



Structural Bioinformatics Approach in Bioactive Peptide Research: Tomato Vicilin Case Study

Burcu Kaplan Türköz, Bahar Bakar, Canan Kartal, Semih Ötleş

Burcu KAPLAN TÜRKÖZ

Laboratory of Molecular Biology

Department of Food Engineering,

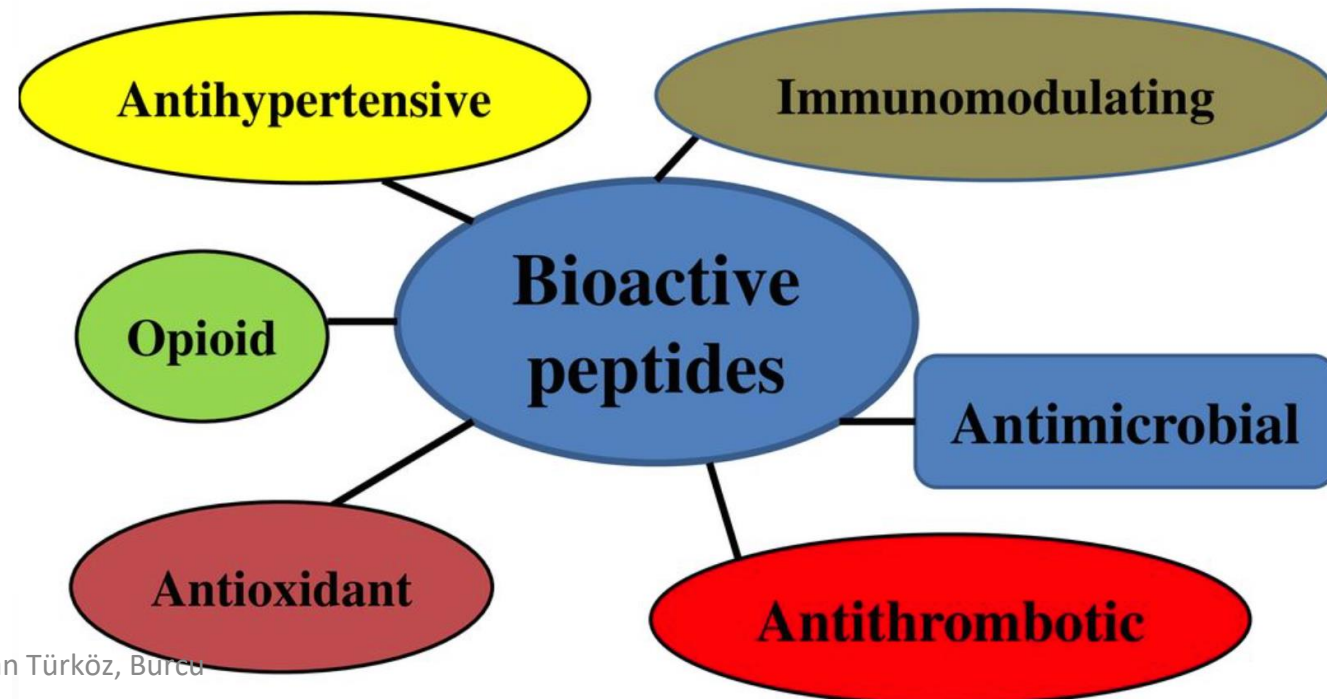
Ege University, İzmir, Turkey

Bioactive Peptides

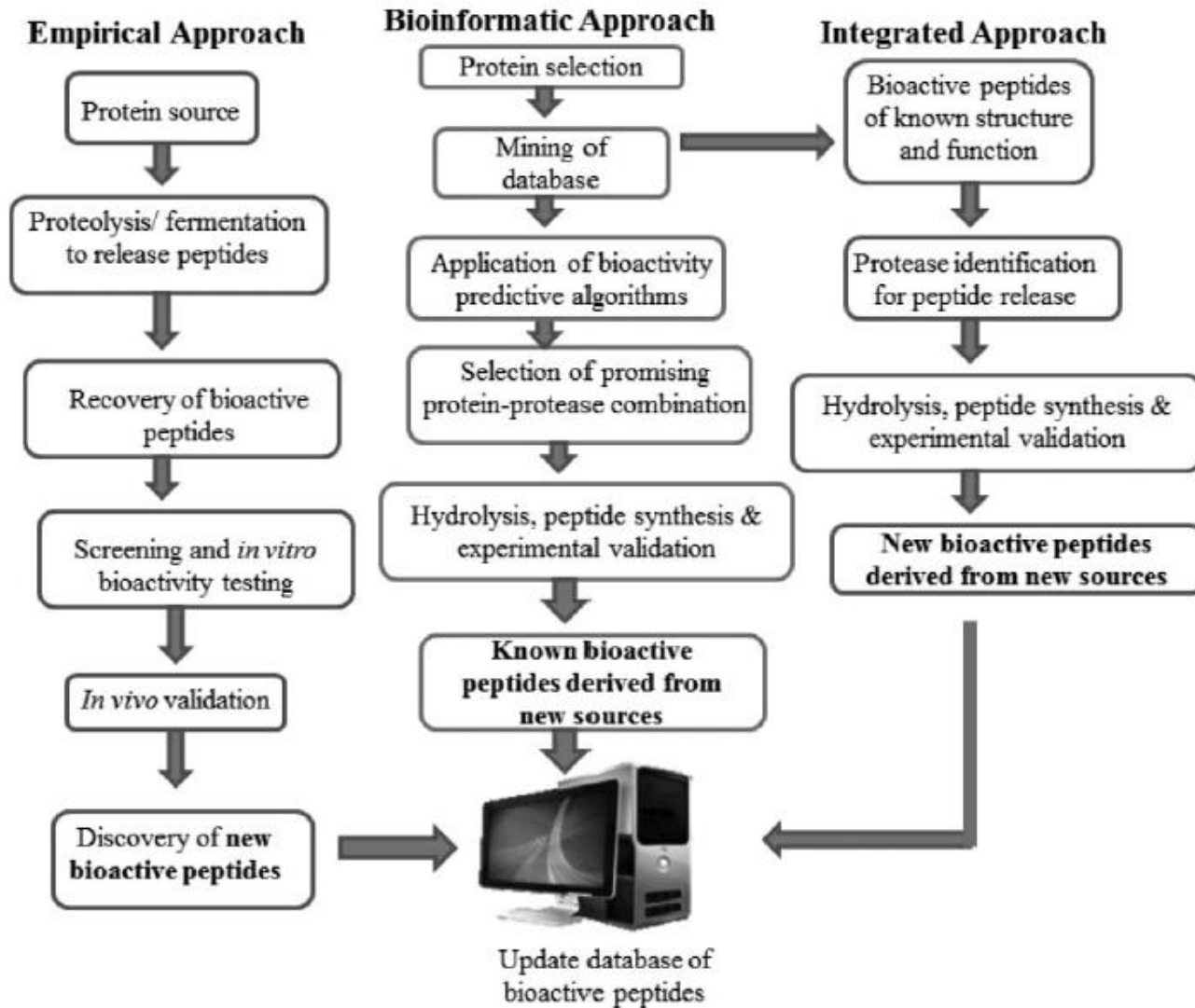


- Bioactive peptides (BAP) are derived from food proteins and exert a positive effect in humans due to its health-promoting properties
- Both animal and plant proteins are known to contain potential bioactive sequences, most studies to date have involved milk proteins.
- BAPs have direct effects on human system & health

- gastrointestinal system
- the cardiovascular system
- the immune system
- the nervous system



BAP Production



Bioinformatic Approach

Start About Chair Staff Didactics BIOPEP-UWM BIOPEP-UWM News

Databases	Current number
Proteins	740
Bioactive peptides	3803
Allergenic proteins with their epitopes	135
Sensory peptides and amino acids	483

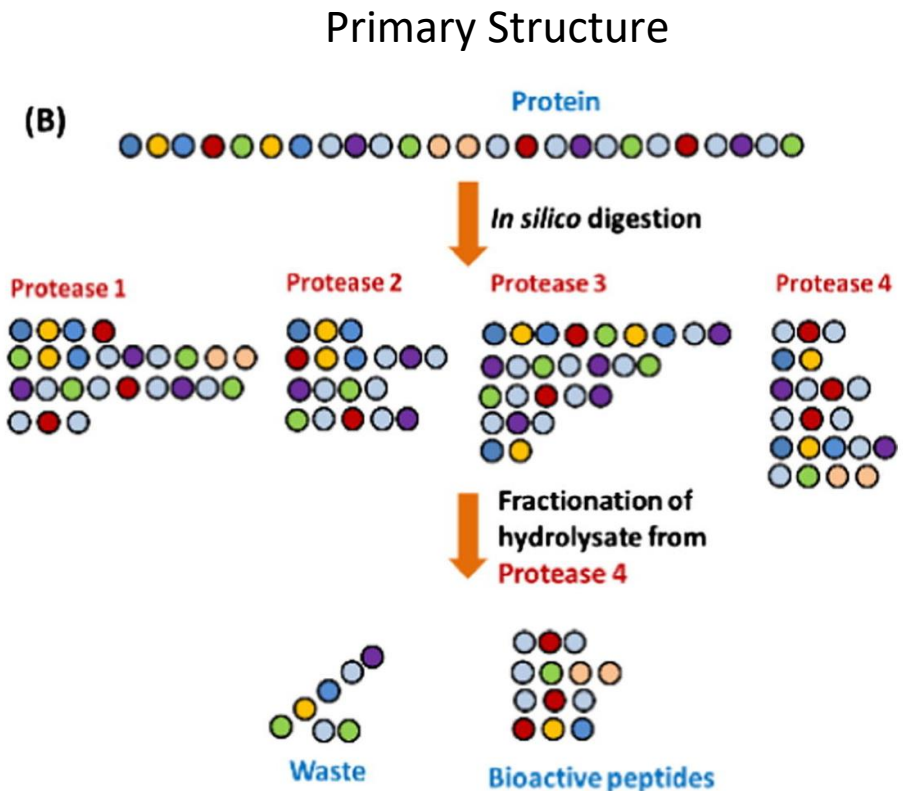
Submit new peptide sequence

Login

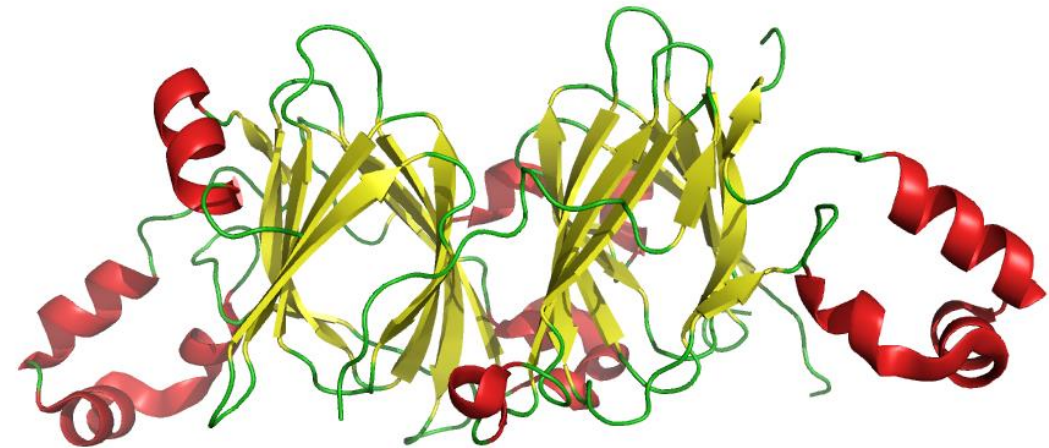
Please cite the following paper if you are using BIOPEP-UWM database:
 Minkiewicz P., Dziuba J., Iwaniak A., Dziuba M., Darewicz M., BIOPEP database and other programs for processing bioactive peptide sequences.
 Journal of AOAC International, 91, 2008, 965-980.

BAP prediction: Protein Structure ?

- There are several bioinformatics tools used in identification of BAPS based on amino acid sequence of the protein and the digestion patterns of different proteases.



Tertiary Structure



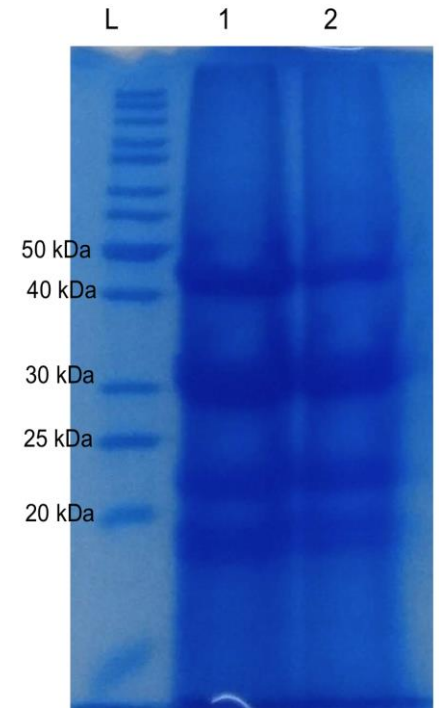
Structural Biology approach

Plant BAPs: Tomato seed proteins

- Tomato seed proteins extracted
- in silico proteolysis
 - experimental proteolysis
 - in vitro activity tests
 - sequencing of bioactive peptides

4 major protein bands ;
vicilin (45 kDa),
legumin α (30 kDa),
14-3-3 (22 kDa)
legumin β (18 kDa)

vicilin was selected for
further analysis.



Analysis of tomato seed proteins on 12% SDS-PAGE. L:ladder, 1-2: different extraction protocols

Structural modeling



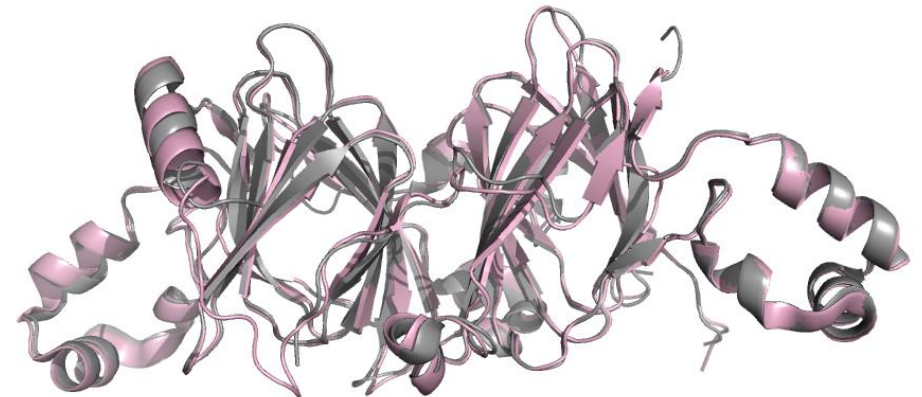
[My Jobs](#) | [Docs](#) | [Download](#) | [Inquiry & Bug](#)

```
>vicilin precursor [Solanum lycopersicum]
```

```
MAIFTKPKLL FIFFLILSLV LVSQCYDQNP RGYQDPQEKL  
RECQQRRCERQ QPGQQKQLCKQRCEQQYRKE QQQQHGGGETG  
EDDLGNRGPD KSYKRLQECQ RRCQSEQQGQ  
RLQECQQRCCQEYQREKGOH QGETNPQWEQ QEKSNNPYLF  
ESQRFRRFR ASHGDFRILE RFNQRSQLLKGIEKYRVAIL  
ELEPQSFVLP HHCDGEAIYV VVKGQGVINI AEQDNKNSFN  
LQKGDVIRLFAGSNVYLLNK DNNEKLFVYV LAKSVNAPGN  
LQEYFSAGGQ NPESFYRAFS SDILESAFNNPRDKLERLFG  
QHKEGIIKA SEEQIRAISE HASRSTQQTR GRTQGPFNLM  
KERPVFESRFGQFFEARPER YEQLRDLDA A VGFMNINQGG  
MVLPHYNTKS TKLVMVIEGN ARFEMACPHLGRQSQSPWSR  
GQGREQEREQ EQEQEEGDVH YQKIRGNLNV GDVLVIPAGH  
PITFVATGNSNLRIVGFGVD AENNKKNFLA GKQNIWRNID  
REAKELSFMS PGREVEEIFQ RQDQSYFVPEHRQQRERG  
EEGRRGQDQY LSSILDFVF
```

RaptorX: Protein Structure and Function Prediction Powered by Deep Learning

RaptorX is developed by Xu group, excelling at secondary, tertiary and contact prediction for protein sequences without close homologs in the Protein Data Bank (PDB). RaptorX predicts protein secondary and tertiary structures, contact and distance map, solvent accessibility, disordered regions, functional annotation and binding sites. RaptorX also assigns confidence scores to predicted structures. See details below and [HERE](#).



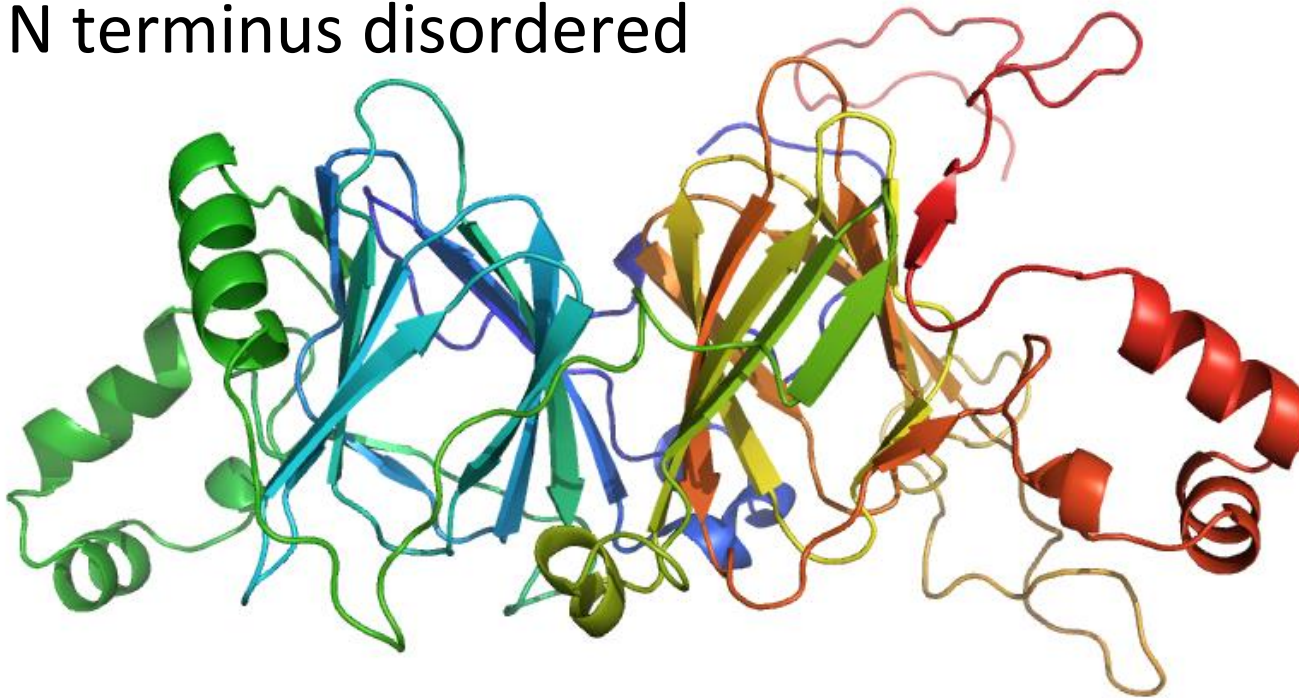
Vicilin experimental structures:

Gray: Crystal structure of the vicilin from *S. melongena* 5vf5

Pink: Structure of vicilin from *Capsicum annuum*, 5yjs

Tomato seed vicilin structural model

- Region Pro147-Phe567 modeled
- N terminus disordered



Model shown as cartoon colored from N terminus to C terminus (blue to red)



Trypsin, sequence based vicilin domain

Results of enzyme action

NPYLF ESQR - FR - SR - FR - ASHGDFR - ILE R - FNQR - SQLLK - GIEK - YR - VAIL ELEPQSFVLP HHCDGEAIYV VVK - GQGVINI AEQDNK - NSFN LQK - GDVIR - LFAGSNVYLLNK - DNNEK - LRVYV LAK - SVNAPGN LQEYFSAGGQ NPESFYR - AFS SDILESANFNPR - DK - LER - LFG QHK - EGIIK - A SEEQIR - AISE HASR - STQQTR - GR - TQGPFLNM K - ER - PVFESR - FGQFFEAR - PER - YEQLR - DLDAV VGFMINQGG MVLPPYNTK - S TK - LVMVIEGN AR - FEMACPHLGR - QSQSPWSR - GQGR - EQER - EQ EQEQEEGDVH YQK - IR - GNLNV GDVLPVAGH PITFVATGNSNLR - IVGFGVD AENNK - K - NFLA GK - QNIWR - NID R - EAK - ELSFSM PGR - EVEEIFQ R - QDQSYFVPEHR - QQR - ER - G EEGR - R - GQDQY LSSILDFVF

Location of released peptides

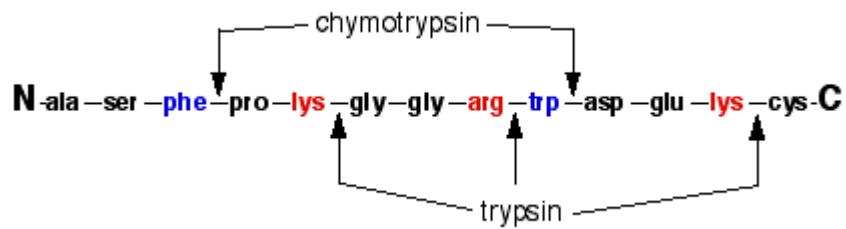
[1-10],[11-12],[13-14],[15-16],[17-24],[25-29],[30-33],[34-38],[39-42],[43-44],[45-74],[75-88],[89-96],[97-101],[102-113],[114-119],[120-128],[129-154],[155-170],[171-172],[173-175],[176-182],[183-188],[189-196],[197-205],[206-211],[212-214],[215-224],[225-226],[227-232],[233-240],[241-243],[244-249],[250-275],[276-279],[280-290],[291-300],[301-308],[309-313],[314-317],[318-334],[335-336],[337-366],[367-379],[380-380],[381-387],[388-392],[393-397],[398-400],[401-410],[411-419],[420-430],[431-433],[434-435],[436-441],[442-442],[443-457]

BIOPEP: Search for active fragments

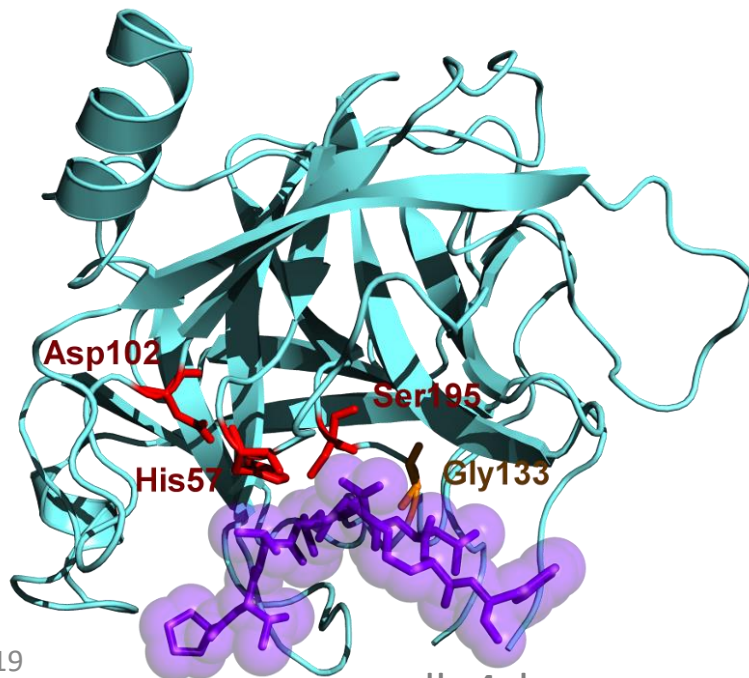
No	Peptide ID	Sequence	Location	Name	Function	Activity	Monoisotopic mass	Chemical mass
1	3258	IR	[335-336]	beta-lactokinin	Inhibitor of Angiotensin-Converting Enzyme (ACE) (EC 3.4.15.1) (MEROPS ID: M02-001)	ACE inhibitor	287.1850	287.3480
2	7592	FR	[11-12]	ACE inhibitor		ACE inhibitor	321.1690	321.3650
3	7592	FR	[15-16]	ACE inhibitor		ACE inhibitor	321.1690	321.3650
4	8130	EAK	[398-400]	peptide derived from dried bonito		antioxidative	346.1730	346.3690
5	8215	IR	[335-336]	Antioxidative peptide	Oxygen radical scavenging	antioxidative	287.1850	287.3480
6	8246	IR	[335-336]	renin inhibitor	Inhibitor of Renin (EC 3.4.23.15) (MEROPS ID A01.007)	renin inhibitor	287.1850	287.3480
7	8247	IR	[335-336]	CaMPDE inhibitor	Inhibitor of calmodulin-dependent phosphodiesterase 1 (abbrev. CaMPDE) (EC 3.1.4.17).	CaMPDE inhibitor	287.1850	287.3480
8	8780	FR	[11-12]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)	Inhibitor of Dipeptidyl Peptidase IV (EC 3.4.14.5) (MEROPS ID: S09.003)	dipeptidyl peptidase IV inhibitor	321.1690	321.3650
9	8780	FR	[15-16]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)	Inhibitor of Dipeptidyl Peptidase IV (EC 3.4.14.5) (MEROPS ID: S09.003)	dipeptidyl peptidase IV inhibitor	321.1690	321.3650
10	8806	IR	[335-336]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)	Inhibitor of Dipeptidyl Peptidase IV (EC 3.4.14.5) (MEROPS ID: S09.003)	dipeptidyl peptidase IV inhibitor	287.1850	287.3480
11	8944	YR	[43-44]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)	Inhibitor of Dipeptidyl Peptidase IV (EC 3.4.14.5) (MEROPS ID: S09.003)	dipeptidyl peptidase IV inhibitor	337.1640	337.3580
12	9484	YR	[43-44]	DPP-III inhibitor	Inhibitor of Dipeptidyl peptidase-III (DPP-III) (EC 3.4.14.4) (MEROPS ID: M49.001)	dipeptidyl peptidase III inhibitor	337.1640	337.3580
13	9501	FR	[11-12]	DPP-III inhibitor	Inhibitor of Dipeptidyl peptidase-III (DPP-III) (EC 3.4.14.4) (MEROPS ID: M49.001)	dipeptidyl peptidase III inhibitor	321.1690	321.3650
14	9501	FR	[15-16]	DPP-III inhibitor	Inhibitor of Dipeptidyl peptidase-III (DPP-III) (EC 3.4.14.4)	dipeptidyl peptidase	321.1690	321.3650

[Useful links](#)
[List of publications concerning BIOPEP-IWM database](#)
[About BIOPEP-IWM](#)
[Citing BIOPEP-IWM](#)
[Contact](#)

Proteolysis prediction based on structure



Chymotrypsin structure

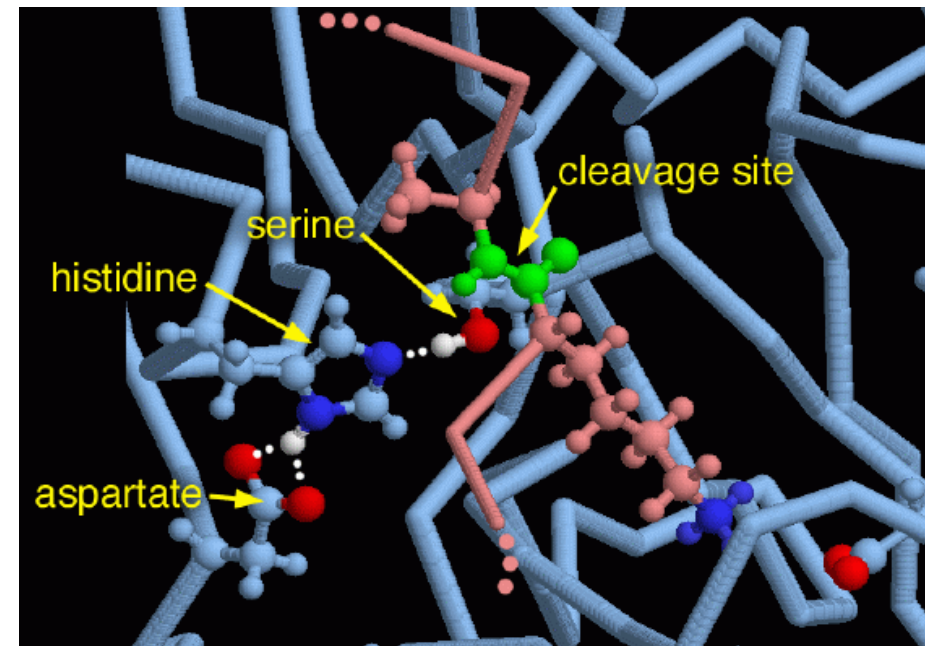


30/10/2019

pdb:4cha

Trypsin only cleaves the peptide bonds after (on the C-terminal side of) the basic amino acids **lysine** and **arginine**
Chymotrypsin prefers cleaving after large **hydrophobic** amino acids such as **phenylalanine**, **tyrosine** and **tryptophan**, and also **leucine** and **methionine**

Trypsin structure



Kaplan Türköz, Burcu

pdb:2ptc

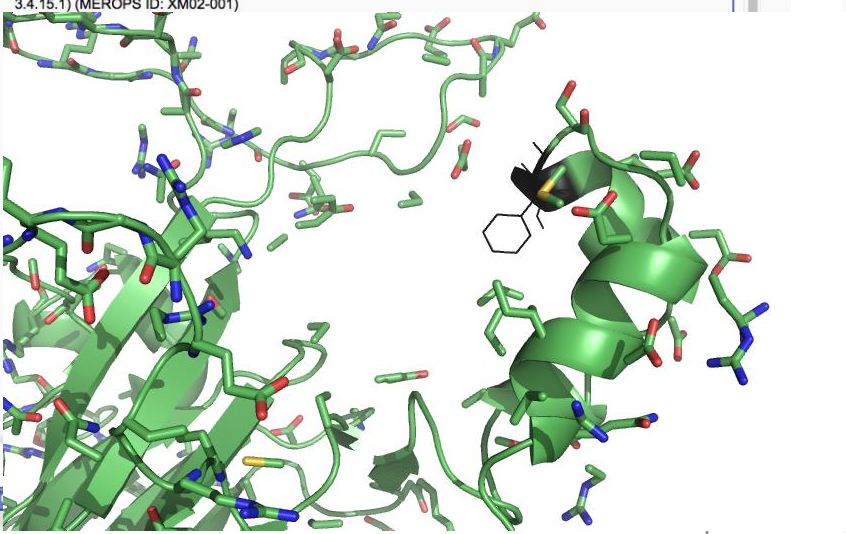
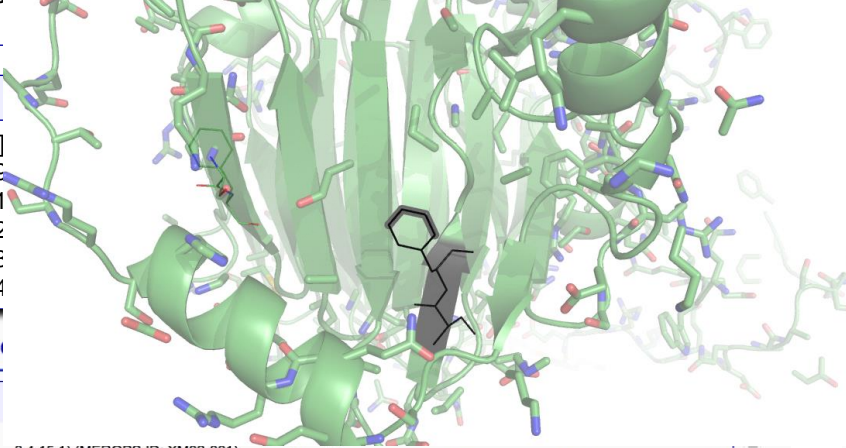
Chymotrypsin, sequence based vicilin domain

Results of enzyme action

N - PY - L - F - ESQRF - RSRF - R ASH - GDF - RIL - E RF - N - QRSQL - L - KGIEKY - RVAIL - EL - EPQSF - VL - P H - H - CDGEAIY - V VVKGGQGVIN - I AEQDN - KN - SF - N - L - QKGDVIRL - F - AGSN - VY - L - L - N - K DN - N - EKL - F - VY - V L - AKSVN - APGN - L - QEY - F - SAGGQ N - PESF - Y - RAF - S SDIL - ESAF - N - N - PRDKL - ERL - F - G QH - KEGIIKA
 SEEQIRAISE H - ASRSTQQTR GRTQGP - N - L - M - KERPVF - ESRF - GQF - F - EARPER Y - EQL - RDL - DAA VGF - M - N - IN - QGG M - VL - PY - Y - N - TKS TKL - VM - VIEGN - ARF - EM - ACPH - L - GRQSQSPW - SR GQGREQEREQ EQEQEEGDVH - Y - QKIRGN - L - N - V GDVL - VIPAGH - PITF - VATGN - SN - L - RIVGF - GVD AEN - N - KKN - F - L - A GKQN - IW - RN - ID REAKEL - SF - SM - PGREVEEIF - Q RODOSY - E - VPEH - ROORERF FEGRRGODOY - L - SSIL - DF - VF

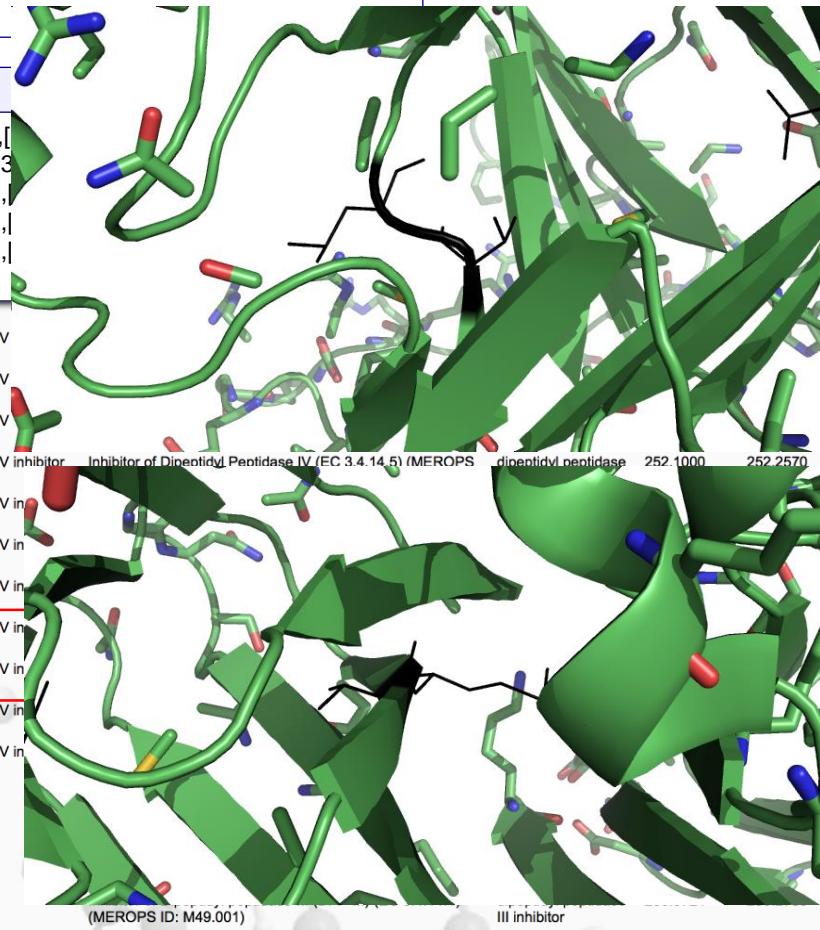
Location of released peptides

[1-1],[2-3],[4-4],[5-5],[6-11],[12-15],[16-20]
 94],[95-102],[103-103],[104-107],[108-109]
 148],[149-152],[153-153],[154-156],[157-1
 236],[237-237],[238-245],[246-248],[249-2
 297],[298-298],[299-306],[307-330],[331-3
 383],[384-389],[390-391],[392-393],[394-4



8],[59-61],[62-62],[63-69],[70-80],[
],[124-126],[127-131],[132-135],[13
 02],[203-219],[220-220],[221-221],
 72],[273-273],[274-280],[281-282],
 34],[365-365],[366-370],[371-377],
 49],[450-453],[454-455],[456-457]

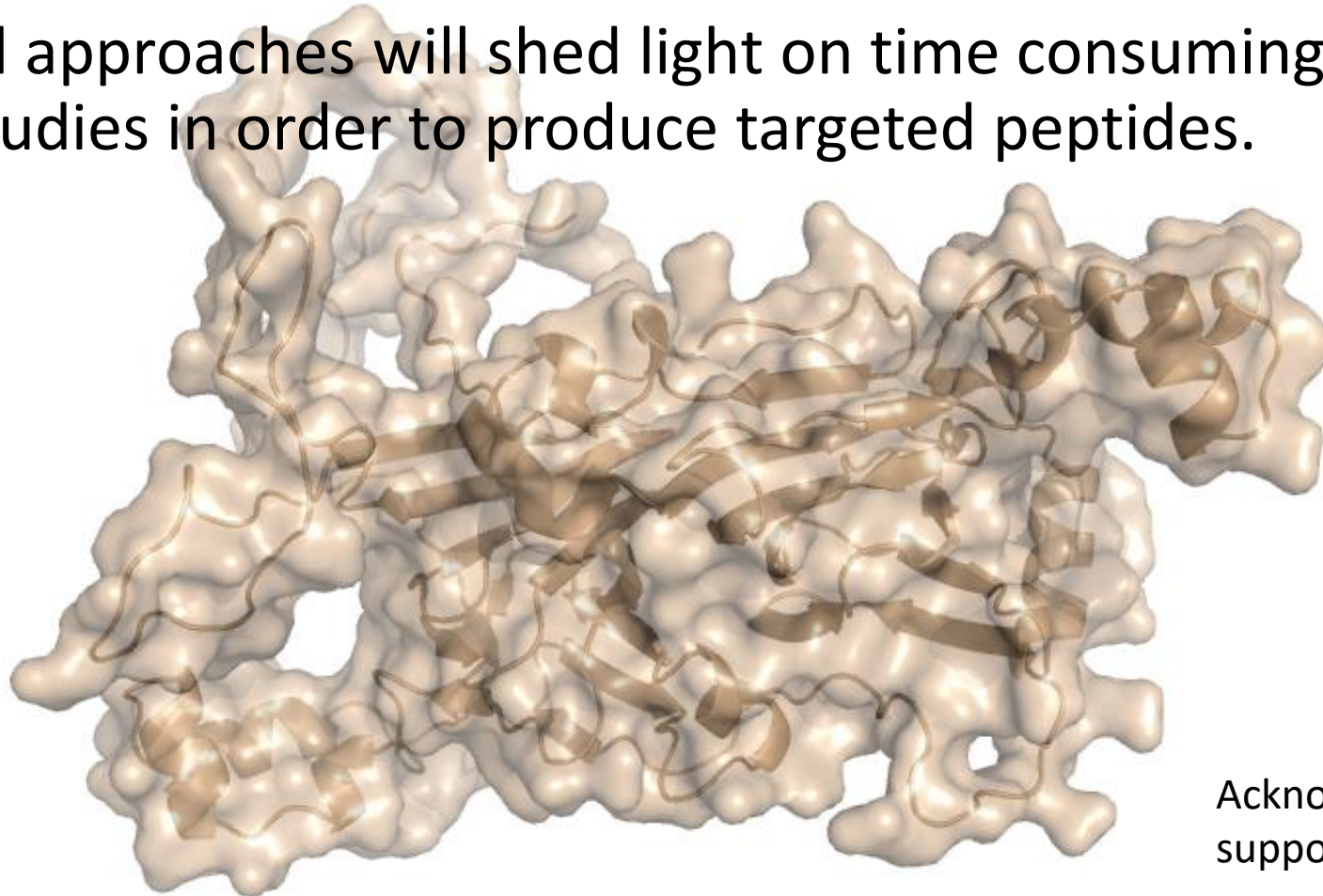
No	Peptide ID	Sequence	Location	Name
1	3384	VF	[456-457]	ACE inhibitor
2	3492	VY	[108-109]	ACE inhibitor from sake
3	3492	VY	[122-123]	ACE inhibitor from sake
4	7544	IW	[390-391]	ACE inhibitor
5	7685	SF	[90-91]	ACE inhibitor from garlic
6	7685	SF	[403-404]	ACE inhibitor from garlic
7	8224	VY	[108-109]	antioxidative peptide
8	8224	VY	[122-123]	antioxidative peptide
9	8320	VL	[57-58]	Glucose uptake stimulating peptide
10	8320	VL	[268-269]	Glucose uptake stimulating peptide
11	8804	IN	[261-262]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
12	8807	IW	[390-391]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
13	8866	PY	[2-3]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
14	8866	PY	[270-271]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
15	8888	RN	[392-393]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
16	8891	SF	[90-91]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
17	8891	SF	[403-404]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
18	8917	VF	[456-457]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
19	8922	VL	[57-58]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
20	8922	VL	[268-269]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
21	8923	VM	[281-282]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
22	8929	VY	[108-109]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
23	8929	VY	[122-123]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
24	9074	DF	[454-455]	ACE inhibitor
25	9432	SF	[90-91]	Renin inhibitor
26	9432	SF	[403-404]	Renin inhibitor
27	9507	SM	[405-406]	DPP-III inhibitor



No	Peptide ID	Sequence	Location	Name
1	3384	VF	[456-457]	ACE inhibitor
2	3492	VY	[108-109]	ACE inhibitor from sake
3	3492	VY	[122-123]	ACE inhibitor from sake
4	7544	IW	[390-391]	ACE inhibitor
5	7685	SF	[90-91]	ACE inhibitor from garlic
6	7685	SF	[403-404]	ACE inhibitor from garlic
7	8224	VY	[108-109]	antioxidative peptide
8	8224	VY	[122-123]	antioxidative peptide
9	8320	VL	[57-58]	Glucose uptake stimulating peptide
10	8320	VL	[268-269]	Glucose uptake stimulating peptide
11	8804	IN	[261-262]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
12	8807	IW	[390-391]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)
13	8866	PY	[2-3]	dipeptidyl peptidase IV inhibitor (DPP IV inhibitor)

Conclusions

- BAPs are very useful considering their therapeutic potential.
- Sequence based prediction does not take into account the buried residues.
- Structure based approaches will shed light on time consuming experimental studies in order to produce targeted peptides.



Acknowledgements: This work is supported by TÜBİTAK(117O319)

Active Projects and Collaborators

- Identification and characterization of probiotic TIR domain proteins
- **Characterization and Determination of Functional Properties of Bioactive peptides from plant food by-products ; Prof. Dr. Semih Ötleş**
- Structural characterization of levansucrase and its use in functional food production; Prof. Dr. Taner Baysal, Prof. Dr. Yekta Göksungur
- Purification and characterization of collagen and related molecules from food industry by-products; Prof. Dr. Şebnem Tavman

Burcu KAPLAN TÜRKÖZ Laboratory of Molecular Biology, Department of Food Engineering, Ege University, İzmir, Turkey