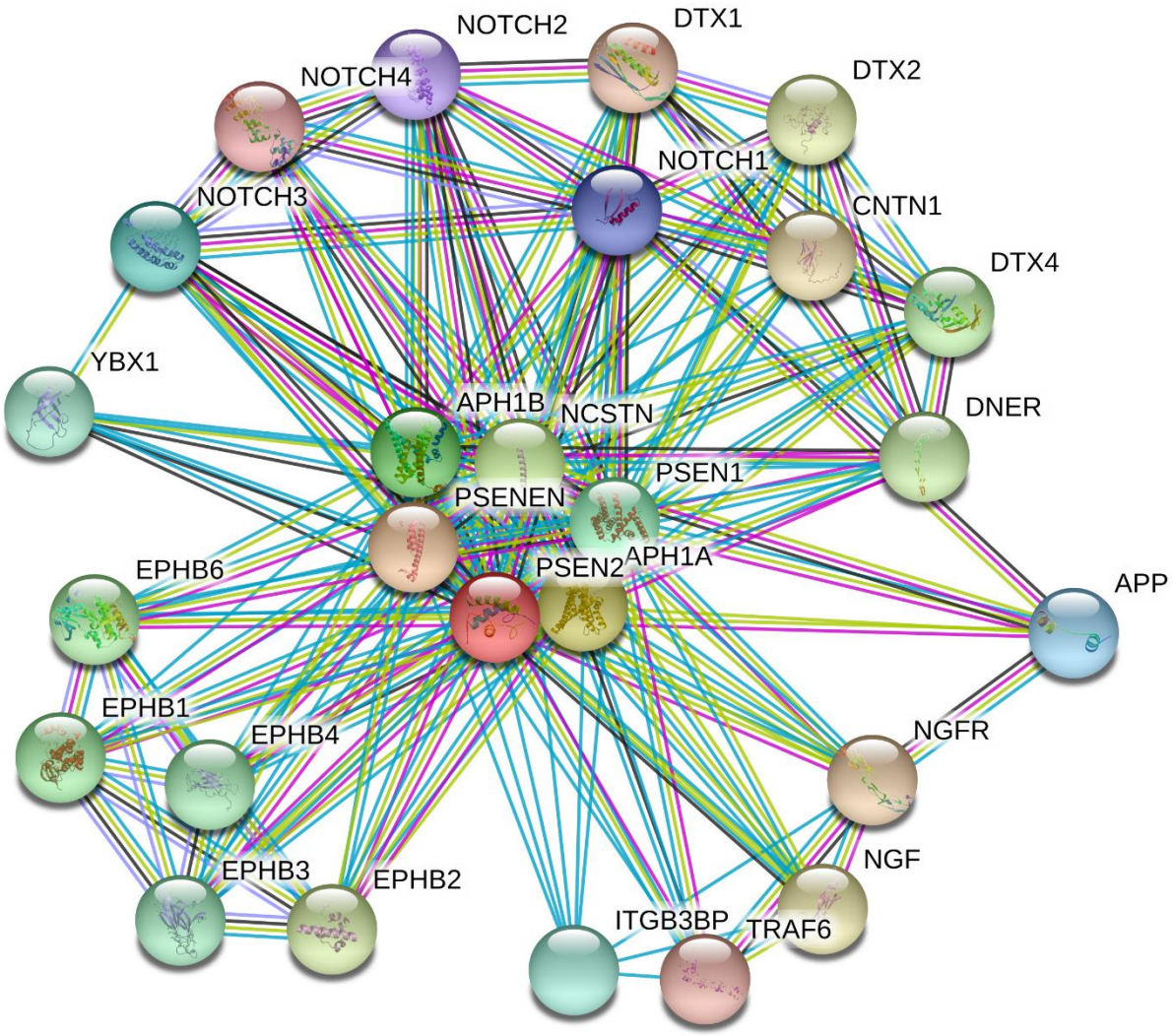

NOTCH2
DTX1
NOTCH4
DTX2
NOTCH1
NOTCH3
CNTN1
DTX4
YBX1
APH1B
NCSTN
DNER
PSEN1
PSENEN
APH1A
EPHB6
PSEN2
APP
EPHB1
EPHB4
NGFR
NGF
EPHB3
EPHB2
ITGB3BP
TRAF6

Figure S1E

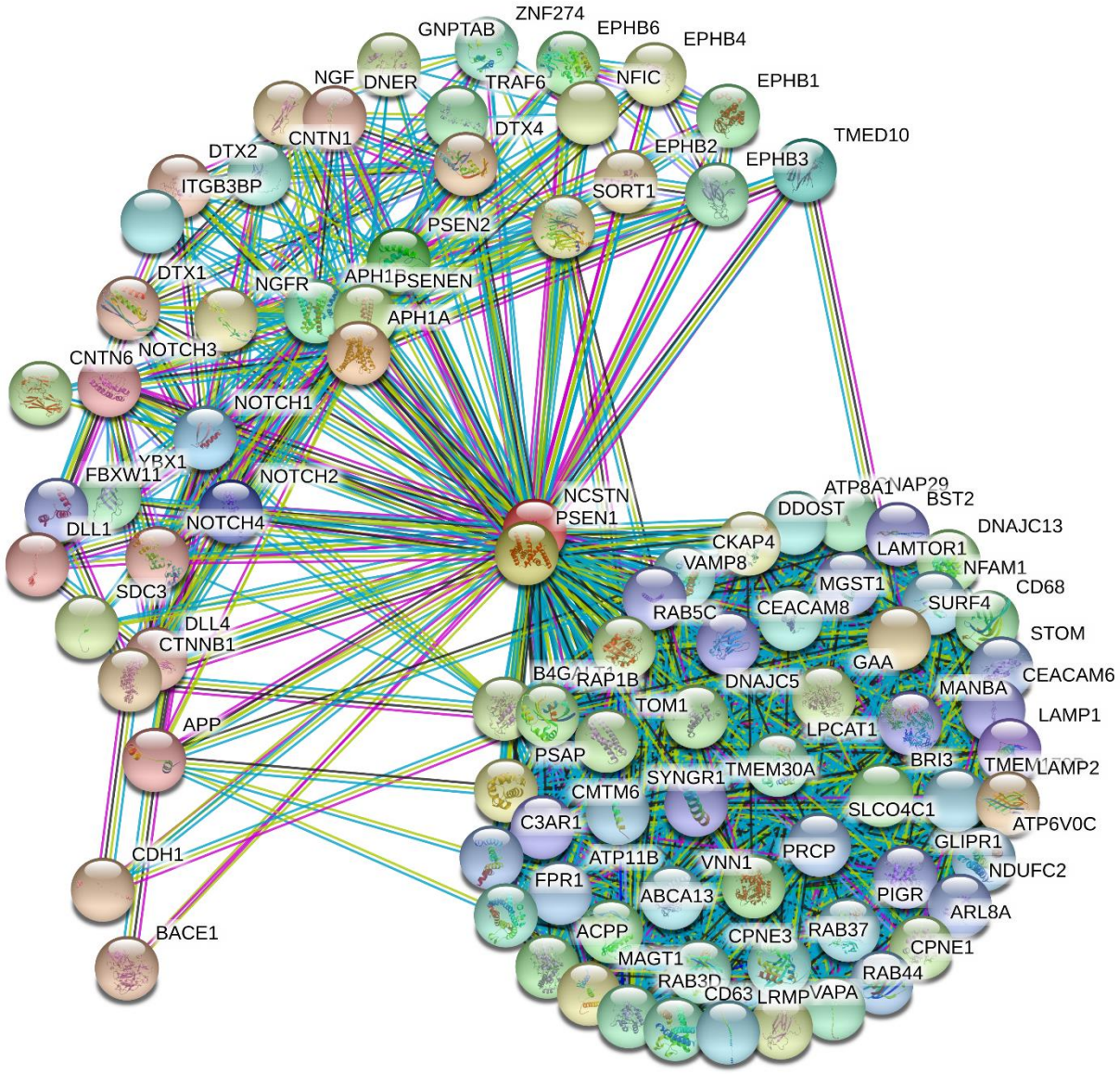

ZNF274
GNPTAB
EPHB6
EPHB4
NGF
DNER
TRAF6
NFIC
EPHB1
CNTN1
DTX4
TMED10
DTX2
EPHB2
EPHB3
ITGB3BP
SORT1
PSEN2
DTX1
NGFR
APH1B
PSENEN
APH1A
CNTN6
NOTCH3
NOTCH1
YBX1
FBXW11
NOTCH2
NCSTN
PSEN1
DLL1
NOTCH4
ATP8A1
DDOST
BST2
DNAJC13
CKAP4
LAMTOR1
VAMP8
NFAM1
SDC3
MGST1
CD68
RAB5C
CEACAM8
SURF4
DLL4
STOM
CTNNB1
GAA
B4GALT1
RAP1B
DNAJC5
CEACAM6
MANBA
TOM1
LAMP1
APP
LPCAT1
PSAP
BRI3
SYNGR1
TMEM30A
LAMP2
CMTM6
SLCO4C1
C3AR1
ATP6V0C
CDH1
GLIPR1
PRCP
ATP11B
VNN1
NDUFC2
FPR1
ABCA13
PIGR
ARL8A
BACE1
ACPP
CPNE3
RAB37
CPNE1
MAGT1
RAB44
RAB3D
CD63
LRMP
VAPA

Figure S1F

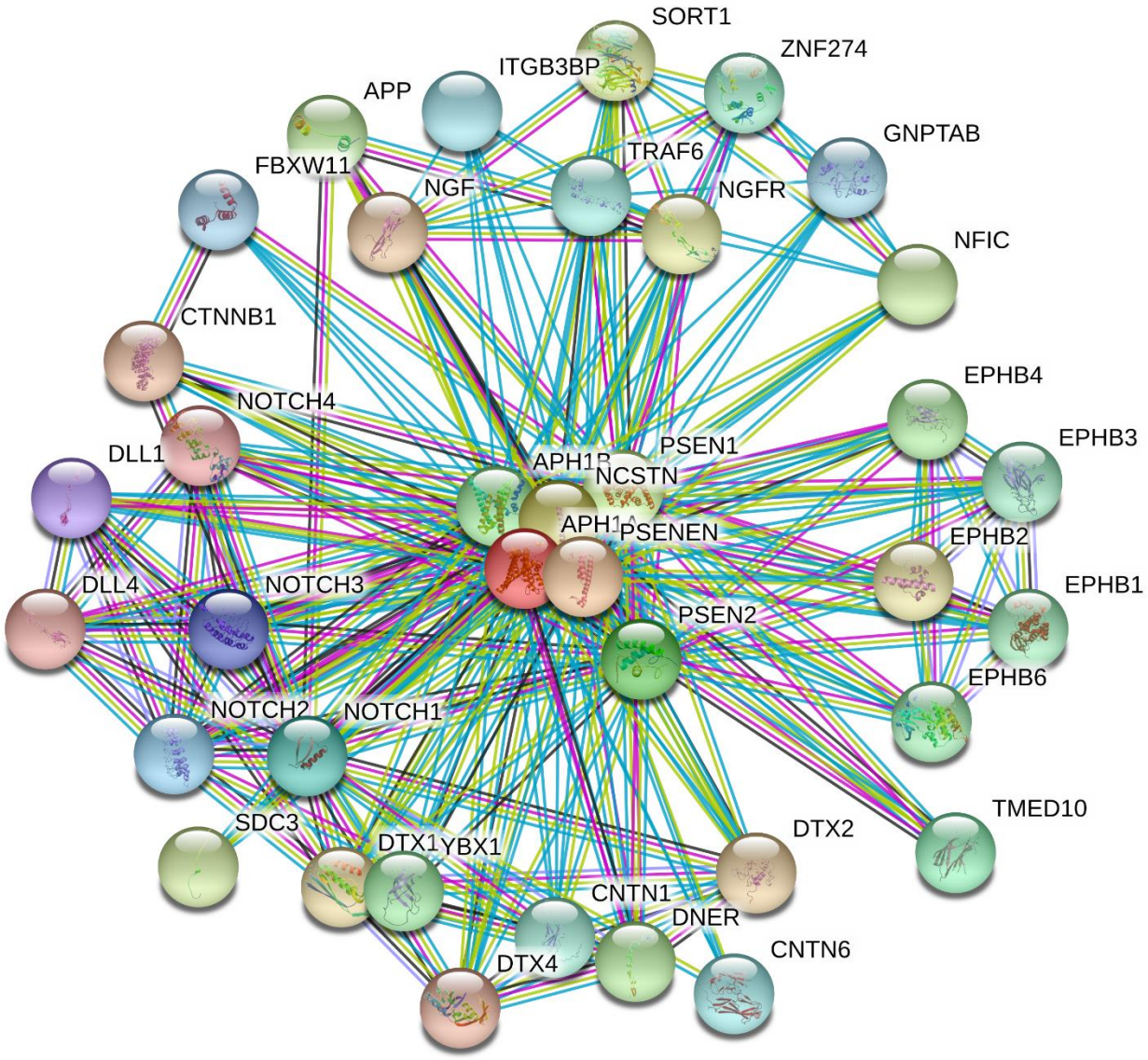

SORT1
ZNF274
ITGB3BP
APP
GNPTAB
FBXW11
TRAF6
NGF
NGFR
NFIC
CTNNB1
EPHB4
NOTCH4
EPHB3
DLL1
PSEN1
NCSTN
APH1A
PSENEN
EPHB2
DLL4
NOTCH3
EPHB1
PSEN2
EPHB6
NOTCH2
NOTCH1
TMED10
SDC3
DTX1
YBX1
DTX2
CNTN1
DNER
CNTN6
DTX4

Figure S1G

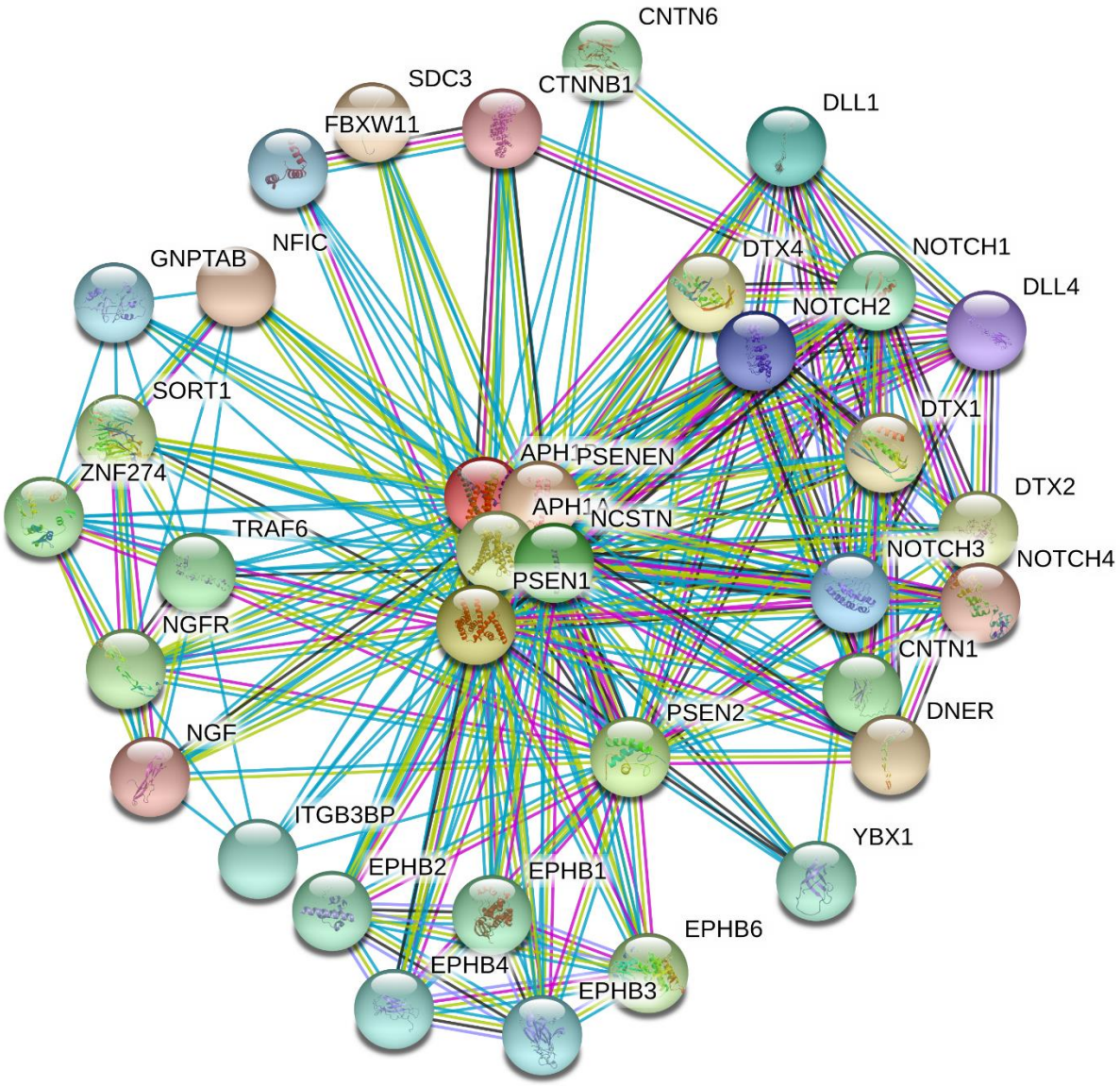

CNTN6
SDC3
CTNNB1
FBXW11
DLL1
NFIC
GNPTAB
DTX4
NOTCH1
DLL4
NOTCH2
SORT1
DTX1
APH1
PSENEN
ZNF274
DTX2
APH1
NCSTN
TRAF6
NOTCH3
NOTCH4
PSEN1
NGFR
CNTN1
PSEN2
DNER
NGF
ITGB3BP
YBX1
EPHB2
EPHB1
EPHB6
EPHB4
EPHB3

Figure S1H

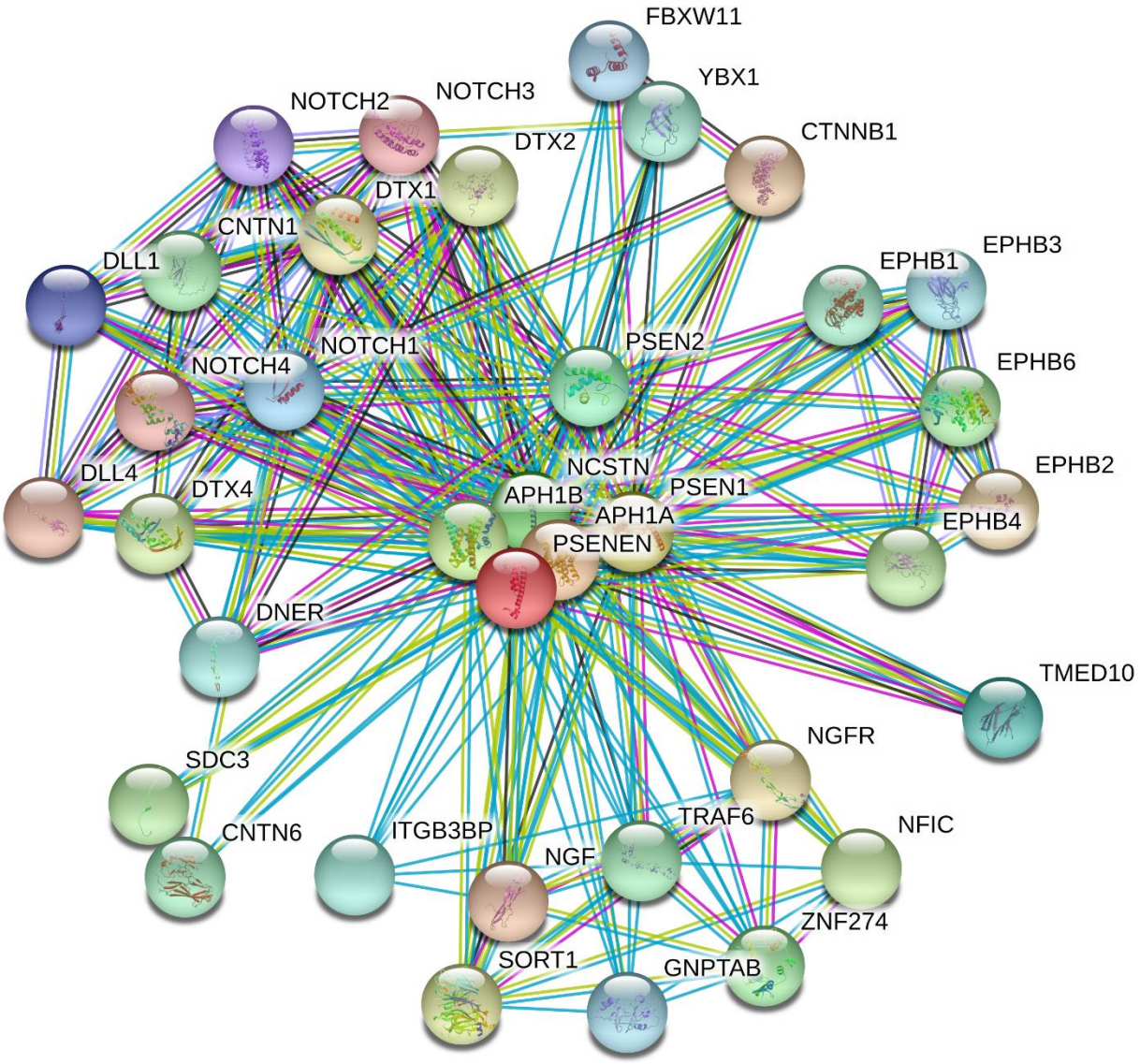

FBXW11
YBX1
NOTCH2
NOTCH3
CTNNB1
DTX2
DTX1
CNTN1
DLL1
EPHB3
EPHB1
NOTCH1
PSEN2
NOTCH4
EPHB6
NCSTN
EPHB2
DLL4
DTX4
APH1B
PSEN1
APH1A
EPHB4
PSENEN
DNER
TMED10
NGFR
SDC3
TRAF6
CNTN6
ITGB3BP
NFIC
NGF
ZNF274
SORT1
GNPTAB

Figure S1I

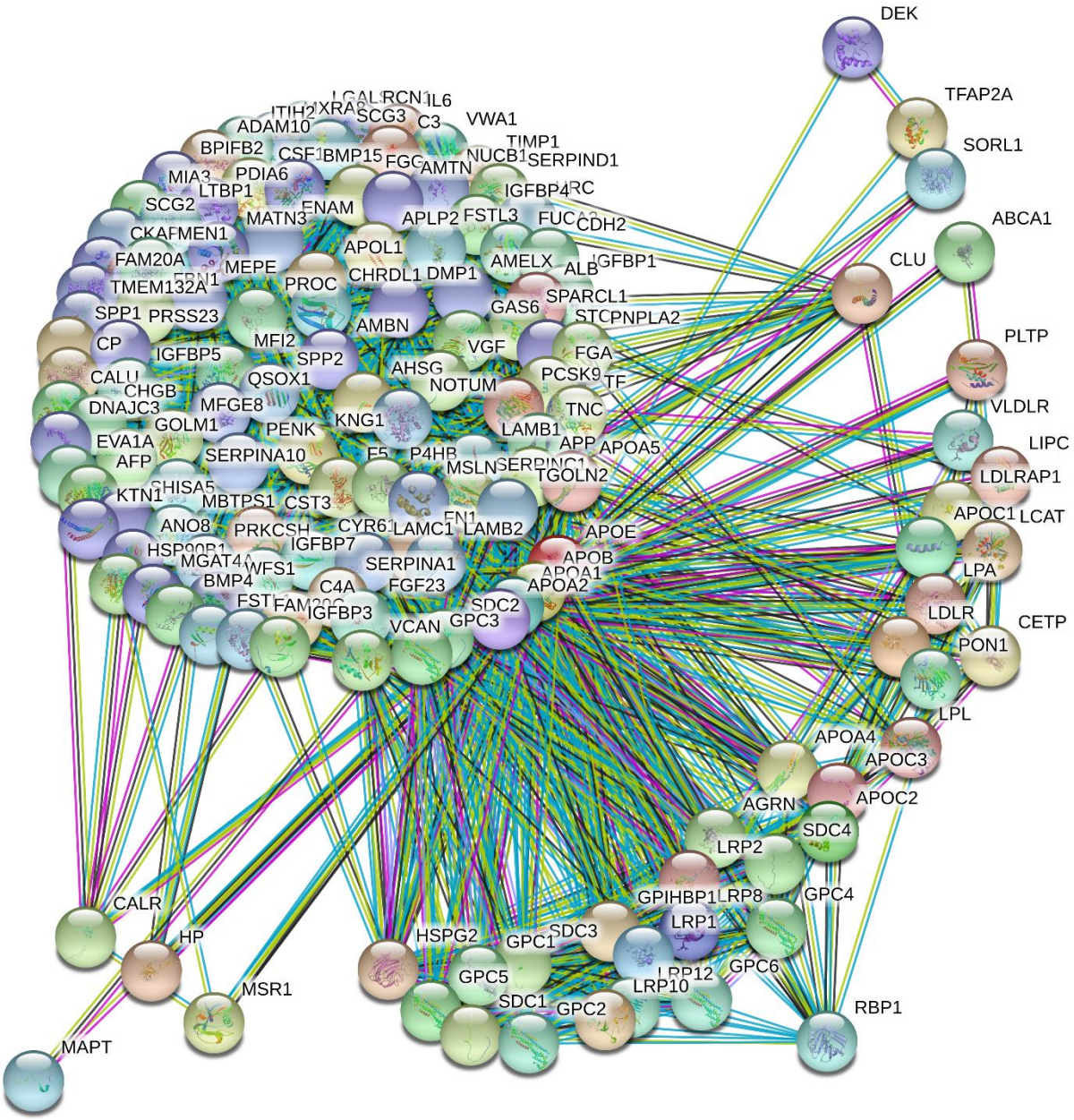

DEK
TFAP2A
SORL1
ABCA1
CLU
PLTP
VLDLR
LIPC
LDLRAP1
APOC1
LCAT
LPA
CETP
LDLR
PON1
LPL
APOA4
APOC3
APOC2
AGRN
SDC4
LRP2
GPC4
LRP8
GPIHBP1
LRP1
SDC3
GPC1
HSPG2
GPC5
LRP12
GPC6
LRP10
SDC1
GPC2
RBP1
CALR
HP
MSR1
MAPT
ADAM10
BPIFB2
MIA3
PDIA6
LTBP1
SCG2
MATN3
ENAM
CKAP4
MEN1
FAM20A
MEPE
TMEM132A
PROC
SPP1
PRSS23
CP
IGFBP5
CALU
CHGB
DNAJC3
MFGE8
GOLM1
EVA1A
AFP
SERPINA10
KTN1
CHISA5
MBTPS1
CST3
ANO8
PRKCSH
HSP90B1
IGFBP7
MGAT4A
WFS1
BMP4
FST
FAM20C
IGFBP3
SCG3
C3
VWA1
BMP15
FGG
AMTN
NUCB1
SERPIND1
IGFBP4
APLP2
FSTL3
FUCA2
CDH2
APOL1
AMELX
ALB
IGFBP1
CHRDL1
DMP1
AMBN
GAS6
SPARCL1
PNPLA2
MFI2
SPP2
VGF
FGA
AHSG
PCSK9
TF
QSOX1
NOTUM
TNC
KNG1
PENK
LAMB1
APP
APOA5
F5
P4HB
MSLN
SERPINC1
TGOLN2
CYR61
LAMC1
LAMB2
EN1
APOE
APOB
APOA1
APOA2
SERPINA1
C4A
FGF23
SDC2
GPC3
VCAN

Figure S1J

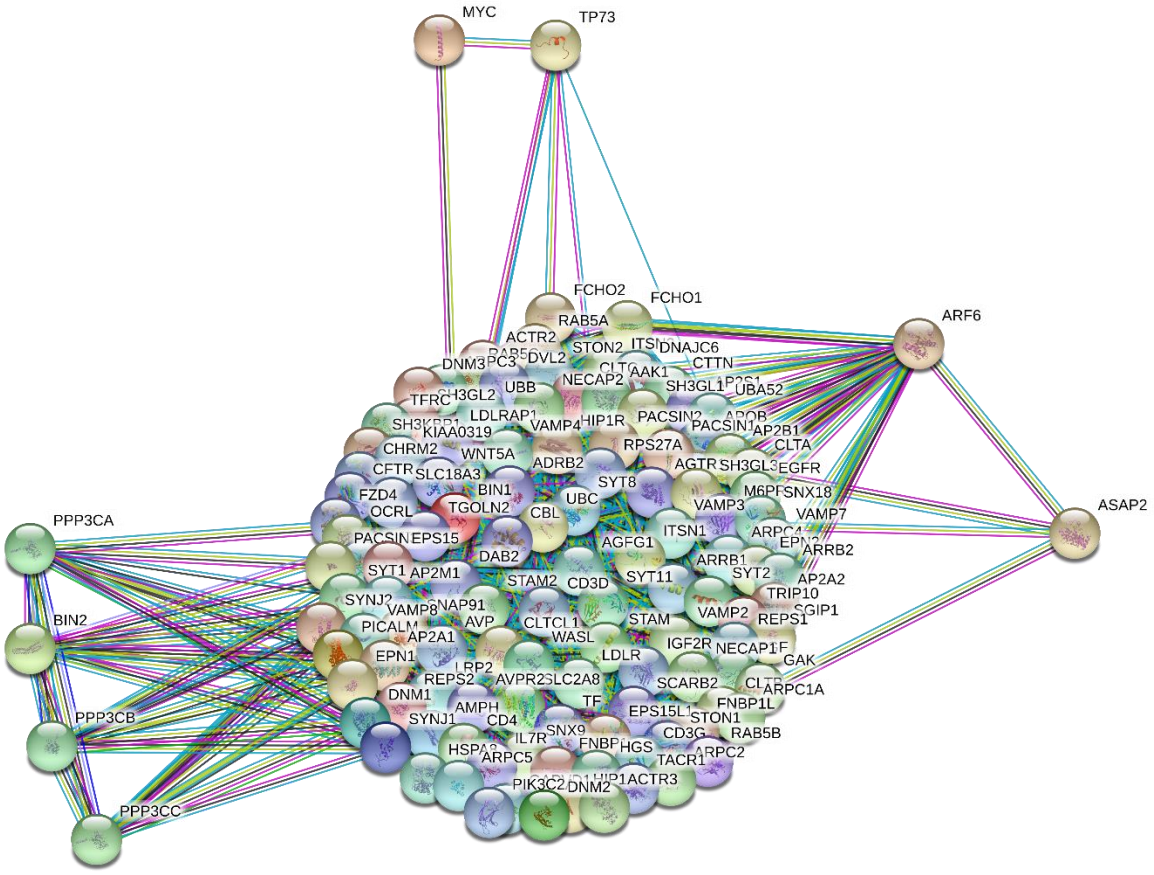

MYC
TP73
FCHO2
FCHO1
ARF6
RAB5A
ACTR2
STON2
DNAJC6
DNM3
DVL2
CTTN
AAK1
NECAP2
SH3GL1
UBA52
UBB
TFRC
LDLRAP1
VAMP4
HIP1R
PACSIN2
PACSIN1
AP2B1
KIAA0319
CHRM2
WNT5A
RPS27A
CLTA
CFTR
SLC18A3
ADRB2
AGTR
SH3GL3
EGFR
FZD4
BIN1
SYT8
M6PR
SNX18
OCRL
TGOLN2
UBC
CBL
VAMP3
VAMP7
ASAP2
PPP3CA
PACSIN
EPS15
ITSN1
AGFG1
ARRB2
DAB2
SYT1
AP2M1
STAM2
CD3D
SYT11
SYT2
AP2A2
TRIP10
SYNJ2
VAMP8
NAP91
VAMP2
SGIP1
BIN2
PICALM
AVP
CLTCL1
STAM
REPS1
AP2A1
WASL
IGF2R
NECAP1
EPN1
LDLR
GAK
REPS2
AVPR2
SLC2A8
SCARB2
ARPC1A
DNM1
TF
FNBP1L
AMPH
EPS15L1
SYNJ1
CD4
STON1
SNX9
CD3G
RAB5B
PPP3CB
HSPA8
IL7R
FNBP1
HGS
ARPC2
ARPC5
TACR1
PIK3C2
DNM2
ACTR3
PPP3CC

Figure S1K

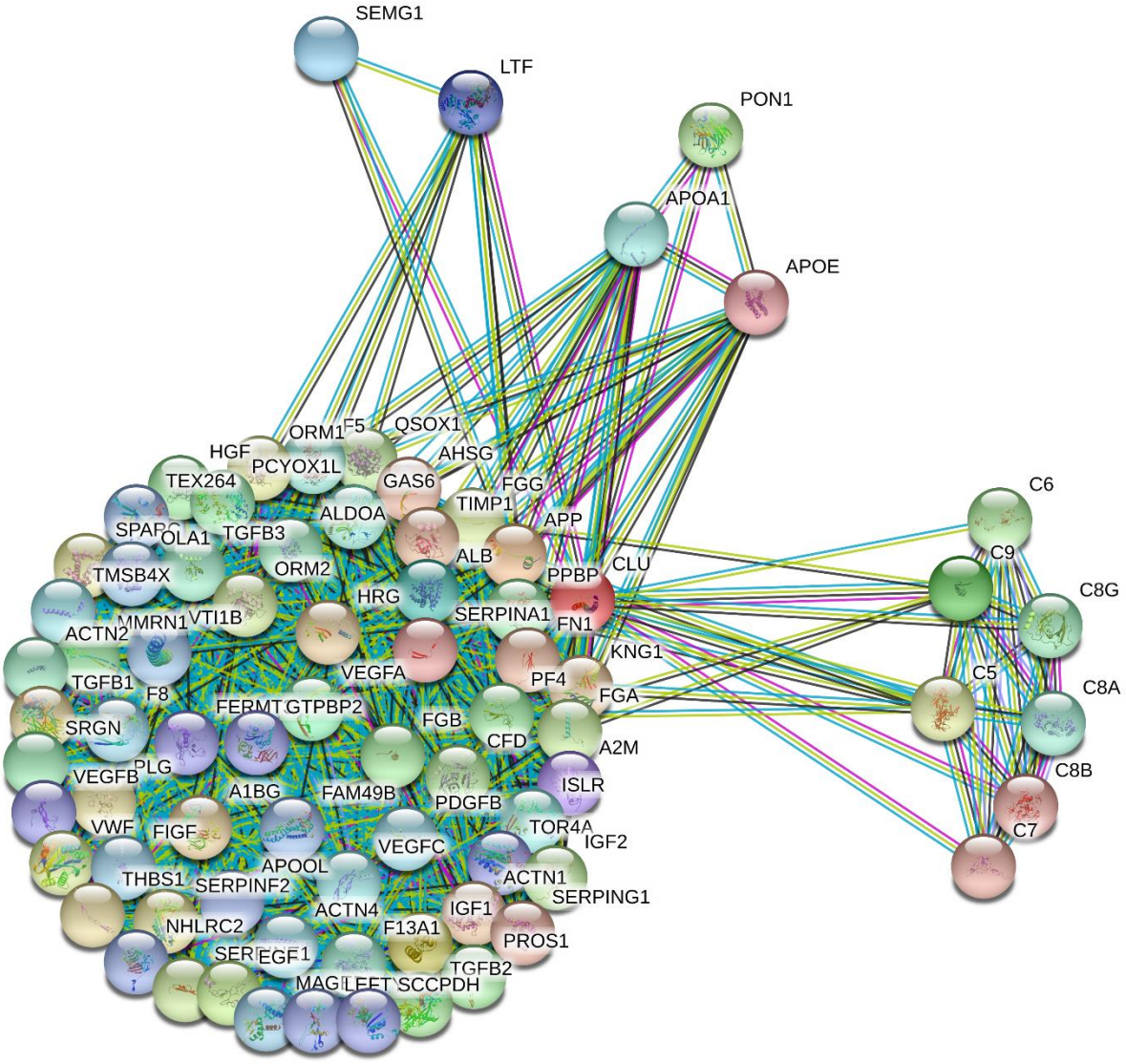

SEMG1
LTF
PON1
APOA1
APOE
ORM1
F5
QSOX1
HGF
PCYOX1L
AHSG
TEX264
GAS6
FGG
TIMP1
ALDOA
APP
SPARC
OLA1
TGFB3
ALB
C6
C9
TMSB4X
ORM2
PPBP
CLU
C8G
HRG
SERPINA1
FN1
MMRN1
VTI1B
ACTN2
KNG1
VEGFA
PF4
C5
C8A
TGFB1
F8
FERMT3
GTPBP2
FGA
SRGN
FGB
CFD
A2M
PLG
C8B
VEGFB
A1BG
FAM49B
ISLR
PDGFB
VWF
FIGF
TOR4A
C7
IGF2
VEGFC
APOOL
THBS1
SERPINF2
ACTN1
SERPING1
ACTN4
IGF1
NHLRC2
F13A1
PROS1
SERPINF1
TGFB2
MAGEL2
SCCPDH

Figure S1L

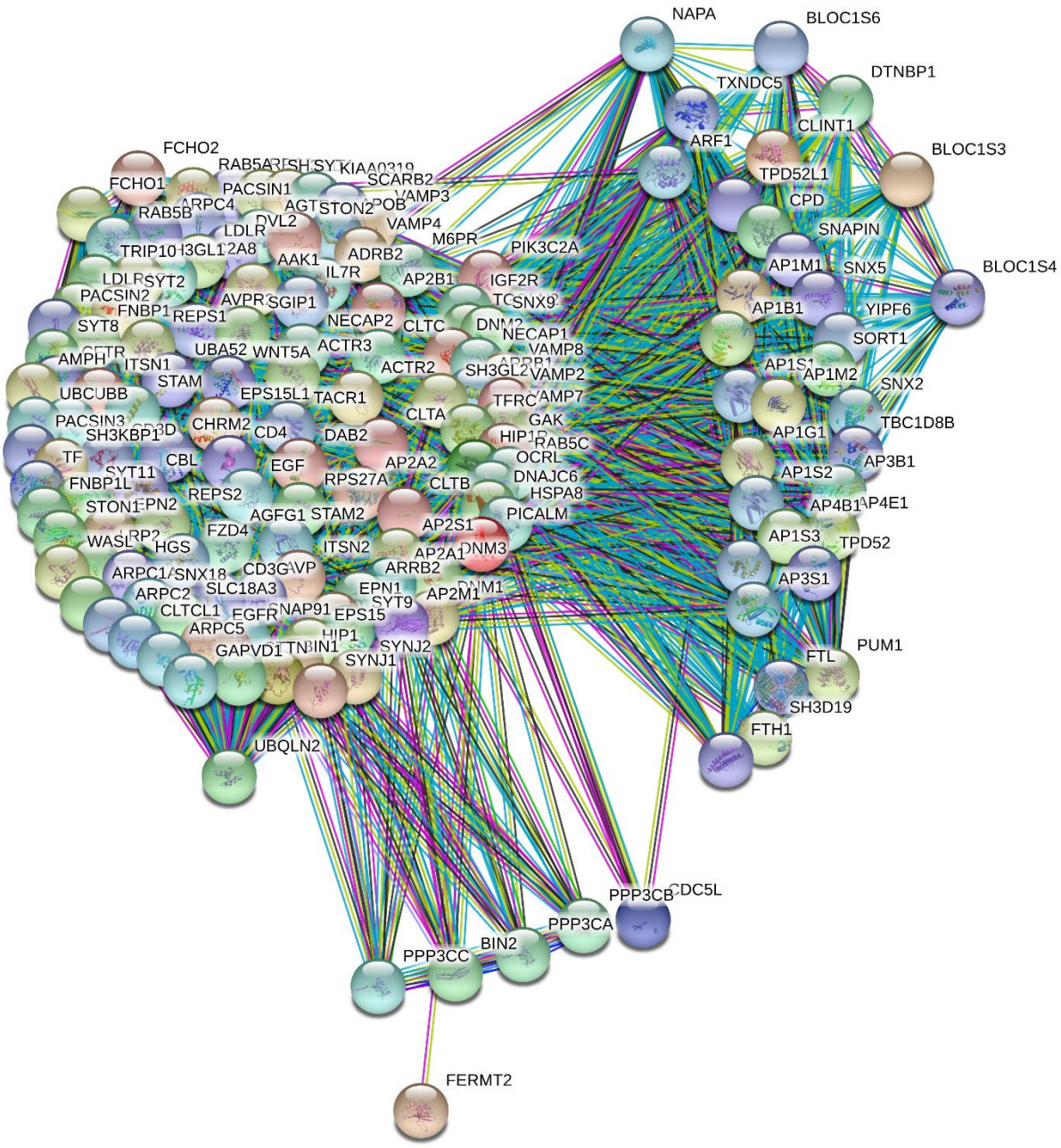

NAPA
BLOC1S6
DTNBP1
TXNDC5
CLINT1
ARF1
BLOC1S3
FCHO2
FCHO1
PACSIN1
SCARB2
VAMP3
RAB5B
ARPC4
AGTL
STON2
POB
TPD52L1
CPD
DVL2
VAMP4
LDLR
SNAPIN
TRIP10
M6PR
PIK3C2A
AAK1
ADRB2
AP1M1
SNX5
BLOC1S4
IL7R
AP2B1
IGF2R
LDLR
SYT2
PACSIN2
AVPR
SGIP1
SNX9
AP1B1
YIPF6
FNBP1
REPS1
NECAP2
CLTC
DNM2
SYT8
NECAP1
SORT1
UBA52
WNT5A
ACTR3
VAMP8
AMPH
ITSN1
ACTR2
SH3GL2
VAMP2
AP1M2
SNX2
STAM
UBCUBB
EPS15L1
TACR1
TFRC
VAMP7
CLTA
GAK
TBC1D8B
PACSIN3
CHRM2
CD4
DAB2
SH3KBP1
HIP1R
RAB5C
AP1G1
TF
CBL
EGF
AP2A2
OCRL
AP3B1
SYT11
RPS27A
CLTB
DNAJC6
AP1S2
FNBP1L
REPS2
HSPA8
STON1
EPN2
AGFG1
STAM2
PICALM
AP4B1
AP4E1
AP2S1
FZD4
WASL
RP2
HGS
ITSN2
AP2A1
DNM3
AP1S3
TPD52
ARPC1A
SNX18
CD3G
VP
ARRB2
AP3S1
ARPC2
SLC18A3
EPN1
DNM1
AP2M1
SYT9
CLTCL1
EGFR
NAP91
EPS15
ARPC5
HIP1
GAPVD1
TNIN1
SYNJ2
SYNJ1
PUM1
FTL
SH3D19
FTH1
UBQLN2
PPP3CB
CDC5L
PPP3CA
BIN2
PPP3CC
FERMT2

Figure S1M

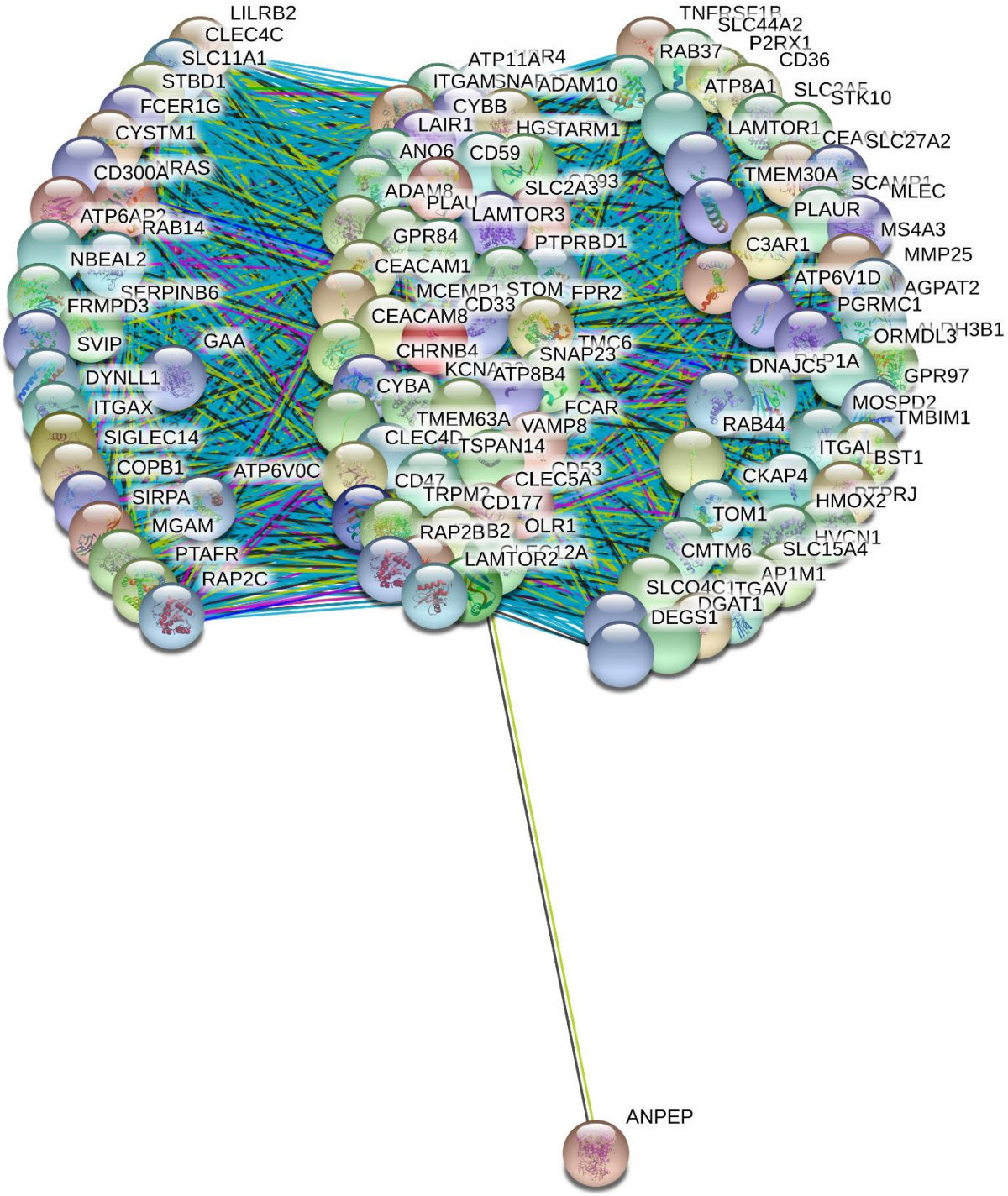

LILRB2
CLEC4C
SLC11A1
STBD1
FCER1G
CYSTM1
CD300A
RAS
ATP6AP2
RAB14
NBEAL2
SERPINB6
FRMPD3
SVIP
GAA
DYNLL1
ITGAX
SIGLEC14
COPB1
ATP6V0C
SIRPA
MGAM
PTAFR
RAP2C
ATP11A
R4
ITGAM
SNAI
ADAM10
CYBB
LAIR1
HGS
TARM1
ANO6
CD59
ADAM8
PLAU
SLC2A3
LAMTOR3
GPR84
PTPRB
D1
CEACAM1
MCEMP1
STOM
FPR2
CEACAM8
CD33
CHRNB4
TMC6
SNAP23
KCN
ATP8B4
CYBA
FCAR
TMEM63A
VAMP8
CLEC4D
TSPAN14
CD47
CLEC5A
TRPM2
CD177
RAP2B
B2
OLR1
LAMTOR2
A
TNFRSF1B
SLC44A2
P2RX1
RAB37
CD36
ATP8A1
SLC
STK10
LAMTOR1
CEA
SLC27A2
TMEM30A
SCAMP1
MLEC
PLAUR
MS4A3
C3AR1
MMP25
ATP6V1D
AGPAT2
PGRMC1
ALDH3B1
ORMDL3
DNAJC5
P1A
GPR97
MOSPD2
TMBIM1
RAB44
ITGAL
BST1
CKAP4
HMOX2
TOM1
HVCN1
CMTM6
SLC15A4
AP1M1
SLCO4C
ITGAV
DGAT1
DEGS1
ANPEP

Figure S1N

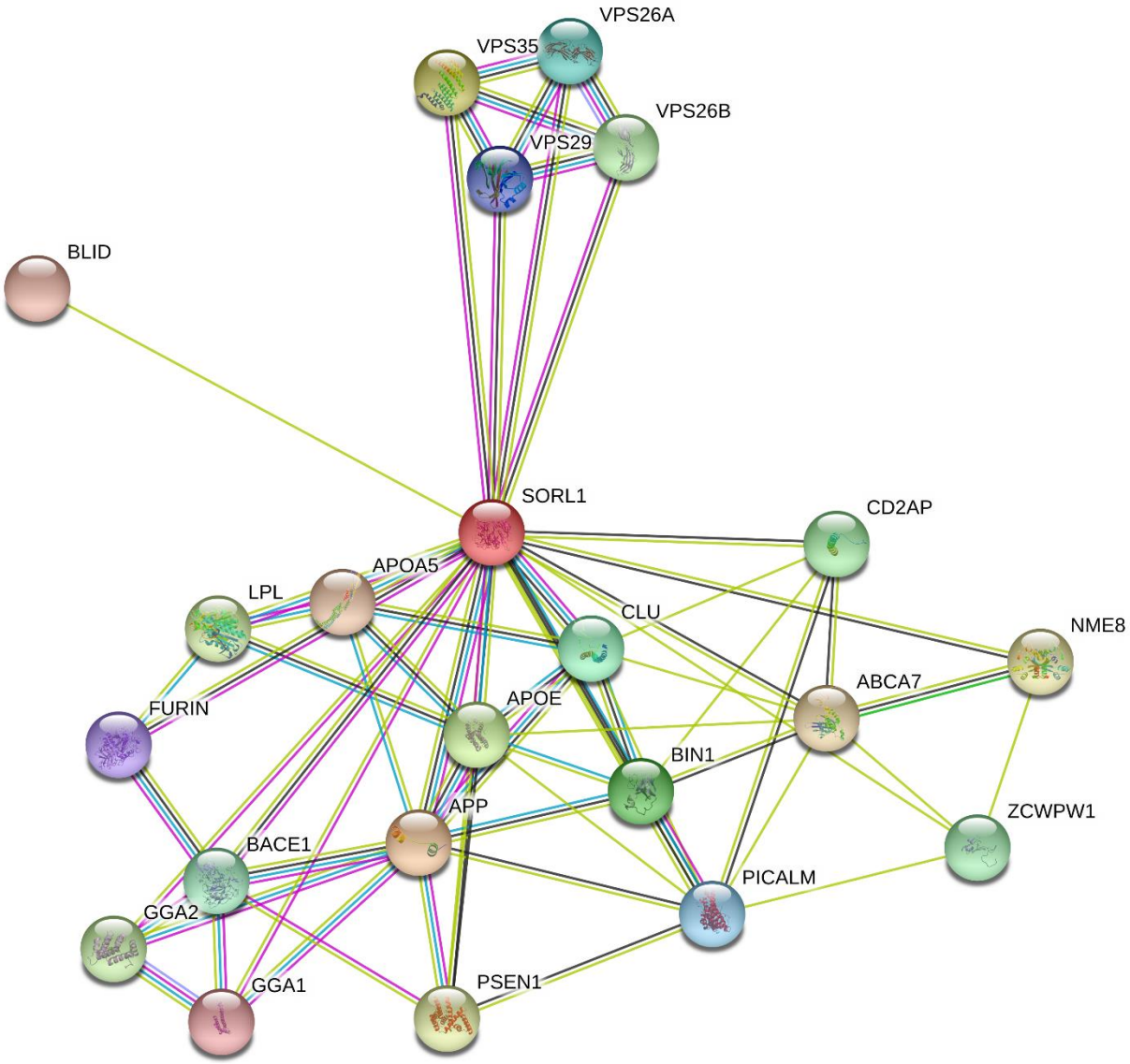

VPS26A
VPS35
VPS26B
VPS29
BLID
SORL1
CD2AP
APOA5
LPL
CLU
NME8
ABCA7
FURIN
APOE
BIN1
APP
ZCWPW1
BACE1
PICALM
GGA2
GGA1
PSEN1

Figure S1O

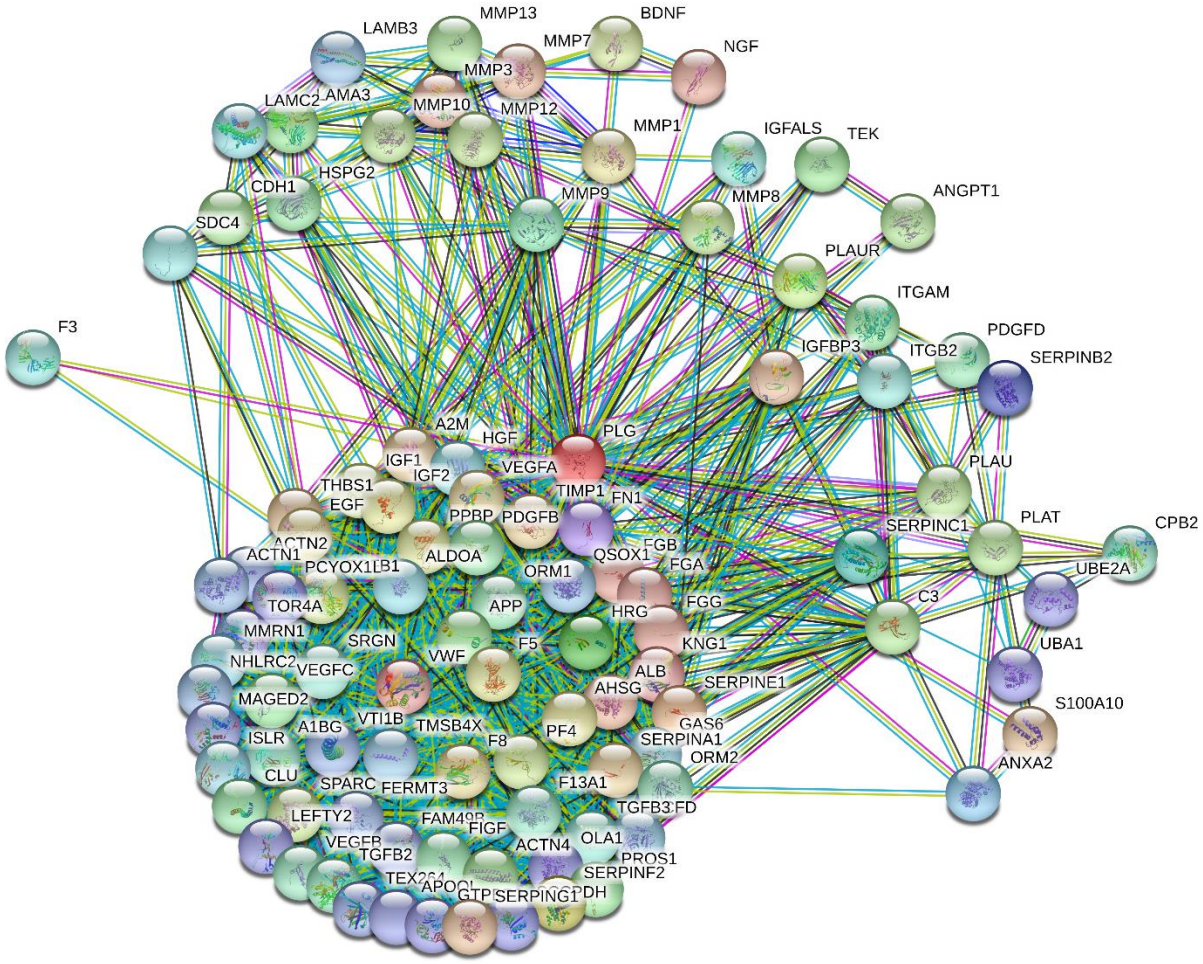
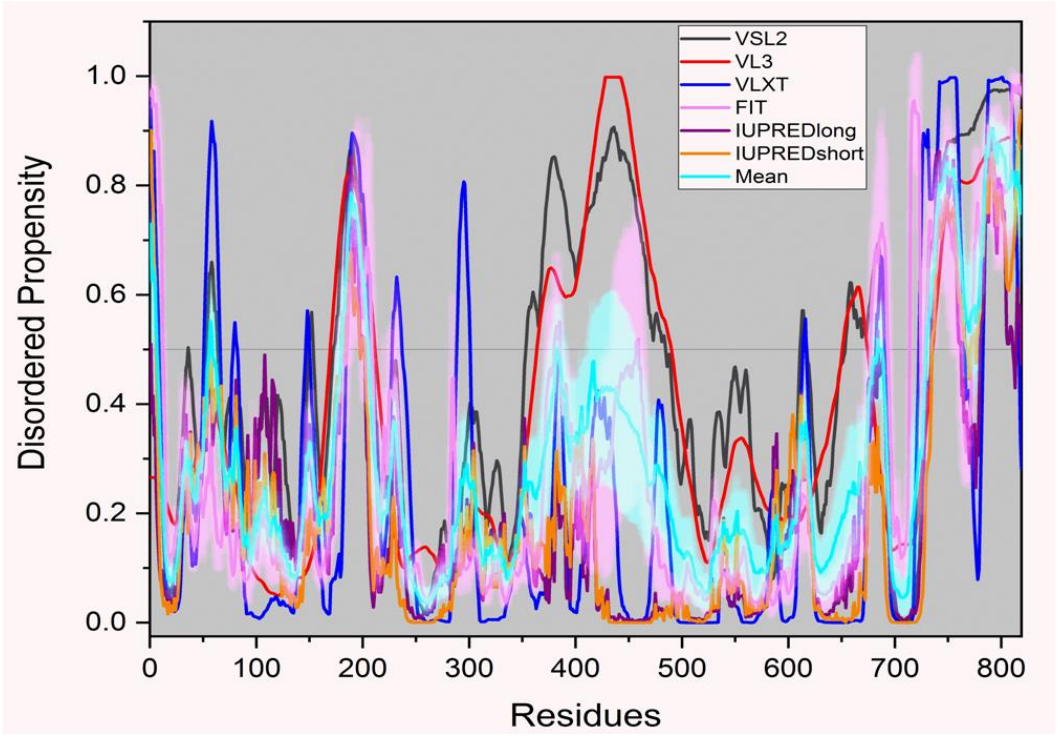


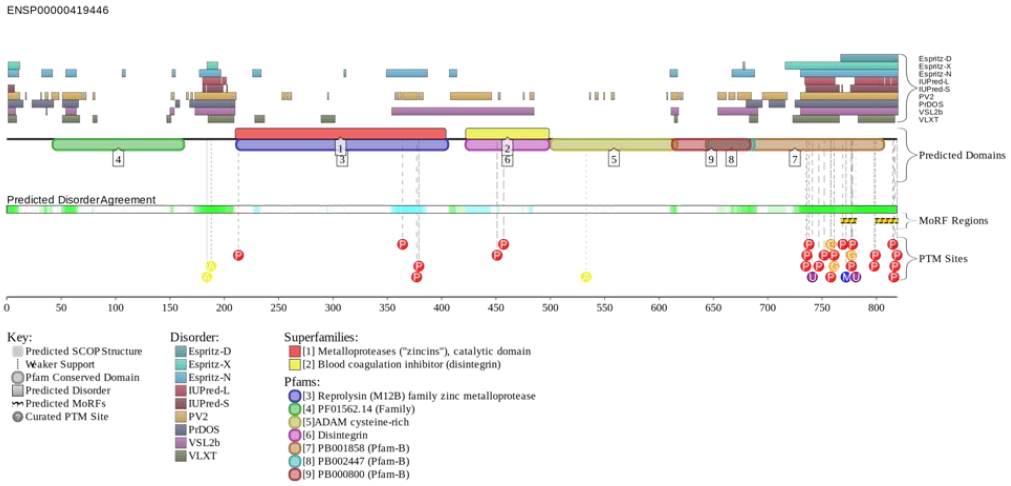
Figure S1P

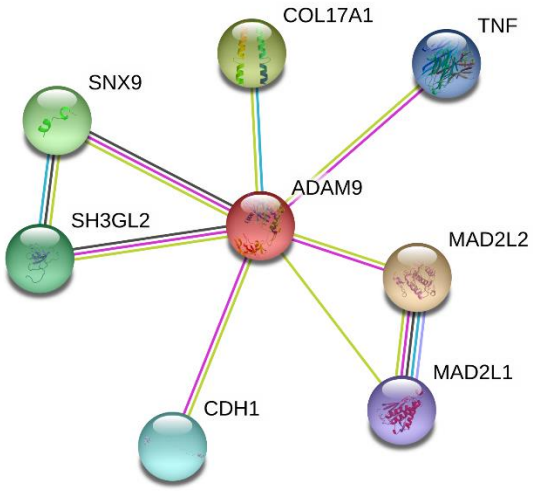
Supplementary Figure S1. Analysis of the interactivity of proteins involved in the amyloid cascade hypothesis. The STRING resource was used to study the interactions of each protein involved in the amyloid cascade hypothesis. (A) APP, (B) ADAM17, (C) BACE1, (D) PSEN1, (E) PSEN2, (F) Nicastrin, (G) APH1A, (H) APH1B, (I) PEN2, (J) APOE4, (K) BIN1, (L) Clusterin, (M) PICALM, (N) CD33, (O) SORL1, and (P) PLG.

(A)


VSL2
VL3
VLXT
FIT
IUPREDlong
IUPREDshort
Mean
Disordered Propensity
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
400
500
600
700
800
Residues

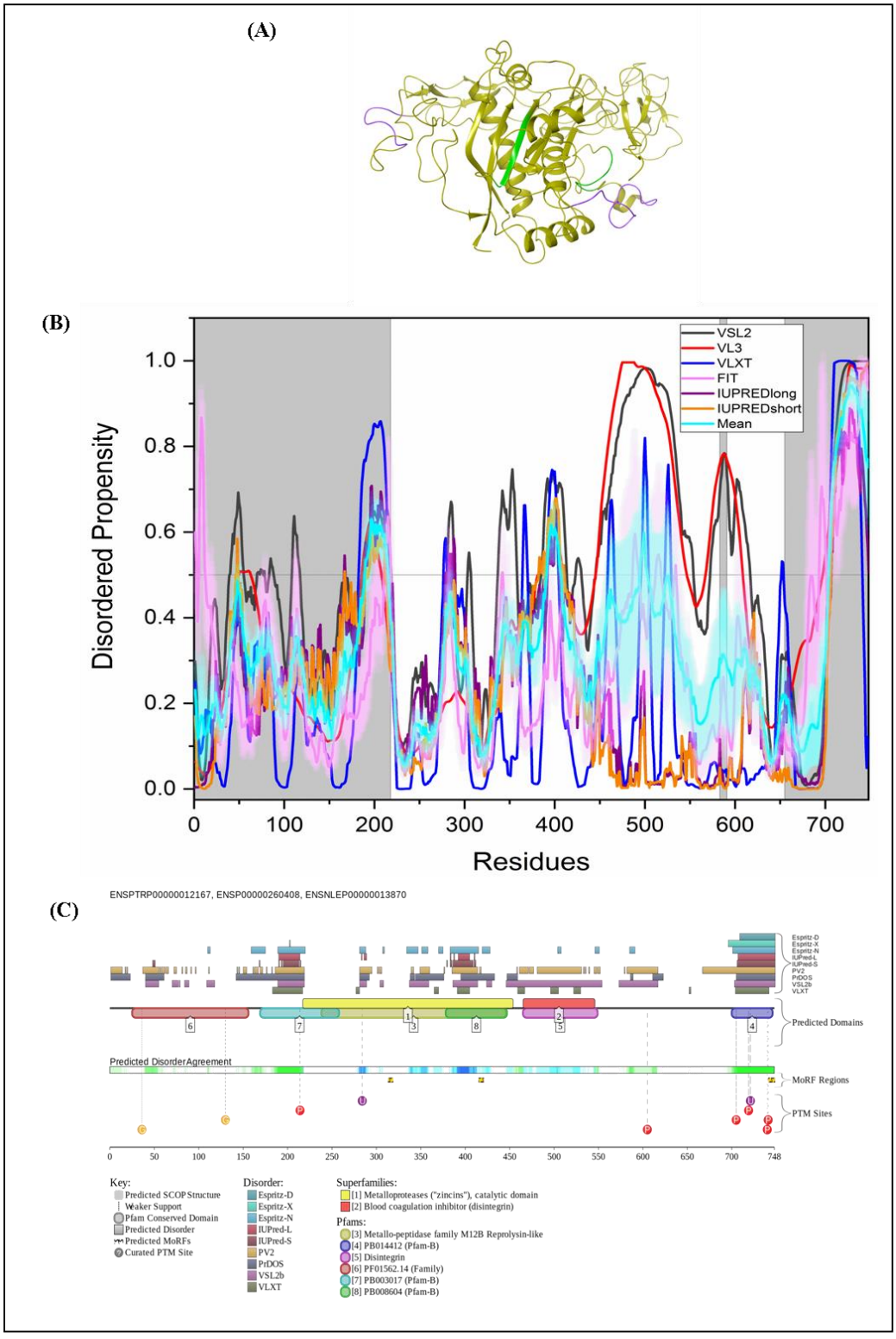
(B)

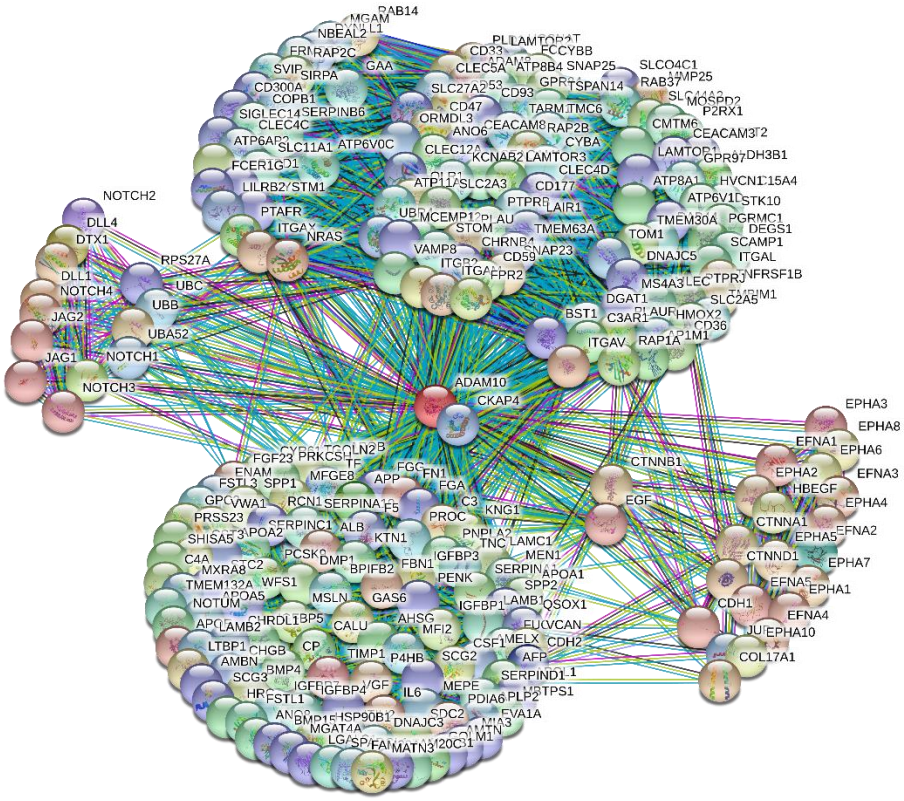

ENSP00000419446
Espritz-D
Espritz-X
Espritz-N
IUPred-L
IUPred-S
PV2
PrDOS
VSL2b
VLXT
Predicted Domains
Predicted DisorderAgreement
MoRF Regions
PTM Sites
0
50
100
150
200
250
300
350
400
450
500
550
600
650
700
750
800
Key:
Predicted SCOP Structure
Weaker Support
Pfam Conserved Domain
Predicted Disorder
Predicted MoRFs
Curated PTM Site
Disorder:
Espritz-D
Espritz-X
Espritz-N
IUPred-L
IUPred-S
PV2
PrDOS
VSL2b
VLXT
Superfamilies:
[1] Metalloproteases ("zincins"), catalytic domain
[2] Blood coagulation inhibitor (disintegrin)
Pfams:
[3] Reprolysin (M12B) family zinc metalloprotease
[4] PF01562.14 (Family)
[5]ADAM cysteine-rich
[6] Disintegrin
[7] PB001858 (Pfam-B)
[8] PB002447 (Pfam-B)
[9] PB000800 (Pfam-B)



(C)

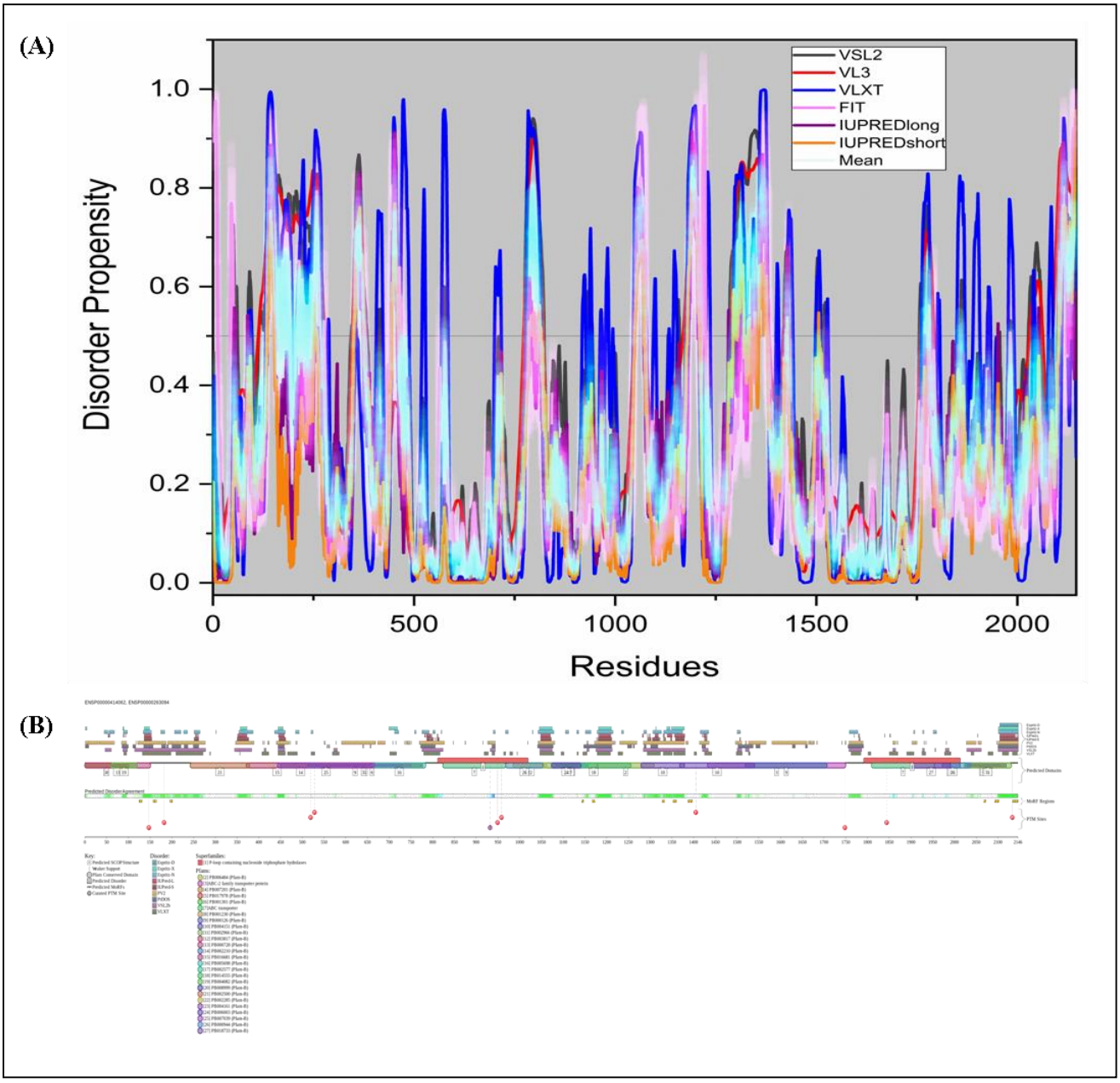
Supplementary Figure S2. Evaluation of intrinsic disorder propensity of ADAM9 (UniProt ID: Q13443): (A) Disorder profile generated by PONDR® VSL2 (grey line), PONDR® VL3 (red line), PONDR® VLXT (blue line), PONDR® FIT (magenta line), IUPRED long (purple line) and IUPRED short (orange line). The cyan line represents the mean disorder propensity calculated using the disorder scores from all six predictors. The light magenta shadow region represents the PONDR® FIT error distribution and the light cyan region represents the mean error distribution. **(B)** Evaluation of MoRF sites and PTMs in ADAM9 by D^2P^2. **(C)** STRING-generated PPI network using the high confidence level of 0.7 includes 8 nodes connected by 9 edges. This network is characterized by the average node degree of 2.25 and an average clustering coefficient of 0.887. The most common biological processes include positive regulation of membrane protein ectodomain proteolysis, regulation of protein catabolic process, lipid tube assembly, negative regulation of ubiquitin protein ligase activity, and regulation of transferase activity, whereas the most common molecular functions associated with the this PPI network are identical protein binding, enzyme binding, protein kinase binding, cell adhesion molecule binding, protein homodimerization activity.

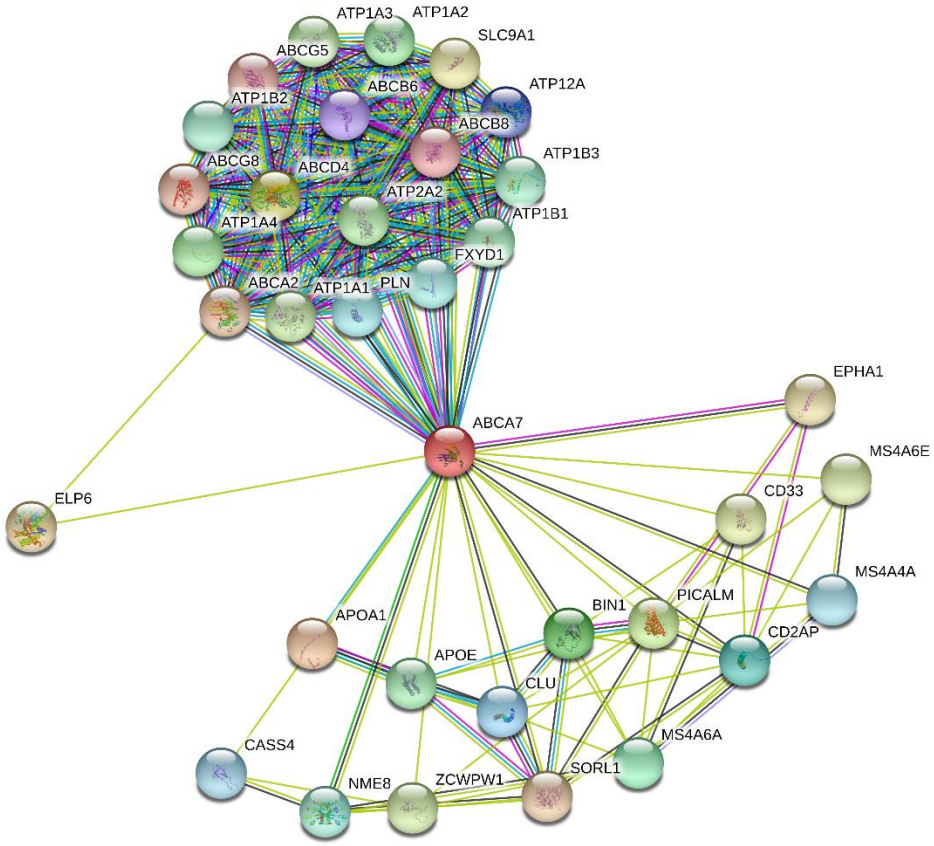

(A)
(B)
VSL2
VL3
VLXT
FIT
IUPREDlong
IUPREDshort
Mean
Disordered Propensity
Residues
(C)
ENSPTRP00000012167, ENSP00000260408, ENSNLEP00000013870
Predicted Domains
Predicted Disorder Agreement
MoRF Regions
PTM Sites
Key:
Predicted SCOP Structure
Weaker Support
Pfam Conserved Domain
Predicted Disorder
Predicted MoRFs
Curated PTM Site
Disorder:
Espritz-D
Espritz-X
Espritz-N
IUPred-L
IUPred-S
PV2
PrDOS
VSL2b
VLXT
Superfamilies:
[1] Metalloproteases ("zincins"), catalytic domain
[2] Blood coagulation inhibitor (disintegrin)
Pfams:
[3] Metallo-peptidase family M12B Reprolysin-like
[4] PB014412 (Pfam-B)
[5] Disintegrin
[6] PF01562.14 (Family)
[7] PB003017 (Pfam-B)
[8] PB008604 (Pfam-B)



(D)

Supplementary Figure S3. Evaluation of intrinsic disorder propensity of ADAM10 (UniProt ID: O14672). (A) 2.80 Å resolution structure of ADAM10 residues 214-654 (PDB ID: 6BE6). The IDPRs (violet color) and MoRF regions (green color) are mapped on the protein structure (olive color). **(B)** Disorder profile generated by PONDR® VSL2 (grey line), PONDR® VL3 (red line), PONDR® VLXT (blue line), PONDR® FIT (magenta line), IUPRED long (purple line) and IUPRED short (orange line). The cyan line represents the mean disorder propensity calculated using the disorder scores from all six predictors. The light magenta shadow region represents the PONDR® FIT error distribution and the light cyan region represents the mean error distribution. The grey shaded area of the graph represents residues which are either missing in the PDB structure or the residues for which no structure is available. **(C)** Evaluation of MoRF sites and PTMs in ADAM10 by D^2P^2. **(D)** STRING-generated PPI network using the highest confidence level of 0.9 includes 253 nodes connected by 11,370 edges, which significantly exceeds the number of edges (1,391) expected for a random set of proteins of similar size. This network is characterized by the average node degree of 89.9, an average clustering coefficient of 0.928, and PPI enrichment p-value < 10^{-16}. Most common biological processes associated with the this network include regulated exocytosis, neutrophil mediated immunity, neutrophil degranulation, exocytosis, and cell activation involved in immune response, and top 5 molecular functions are signaling receptor binding, ephrin receptor activity, cell adhesion molecule binding, protein binding, and peptidase inhibitor activity.

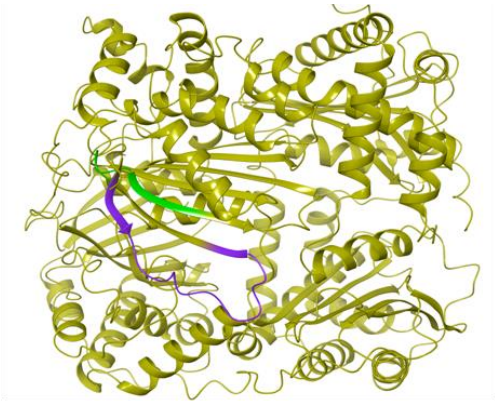

(A)
VSL2
VL3
VLXT
FIT
IUPREDlong
IUPREDshort
Mean
Disorder Propensity
1.0
0.8
0.6
0.4
0.2
0.0
0
500
1000
1500
2000
Residues
(B)



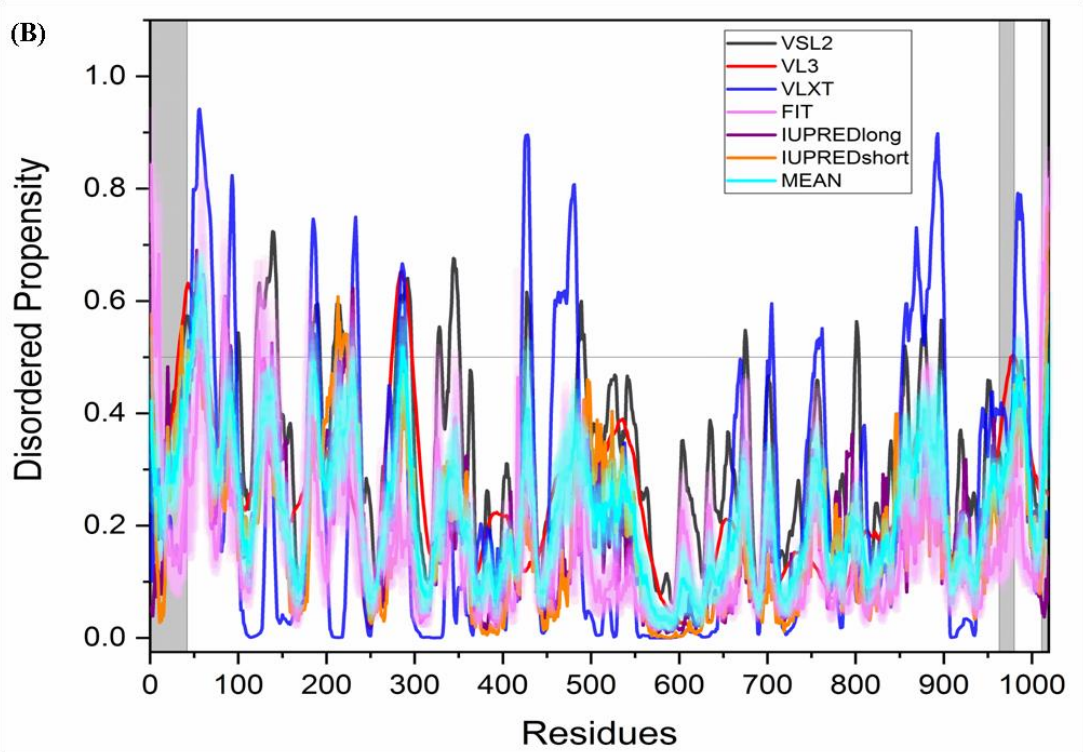
(C)

Supplementary figure 4. Evaluation of intrinsic disorder propensity of ABCA7 (UniProt ID; Q8IZY2): (A**)** Disorder profile generated by PONDR® VSL2 (grey line), PONDR® VL3 (red line), PONDR® VLXT (blue line), PONDR® FIT (magenta line), IUPRED long (purple line) and IUPRED short (orange line). The cyan line represents the mean disorder propensity calculated using the disorder scores from all six predictors. The light magenta shadow region represents the PONDR® FIT error distribution and the light cyan region represents the mean error distribution. **(B)** Evaluation of MoRF sites and PTMs in ABCA7 by D^2P^2. **(C)** STRING-generated PPI network using high confidence level of 0.7 includes 35 nodes connected by 237 edges, which significantly exceeds the number of edges (36) expected for a random set of proteins of similar size. This network is characterized by the average node degree of 13, an average clustering coefficient of 0.882, and PPI enrichment p-value < 10^{-16}. Most common biological processes associated with the this network include establishment or maintenance of transmembrane electrochemical gradient, sodium ion export across plasma membrane, transmembrane transport, cellular sodium ion homeostasis, and ATP hydrolysis coupled ion transmembrane transport, and top 5 molecular functions are ATPase activity, coupled to transmembrane movement of substances, active transmembrane transporter activity, sodium:potassium-exchanging ATPase activity, ATPase activity, coupled to transmembrane movement of ions, phosphorylative mechanism, and transporter activity.

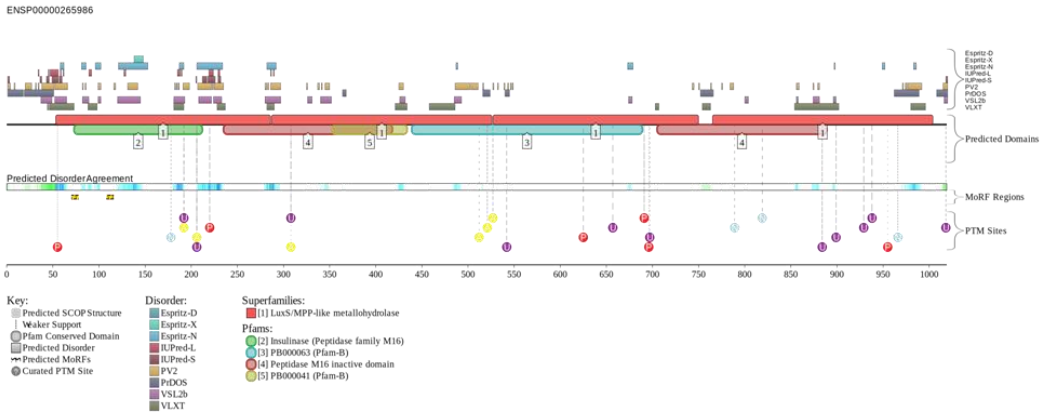
(A)

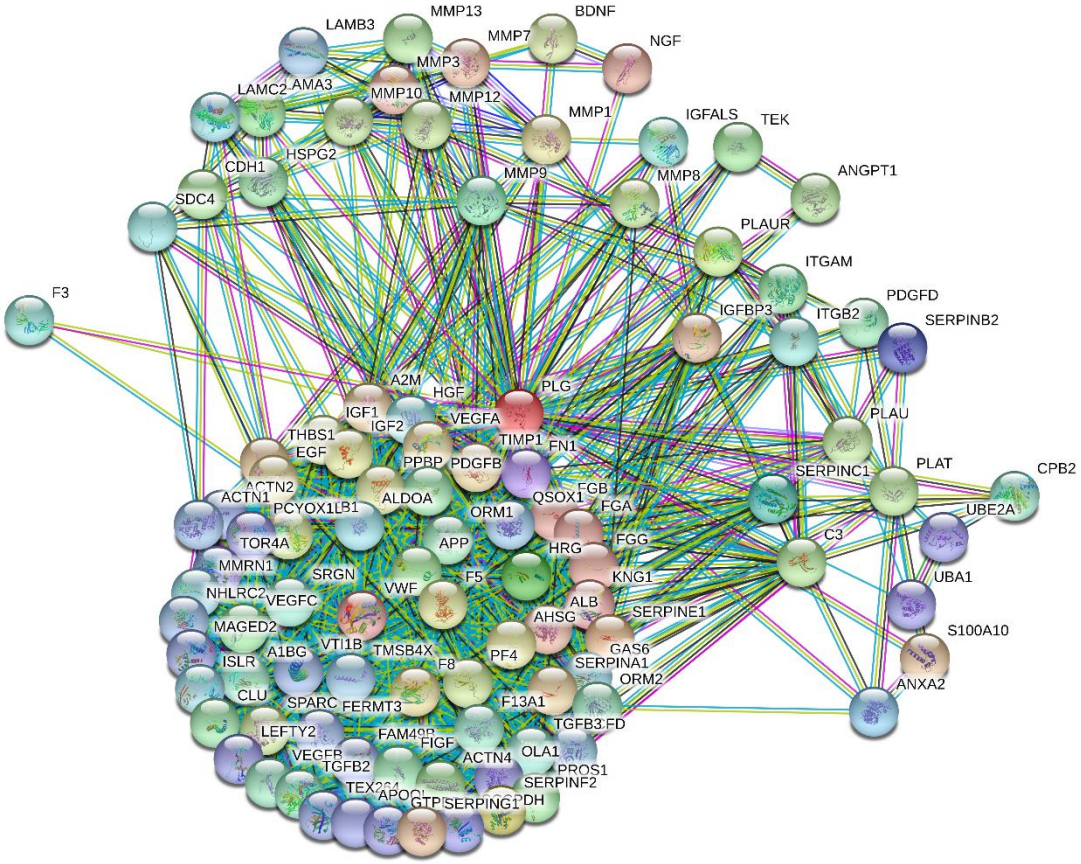


(B)


VSL2
VL3
VLXT
FIT
IUPREDlong
IUPREDshort
MEAN
Disordered Propensity
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
400
500
600
700
800
900
1000
Residues

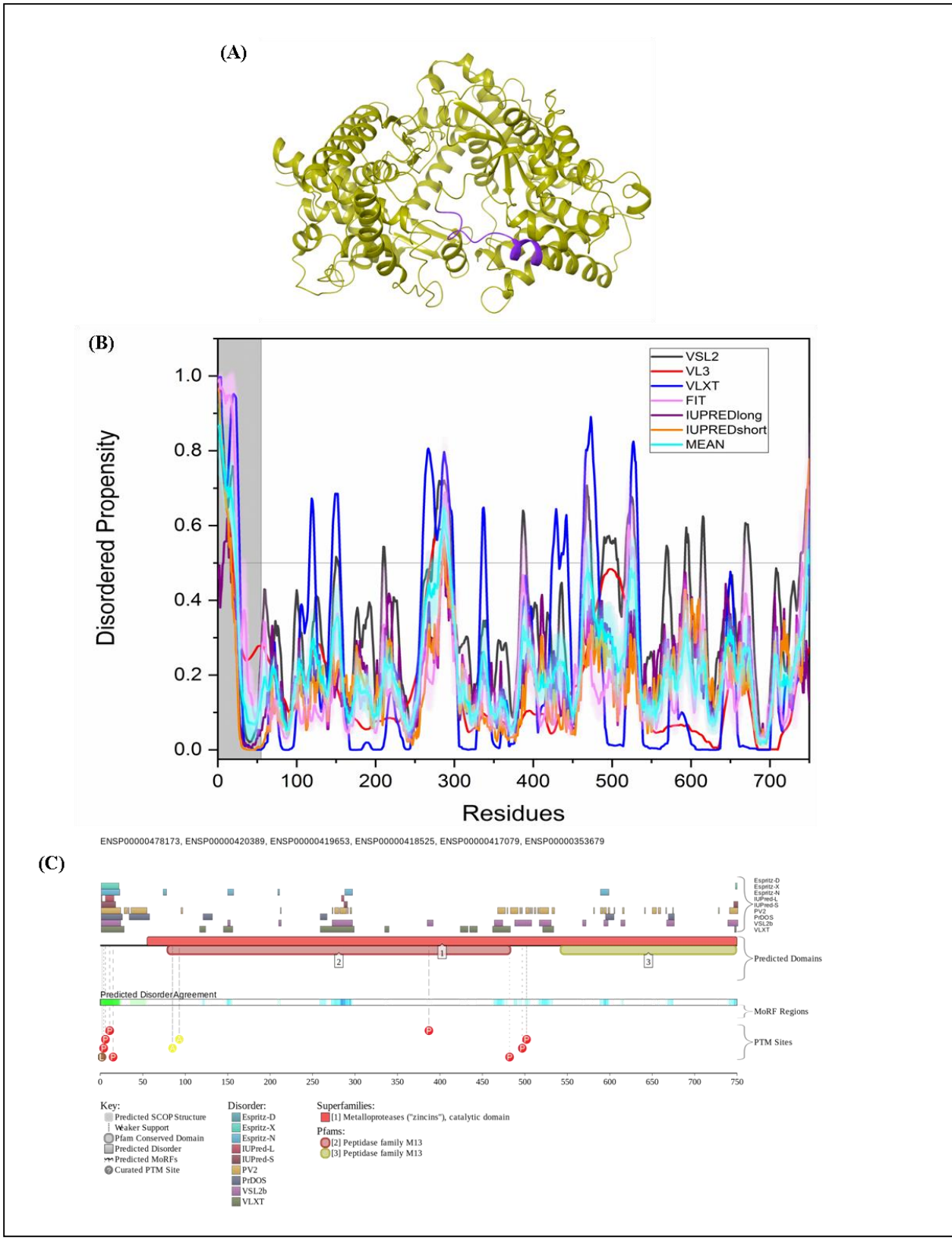
(C)

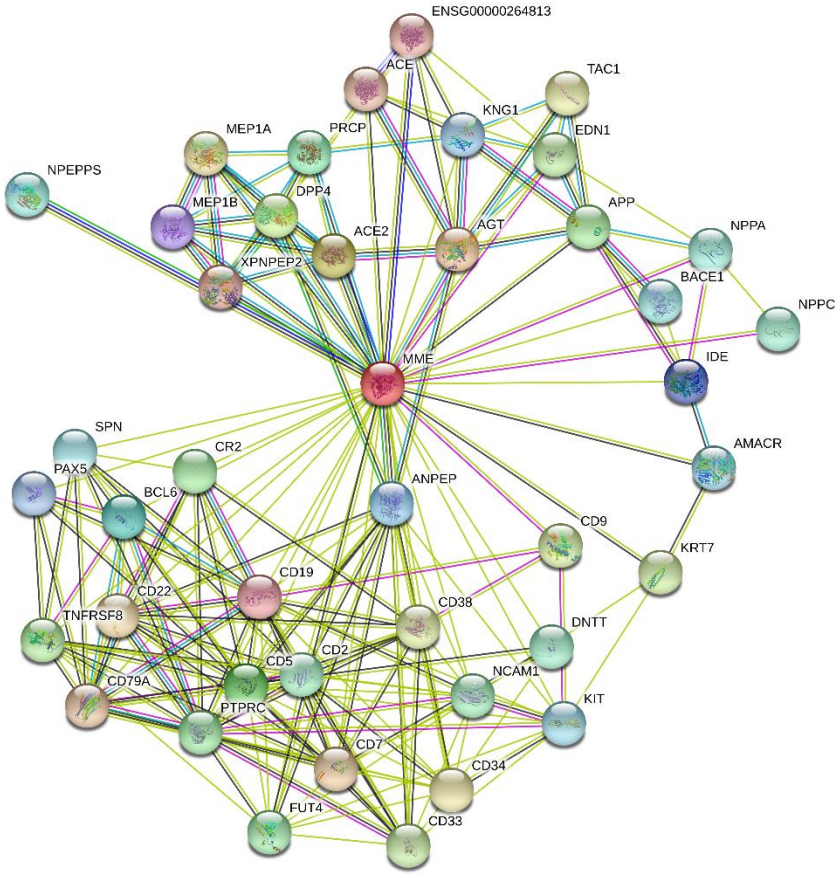

ENSP00000265986
Predicted Domains
Predicted Disorder Agreement
MoRF Regions
PTM Sites
Key:
Disorder:
Superfamilies:
Pfams:



(D)

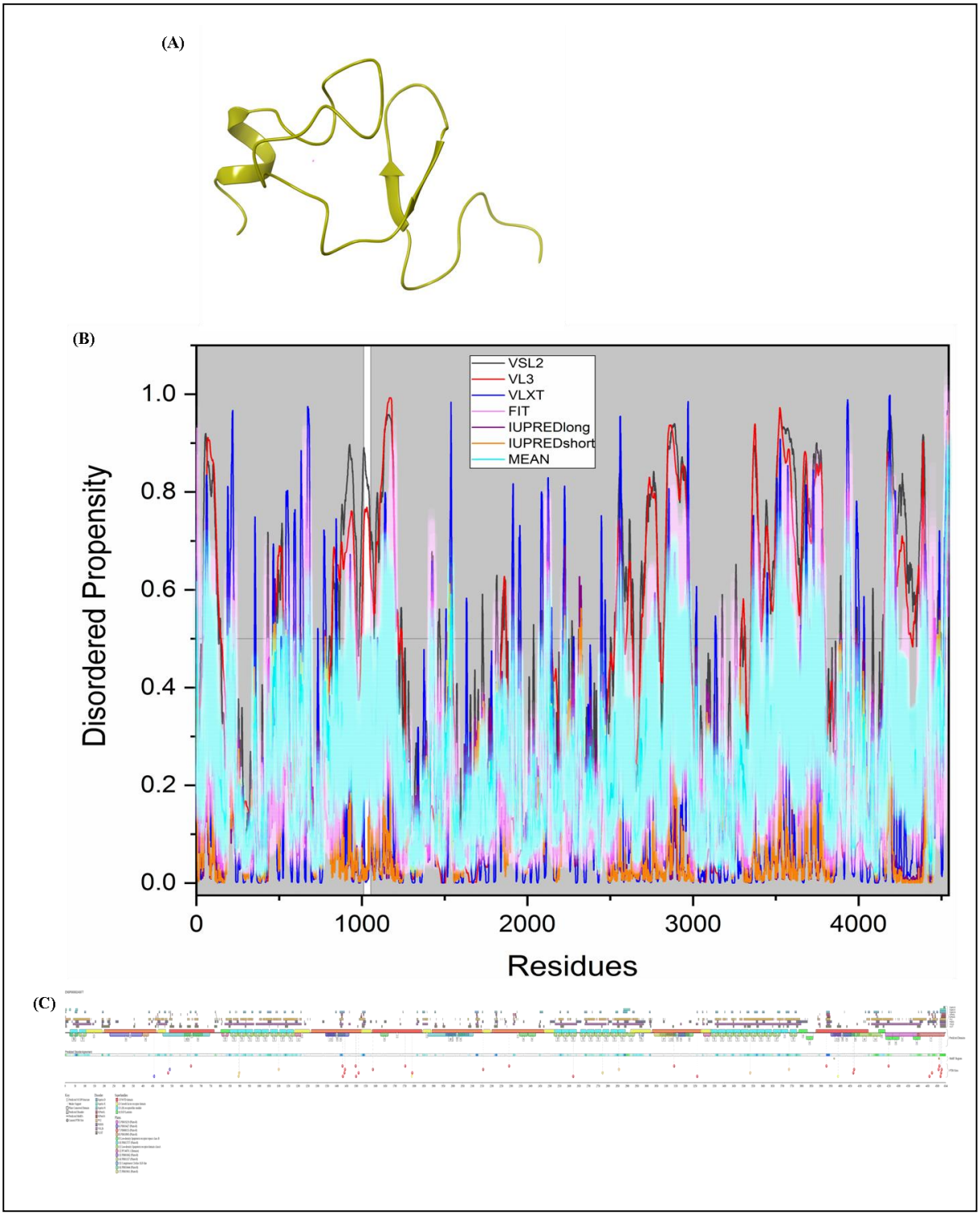
Supplementary Figure S5. Evaluation of intrinsic disorder propensity of IDE (UniProt ID: P14735): (A) 1.96 Å resolution structure of IDE residues 42-1019 (PDB ID: 3CWW). The IDPRs (violet color) and MoRF regions (green color) are mapped on the protein structure (olive color). **(B)** Disorder profile generated by PONDR® VSL2 (grey line), PONDR® VL3 (red line), PONDR® VLXT (blue line), PONDR® FIT (magenta line), IUPRED long (purple line) and IUPRED short (orange line). The cyan line represents the mean disorder propensity calculated using the disorder scores from all six predictors. The light magenta shadow region represents the PONDR® FIT error distribution and the light cyan region represents the mean error distribution. The grey shaded area of the graph represents residues which are either missing in the PDB structure or the residues for which no structure is available. **(C)** Evaluation of MoRF sites and PTMs in IDE by D^2P^2. **(D)** STRING-generated PPI network using the highest confidence level of 0.9 includes 66 nodes connected by 1,357 edges, which significantly exceeds the number of edges (137) expected for a random set of proteins of similar size. This network is characterized by the average node degree of 41.1, an average clustering coefficient of 0.975, and PPI enrichment p-value < 10^{-16}. Most common biological processes associated with the this network include protein targeting to peroxisome, protein targeting, establishment of protein localization to organelle, intracellular transport, and nitrogen compound transport, and top 5 molecular functions are signaling receptor binding, coenzyme binding, protein binding, catalytic activity, and CoA hydrolase activity.

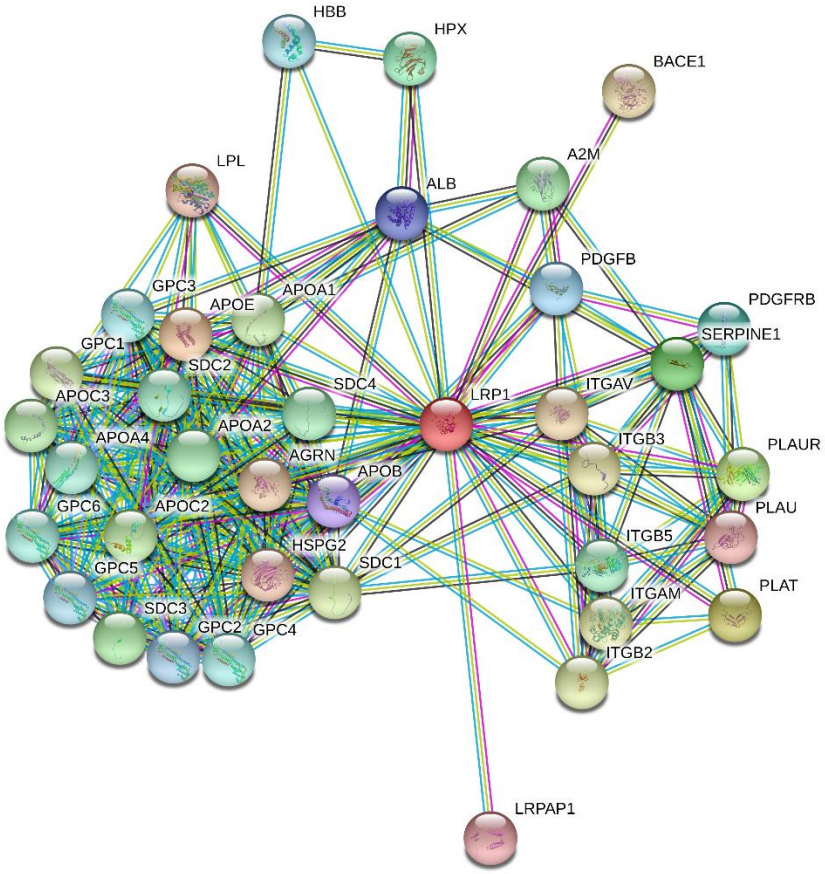

(A)
(B)
Disordered Propensity
Residues
VSL2
VL3
VLXT
FIT
IUPREDlong
IUPREDshort
MEAN
ENSP00000478173, ENSP00000420389, ENSP00000419653, ENSP00000418525, ENSP00000417079, ENSP00000353679
(C)
Predicted Domains
Predicted DisorderAgreement
MoRF Regions
PTM Sites
Key:
Predicted SCOP Structure
Weaker Support
Pfam Conserved Domain
Predicted Disorder
Predicted MoRFs
Curated PTM Site
Disorder:
Espritz-D
Espritz-X
Espritz-N
IUPred-L
IUPred-S
PV2
PrDOS
VSL2b
VLXT
Superfamilies:
[1] Metalloproteases ("zincins"), catalytic domain
Pfams:
[2] Peptidase family M13
[3] Peptidase family M13



(D)

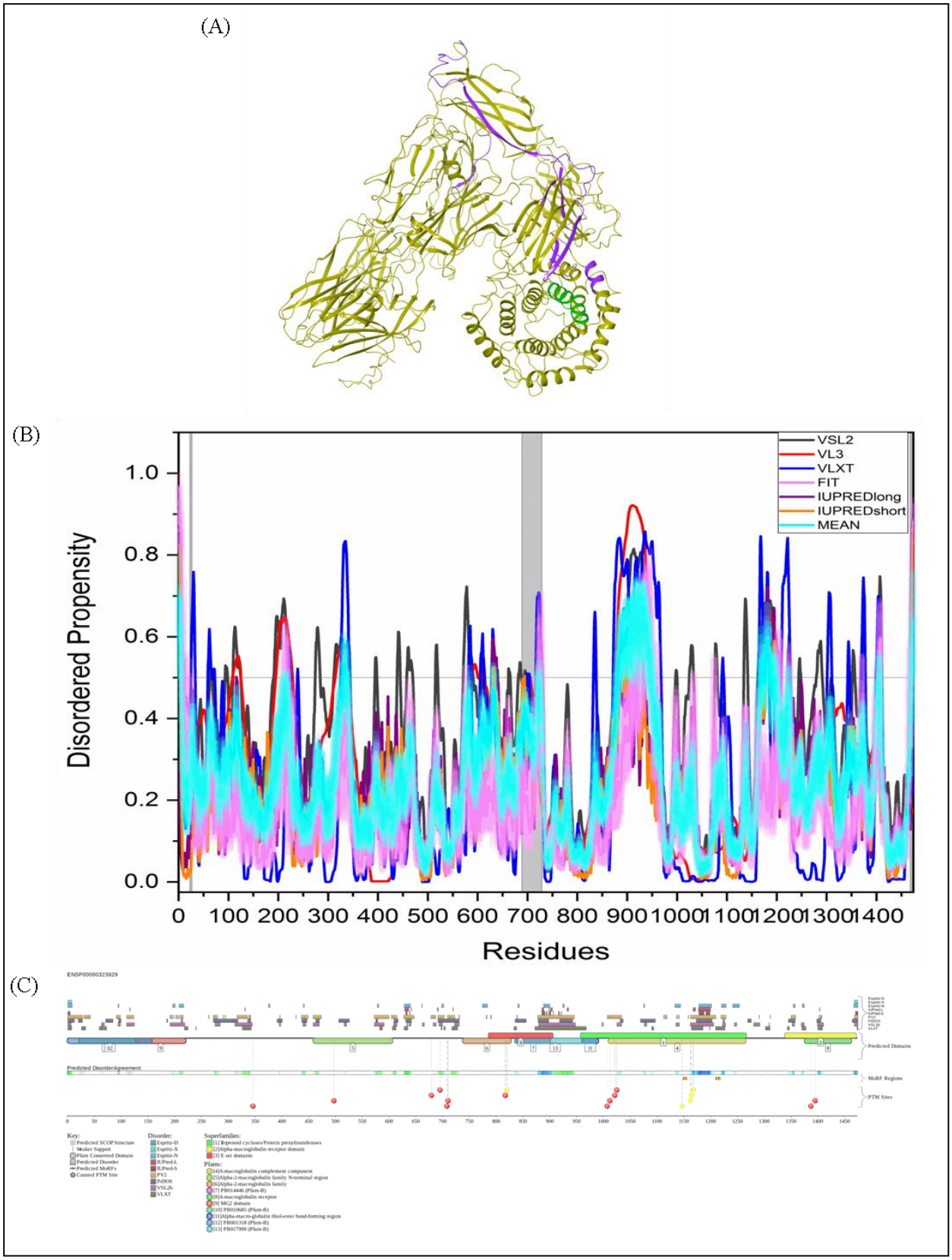
Supplementary Figure S6. Evaluation of intrinsic disorder propensity of Neprilysin (UniProt ID: P08473): (A) 1.90 Å resolution structure of Neprilysin residues 55-750 (PDB ID: 6GID). The IDPRs (violet color) are mapped on the protein structure (olive color). **(B)** Disorder profile generated by PONDR® VSL2 (grey line), PONDR® VL3 (red line), PONDR® VLXT (blue line), PONDR® FIT(magenta line), IUPRED long (purple line)and IUPRED short (orange line). The cyan line represents the mean disorder propensity calculated using the disorder scores from all six predictors. The light magenta shadow region represents the PONDR® FIT error distribution and the light cyan region represents the mean error distribution. The grey shaded area of the graph represents residues for which no structure is available. **(C)** Evaluation of MoRF sites and PTMs in Neprilysin by D^2P [2]. **(D)** STRING-generated PPI network using high confidence level of 0.7 includes 42 nodes connected by 203 edges, which significantly exceeds the number of edges (50) expected for a random set of proteins of similar size. This network is characterized by the average node degree of 9.67, an average clustering coefficient of 0.77, and PPI enrichment p-value $< 10^{-16}$. Most common biological processes associated with the this network include immune system process, blood circulation, regulation of blood pressure, regulation of immune system process, and regulation of systemic arterial blood pressure, and top 5 molecular functions are exopeptidase activity, metallopeptidase activity, peptidase activity, acting on L-amino acid peptides, virus receptor activity, and aminopeptidase activity.

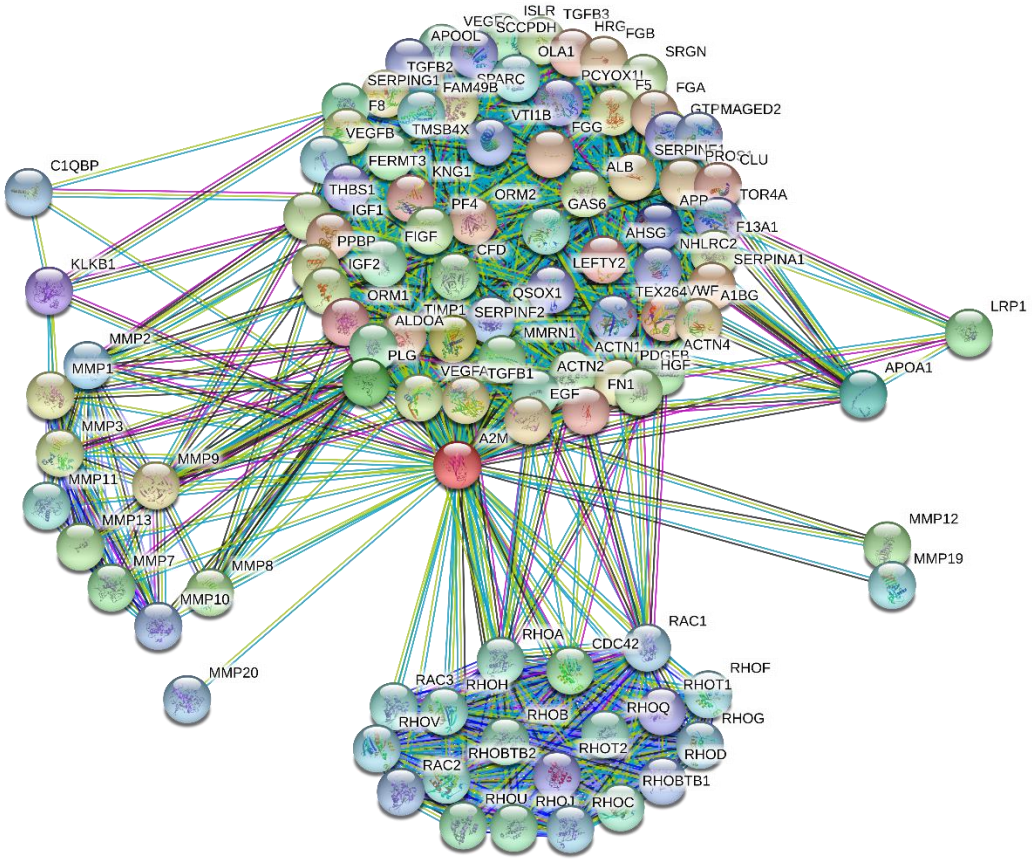

(A)
(B)
VSL2
VL3
VLXT
FIT
IUPREDlong
IUPREDshort
MEAN
Disordered Propensity
1.0
0.8
0.6
0.4
0.2
0.0
0
1000
2000
3000
4000
Residues
(C)



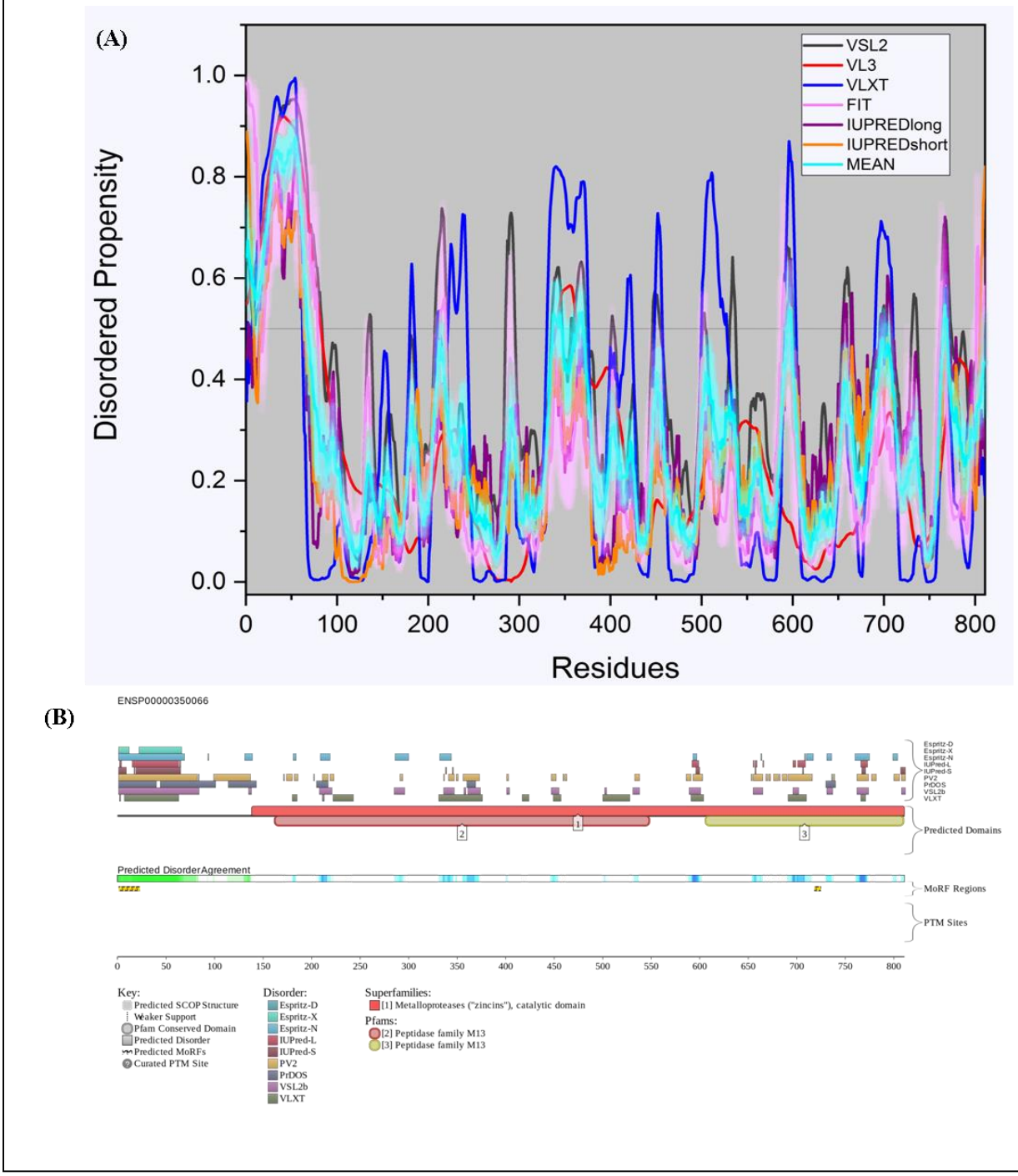
(D)

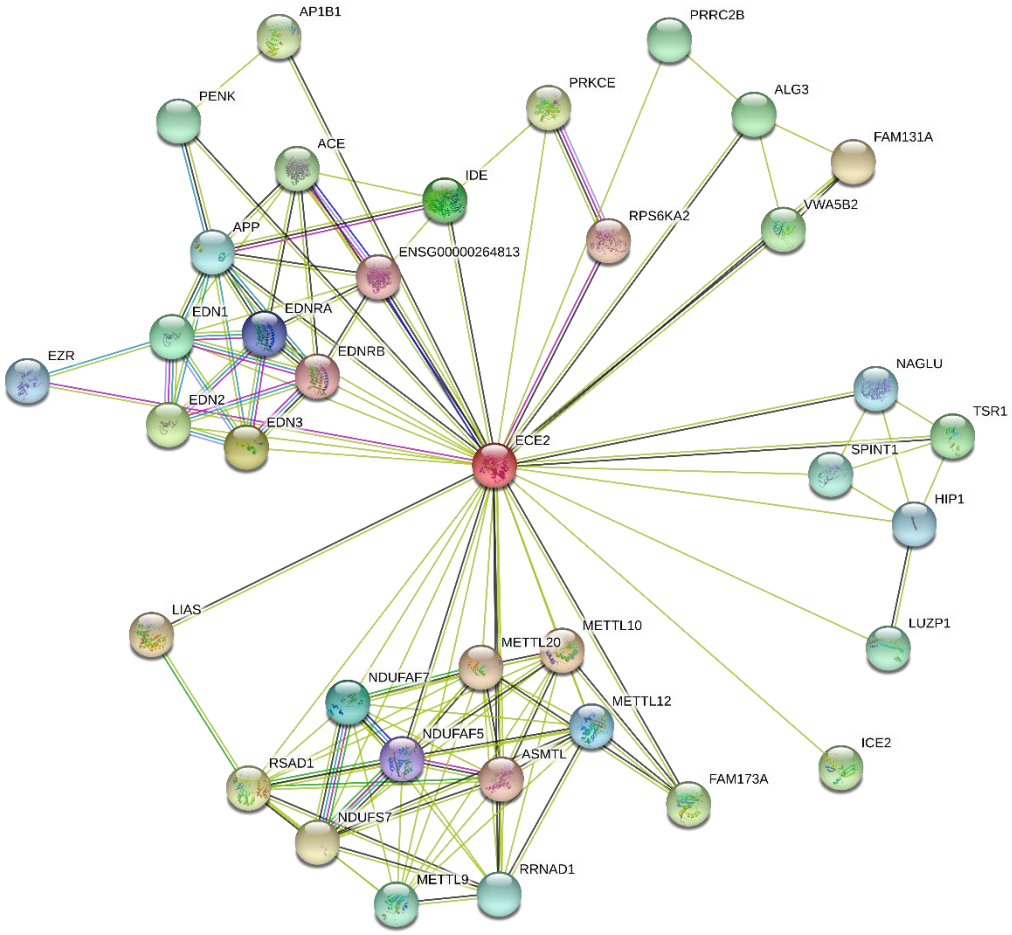
Supplementary Figure S7. Evaluation of intrinsic disorder propensity of LRP1 (UniProt ID: Q07954): (A) 1.85 Å resolution structure (olive color) of LRP1 residues 1011-1054 (PDB ID: 1J8E). **(B)** Disorder profile generated by PONDR® VSL2 (grey line), PONDR® VL3 (red line), PONDR® VLXT (blue line), PONDR® FIT (magenta line), IUPRED long (purple line) and IUPRED short (orange line). The cyan line represents the mean disorder propensity calculated using the disorder scores from all six predictors. The light magenta shadow region represents the PONDR® FIT error distribution and the light cyan region represents the mean error distribution. The grey shaded area of the graph represents residues for which no structure is available. **(C)** Evaluation of MoRF sites and PTMs in LRP1 by D^2P^2. **(D)** STRING-generated PPI network using the highest confidence level of 0.9 includes 38 nodes connected by 275 edges, which significantly exceeds the number of edges (49) expected for a random set of proteins of similar size. This network is characterized by the average node degree of 14.5, an average clustering coefficient of 0.812, and PPI enrichment p-value < 10^{-16}. Most common biological processes associated with the this network include retinoid metabolic process, extracellular structure organization, cellular lipid metabolic process, glycosaminoglycan catabolic process, and glycosaminoglycan biosynthetic process, and top 5 molecular functions are lipoprotein particle receptor binding, cholesterol transporter activity, signaling receptor binding, phosphatidylcholine-sterol O-acyltransferase activator activity, and protein binding.


(A)
(B)
VSL2
VL3
VLXT
FIT
IUPREDlong
IUPREDshort
MEAN
Disordered Propensity
Residues
(C)
ENSP00000323929
Predicted Disorder Agreement
Predicted Domains
MoRF Regions
PTM Sites



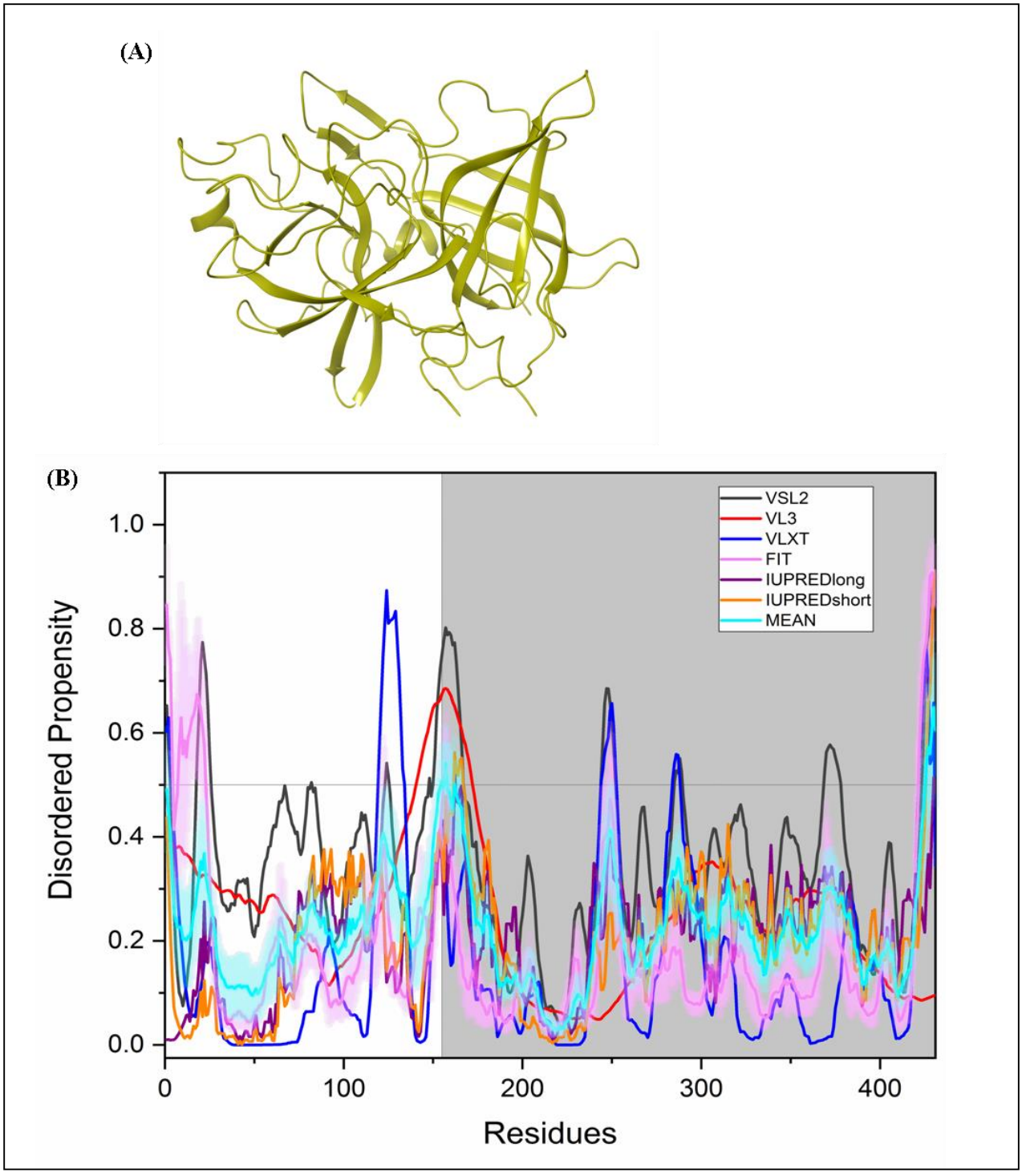
Supplementary Figure S8. Evaluation of intrinsic disorder propensity of A2M (UniProt ID: P01023): (A) 4.30 Å resolution structure of A2M residues 24-1474 (PDB ID: 4ACQ). The IDPRs (violet color) and MoRF regions (green color) are mapped on the protein structure (olive color). **(B)** Disorder profile generated by PONDR® VSL2 (grey line), PONDR® VL3 (red line), PONDR® VLXT (blue line), PONDR® FIT (magenta line), IUPRED long (purple line) and IUPRED short (orange line). The cyan line represents the mean disorder propensity calculated using the disorder scores from all six predictors. The light magenta shadow region represents the PONDR® FIT error distribution and the light cyan region represents the mean error distribution. The grey shaded area of the graph represents residues for which no structure is available. **(C)** Evaluation of MoRF sites and PTMs in A2M by D^2P^2. **(D)** STRING-generated PPI network using the highest confidence level of 0.9 includes 100 nodes connected by 2,385 edges, which significantly exceeds the number of edges (236) expected for a random set of proteins of similar size. This network is characterized by the average node degree of 47.7, an average clustering coefficient of 0.943, and PPI enrichment p-value < 10^{-16}. Most common biological processes associated with the this network include platelet degranulation, regulated exocytosis, vesicle-mediated transport, transport, and extracellular structure organization, and top 5 molecular functions are signaling receptor binding, GTPase activity, nucleoside binding, GTP binding, and growth factor activity.

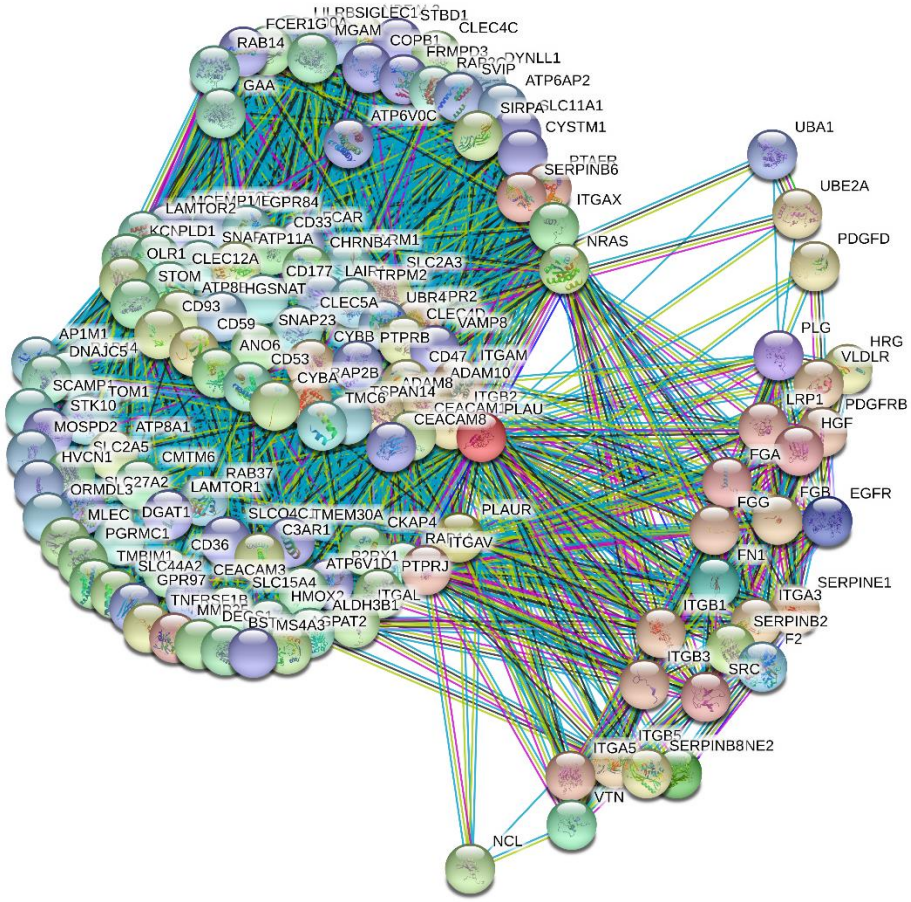

(A)
VSL2
VL3
VLXT
FIT
IUPREDlong
IUPREDshort
MEAN
Disordered Propensity
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
400
500
600
700
800
Residues
(B)
ENSP00000350066
Espritz-D
Espritz-X
Espritz-N
IUPred-L
IUPred-S
PV2
PrDOS
VSL2b
VLXT
Predicted Domains
Predicted Disorder Agreement
MoRF Regions
PTM Sites
0
50
100
150
200
250
300
350
400
450
500
550
600
650
700
750
800
Key:
Predicted SCOP Structure
Weaker Support
Pfam Conserved Domain
Predicted Disorder
Predicted MoRFs
Curated PTM Site
Disorder:
Espritz-D
Espritz-X
Espritz-N
IUPred-L
IUPred-S
PV2
PrDOS
VSL2b
VLXT
Superfamilies:
[1] Metalloproteases ("zincins"), catalytic domain
Pfams:
[2] Peptidase family M13
[3] Peptidase family M13



(C)

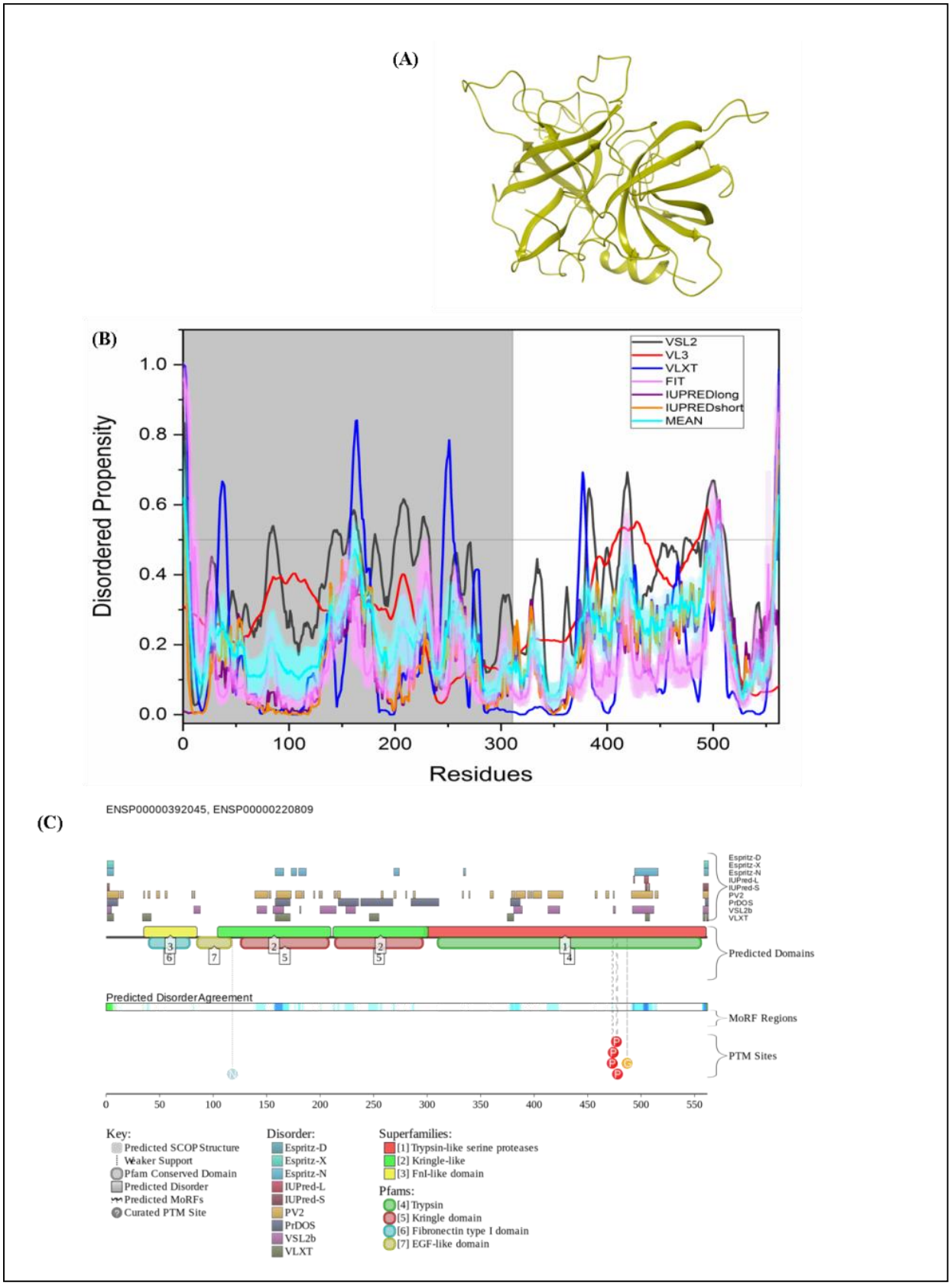
Supplementary Figure S9. Evaluation of intrinsic disorder propensity of ECE2 (UniProt ID: P0DPD6): (A) Disorder profile generated by PONDR® VSL2 (grey line), PONDR® VL3 (red line), PONDR® VLXT (blue line), PONDR® FIT (magenta line), IUPRED long (purple line) and IUPRED short (orange line). The cyan line represents the mean disorder propensity calculated using the disorder scores from all six predictors. The light magenta shadow region represents the PONDR® FIT error distribution and the light cyan region represents the mean error distribution. **(B)** Evaluation of MoRF sites and PTMs in ECE2 by D^2P^2. **(C)** STRING-generated PPI network using the nedium confidence level of 0.4 includes 37 nodes connected by 128 edges, which significantly exceeds the number of edges (45) expected for a random set of proteins of similar size. This network is characterized by the average node degree of 6.92, an average clustering coefficient of 0.885, and PPI enrichment p-value < 10^{-16}. Most common biological processes associated with the this network include vascular smooth muscle contraction, vein smooth muscle contraction, methylation, regulation of blood vessel diameter, and regulation of systemic arterial blood pressure by endothelin, and top 5 molecular functions are methyltransferase activity, endothelin B receptor binding, N-methyltransferase activity, G protein-coupled receptor binding, and S-adenosylmethionine-dependent methyltransferase activity.

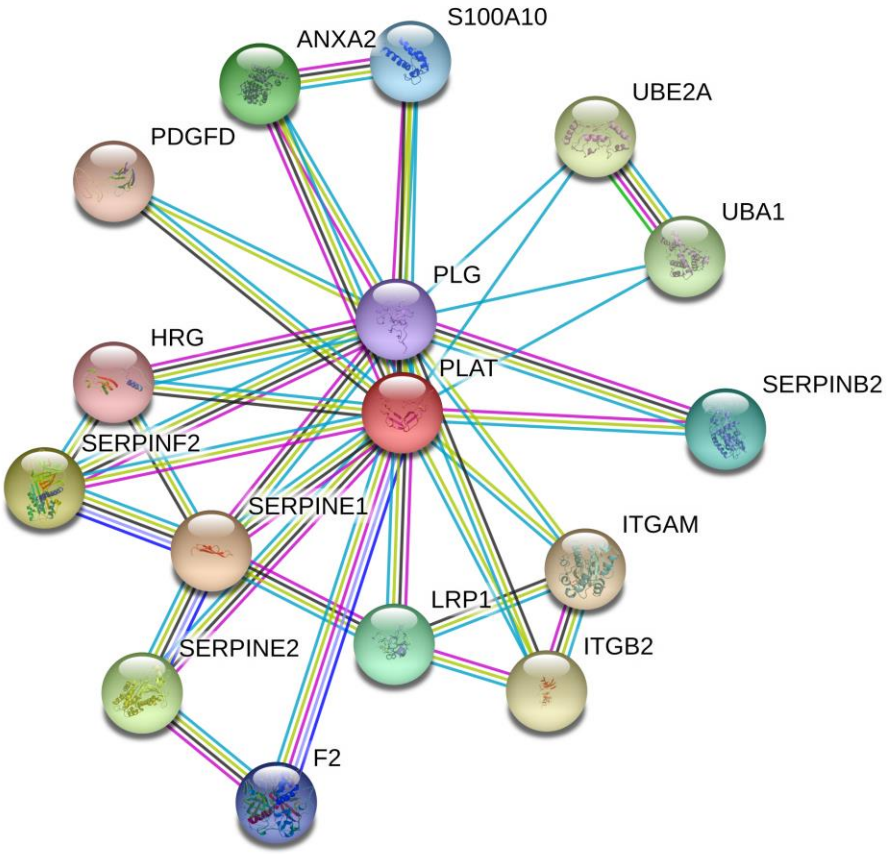

(A)
(B)
VSL2
VL3
VLXT
FIT
IUPREDlong
IUPREDshort
MEAN
Disordered Propensity
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
300
400
Residues



(C)

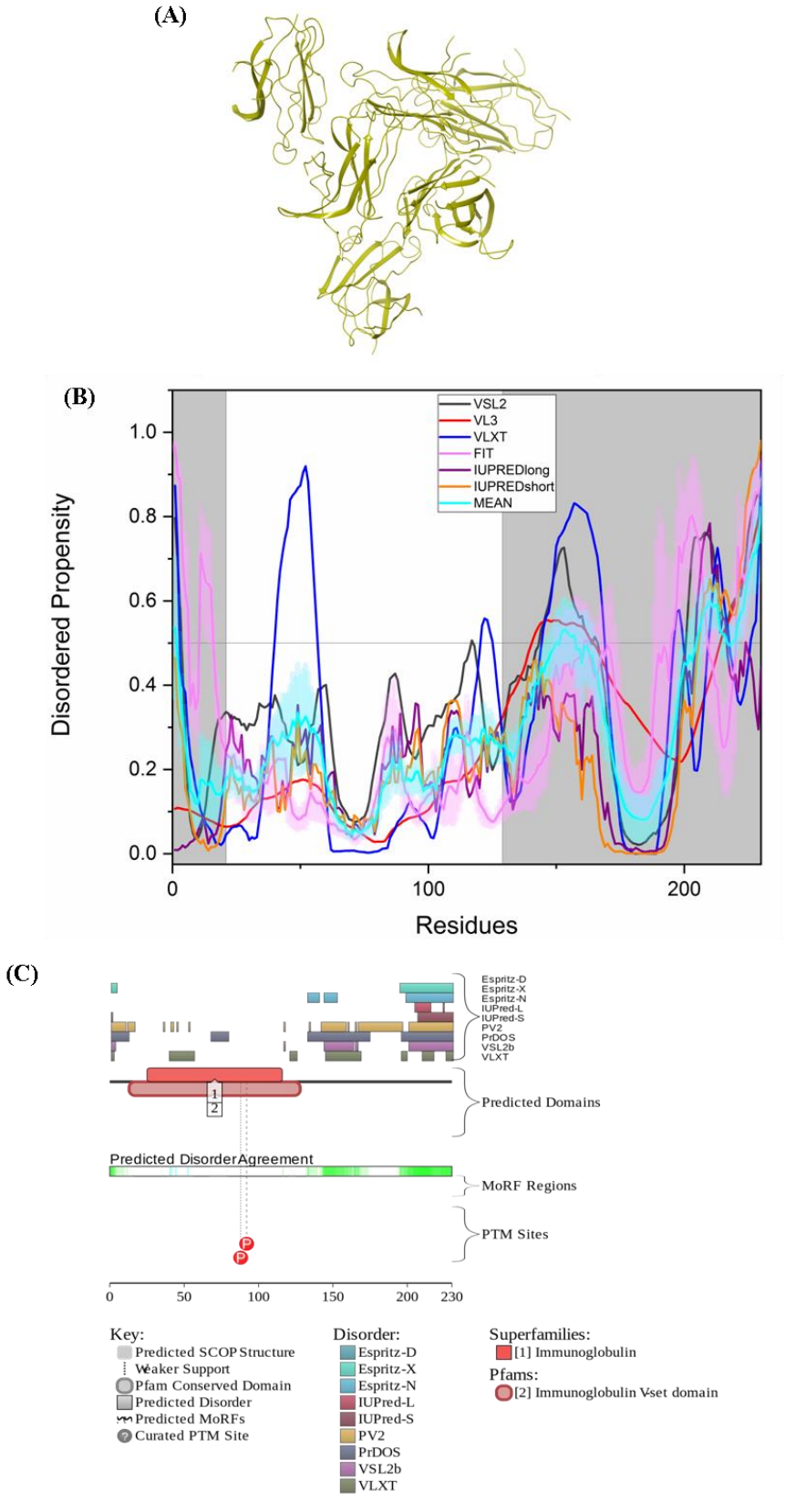
Supplementary Figure S10. Evaluation of intrinsic disorder propensity of PLAU (UniProt ID: P00749): (A) 1.75 Å resolution structure (olive color) of PLAU residues 156-431 (PDB ID: 1C5X). **(B)** Disorder profile generated by PONDR® VSL2 (grey line), PONDR® VL3 (red line), PONDR® VLXT (blue line), PONDR® FIT (magenta line), IUPRED long (purple line) and IUPRED short (orange line). The cyan line represents the mean disorder propensity calculated using the disorder scores from all six predictors. The light magenta shadow region represents the PONDR® FIT error distribution and the light cyan region represents the mean error distribution. The grey shaded area of the graph represents residues for which no structure is available. (No results for PLAU were found in D^2P^2). **(C)** STRING-generated PPI network using the highest confidence level of 0.9 includes 140 nodes connected by 5,552 edges, which significantly exceeds the number of edges (419) expected for a random set of proteins of similar size. This network is characterized by the average node degree of 79.3, an average clustering coefficient of 0.873, and PPI enrichment p-value < 10^{-16}. Most common biological processes associated with the this network include regulated exocytosis, neutrophil mediated immunity, neutrophil degranulation, cell activation involved in immune response, and secretion by cell, and top 5 molecular functions are cell adhesion molecule binding, carbohydrate binding, integrin binding, signaling receptor activity, and signaling receptor binding.

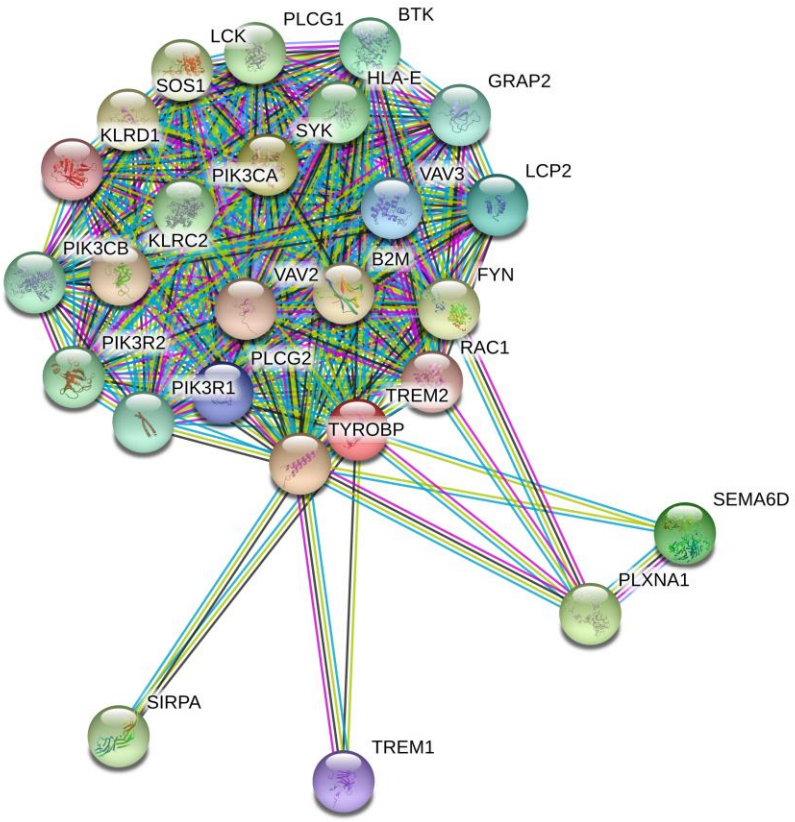

(A)
(B)
VSL2
VL3
VLXT
FIT
IUPREDlong
IUPREDshort
MEAN
Disordered Propensity
Residues
(C)
ENSP00000392045, ENSP00000220809
Predicted Domains
Predicted Disorder Agreement
MoRF Regions
PTM Sites
Key:
Predicted SCOP Structure
Weaker Support
Pfam Conserved Domain
Predicted Disorder
Predicted MoRFs
Curated PTM Site
Disorder:
Espritz-D
Espritz-X
Espritz-N
IUPred-L
IUPred-S
PV2
PrDOS
VSL2b
VLXT
Superfamilies:
[1] Trypsin-like serine proteases
[2] Kringle-like
[3] FnI-like domain
Pfams:
[4] Trypsin
[5] Kringle domain
[6] Fibronectin type I domain
[7] EGF-like domain



(D)

Supplementary Figure S11. Evaluation of intrinsic disorder propensity of PLAT (UniProt ID: P00750): (A**)** 2.90 Å resolution structure (olive color) of PLAT residues 311 – 562 (PDB ID: 1A5H). **(B)** Disorder profile generated by PONDR® VSL2 (grey line), PONDR® VL3 (red line), PONDR® VLXT (blue line), PONDR® FIT (magenta line), IUPRED long (purple line) and IUPRED short (orange line). The cyan line represents the mean disorder propensity calculated using the disorder scores from all six predictors. The light magenta shadow region represents the PONDR® FIT error distribution and the light cyan region represents the mean error distribution. The grey shaded area of the graph represents residues for which no structure is available. **(C)** Evaluation of MoRF sites and PTMs in PLAT by D^2P^2. **(D)** STRING-generated PPI network using the highest confidence level of 0.9 includes 16 nodes connected by 37 edges, which significantly exceeds the number of edges (17) expected for a random set of proteins of similar size. This network is characterized by the average node degree of 4.62, an average clustering coefficient of 0.813, and PPI enrichment p-value < 1.49×10^{-5}. Most common biological processes associated with the this network include fibrinolysis, negative regulation of blood coagulation, regulation of response to external stimulus, negative regulation of fibrinolysis, and regulation of body fluid levels, and top 5 molecular functions are serine-type endopeptidase inhibitor activity, enzyme inhibitor activity, protease binding, enzyme regulator activity, and signaling receptor binding.


(A)
(B)
VSL2
VL3
VLXT
FIT
IUPREDlong
IUPREDshort
MEAN
Disordered Propensity
1.0
0.8
0.6
0.4
0.2
0.0
0
100
200
Residues
(C)
Espritz-D
Espritz-X
Espritz-N
IUPred-L
IUPred-S
PV2
PrDOS
VSL2b
VLXT
1
2
Predicted Domains
Predicted Disorder Agreement
MoRF Regions
PTM Sites
0
50
100
150
200
230
Key:
Predicted SCOP Structure
Weaker Support
Pfam Conserved Domain
Predicted Disorder
Predicted MoRFs
Curated PTM Site
Disorder:
Espritz-D
Espritz-X
Espritz-N
IUPred-L
IUPred-S
PV2
PrDOS
VSL2b
VLXT
Superfamilies:
[1] Immunoglobulin
Pfams:
[2] Immunoglobulin V-set domain



(D)

Supplementary Figure S12. Evaluation of intrinsic disorder propensity of TREM2 (UniProt ID: P00750): (A) 2.20 Å resolution structure (olive color) of TREM2 residues 19-174 (PDB ID: 6B8O). **(B)** Disorder profile generated by PONDR® VSL2 (grey line), PONDR® VL3 (red line), PONDR® VLXT (blue line), PONDR® FIT (magenta line), IUPRED long (purple line) and IUPRED short (orange line). The cyan line represents the mean disorder propensity calculated using the disorder scores from all six predictors. The light magenta shadow region represents the PONDR® FIT error distribution and the light cyan region represents the mean error distribution. The grey shaded area of the graph represents residues for which no structure is available. **(C)** Evaluation of MoRF sites and PTMs in TREM2 by D^2P^2. **(D)** STRING-generated PPI network using the highest confidence level of 0.9 includes 26 nodes connected by 242 edges, which significantly exceeds the number of edges (38) expected for a random set of proteins of similar size. This network is characterized by the average node degree of 18.6, an average clustering coefficient of 0.965, and PPI enrichment p-value < 1.49×10^{-5}. Most common biological processes associated with the this network include Fc-epsilon receptor signaling pathway, regulation of immune response, Fc receptor signaling pathway, immune response-regulating cell surface receptor signaling pathway, and immune response-activating cell surface receptor signaling, and top 5 molecular functions are phosphotyrosine residue binding, phosphatidylinositol-4,5-bisphosphate 3-kinase activity, protein binding, signaling receptor binding, and non-membrane spanning protein tyrosine kinase activity.

Supplementary Table 1. SLiMs associated with MoRF residues and their annotation.

Protein	MoRF residues (D2P2)	SLiM Position	SLiM Sequence	Position w.r.t. MoRF	Type of ELM annotation	SLiM Description
BIN1	67-72 (6)	335-337	LRK	Close proximity	ELM_CLV	N-Arg dibasic convertase (NRD/Nardilysin) cleavage site (X-|-R-K or R-|-R-X).
	234-246 (13)	351-355	KEVKQ	Overlap	ELM_CLV	Subtilisin/kexin isozyme-1 (SKI1) cleavage site ([RK]-X-[hydrophobic]-[LTKF]-|-X).
	260-286 (27)	360-370	SLFEDTFVPEI	Embedded	ELM_DEG	A destruction motif interacts with the COP1 WD 40 domain for target ubiquitination and degradation.
	304-336 (33)	363-370	EDTFVPEI	Embedded	ELM_DEG	
	353-402 (50)	345-352	PKHTPSKE	Close proximity	ELM_DEG	The TPxxE phospho-dependent degron binds the FBW7 F box proteins of the SCF (Skp1_Cullin-Fbox) complex.
	411-452 (42)	485-489	AASSS	Embedded	ELM_DEG	The S/T rich motif known as the SPOP-binding consensus (SBC) of the MATH-BTB protein, SPOP, is present in substrates that undergo SPOP/Cul3-dependant ubiquitination.
	456-475 (20)	372-377	VTTPSQ	Embedded	ELM_DOC	Phospho-dependent motif that mediates docking of CDK substrates and regulators to cyclin-CDK-bound Cks1
	478-503 (26)	351-361	KEVKQEQILSL	Overlap	ELM_DOC	MAPK interacting molecules (e.g. MAPKKs, substrates, phosphatases) carry docking motif that help to regulate specific interaction in the MAPK cascade. The classic motif approximates (R/K)xxxx#x# where # is a hydrophobic residue.
	513-528 (16)	542-550	KAGDVVLVI	Embedded	ELM_DOC	A kinase docking motif that mediates interaction towards the ERK1/2 and p38 subfamilies of MAP kinases.
	541-552 (12)	515-518	LDLP	Embedded	ELM_DOC	Docking motif in calcineurin substrates that binds at the interface of

						the catalytic CNA and regulatory CNB subunits.
		307-313	TPEIRVN	Embedded	ELM_DOC	Calcineurin substrate docking site, leads to the effective dephosphorylation of serine/threonine phosphorylation sites.
		367-373	VPEISVT	Embedded	ELM_DOC	
		283-287	AQPSD	Overlap	ELM_DOC	The USP7 MATH domain binding motif variant based on the MDM2 and p53 interactions.
		408-412	PTPSG	Overlap		
		426-430	PAGSL	Embedded		
		459-463	AEASE	Embedded		
		485-489	AASSS	Embedded		
		486-490	ASSSL	Embedded		
		300-305	PDGSPA	Overlap	ELM_DOC	The Class IV WW domain interaction motif is recognised primarily by the Pin1 phosphorylation-dependent prolyl isomerase.
		304-309	PAATPE	Embedded		
		320-325	GGATPG	Embedded		
		328-333	LPKSPS	Embedded		
		371-376	SVTTPS	Embedded		
		400-405	PVTSPV	Overlap		
		406-411	KAPTPS	Overlap		
		422-427	PTESPA	Embedded		
		362-366	FEDTF	Embedded	ELM_LIG	FxDxF motif responsible for the binding of accessory endocytic proteins to the appendage of the alpha-subunit of adaptor protein complex AP-2
		359-364	FEDTF	Embedded	ELM_LIG	Amphipathic motif that is involved in APC/C inhibition by binding of CDH1/CDC20. In metazoan cyclin A, the motif also acts as a degron, enabling the cyclin's degradation in prometaphase.
		275-279	GSNTF	Embedded	ELM_LIG	Phosphopeptide motif which directly interacts with the BRCT (carboxy-terminal) domain of the Breast Cancer Gene BRCA1 with low affinity
		390-394	LLDLD	Embedded	ELM_LIG	Clathrin box motif found on cargo adaptor proteins, it interacts with the beta propeller structure located at the N-terminus of Clathrin heavy chain.

		415-420	IPWDLW	Embedded	ELM_LIG	Clathrin box motif found on cargo adaptor proteins, it mediates binding to the N-terminal beta propeller of clathrin heavy chain. Also called W box, it is found in the central region of Amphiphysins where it coexists with a "classical" clathrin box.
		276-282	SNTFTVK	Embedded	ELM_LIG	Phosphothreonine motif binding a subset of FHA domains that show a preference for a large aliphatic amino acid at the pT+3 position.
		305-311	AATPEIR	Embedded		
		400-406	PVTSPVK	Overlap		
		439-445	EGTFAVS	Embedded		
		534-540	TDTDELQ	Close proximity		
		415-422	IPWDLWEP	Embedded	ELM_LIG	This short WD or WE motif is found in cargo proteins and mediates kinesin-1-dependent microtubule transport by binding to the KLC TPR region.
		392-398	DLDFDPL	Embedded	ELM_LIG	Canonical LIR motif that binds to Atg8 protein family members to mediate processes involved in autophagy.
		394-398	DFDPL	Embedded		
		362-366	FEDTF	Embedded	ELM_LIG	Fxxx[WF] motifs are present in Pex19 and S. cerevisiae Pex5 cytosolic receptors that bind to peroxisomal membrane docking member, Pex14
		442-446	FAVSW	Embedded		
		336-342	RKGPPVP	Overlap	ELM_LIG	This is the motif recognized by class I SH3 domains
		337-343	KGPPVPP	Close proximity		
		326-332	ATLPKSP	Embedded	ELM_LIG	This is the motif recognized by those SH3 domains with a non-canonical class I recognition specificity
		336-342	RKGPPVP	Overlap		
		337-343	KGPPVPP	Close proximity		
		338-344	GPPVPPP	Close proximity		
		369-375	EISVTTP	Embedded		
		394-400	DFDPLPP	Embedded		
		398-404	LPPVTSP	Overlap		
		402-408	TSPVKAP	Overlap		
		441-447	TFAVSWP	Embedded		
		452-458	EPGPAQP	Overlap		
		356-365	EQILSLFEDT	Embedded	ELM_LIG	

		357-365	QILSLFEDT	Embedded		Motif for the parallel beta augmentation mode of non-covalent binding to SUMO protein.
		389-394	SLLDLD	Embedded		
		456-464	AQPAEASEV	Embedded	ELM_LIG	TRAF6 binding site. Members of the tumor necrosis factor receptor (TNFR) superfamily initiate intracellular signaling by recruiting the C-domain of the TNFR-associated factors (TRAFs) through their cytoplasmatic tails
		553-561	QNPEEQDEG	Close proximity		
		345-351	PKHTPSK	Close proximity	ELM_MOD	Canonical version of the CDK phosphorylation site which shows specificity towards a lysine/arginine residue at the [ST]+3 position.
		400-406	PVTSPVK	Overlap		
		304-311	PAATPEIR	Embedded	ELM_MOD	Longer version of the CDK phosphorylation site which shows specificity towards a lysine/arginine residue at position +4 after the phospho-Ser/Thr
		371-377	SVTTPSQ	Embedded	ELM_MOD	CK1 phosphorylation site
		411-417	SGQSIPW	Embedded		
		429-435	SLPSGEP	Embedded		
		445-451	SWPSQTA	Embedded		
		357-363	QILSLFE	Embedded	ELM_MOD	CK2 phosphorylation site
		373-379	TTPSQFE	Embedded		
		433-439	GEPSAAE	Embedded		
		417-420	WDLW	Embedded	ELM_MOD	Motif for attachment of a mannosyl residue to a tryptophan
		410-413	PSGQ	Overlap	ELM_MOD	Glycosaminoglycan attachment site
		431-434	PSGE	Embedded		
		434-438	EPSAA	Embedded		
		435-438	PSAA	Embedded		
		509-512	GSGA	Close proximity		
		273-280	QHGSNTFT	Embedded	ELM_MOD	GSK3 phosphorylation recognition site
		300-307	PDGSPAAT	Overlap		
		320-327	GGATPGAT	Embedded		
		324-331	PGATLPKS	Embedded		
		382-389	GPFSEQAS	Embedded		
		422-429	PTESPAGS	Embedded		

		429-436	SLPSGEPS	Embedded		
		438-445	AEGTFAVS	Embedded		
		480-487	TAASEAAS	Embedded		
		494-501	VVETFPAT	Embedded		
		498-505	FPATVNGT	Overlap		
		529-536	HDYTATDT	Close proximity		
		502-507	VNGTVE	Overlap	ELM_MOD	Generic motif for N-glycosylation. It was shown that Trp, Asp, and Glu are uncommon before the Ser/Thr position. Efficient glycosylation usually occurs when ~60 residues or more separate the glycosylation acceptor site from the C-terminus.
		362-367	FEDTFV	Embedded	ELM_MOD	NEK2 phosphorylation motif with preferred Phe, Leu or Met in the -3 position to compensate for less favorable residues in the +1 and +2 position.
		442-447	FAVSWP	Embedded		
		330-336	KSPSQLR	Embedded	ELM_MOD	(ST)Q motif which is phosphorylated by PIKK family members.
		373-379	TTPSQFE	Embedded		
		445-451	SWPSQTA	Embedded		
		465-471	AGGTQPA	Embedded		
		362-368	FEDTFVP	Embedded	ELM_MOD	Ser/Thr residue phosphorylated by the Plk1 kinase
		368-374	PEISVTT	Embedded		
		386-392	EQASLLD	Embedded	ELM_MOD	Ser/Thr residue phosphorylated by Plk2 and Plk3
		357-363	QILSLFE	Embedded	ELM_MOD	Ser/Thr residue phosphorylated by Plk4
		362-368	FEDTFVP	Embedded		
		386-392	EQASLLD	Embedded		
		300-306	PDGSPAA	Overlap	ELM_MOD	Proline-Directed Kinase (e.g. MAPK) phosphorylation site in higher eukaryotes
		304-310	PAATPEI	Embedded		
		320-326	GGATPGA	Embedded		
		328-334	LPKSPSQ	Embedded		
		345-351	PKHTPSK	Close proximity		
		371-377	SVTTPSQ	Embedded		
		400-406	PVTSPVK	Overlap		
		406-412	KAPTPSG	Overlap		
		422-428	PTESPAG	Embedded		

		353-356	VKQE	Embedded	ELM_MOD	Motif recognised for modification by SUMO-1
		386-391	EQASLL	Embedded	ELM_TRG	Sorting and internalisation signal found in the cytoplasmic juxta-membrane region of type I transmembrane proteins. Targets them from the Trans Golgi Network to the lysosomal-endosomal-melanosomal compartments. Interacts with adaptor protein (AP) complexes
APOE4	278-283 (6)	279-285	LKSWFEP	Overlap	ELM_DOC	docking site required for the regulatory subunit B56 of PP2A for protein dephosphorylation.
		278-283	RLKSWF	Embedded	ELM_LIG	Canonical Arg-containing phospho-motif mediating a strong interaction with 14-3-3 proteins.
		278-286	RLKSWFEPL	Overlap	ELM_LIG	A motif present in the BRCA2 protein which binds to the WD 40 repeat (blade 4,5) domain of PALB2 which is required for the recognition of DNA double strand breaks and repair.
		281-286	SWFEPL	Overlap	ELM_LIG	Canonical LIR motif that binds to Atg8 protein family members to mediate processes involved in autophagy.
		278-284	RLKSWFE	Overlap	ELM_MOD	CK2 phosphorylation site
		278-284	RLKSWFE	Overlap	ELM_MOD	Ser/Thr residue phosphorylated by Plk4
APP	181-190 (10)	191-195	EESDN	Close proximity	ELM_CLV	Caspase-3 and Caspase-7 cleavage site.
	205-243 (39)	216-220	DYADG	Embedded		
	251-275 (25)	285-289	EVVRE	Embedded	ELM_CLV	Separase cleavage site, best known in sister chromatid separation.
	283-291 (9)	273-277	ATTTT	Embedded	ELM_DEG	The S/T rich motif known as the SPOP-binding consensus (SBC) of the MATH-BTB protein, SPOP, is present in substrates that undergo SPOP/Cul3-dependant ubiquitination.
	301-322 (22)	218-222	ADGSE	Embedded	ELM_DOC	The USP7 MATH domain binding motif variant

						based on the MDM2 and p53 interactions.
	336-346 (11)	267-273	ERTTSIA	Embedded	ELM_LIG	Phosphothreonine motif binding a subset of FHA domains that show a preference for a large aliphatic amino acid at the pT+3 position.
	391-396 (6)	278-284	TTTESVE	Close proximity		
	426-437 (12)	601-607	KTTVELL	Overlap		
	471-479 (9)	276-282	TTTTTES	Close proximity	ELM_LIG	Phosphothreonine motif binding a subset of FHA domains that have a preference for an acidic amino acid at the pT+3 position.
	491-497 (7)	600-606	TKTTVEL	Overlap		
	545-550 (6)	217-220	YADG	Embedded	ELM_LIG	Src-family Src Homology 2 (SH2) domains binding motif.
	606-626 (21)	262-265	YEEA	Embedded		
		602-608	TTVELLP	Overlap	ELM_LIG	Motif for the antiparallel beta augmentation mode of non-covalent binding to SUMO protein.
		603-608	TVELLP	Overlap		
		221-233	SEDKVVEVAEEEE	Embedded	ELM_LIG	Motif for the parallel beta augmentation mode of non-covalent binding to SUMO protein.
		222-233	EDKVVEVAEEEE	Embedded		
		223-233	DKVVEVAEEEE	Embedded		
		224-233	KVVEVAEEEE	Embedded		
		229-232	AEEE	Embedded	ELM_LIG	Major TRAF2-binding consensus motif. Members of the tumor necrosis factor receptor (TNFR) superfamily initiate intracellular signaling by recruiting the C-domain of the TNFR-associated factors (TRAFs) through their cytoplasmic tails.
		261-264	PYEE	Embedded		
		282-285	SVEE	Overlap		
		271-277	SIATTTT	Overlap	ELM_MOD	CK1 phosphorylation site
		343-349	SAMSQSL	Overlap		
		275-281	TTTTTTE	Overlap	ELM_MOD	CK2 phosphorylation site
		279-285	TTESVEE	Overlap		
		349-355	LLKTTQE	Close proximity		
		599-605	ETKTTVE	Close proximity		
		342-345	GSAM	Embedded	ELM_MOD	Glycosaminoglycan attachment site
		263-270	EEATERTT	Embedded	ELM_MOD	GSK3 phosphorylation recognition site
		267-274	ERTTSIAT	Embedded		
		268-275	RTTSIATT	Embedded		
		271-278	SIATTTTT	Overlap		

		272-279	IATTTTTT	Overlap		
		273-280	ATTTTTTT	Overlap		
		275-282	TTTTTTES	Overlap		
		345-352	MSQSLLKT	Overlap		
		345-350	MSQSLL	Overlap	ELM_MOD	NEK2 phosphorylation motif with preferred Phe, Leu or Met in the -3 position to compensate for less favorable residues in the +1 and +2 position.
		349-354	LLKTTQ	Close proximity		
		343-349	SAMSQSL	Overlap	ELM_MOD	(ST)Q motif which is phosphorylated by PIKK family members.
		611-617	GEFSLDD	Embedded	ELM_MOD	Ser/Thr residue phosphorylated by the Plk1 kinase
		345-351	MSQSLLK	Overlap	ELM_MOD	Ser/Thr residue phosphorylated by Plk4
		216-226	DYADGSEDKVV	Embedded	ELM_MOD	Inverted version of SUMOylation motif recognized for modification by SUMO-1
		219-226	DGSEDKVV	Embedded		
		221-226	SEDKVV	Embedded		
		310-317	DVTEGKCA	Embedded		
		596-603	SLTETKTT	Close proximity		
		605-617	ELLPVNGEFSLDD	Overlap	ELM_TRG	Some proteins re-exported from the nucleus contain a Leucine-rich nuclear export signal (NES) binding to the CRM1 exportin protein.
PICALM	308-314 (7)	538-542	DDLDS	Overlap	ELM_CLV	Caspase-3 and Caspase-7 cleavage site.
	335-344 (10)	376-383	IFSTPSSS	Overlap	ELM_DEG	The TPxxS phospho-dependent degron binds the FBW7 F box proteins of the SCF (Skp1_Cullin-Fbox) complex.
	366-378 (13)	360-364	PVSTS	Close proximity	ELM_DEG	The S/T rich motif known as the SPOP-binding consensus (SBC) of the MATH-BTB protein, SPOP, is present in substrates that undergo SPOP/Cul3-dependant ubiquitination.
	392-397 (6)	377-382	FSTPSS	Overlap	ELM_DOC	Phospho-dependent motif that mediates docking of CDK substrates and regulators to cyclin-CDK-bound Cks1.
	506-514 (9)	371-377	APAIDIF	Embedded	ELM_DOC	Calcineurin substrate docking site, leads to the

						effective dephosphorylation of serine/threonine phosphorylation sites.
	540-553 (14)	305-309	AVSSL	Embedded	ELM_DOC	The USP7 MATH domain binding motif variant based on the MDM2 and p53 interactions.
		380-384	PSSSN	Close proximity		
		376-381	IFSTPS	Overlap	ELM_DOC	The Class IV WW domain interaction motif is recognised primarily by the Pin1 phosphorylation-dependent prolyl isomerase.
		503-507	DSGGF	Overlap	ELM_LIG	Phosphopeptide motif which directly interacts with the BRCT (carboxy-terminal) domain of the Breast Cancer Gene BRCA1 with low affinity
		384-390	NSTSKLP	Close proximity	ELM_LIG	Phosphothreonine motif binding a subset of FHA domains that show a preference for a large aliphatic amino acid at the pT+3 position.
		504-510	SGGFDEL	Overlap	ELM_LIG	Canonical LIR motif that binds to Atg8 protein family members to mediate processes involved in autophagy.
		509-515	ELGGLLK	Embedded	ELM_LIG	The nuclear receptor box motif (LXXLL) confers binding to nuclear receptors.
		308-314	SLASTGL	Embedded	ELM_MOD	CK1 phosphorylation site
		359-365	SPVSTSA	Close proximity		
		378-384	STPSSSN	Overlap		
		363-366	TSAG	Overlap	ELM_MOD	Glycosaminoglycan attachment site
		503-506	DSGG	Overlap		
		300-307	TTLSNAVS	Close proximity	ELM_MOD	GSK3 phosphorylation recognition site
		304-311	NAVSSLAS	Overlap		
		305-312	AVSSLAST	Overlap		
		308-315	SLASTGLS	Overlap		
		375-382	DIFSTPSS	Overlap		
		376-383	IFSTPSSS	Overlap		
		378-385	STPSSSNS	Overlap		
		379-386	TPSSSNST	Close proximity		

		380-387	PSSSNSTS	Close proximity		
		397-404	QQPTFHPS	Overlap		
		554-559	GNGTTK	Close proximity	ELM_MOD	Generic motif for N-glycosylation. It was shown that Trp, Asp, and Glu are uncommon before the Ser/Thr position. Efficient glycosylation usually occurs when ~60 residues or more separate the glycosylation acceptor site from the C-terminus.
		309-314	LASTGL	Embedded	ELM_MOD	NEK2 phosphorylation motif with preferred Phe, Leu or Met in the -3 position to compensate for less favorable residues in the +1 and +2 position.
		540-545	LDSSLA	Embedded		
		540-546	LDSSLAN	Embedded	ELM_MOD	Ser/Thr residue phosphorylated by the Plk1 kinase
		376-382	IFSTPSS	Overlap	ELM_MOD	Proline-Directed Kinase (e.g. MAPK) phosphorylation site in higher eukaryotes.
		509-514	ELGGLL	Embedded	ELM_TRG	Sorting and internalisation signal found in the cytoplasmic juxta-membrane region of type I transmembrane proteins. Targets them from the Trans Golgi Network to the lysosomal-endosomal-melanosomal compartments. Interacts with adaptor protein (AP) complexes
CD33	334-348 (15)	341-345	ASLNF	Embedded	ELM_LIG	Phosphopeptide motif which directly interacts with the BRCT (carboxy-terminal) domain of the Breast Cancer Gene BRCA1 with low affinity
		348-355	MNPSKDTS	Overlap	ELM_MOD	GSK3 phosphorylation recognition site
		340-343	YASL	Embedded	ELM_TRG	Tyrosine-based sorting signal responsible for the interaction with mu subunit of AP (Adaptor Protein) complex
PSEN1	1-26 (26)	28-32	SQNDN	Close proximity	ELM_CLV	Caspase-3 and Caspase-7 cleavage site.
		4_7	LPAP	Embedded	ELM_DOC	docking motif in calcineurin substrates that

						binds at the interface of the catalytic CNA and regulatory CNB subunits.
		14-18	AQMSE	Embedded	ELM_DOC	The USP7 MATH domain binding motif variant based on the MDM2 and p53 interactions.
		22-27	LSNTVR	Overlap	ELM_MOD	NEK2 phosphorylation motif with preferred Phe, Leu or Met in the -3 position to compensate for less favorable residues in the +1 and +2 position.
		25-31	TVRSQND	Overlap	ELM_MOD	(ST)Q motif which is phosphorylated by PIKK family members.
		6_12	APLSYFQ	Embedded	ELM_MOD	Ser/Thr residue phosphorylated by Plk4
ADAM17	225-236 (12)	783-787	PNSST	Embedded	ELM_DEG	The S/T rich motif known as the SPOP-binding consensus (SBC) of the MATH-BTB protein, SPOP, is present in substrates that undergo SPOP/Cul3-dependant ubiquitination.
	433-439 (7)	759-764	MDTIQE	Overlap	ELM_DOC	docking site required for the regulatory subunit B56 of PP2A for protein dephosphorylation.
	745-761 (17)	783-787	PNSST	Embedded	ELM_DOC	The USP7 MATH domain binding motif variant based on the MDM2 and p53 interactions.
	774-815 (42)	788-792	AAKSF	Embedded		
		780-782	DPF	Embedded	ELM_LIG	DPF/W motif binds alpha and beta subunits of AP2 adaptor complex.
		759-765	MDTIQED	Overlap	ELM_LIG	Phosphothreonine motif binding a subset of FHA domains that have a preference for an acidic amino acid at the pT+3 position.
		799-805	PVTRSEK	Embedded		
		791-795	SFEDL	Embedded	ELM_LIG	Canonical LIR motif that binds to Atg8 protein family members to mediate processes involved in autophagy.
		745-750	IPSAPA	Embedded	ELM_LIG	PTAP motif binds the N-terminal UEV domain of Tsg101.
		433-436	YVMY	Embedded	ELM_LIG	STAT5 Src Homology 2 (SH2) domain binding motif.
		740-746	QPAPVIP	Overlap	ELM_LIG	

		743-749	PVIPSAP	Overlap		This is the motif recognized by those SH3 domains with a non-canonical class I recognition specificity
		746-752	PSAPAAP	Embedded		
		436-441	YPIAVS	Overlap	ELM_LIG	Motif for the parallel beta augmentation mode of non-covalent binding to SUMO protein.
		761-764	TIQE	Overlap	ELM_LIG	Major TRAF2-binding consensus motif. Members of the tumor necrosis factor receptor (TNFR) superfamily initiate intracellular signaling by recruiting the C-domain of the TNFR-associated factors (TRAFs) through their cytoplasmic tails.
		767-773	STDSHMD	Close proximity	ELM_MOD	CK1 phosphorylation site
		758-764	RMDTIQE	Overlap	ELM_MOD	CK2 phosphorylation site
		798-804	HPVTRSE	Embedded		
		440-443	VSGD	Overlap	ELM_MOD	Glycosaminoglycan attachment site
		746-749	PSAP	Embedded		
		784-791	NSSTAAKS	Embedded	ELM_MOD	GSK3 phosphorylation recognition site
		756-762	HQRMDTI	Overlap	ELM_MOD	The LATS phosphorylation motif is recognised by the LATS kinases for Ser/Thr phosphorylation. Substrates are often found toward the end of the Hippo signalling pathway.
		783-788	PNSSTA	Embedded	ELM_MOD	Generic motif for N-glycosylation. It was shown that Trp, Asp, and Glu are uncommon before the Ser/Thr position. Efficient glycosylation usually occurs when ~60 residues or more separate the glycosylation acceptor site from the C-terminus.
		221-226	MKNTCK	Overlap	ELM_MOD	NEK2 phosphorylation motif with preferred Phe, Leu or Met in the -3 position to compensate for less favorable residues in the +1 and +2 position.
Clusterin	0					

PSEN2	1-25 (25)	6_10	ASDSE	Embedded	ELM_DOC	The USP7 MATH domain binding motif variant based on the MDM2 and p53 interactions.
		22-27	SAESPT	Overlap	ELM_DOC	The Class IV WW domain interaction motif is recognised primarily by the Pin1 phosphorylation-dependent prolyl isomerase.
		24-29	ESPTPR	Overlap		
		17-21	RTSLM	Embedded	ELM_LIG	Canonical Arg-containing phospho-motif mediating a strong interaction with 14-3-3 proteins.
		24-29	ESPTPR	Overlap	ELM_MOD	Short version of the CDK phosphorylation site which shows specificity towards a lysine/arginine residue at the [ST] +2 position.
		22-29	SAESPTPR	Overlap	ELM_MOD	Longer version of the CDK phosphorylation site which shows specificity towards a lysine/arginine residue at position +4 after the phospho-Ser/Thr
		19-25	SLMSAES	Embedded	ELM_MOD	CK1 phosphorylation site
		22-28	SAESPTP	Overlap		
		4-10	FMASDSE	Embedded	ELM_MOD	CK2 phosphorylation site
		6_12	ASDSEEE	Embedded		
		27-33	TPRSCQE	Close proximity		
		15-22	DERTSLMS	Embedded	ELM_MOD	GSK3 phosphorylation recognition site
		16-22	ERTSLMS	Embedded	ELM_MOD	Secondary preference for PKA-type AGC kinase phosphorylation.
		15-21	DERTSLM	Embedded	ELM_MOD	Ser/Thr residue phosphorylated by the Plk1 kinase
		16-22	ERTSLMS	Embedded	ELM_MOD	Ser/Thr residue phosphorylated by Plk4
		22-28	SAESPTP	Overlap	ELM_MOD	Proline-Directed Kinase (e.g. MAPK) phosphorylation site in higher eukaryotes
		24-30	ESPTPRS	Overlap		
		21-24	MSAE	Embedded	ELM_MOD	Glycosaminoglycan attachment site
PLG	463-470 (8)	472-477	DVETPS	Close proximity	ELM_DOC	The Class IV WW domain interaction motif is recognised primarily by the Pin1 phosphorylation-dependent prolyl isomerase

		466-470	PVVLL	Embedded	ELM_LIG	LIGand to interface formed by dimerisation of two chromoshadow domains in HP1 proteins.
		458-464	EASVVAP	Overlap	ELM_LIG	This is the motif recognized by those SH3 domains with a non-canonical class I recognition specificity
		459-465	ASVVAPP	Overlap		
		465-471	PPVVLLP	Overlap		
		470-476	LPDVETP	Overlap		
		466-472	PVVLLPD	Overlap	ELM_LIG	Motif for the parallel beta augmentation mode of non-covalent binding to SUMO protein.
		472-478	DVETPSE	Close proximity	ELM_MOD	CK2 phosphorylation site
		457-463	TEASVVA	Overlap	ELM_MOD	Ser/Thr residue phosphorylated by the Plk1 kinase
		457-463	TEASVVA	Overlap	ELM_MOD	Ser/Thr residue phosphorylated by Plk4
		472-478	DVETPSE	Close proximity	ELM_MOD	Proline-Directed Kinase (e.g. MAPK) phosphorylation site in higher eukaryotes.
ADAM9	767-781 (15)	759-765	PVTPPRE	Close proximity	ELM_DEG	The TPxxE phospho-dependent degron binds the FBW7 F box proteins of the SCF (Skp1_Cullin-Fbox) complex.
	799-819 (21)	760-765	VTPPRE	Close proximity		
		772-776	RFAVP	Embedded	ELM_DOC	SPAK/OSR1 kinase binding motif acts as a docking site which aids the interaction with their binding partners including the upstream activators and the phosphorylated substrates.
		795-799	PKVSS	Overlap	ELM_DOC	The USP7 MATH domain binding motif variant based on the MDM2 and p53 interactions.
		815-819	YSSLT	Embedded	ELM_LIG	CRK family SH2 domain binding motif.
		769-772	YANR	Embedded	ELM_LIG	GRB2-like Src Homology 2 (SH2) domains binding motif
		807-813	RPAPAPP	Embedded	ELM_LIG	This is the motif recognized by class I SH3 domains
		802-808	NLIPARP	Embedded	ELM_LIG	This is the motif recognized by those SH3 domains with a non-canonical class I recognition specificity
		807-813	RPAPAPP	Embedded		

		812-815	PPLY	Embedded	ELM_LIG	PPXY is the motif recognized by WW domains of Group I
		814-819	LYSSLT	Embedded	ELM_MOD	NEK2 phosphorylation motif with preferred Phe, Leu or Met in the -3 position to compensate for less favorable residues in the +1 and +2 position.
		796-802	KVSSQGN	Overlap	ELM_MOD	(ST)Q motif which is phosphorylated by PIKK family members.
		815-818	YSSL	Embedded	ELM_TRG	Tyrosine-based sorting signal responsible for the interaction with mu subunit of AP (Adaptor Protein) complex
ADAM10	313-318 (6)	737-741	PRESY	Overlap	ELM_DOC	The USP7 MATH domain binding motif variant based on the MDM2 and p53 interactions.
	415-420 (6)	737-743	PRESYQM	Overlap	ELM_MOD	Secondary preference for PKA-type AGC kinase phosphorylation.
	741-748 (8)	746-748	MRR	Embedded	ELM_TRG	The di-Arg ER retention motif is defined by two consecutive arginine residues (RR) or with a single residue insertion (RXR). The motif is completed by an adjacent hydrophobic/arginine residue which may be on either side of the Arg pair.
ABCA7	124-131 (8)	1353-1355	RRL	Embedded	ELM_CLV	N-Arg dibasic convertase (NRD/Nardilysin) cleavage site (X-|-R-K or R-|-R-X).
	156-164 (9)	2130-2132	KRV	Close proximity	ELM_CLV	NEC1/NEC2 cleavage site (K-R-|-X).
	195-201 (7)	1349-1355	RPGARRL	Overlap	ELM_CLV	Proprotein convertase 7 (PC7, PCSK7) cleavage site (R-X-X-X-[RK]-R-|-X).
	1142-1147(6)	1353-1357	KNLTA	Embedded	ELM_CLV	Subtilisin/kexin isozyme-1 (SKI1) cleavage site ([RK]-X-[hydrophobic]-[LTKF]-|-X).
	1166-1173(8)	2064-2068	RCALA	Overlap		
	1327-1333 (7)	2063-2071	GRCALARVF	Overlap	ELM_DEG	An RxxL-based motif that binds to the Cdh1 and Cdc20 components of APC/C thereby targeting the protein for destruction in a cell cycle dependent manner

	1352-1360 (9)	1349-1358	GARRLLPD	Overlap	ELM_DOC	This motif is mainly based on cyclin A binding peptides and may not apply to all cyclins.
	1386-1397 (12)	1351-1358	GARRLLPD	Overlap		
	2067-2072 (6)	165-173	RTESLGLAL	Close proximity	ELM_DOC	A kinase docking motif that mediates interaction towards the ERK1/2 and p38 subfamilies of MAP kinases.
	2092-2102 (11)	197-203	LRSLVEL	Overlap	ELM_DOC	Docking site required for the regulatory subunit B56 of PP2A for protein dephosphorylation.
	2133-2146 (14)	1144-1150	LKVVEEC	Overlap		
		196-200	ALRSL	Embedded	ELM_DOC	The USP7 MATH domain binding motif variant based on the MDM2 and p53 interactions.
		1334-1339	GNWTPE	Close proximity	ELM_DOC	The Class IV WW domain interaction motif is recognised primarily by the Pin1 phosphorylation-dependent prolyl isomerase.
		133-143	RAARSTAQPQP	Close proximity	ELM_LIG	Canonical Arg-containing phospho-motif mediating a strong interaction with 14-3-3 proteins.
		1385-1392	RNLSDFLV	Overlap		
		2131-2139	RVSQFLDDP	Overlap		
		118-135	AHRTLAGLGKLIATLRAA	Overlap	ELM_LIG	The WH2 motif is of variable length (16-19 amino acids) binding to the hydrophobic cleft formed by actin's subdomains 1 and 3. At the N-terminus it forms an alpha-helix followed by a flexible loop stabilised upon actin binding.
		159-165	LLTSLLR	Overlap	ELM_LIG	Phosphothreonine motif binding a subset of FHA domains that show a preference for a large aliphatic amino acid at the pT+3 position.
		164-170	LRTESLG	Overlap		
		1173-1179	DVTLRLK	Overlap		
		2089-2095	SQTMLEE	Overlap	ELM_LIG	Phosphothreonine motif binding a subset of FHA domains that have a preference for an acidic amino acid at the pT+3 position.

		156-164	VAELLTSLL	Embedded	ELM_LIG	Amphipatic alpha helix that binds the GTPase-binding domain (GBD) in WASP and N-WASP.
		1140-1146	EEIFLKV	Overlap	ELM_LIG	Canonical LIR motif that binds to Atg8 protein family members to mediate processes involved in autophagy
		1141-1146	EIFLKV	Overlap		
		1394-1398	TYPRL	Overlap		
		159-165	LLTSLLR	Overlap	ELM_LIG	The nuclear receptor box motif (LXXLL) confers binding to nuclear receptors
		202-208	ELRALLQ	Close proximity		
		2141-2146	TAETVL	Embedded	ELM_LIG	The C-terminal class 1 PDZ-binding motif is classically represented by a pattern like (ST)X(VIL)*
		2099-2102	YFSK	Embedded	ELM_LIG	STAT5 Src Homology 2 (SH2) domain binding motif.
		1354-1360	RLLPDCP	Embedded	ELM_LIG	This is the motif recognized by class I SH3 domains
		1335-1341	NWTPESP	Close proximity	ELM_LIG	This is the motif recognized by those SH3 domains with a non-canonical class I recognition specificity
		1354-1360	RLLPDCP	Embedded		
		152-158	PMLDVAE	Overlap	ELM_LIG	Motif for the parallel beta augmentation mode of non-covalent binding to SUMO protein.
		193-199	ELLALRS	Overlap		
		1138-1141	SLEE	Close proximity	ELM_LIG	Major TRAF2-binding consensus motif. Members of the tumor necrosis factor receptor (TNFR) superfamily initiate intracellular signaling by recruiting the C-domain of the TNFR-associated factors (TRAFs) through their cytoplasmic tails.
		150-156	EPPMLDV	Overlap	ELM_LIG	This WDR5-binding motif binds to a cleft between blades 5 and 6 of the WD40 repeat domain of WDR5, opposite of the Win motif-binding site, to mediate assembly of histone modification complexes.
		1135-1141	SDTSLEE	Close proximity	ELM_MOD	CK1 phosphorylation site

		196-202	ALRSLVE	Overlap	ELM_MOD	CK2 phosphorylation site
		1134-1140	ISDTSLE	Close proximity		
		1135-1141	SDTSLEE	Close proximity		
		2088-2094	VSQTMLE	Overlap		
		1332-1335	ASGN	Overlap		
		159-166	LLTSLLRT	Overlap	ELM_MOD	GSK3 phosphorylation recognition site
		1330-1337	VLASGNWT	Overlap		
		2084-2091	EDFSVSQT	Close proximity		
		2137-2144	DDPSTAET	Embedded		
		2128-2134	HPKRVSQ	Overlap	ELM_MOD	The LATS phosphorylation motif is recognised by the LATS kinases for Ser/Thr phosphorylation. Substrates are often found toward the end of the Hippo signalling pathway.
		1334-1339	GNWTPE	Close proximity	ELM_MOD	Generic motif for N-glycosylation. It was shown that Trp, Asp, and Glu are uncommon before the Ser/Thr position. Efficient glycosylation usually occurs when ~60 residues or more separate the glycosylation acceptor site from the C-terminus.
		1380-1385	QNLTGR	Close proximity		
		1385-1390	RNLSDF	Overlap		
		128-133	LIATLR	Overlap	ELM_MOD	NEK2 phosphorylation motif with preferred Phe, Leu or Met in the -3 position to compensate for less favorable residues in the +1 and +2 position.
		159-164	LLTSLL	Embedded		
		1172-1177	LDVTLR	Embedded		
		1391-1396	LVKTYP	Embedded		
		2086-2091	FSVSQT	Close proximity		
		2086-2092	FSVSQTM	Overlap	ELM_MOD	(ST)Q motif which is phosphorylated by PIKK family members.
		2130-2136	KRVSQFL	Overlap		
		2130-2136	KRVSQFL	Overlap	ELM_MOD	Main preference for PKA-type AGC kinase phosphorylation.
		2130-2136	KRVSQFL	Overlap	ELM_MOD	Secondary preference for PKA-type AGC kinase phosphorylation.

		1135-1141	SDTSLEE	Close proximity	ELM_MOD	Ser/Thr residue phosphorylated by the Plk1 kinase
		1172-1178	LDVTLRL	Overlap		
		2084-2090	EDFSVSQ	Close proximity		
		159-165	LLTSLLR	Overlap	ELM_MOD	Ser/Thr residue phosphorylated by Plk4
		2088-2094	VSQTMLE	Overlap		
		1334-1340	GNWTPES	Close proximity	ELM_MOD	Proline-Directed Kinase (e.g. MAPK) phosphorylation site in higher eukaryotes.
		1395-1398	YPRL	Overlap	ELM_TRG	Tyrosine-based sorting signal responsible for the interaction with mu subunit of AP (Adaptor Protein) complex
		1353-1355	RRL	Embedded	ELM_TRG	The di-Arg ER retention motif is defined by two consecutive arginine residues (RR) or with a single residue insertion (RXR). The motif is completed by an adjacent hydrophobic/arginine residue which may be on either side of the Arg pair.
		155-160	DVAELL	Overlap	ELM_TRG	Sorting and internalisation signal found in the cytoplasmic juxta-membrane region of type I transmembrane proteins. Targets them from the Trans Golgi Network to the lysosomal-endosomal-melanosomal compartments. Interacts with adaptor protein (AP) complexes
LRP1	3965-3970 (6)	4501-4507	NFTNPVY	Overlap	ELM_LIG	Phosphothreonine motif binding a subset of FHA domains that show a preference for a large aliphatic amino acid at the pT+3 position.
	4505-4513 (9)	4501-4508	NFTNPVYA	Overlap	ELM_LIG	These phosphorylation-independent motifs bind to Dab-like PTB domains. Binding is not driven by contacts at the 0 or FY position, but instead is dependent upon the large number of hydrophobic and hydrogen bond

						contacts between motif and domain.
		4501-4507	NFTNPVY	Overlap	ELM_LIG	This phosphorylation-dependent motif binds to Shc-like and IRS-like PTB domains. The pTyr is positioned within a highly basic-charged anchoring pocket. A hydrophobic residue -5 (compared to pY) increases the affinity of the interaction.
		4507-4511	YATLY	Embedded	ELM_LIG	CRK family SH2 domain binding motif.
		4500-4505	TNFTNP	Overlap	ELM_MOD	Generic motif for N-glycosylation. It was shown that Trp, Asp, and Glu are uncommon before the Ser/Thr position. Efficient glycosylation usually occurs when ~60 residues or more separate the glycosylation acceptor site from the C-terminus.
		4506-4512	VYATLYM	Embedded	ELM_MOD	Ser/Thr residue phosphorylated by Plk4
		4507-4510	YATL	Embedded	ELM_TRG	Tyrosine-based sorting signal responsible for the interaction with mu subunit of AP (Adaptor Protein) complex
TREM2	0					
ECE2	1-22 (22)	18-20	KRA	Embedded	ELM_CLV	NEC1/NEC2 cleavage site (K-R-|-X).
	718-724 (7)	18-22	KRATL	Embedded	ELM_CLV	Subtilisin/kexin isozyme-1 (SKI1) cleavage site ([RK]-X-[hydrophobic]-[LTKF]-|-X).
		1_3	MNV	Embedded	ELM_DEG	N-terminal motif that initiates protein degradation by binding to the UBR-box of N-recognins. This N-degron variant comprises N-terminal Asn or Gln as destabilizing residue.
		3_19	VALQELGAGSNMVEYKR	Embedded	ELM_DOC	Reverse (C to N direction) of the classical MAPK docking motif ELM:DOC_MAPK_gen_1 with an often extended linker region of the bipartite motif.
		19-23	RATLR	Overlap	ELM_LIG	Canonical Arg-containing phospho-motif mediating a

						strong interaction with 14-3-3 proteins.
		19-25	RATLRDE	Overlap	ELM_LIG	Phosphothreonine motif binding a subset of FHA domains that have a preference for an acidic amino acid at the pT+3 position.
		4_7	ALQE	Embedded	ELM_LIG	Major TRAF2-binding consensus motif. Members of the tumor necrosis factor receptor (TNFR) superfamily initiate intracellular signaling by recruiting the C-domain of the TNFR-associated factors (TRAFs) through their cytoplasmic tails.
		18-24	KRATLRD	Overlap	ELM_MOD	Main preference for PKA-type AGC kinase phosphorylation
		18-24	KRATLRD	Overlap	ELM_MOD	Secondary preference for PKA-type AGC kinase phosphorylation.
		12-20	SNMVEYKRA	Embedded	ELM_MOD	Inverted version of SUMOylation motif recognized for modification by SUMO-1
SORL1	0					
A2M	1148-1155(8)	None				
	1209-1218 (10)					
BACE1	0					
Neprilysin	0					
Nicastrin	0					
PEN2	0					
IDE	70-77 (8)	74-78	KVLLI	Embedded	ELM_CLV	Subtilisin/kexin isozyme-1 (SKI1) cleavage site ([RK]-X-[hydrophobic]-[LTKF]-|-X).
	108-115 (8)	61-69	KREYRGLEL	Close proximity	ELM_DOC	MAPK interacting molecules (e.g. MAPKKs, substrates, phosphatases) carry docking motif that help to regulate specific interaction in the MAPK cascade. The classic motif approximates (R/K)xxxx#x# where # is a hydrophobic residue.

		69-77	LANGIKVLL	Overlap	ELM_LIG	Amphipatic alpha helix that binds the GTPase-binding domain (GBD) in WASP and N-WASP.
		74-79	KVLLIS	Overlap	ELM_LIG	A binding site for IRF-3 protein present in various innate adaptor proteins and the viral protein NSP1to trigger the innate immune responsive pathways
		74-80	KVLLISD	Overlap	ELM_LIG	Motif for the parallel beta augmentation mode of non-covalent binding to SUMO protein.
		113-120	MLFLGTKK	Overlap	ELM_LIG	UBA3 adenylation domain binding motif variant based on the UBE2M and UBE2F interactions.
		79-85	SDPTTDK	Close proximity	ELM_MOD	CK1 phosphorylation site
		76-83	LLISDPTT	Overlap	ELM_MOD	GSK3 phosphorylation recognition site
		79-86	SDPTTDKS	Close proximity		
		115-120	FLGTKK	Overlap	ELM_MOD	NEK2 phosphorylation motif with preferred Phe, Leu or Met in the -3 position to compensate for less favorable residues in the +1 and +2 position.
		68-80	ELANGIKVLLISD	Overlap	ELM_TRG	Some proteins re-exported from the nucleus contain a Leucine-rich nuclear export signal (NES) binding to the CRM1 exportin protein.
PLAU	No result found					
PLAT	0					
APH1A	0					
APH1B	0					