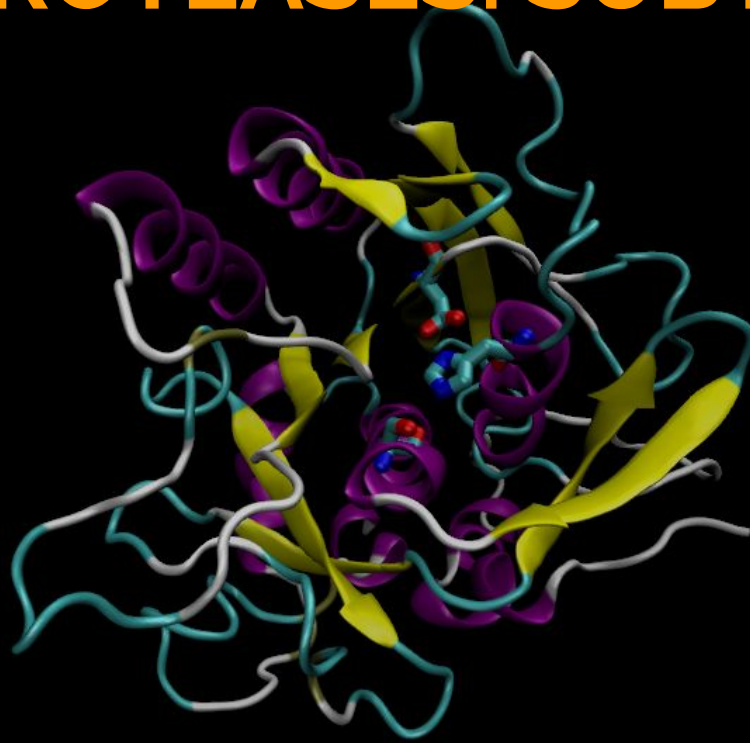


SERIN-PROTEASES: SUBTILISINS



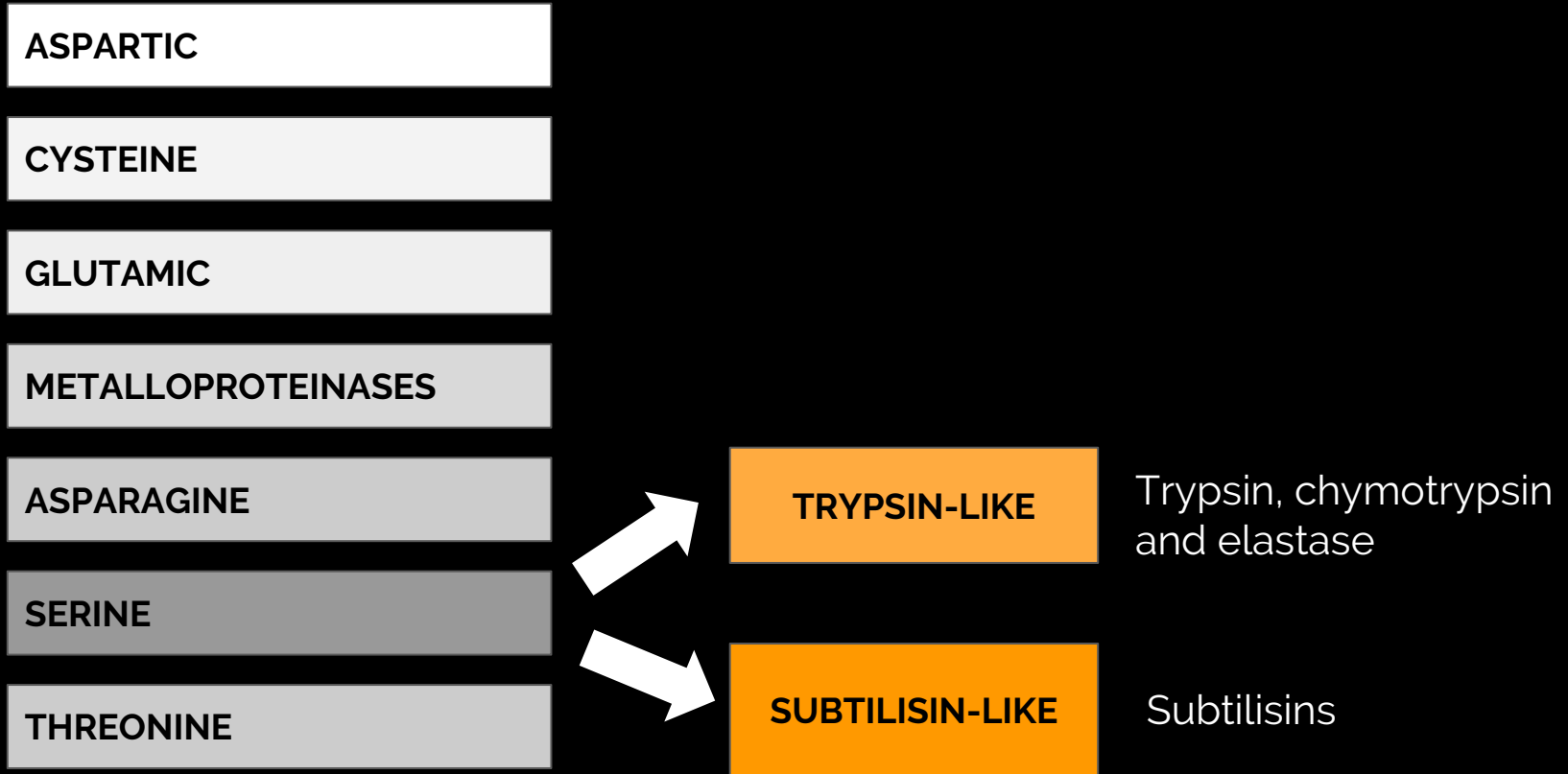
Júlia Garcia, Gemma Núñez, Marina Solà, Marc Ten
24th February 2017

PRESENTATION

- Introduction serine proteases
- Subtilisins
- Sequences analysis
- Subtilisins vs chymotrypsin
- Conclusions
- Questions



Introduction - Proteases



Introduction - SCOP Classification

Class	Fold	Superfamilies	Families
All beta proteins	Closed barrel with greek-key motifs	Trypsin-like serine proteases	<ul style="list-style-type: none">- prokaryotic proteases- eukaryotic proteases (Trypsin, chymotrypsin and elastase)- viral proteases- viral cysteine protease of trypsin fold
Alpha and beta proteins (a/b)	Beta Sheet $\alpha/\beta/\alpha$ Left-Handed Crossover	Subtilisin-like	<ul style="list-style-type: none">- Subtilases (subtilisin)- Serine-carboxyl proteinase (sedolisin)

Introduction - Function

Serine proteinases cleave peptide bonds within a polypeptide to produce two new smaller peptides by forming tetrahedral transition states

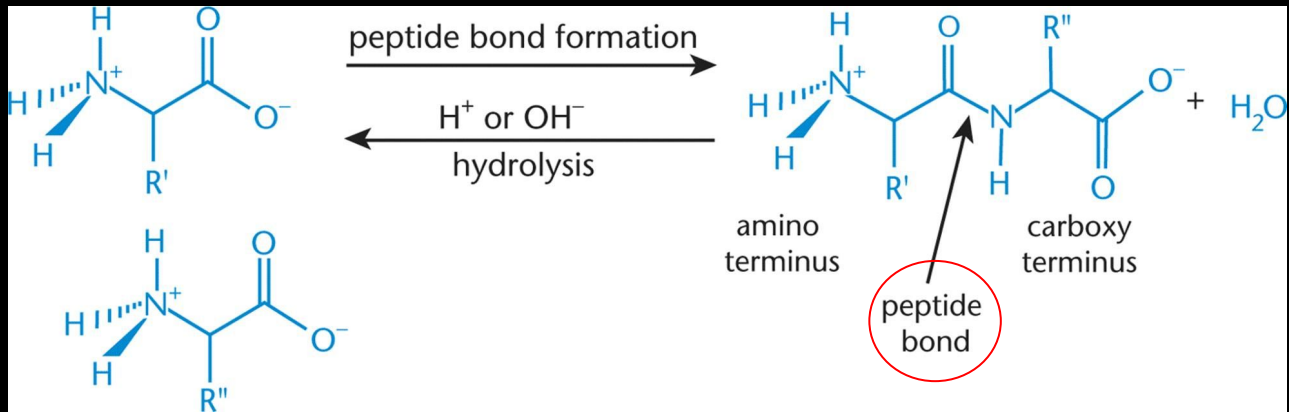
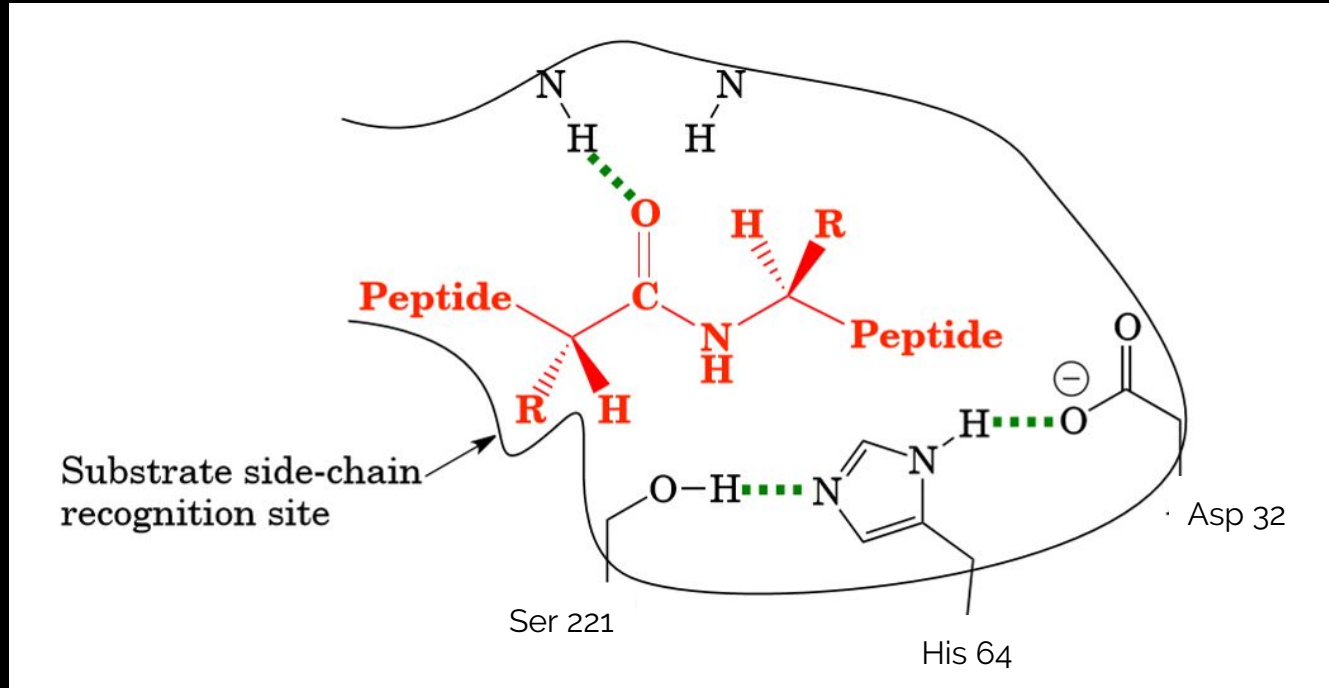


Image obtained from: http://schoolbag.info/chemistry/mcat_biochemistry/mcat_biochemistry_files/image019.jpg

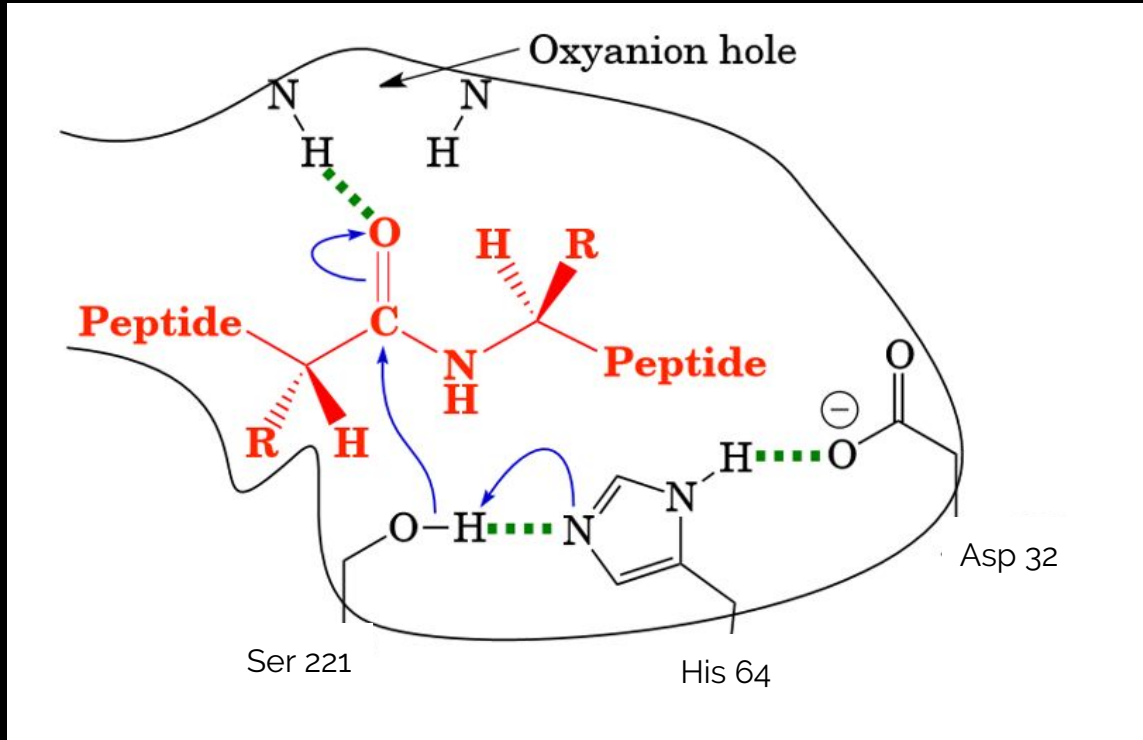
Catalytic mechanism

1. Substrate binding



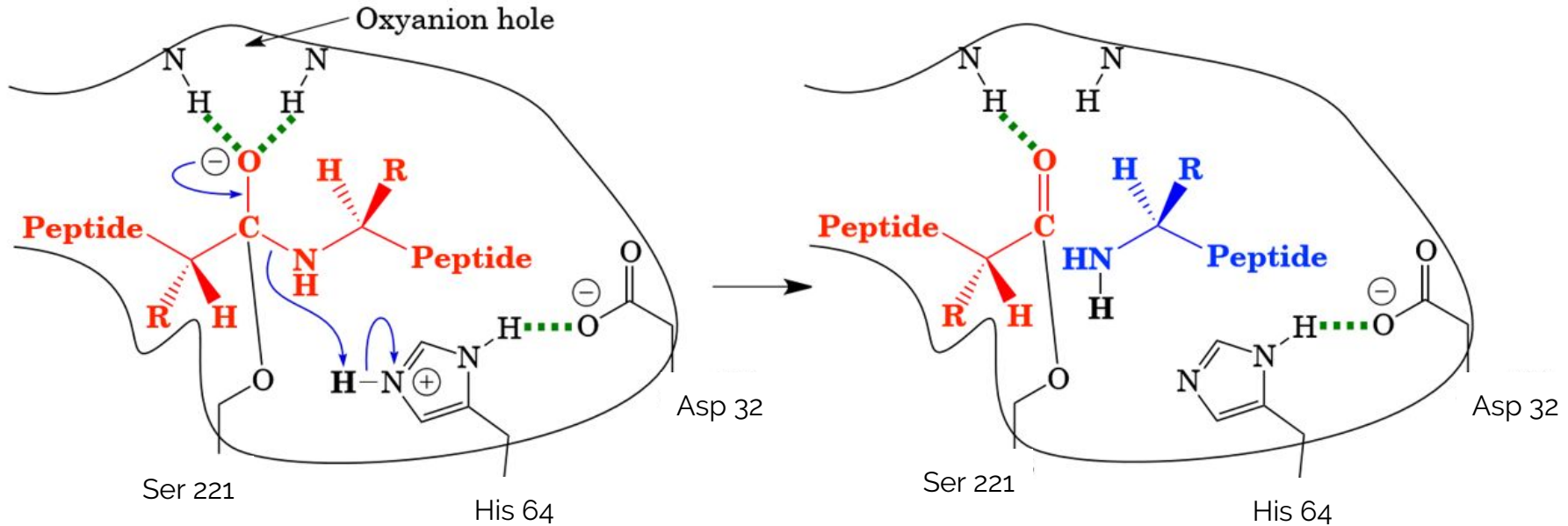
Catalytic mechanism

2. Nucleophilic attack: Ser hydroxyl group attacks the carbonyl group of the substrate.



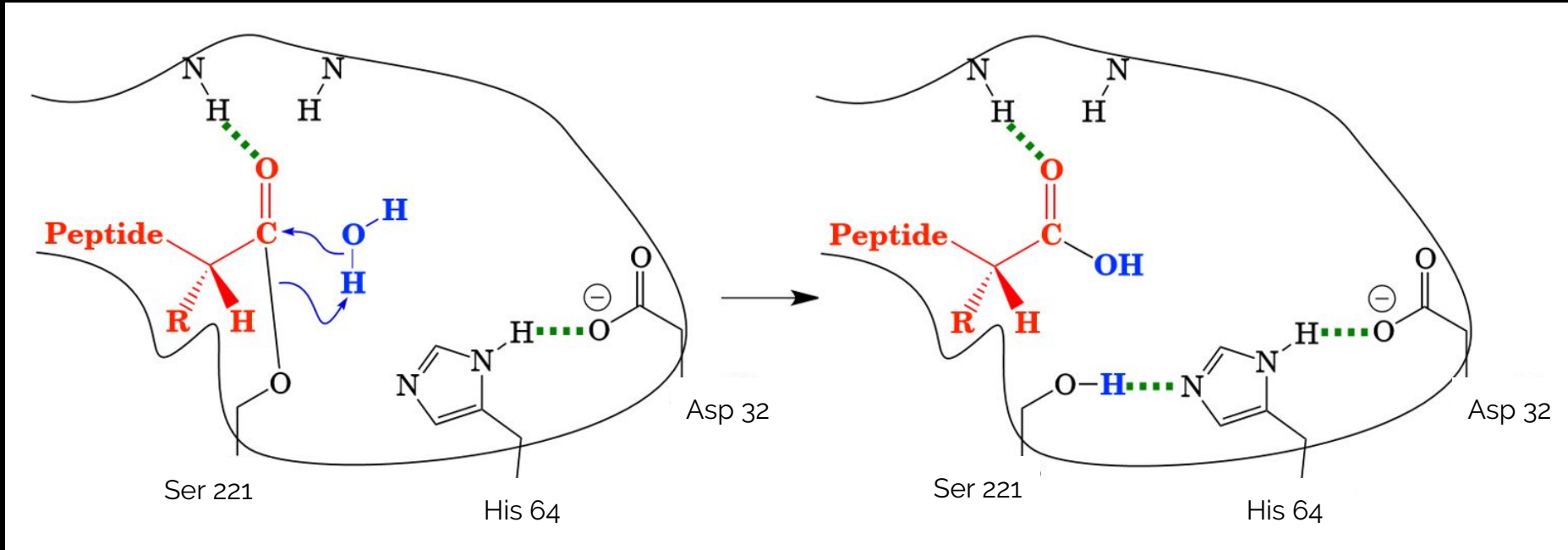
Introduction - Catalytic mechanism

3. Covalent bonding (acylation)



Introduction - Catalytic mechanism

4. Ester hydrolysis (deacylation):



Subtilisins

- SCOP clasification
- Activation mechanism
- Prosubtilisin
- Structure
- Structural analysis
- Essential features
- Sequence analysis



Subtilisins - SCOP Clasification

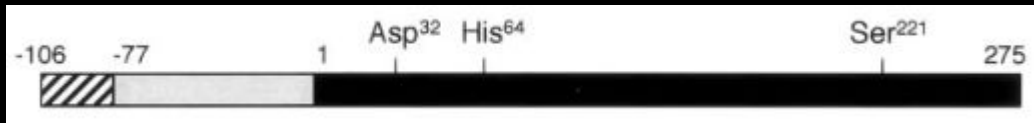
Families	Protein	Species	PDB code
Subtilases	Subtilisin	<ul style="list-style-type: none"> - Bacillus subtilis, carlsberg - Bacillus subtilis, E - Bacillus licheniformis - Bacillus lentus, savinase (TM) - Bacillus lentus - Bacillus amyloliquefaciens, Novo/BPN 	<ul style="list-style-type: none"> - 1sbc, 2ec - 1scj - 1gcil - 1st3 - 1sbt
	Messengericopeptidase	Bacillus mesentericus	1mee
	M-proteinase	Bacillus sp.,	1wsd
	Thermitase	Thermoactinomyces vulgaris	1Thm
	Furin	Mouse (Mus musculus)	1p8j
	Kexin	Saccharomyces cerevisiae	2id4

Subtilisins - Activation mechanism

Subtilisins are activated from pro-subtilisins by 3 steps:

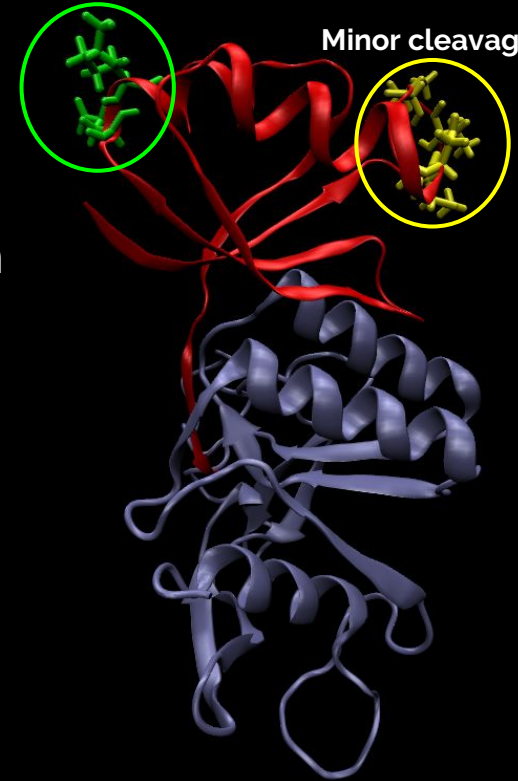
- ① Folding of the subtilisin domain
- ② Autoprocessing between the propeptide and subtilisin domains
- ③ Degradation of the propeptide by activated subtilisin.

Propeptide (77 residues) → intramolecular chaperone, temporal potent inhibitor of subtilisin.

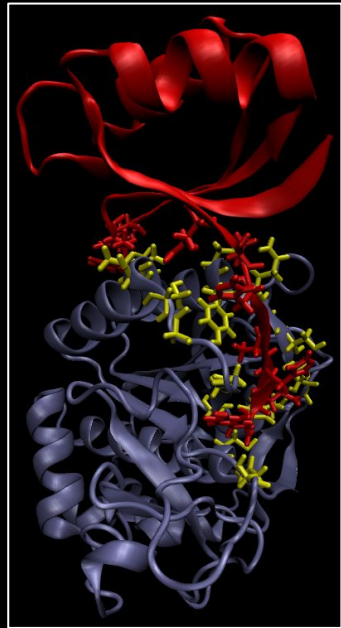


Main cleavage site

Minor cleavage site

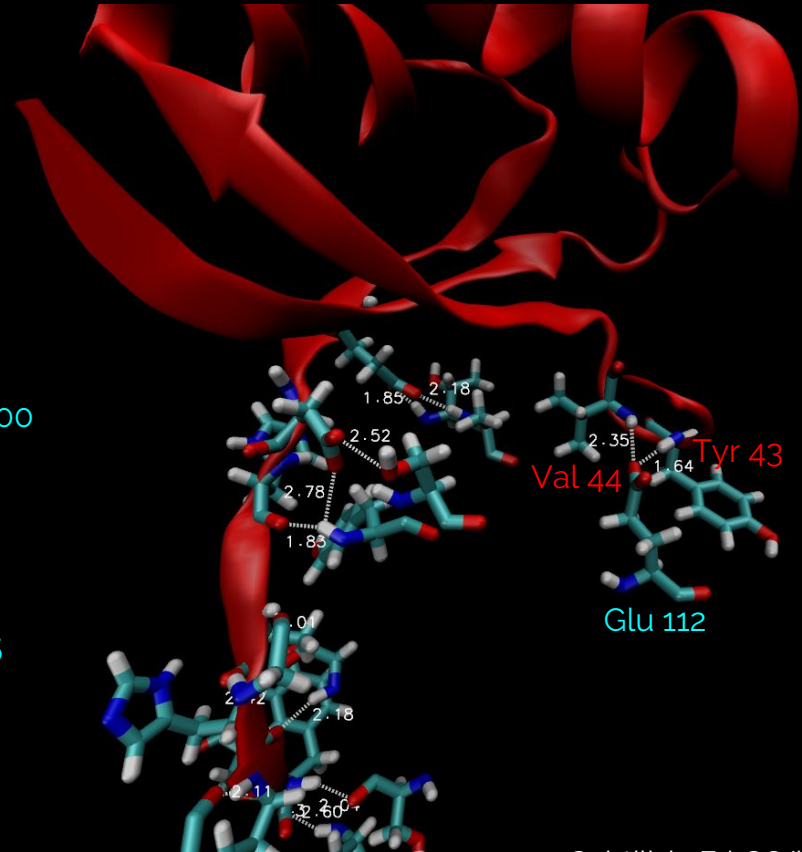
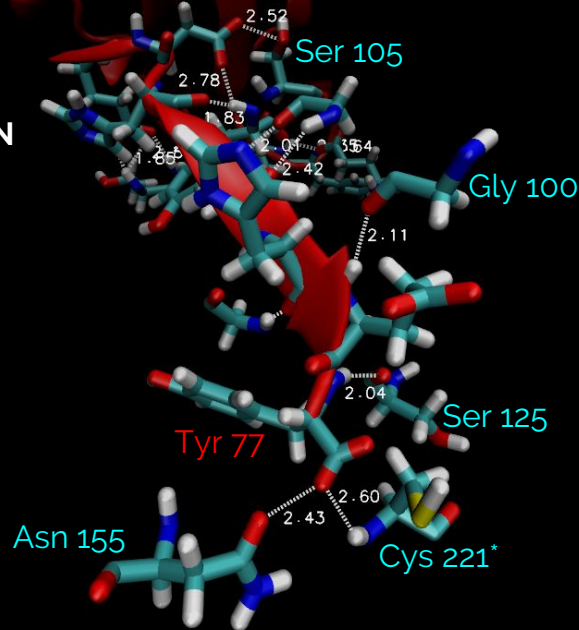


Subtilisins - Interactions between propeptid-subtilisin



PROPEPTIDE DOMAIN

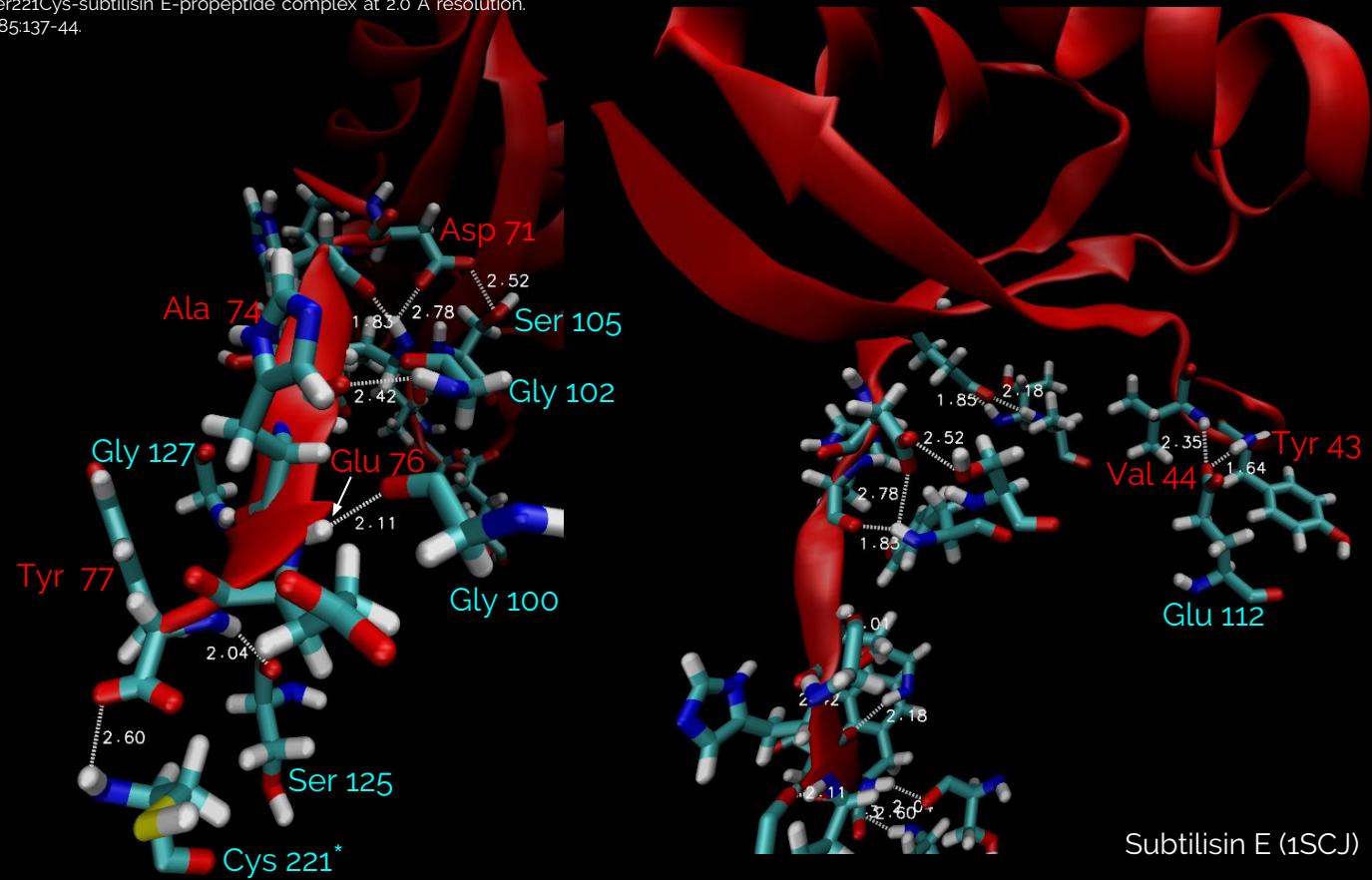
SUBTLISIN



Subtilisins - Interactions between propeptid-subtilisin

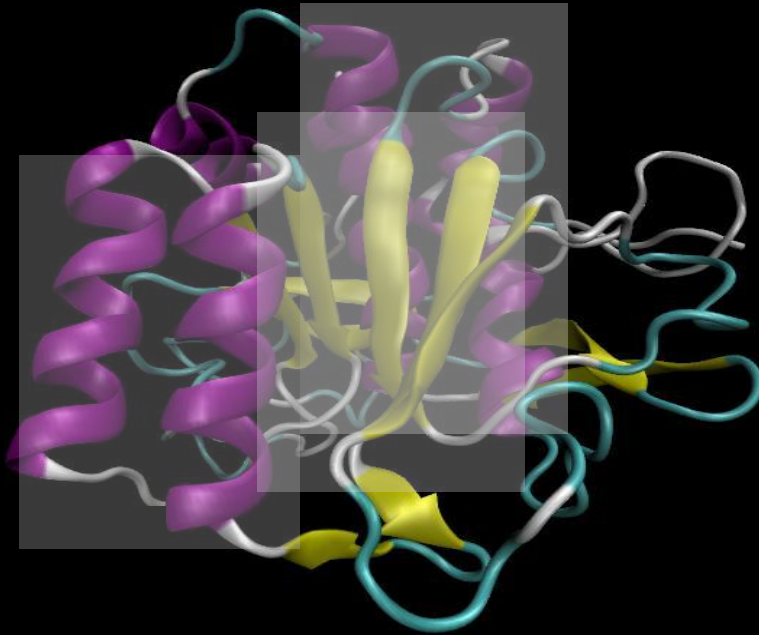
Jain SC, Shinde U, Li Y, Inouye M, Berman HM. The crystal structure of an autoprocessed Ser221Cys-subtilisin E-propeptide complex at 2.0 Å resolution. *J. Mol. Biol.* 1998;285:137-44.

Subtilisin		Propeptide		
Residue no.	Atom	Residue no.	Atom	
<i>Hydrogen bonds</i>				
Ser125	O	Tyr77	N ^a	2.04 Å
Gly128	N	Tyr77	O ^{va}	
Cys221	N	Tyr77	O ^a	2.60 Å
Asn155	O ^{δ2}	Tyr77	OT1	2.43 Å
His64	N ^{ε2}	Tyr77	OT2	
Ser62	O ^γ	Glu76	O ^{ε2a}	
Gly100	O	Glu76	N	2.11 Å
Ser101	O ^γ	His75	N ^{ε2a}	
Gly127	N	His75	O	2.18 Å
Gly127	O	His75	N	
Gly102	N	Ala74	O	2.42 Å
Gly102	O	Ala74	N	2.01 Å
Gly131	O	His72	N ^{ε2}	
Tyr104	N	His72	O	1.83 Å
Tyr104	N	Asp71	O ^{δ2}	2.78 Å
Ser105	N	Asp71	O ^{δ2}	
Ser105	O ^γ	Asp71	O ^{δ2}	2.52 Å
Ser105	O ^γ	Asp71	O ^{δ1}	
Gln103	N ^{ε2}	Asp71	O ^{δ1}	
Ser132	O ^γ	Glu69	O ^{ε1}	
Thr133	N	Glu69	O ^{ε1}	1.85 Å
Thr133	N	Glu69	O ^{ε2}	2.83 Å
Ala134	N	Glu69	O ^{ε1}	2.18 Å
Glu112	O ^{ε2}	Val44	N	2.35 Å
Glu112	O ^{ε2}	Tyr43	N	1.64 Å
Glu112	O ^{ε2}	Lys42	N	1.92 Å
Gln103	O ^{ε1}	Lys9	N ^ε	



Subtilisin E (1SCJ)

Subtilisins - Structure

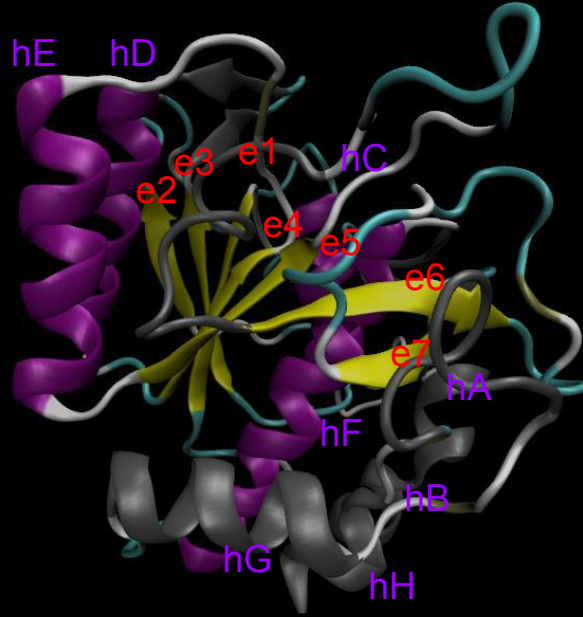
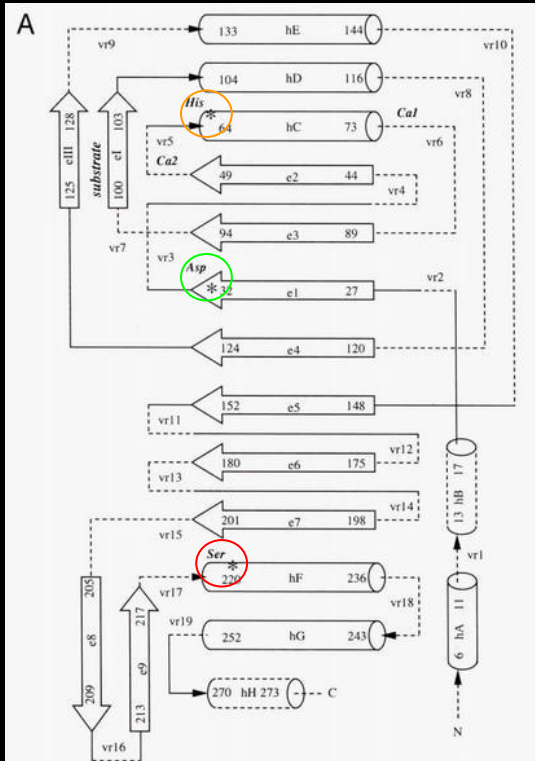


Subtilisin BPN (1SBT)

- Organism: *Bacillus amyloliquefaciens*
- A single polypeptide chain of 275 amino acids, arranged in an α/β structure
- 3 layers: $\alpha/\beta/\alpha$
- Parallel beta-sheet of 7 strands
- Left-handed crossover connection between strands 2 & 3

Subtilisins - Structure

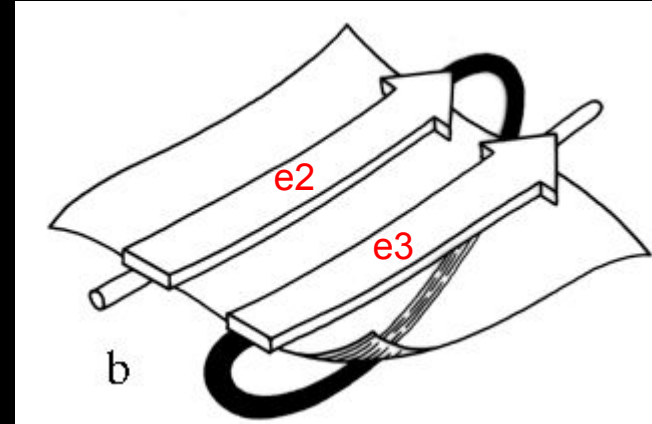
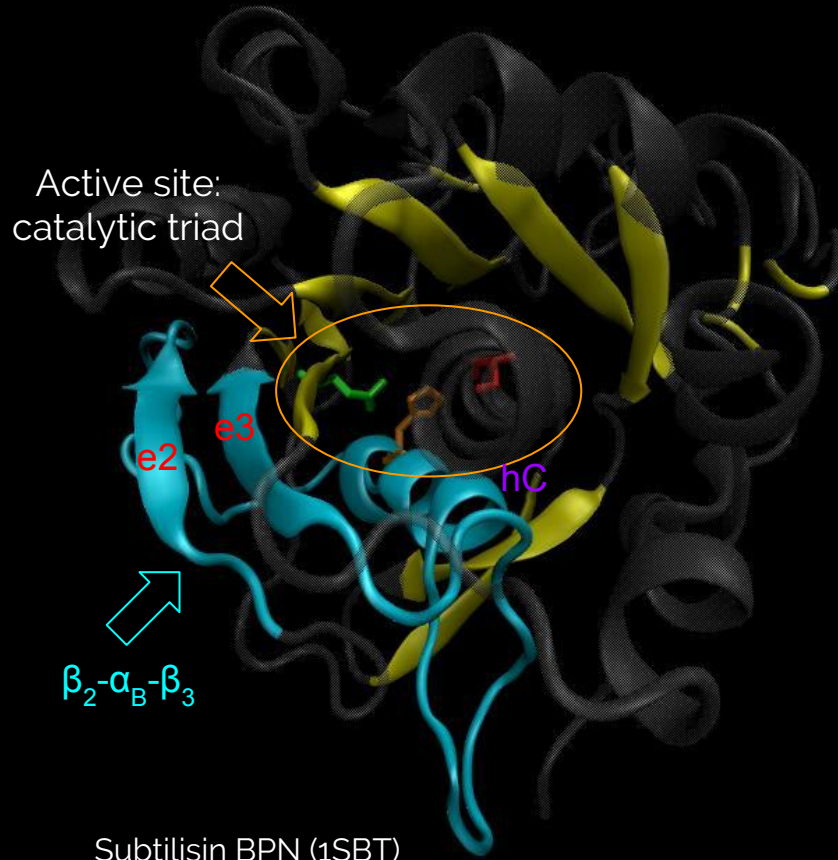
7 parallel β strands surrounded by four helices, two on each side of the parallel β sheet.



Subtilisin BPN1 (1SBT)



Subtilisins - Structure



<http://kinemage.biochem.duke.edu/teaching/anatax/html/anatax.2b.html>

Bacterial subtilisins sequence alignment, HMM

001	sp P29599 SUBB_BACLE	a	q	s	v	p	w	g	i	s	r	v	q	a	p	a	h	n	r	g	l	t	G	S	G	V	K	V	A	V	L	D	T	G	I	-	S	T	H	P	D	L	N	-	I	R	G	G	A	S	
002	sp P00782 SUBT_BACAM	a	q	s	v	p	y	g	v	s	q	i	k	a	p	a	l	h	s	q	g	y	t	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H	P	D	L	K	V	A	G	G	A	S	M
003	sp P04189 SUBT_BACSU	a	q	s	v	p	y	g	i	s	q	i	k	a	p	a	l	h	s	q	g	y	t	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H	P	D	L	N	V	R	G	G	A	S	F
004	sp P00783 SUBT_BACSA	a	q	s	v	p	y	g	i	s	q	i	k	a	p	a	l	h	s	q	g	y	t	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H	P	D	L	N	V	R	G	G	A	S	F
005	sp P07518 SUBT_BACPU	a	q	s	v	p	y	g	i	s	q	i	k	a	p	a	l	h	s	q	g	y	t	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H	P	D	L	N	V	R	G	G	A	S	F
006	sp P35835 SUBN_BACNA	a	q	s	v	p	y	g	i	s	q	i	k	a	p	a	l	h	s	q	g	y	t	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H	P	D	L	N	V	R	G	G	A	S	F
007	sp P29142 SUBT_GEOSE	a	q	s	v	p	y	g	i	s	q	i	k	a	p	a	l	h	s	q	g	y	t	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H	P	D	L	N	V	R	G	G	A	S	F
008	sp P29600 SUBS_BACLE	a	q	s	v	p	w	g	i	s	r	v	q	a	p	a	h	n	r	g	l	t	G	S	G	V	K	V	A	V	L	D	T	G	I	-	S	T	H	P	D	L	N	-	I	R	G	G	A	S	
009	sp P00780 SUBT_BACLI	a	q	t	v	p	y	g	i	p	l	i	k	a	d	k	v	q	a	q	g	f	k	G	A	N	V	K	V	A	V	L	D	T	G	I	Q	A	S	H	P	D	L	N	-	V	V	G	G	A	S

hA

hB

e1 Asp 32

e2

001	sp P29599 SUBB_BACLE	F	V	P	G	E	P	S	T	Q	D	G	N	G	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	E	L	Y	A	V	K	V	L	G	A	D	G
002	sp P00782 SUBT_BACAM	V	P	S	E	T	N	P	F	Q	D	N	N	S	H	G	T	H	V	A	G	T	V	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	S	L	Y	A	V	K	V	L	G	A	D	G
003	sp P04189 SUBT_BACSU	V	P	S	E	T	N	P	Y	Q	D	G	S	S	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	S	L	Y	A	V	K	V	L	D	S	T	G
004	sp P00783 SUBT_BACSA	V	P	S	E	T	N	P	Y	Q	D	G	S	S	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	S	P	S	A	S	L	Y	A	V	K	V	L	D	S	T	G
005	sp P07518 SUBT_BACPU	V	P	S	E	T	N	P	Y	Q	D	G	S	S	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	S	L	Y	A	V	K	V	L	D	S	T	G
006	sp P35835 SUBN_BACNA	V	P	S	E	T	N	P	Y	Q	D	G	S	S	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	S	L	Y	A	V	K	V	L	D	S	T	G
007	sp P29142 SUBT_GEOSE	V	P	S	E	T	N	P	Y	Q	D	G	S	S	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	S	P	S	A	S	L	Y	A	V	K	V	L	D	S	T	G
008	sp P29600 SUBS_BACLE	F	V	P	G	E	P	S	T	Q	D	G	N	G	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	E	L	Y	A	V	K	V	L	G	A	S	G
009	sp P00780 SUBT_BACLI	F	V	A	G	E	A	Y	N	T	D	G	N	G	H	G	T	H	V	A	G	T	V	A	A	L	D	N	T	T	G	V	L	G	V	A	P	S	V	S	L	Y	A	V	K	V	L	N	S	S	G

His 64

hC

e3

001	sp P29599 SUBB_BACLE	R	G	A	-	I	S	S	I	A	Q	G	L	E	W	A	G	N	N	G	M	H	V	A	N	L	S	L	G	-	S	P	S	P	S	A	T	L	E	Q	A	V	N	S	A	T	S	R	G	V	L
002	sp P00782 SUBT_BACAM	S	G	Q	Y	-	S	W	I	I	N	G	I	E	W	A	I	A	N	N	M	D	V	I	N	M	S	L	G	G	P	S	G	S	A	A	L	K	A	-	A	V	D	K	A	V	A	S	G	V	V
003	sp P04189 SUBT_BACSU	S	G	Q	Y	-	S	W	I	I	N	G	I	E	W	A	I	S	N	N	M	D	V	I	N	M	S	L	G	G	P	T	G	S	T	A	L	K	T	-	V	V	D	K	A	V	S	S	G	I	V
004	sp P00783 SUBT_BACSA	S	G	Q	Y	-	S	W	I	I	N	G	I	E	W	A	I	S	N	N	M	D	V	I	N	M	S	L	G	G	P	S	G	S	T	A	L	K	T	-	V	V	D	K	A	V	S	S	G	I	V
005	sp P07518 SUBT_BACPU	S	G	Q	Y	-	S	W	I	I	N	G	I	E	W	A	I	S	N	N	M	D	V	I	N	M	S	L	G	G	P	T	G	S	T	A	L	K	T	-	V	V	D	K	A	V	S	S	G	I	V
006	sp P35835 SUBN_BACNA	S	G	Q	Y	-	S	W	I	I	N	G	I	E	W	A	I	S	N	N	M	D	V	I	N	M	S	L	G	G	P	T	G	S	T	A	L	K	T	-	V	V	D	K	A	V	S	S	G	I	V
007	sp P29142 SUBT_GEOSE	S	G	Q	Y	-	S	W	I	I	N	G	I	E	W	A	I	S	N	N	M	D	V	I	N	M	S	L	G	G	P	S	G	S	T	A	L	K	T	-	V	V	D	K	A	V	S	S	G	I	V
008	sp P29600 SUBS_BACLE	S	G	S	V	-	S	S	I	A	Q	G	L	E	W	A	G	N	N	G	M	H	V	A	N	L	S	L	G	-	S	P	S	P	S	A	T	L	E	Q	A	V	N	S	A	T	S	R	G	V	L
009	sp P00780 SUBT_BACLI	S	G	T	Y	-	S	G	I	V	S	G	I	E	W	A	T	N	N	G	M	D	V	I	N	M	S	L	G	G	P	S	G	S	T	A	M	K	Q	-	A	V	D	N	A	Y	A	R	G	V	V

hD

e4

hE

e5

1	2	3	4	5	6	7	8	9
Variable	Average			Conserved				

Bacterial subtilisins sequence alignment, HMM

001	sp P29599 SUBB_BACLE	V	V	A	A	S	G	N	S	G	-	-	-	-	A	S	S	I	S	Y	P	A	R	Y	A	N	A	M	A	V	G	A	T	D	Q	N	N	N	R	A	S	F	S	Q	Y	G	A	G	L		
002	sp P00782 SUBT_BACAM	V	V	A	A	A	G	N	E	G	T	S	G	S	-	-	S	S	T	V	G	Y	P	G	K	Y	P	S	T	I	A	V	G	A	V	D	S	S	N	Q	R	A	S	F	S	S	V	G	P	E	L
003	sp P04189 SUBT_BACSU	V	A	A	A	A	G	N	E	G	S	S	G	S	-	-	T	S	T	V	G	Y	P	A	K	Y	P	S	T	I	A	V	G	A	V	N	S	S	N	Q	R	A	S	F	S	S	A	G	S	E	L
004	sp P00783 SUBT_BACSA	V	A	A	A	A	G	N	E	G	S	S	G	S	-	-	S	S	T	V	G	Y	P	A	K	Y	P	S	T	I	A	V	G	A	V	N	S	S	N	Q	R	A	S	F	S	S	A	G	S	E	L
005	sp P07518 SUBT_BACPU	V	A	A	A	A	G	N	E	G	S	S	G	S	-	-	T	S	T	V	G	Y	P	A	K	Y	P	S	T	I	A	V	G	A	V	N	S	A	N	Q	R	A	S	F	S	S	A	G	S	E	L
006	sp P35835 SUBN_BACNA	V	A	A	A	A	G	N	E	G	S	S	G	S	-	-	T	S	T	V	G	Y	P	A	K	Y	P	S	T	I	A	V	G	A	V	N	S	S	N	Q	R	A	S	F	S	S	V	G	S	E	L
007	sp P29142 SUBT_GEOSE	V	A	A	A	A	G	N	E	G	S	S	G	S	-	-	S	S	T	V	G	Y	P	A	K	Y	P	S	T	I	A	V	G	A	V	N	S	S	N	Q	R	A	S	F	S	S	A	G	S	E	L
008	sp P29600 SUBS_BACLE	V	V	A	A	S	G	N	S	G	-	-	-	-	A	G	S	I	S	Y	P	A	R	Y	A	N	A	M	A	V	G	A	T	D	Q	N	N	N	R	A	S	F	S	Q	Y	G	A	G	L		
009	sp P00780 SUBT_BACLI	V	V	A	A	A	G	N	S	G	S	-	-	G	N	T	N	T	I	G	Y	P	A	K	Y	D	S	V	I	A	V	G	A	V	D	S	N	S	N	R	A	S	F	S	S	V	G	A	E	L	

e6

e7

001	sp P29599 SUBB_BACLE	D	I	V	A	P	G	V	N	V	Q	S	T	Y	P	G	S	T	Y	A	S	L	N	G	T	S	M	A	T	P	H	V	A	G	A	A	L	V	K	Q	K	N	P	S	W	S	N	V	Q	I
002	sp P00782 SUBT_BACAM	D	V	M	A	P	G	V	S	I	Q	S	T	L	P	G	N	K	Y	G	A	Y	N	G	T	S	M	A	S	P	H	V	A	G	A	A	L	I	L	S	K	H	P	N	W	T	N	T	Q	V
003	sp P04189 SUBT_BACSU	D	V	M	A	P	G	V	S	I	Q	S	T	L	P	G	G	T	Y	G	A	Y	N	G	T	S	M	A	T	P	H	V	A	G	A	A	L	I	L	S	K	H	P	T	W	T	N	A	Q	V
004	sp P00783 SUBT_BACSA	D	V	M	A	P	G	V	S	I	Q	S	T	L	P	G	G	T	Y	G	A	Y	N	G	T	S	M	A	T	P	H	V	A	G	A	A	L	I	L	S	K	H	P	T	W	T	N	A	Q	V
005	sp P07518 SUBT_BACPU	D	V	M	A	P	G	V	S	I	Q	S	T	L	P	G	G	T	Y	G	A	Y	N	G	T	S	M	A	T	P	H	V	A	G	A	A	L	I	L	S	K	H	P	T	W	T	N	A	Q	V
006	sp P35835 SUBN_BACNA	D	V	M	A	P	G	V	S	I	Q	S	T	L	P	G	G	T	Y	G	A	Y	N	G	T	S	M	A	T	P	H	V	A	G	A	A	L	I	L	S	K	H	P	T	W	T	N	A	Q	V
007	sp P29142 SUBT_GEOSE	D	V	M	A	P	G	V	S	I	Q	S	T	L	P	G	G	T	Y	G	A	Y	N	G	T	S	M	A	T	P	H	V	A	G	A	A	L	I	L	S	K	H	P	T	W	T	N	A	Q	V
008	sp P29600 SUBS_BACLE	D	I	V	A	P	G	V	N	V	Q	S	T	Y	P	G	S	T	Y	A	S	L	N	G	T	S	M	A	T	P	H	V	A	G	A	A	L	V	K	Q	K	N	P	S	W	S	N	V	Q	I
009	sp P00780 SUBT_BACLI	E	V	M	A	P	G	A	G	V	Y	S	T	Y	P	T	S	T	Y	A	T	L	N	G	T	S	M	A	S	P	H	V	A	G	A	A	L	I	L	S	K	H	P	N	L	S	A	S	Q	V

Ser 221

hF

001	sp P29599 SUBB_BACLE	R	N	H	L	K	N	T	A	T	S	L	G	S	T	N	L	Y	G	S	G	L	v	n	a	e	a	a	t	r
002	sp P00782 SUBT_BACAM	R	S	S	L	E	N	T	T	T	K	L	G	D	S	F	Y	Y	G	K	G	L	i	n	v	q	a	a	a	q
003	sp P04189 SUBT_BACSU	R	D	R	L	E	S	T	A	T	Y	L	G	N	S	F	Y	Y	G	K	G	L	i	n	v	q	a	a	a	q
004	sp P00783 SUBT_BACSA	R	D	R	L	E	S	T	A	T	Y	L	G	N	S	F	Y	Y	G	K	G	L	i	n	v	q	a	a	a	q
005	sp P07518 SUBT_BACPU	R	D	R	L	E	S	T	A	T	Y	L	G	N	S	F	Y	Y	G	K	G	L	i	n	v	q	a	a	a	q
006	sp P35835 SUBN_BACNA	R	D	R	L	E	S	T	A	T	Y	L	G	N	S	F	Y	Y	G	K	G	L	i	n	v	q	a	a	a	q
007	sp P29142 SUBT_GEOSE	R	D	R	L	E	S	T	A	T	Y	L	G	N	S	F	Y	Y	G	K	G	L	i	n	v	q	a	a	a	q
008	sp P29600 SUBS_BACLE	R	N	H	L	K	N	T	A	T	S	L	G	S	T	N	L	Y	G	S	G	L	v	n	a	e	a	a	t	r
009	sp P00780 SUBT_BACLI	R	N	R	L	S	S	T	A	T	Y	L	G	S	S	F	Y	Y	G	K	G	L	i	n	v	e	a	a	a	q

hG

hH

1	2	3	4	5	6	7	8	9
Variable	Average	Conserved						

Subtilisins - Structure

Methodology:

- PFAM: domain *Peptidase_S8*
- PFAM: extract HMM + select protein target (SUBTILISIN BPN)
- jackhmmer: SUBT_BPN + UNIPROT
- Select templates with best E-value
- hmmanalign: SUBT.hmm + SUBT.mfa
- ConSurf: graphic representation

EMBL-EBI | HOME | SEARCH | BROWSE | FTP | HELP | ABOUT | **Pfam** | keyword search | Go

Family: *Peptidase_S8* (PF00082) | 784 architectures | 17666 sequences | 19 interactions | 2420 spots | 401 structures

Summary | Domain organisation | Class | Alignments | HMM logo | Trees | Curation & model | Species | Interactions | Structures | Jump to... | |

Alignments

We store a range of different sequence alignments for families. As well as the seed alignment from which the family is built, we provide the full alignment, generated by searching the sequence database ([log-odds proteomes](#)) using the family HMM. We also generate alignments using four [representative proteomes](#) (RP) sets, the UniProtKB sequence database, the NCBI sequence database, and our metagenomics sequence database. [More...](#)

View options

We make a range of alignments for each Pfam-A family. You can see a description of each [above](#). You can view these alignments in various ways but please note that some types of alignment are never generated while others may not be available for all families, most commonly because the alignments are too large to handle.

	Seed (st)	Full (7346)	Representative proteomes				UniProt (6640)	NCBI (7252)	Meta (320)
			RP15 (166)	RP35 (166)	RP55 (166)	RP75 (166)			
Jalview	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
HTML	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Pfamlogo	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Format an alignment

	Seed (st)	Full (7346)	Representative proteomes				UniProt (6640)	NCBI (7252)	Meta (320)
			RP15 (166)	RP35 (166)	RP55 (166)	RP75 (166)			
Alignment:	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Format:	Stockholm								
Order:	<input checked="" type="radio"/> Tree <input type="radio"/> Alphabetical								
Sequence:	<input checked="" type="radio"/> Inserts lower case <input type="radio"/> All upper case								
Gaps:	<input type="radio"/> Gaps as "-" or "." (mixed)								
Download/view:	<input type="radio"/> Download <input type="radio"/> View								

Generate

HMM

jackhmmer

hmmanalign

```
@@ Round: 4
@@ Included in MSA: 296 subsequences (query + 295 subseqs from 292 targets)
@@ Model size: 275 positions
@@
```

Scores for complete sequences (score includes all domains):

--- full sequence ---			--- best 1 domain ---			-#dom-		Sequence	Description
E-value	score	bias	E-value	score	bias	exp	N		
8.9e-98	331.8	18.1	1.3e-97	331.2	12.5	1.2	1	sp P00782 SUBT_BACAM	Subtilisin BPN' OS=Bacillus amylolique
9e-97	328.5	17.6	1.4e-96	327.9	12.2	1.3	1	sp P35835 SUBN_BACNA	Subtilisin NAT OS=Bacillus subtilis su
2e-96	327.4	17.8	3.1e-96	326.7	12.4	1.3	1	sp P04189 SUBT_BACSU	Subtilisin E OS=Bacillus subtilis (str
7.9e-96	325.4	17.8	1.2e-95	324.8	12.4	1.3	1	sp P00783 SUBT_BACSA	Subtilisin amylosacchariticus OS=Bacil
1.2e-95	324.8	17.8	1.2e-95	324.8	12.4	1.4	1	sp P29142 SUBT_GEOSE	Subtilisin J OS=Geobacillus stearother
1.4e-95	324.6	17.9	1.6e-95	324.4	12.4	1.0	1	sp P07518 SUBT_BACPU	Subtilisin OS=Bacillus pumilus GN=apr
1.1e-90	308.5	20.7	1.7e-90	307.9	14.4	1.3	1	sp P00780 SUBT_BACLI	Subtilisin Carlsberg OS=Bacillus liche
1.2e-89	305.2	19.4	1.3e-89	305.0	13.4	1.0	1	sp P29600 SUBS_BACLE	Subtilisin Savinase OS=Bacillus lentus
3.1e-89	303.8	18.3	3.4e-89	303.6	12.7	1.0	1	sp P29599 SUBB_BACLE	Subtilisin BL OS=Bacillus lentus PE=1

Subtilisins - Structure

Methodology:

- PDB: download available files (7/9)
- Create domains file
- STAMP: SUBTILISINS
- ALIGNFIT + STAMP (if Sc low)
- TRANSFORM: graphic representation

1SBT
ATOMIC COORDINATES FOR SUBTILISIN BPN (OR NOVO)
DOI: 10.2210/pdb1sbt/pdb
Classification: **HYDROLASE (SERINE PROTEINASE)**
Deposited: 1972-08-11 Released: 1977-01-05
Deposition author(s): Alden, R.A., Birktoft, J.J., Kraut, J., Roberts, J.D., W...
Organism: *Bacillus amyloliquefaciens*
Mutation(s): 10
Structural Biology Knowledgebase: 1SBT (-22 annotations) [SRSB.org](#)

Experimental Data Snapshot
Method: X-RAY DIFFRACTION
Resolution: 2.5 Å
R-Value Work:

wwPDB Validation
Metric Percentile Ranks Value
Clashscore 0.65
Ramachandran outliers 6.6%
Stereom outliers 10.7%

stamp

transform
(rasmol/chimera)

```
STAMP Structural Alignment of Multiple Proteins
Version 4.4 (May 2010)
by Robert B. Russell & Geoffrey J. Barton
Please cite PROTEINS, v14, 309-323, 1992

Running roughfit.

Sc = STAMP score, RMS = RMS deviation, Align = alignment length
Len1, Len2 = length of domain, Nfft = residues fitted
Secs = no. equivalent sec. structs. Eq = no. equivalent residues
SI = seq. identity, NS = sec. str. identity
P(n) = P value (p=1/10) calculated after Murzin (1993), 3MB, 230, 689-694
(MC = P value not calculated - potential FP overflow)


```

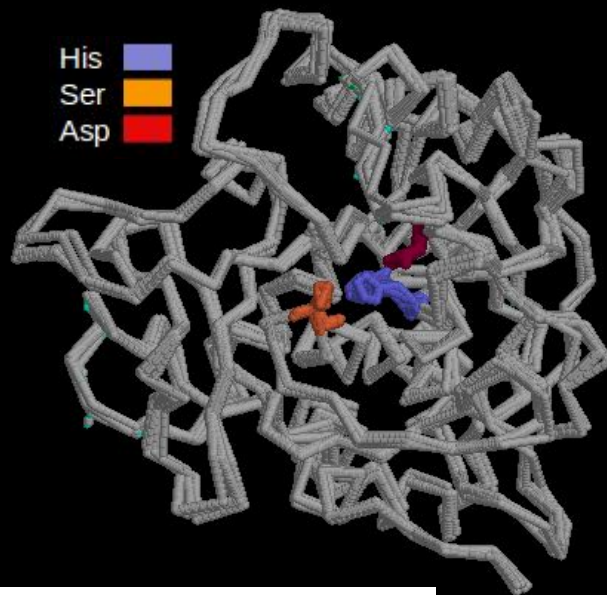
Pair	No.	Domain1	Domain2	Sc	RMS	Len1	Len2	Align	Nfft	Eq.	Secs.	SI	NS	P(n)
Pair 1	1sbt	1gcl	1sbc	8.98	0.94	275	269	275	266	266	0	60.90	100.00	1.76e-91
Pair 2	1sbt	1sbc	1sbc	9.41	0.73	275	274	275	274	274	0	69.34	100.00	0.00e+00
Pair 3	1sbt	3yvv	3yvv	9.40	0.74	275	274	275	274	274	0	65.94	100.00	0.00e+00
Pair 4	1sbt	1scj	1scj	9.45	0.73	275	275	275	275	275	0	84.09	100.00	0.00e+00
Pair 5	1sbt	1s13	1s13	8.99	0.91	275	269	275	266	266	0	60.15	100.00	4.51e-88
Pair 6	1sbt	imee	imee	9.47	0.70	275	275	275	275	275	0	64.36	100.00	0.00e+00
Pair 7	1gcl	1sbc	1sbc	9.13	0.88	269	274	274	266	266	0	62.78	100.00	2.70e-97
Pair 8	1gcl	3yvv	3yvv	9.17	0.70	269	274	275	266	266	0	61.65	100.00	8.73e-94
Pair 9	1gcl	1scj	1scj	9.21	0.71	269	275	275	266	266	0	61.28	100.00	1.76e-91
Pair 10	1gcl	1s13	1s13	9.75	0.25	269	269	269	269	269	0	97.77	100.00	0.00e+00
Pair 11	1gcl	imee	imee	9.19	0.79	269	275	275	268	268	0	61.94	100.00	2.28e-92
Pair 12	1sbc	3yvv	3yvv	9.49	0.57	274	274	275	273	273	0	70.33	100.00	0.00e+00
Pair 13	1sbc	1scj	1scj	9.55	0.55	274	275	275	274	274	0	69.34	100.00	0.00e+00
Pair 14	1sbc	1s13	1s13	9.13	0.87	274	269	274	266	266	0	61.28	100.00	1.76e-91
Pair 15	1sbc	imee	imee	9.54	0.58	274	275	275	274	274	0	70.44	100.00	0.00e+00
Pair 16	3yvv	1scj	1scj	9.70	0.25	274	275	275	274	274	0	99.27	100.00	0.00e+00
Pair 17	3yvv	1s13	1s13	9.17	0.67	274	269	275	265	265	0	61.13	100.00	7.66e-92
Pair 18	3yvv	imee	imee	9.66	0.33	274	275	275	274	274	0	98.91	100.00	0.00e+00
Pair 19	1scj	1s13	1s13	9.23	0.69	275	269	275	266	266	0	60.13	100.00	2.45e-90
Pair 20	1scj	imee	imee	9.72	0.35	275	275	275	275	275	0	98.55	100.00	0.00e+00
Pair 21	1s13	imee	imee	9.20	0.77	269	275	275	268	268	0	61.19	100.00	4.72e-93

```
Reading in matrix file SUBT.mat...
Doing cluster analysis...
Cluster: 1 ( 1gcl & 1s13 ) Sc 9.75 RMS 0.25 Len 269 nfft 269
See file SUBT.1 for the alignment and transformations
Cluster: 2 ( 1scj & imee ) Sc 9.71 RMS 0.35 Len 275 nfft 275
See file SUBT.2 for the alignment and transformations
```


SUPERIMPOSITION OF BACTERIAL SUBTILISINS WITH STAMP

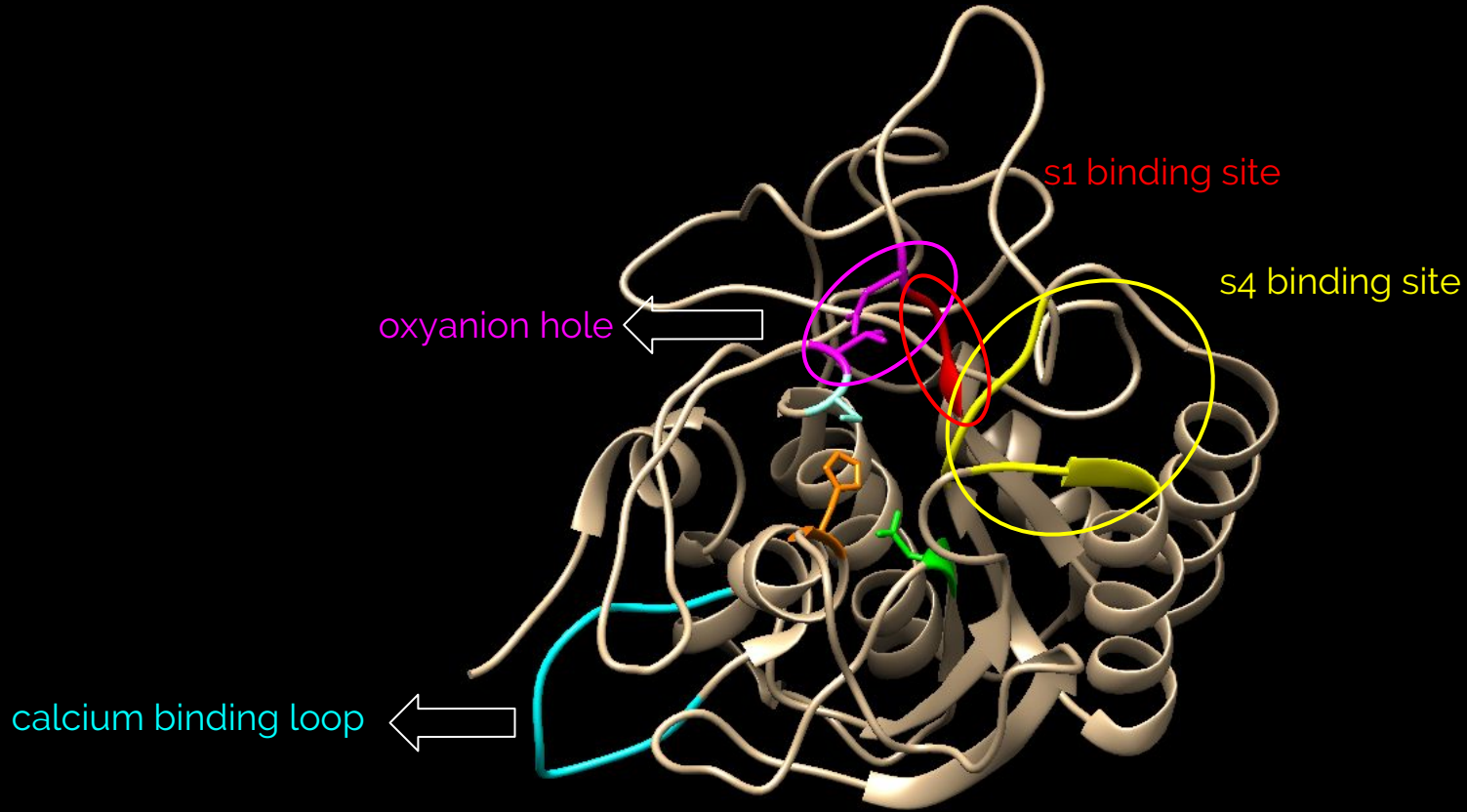


STAMP
Score: 9.36
RMS: 0.69



```
Cluster: 1 ( 1gci & 1st3 ) Sc 9.75 RMS 0.25 Len 269 nfit 269
See file SUBT.1 for the alignment and transformations
Cluster: 2 ( 1scj & 1mee ) Sc 9.71 RMS 0.35 Len 275 nfit 275
See file SUBT.2 for the alignment and transformations
Cluster: 3 ( 3vyv & 1scj 1mee ) Sc 9.58 RMS 0.24 Len 275 nfit 274
See file SUBT.3 for the alignment and transformations
Cluster: 4 ( 1sbc & 3vyv 1scj 1mee ) Sc 9.53 RMS 0.54 Len 275 nfit 273
See file SUBT.4 for the alignment and transformations
Cluster: 5 ( 1sbt & 1sbc 3vyv 1scj 1mee ) Sc 9.52 RMS 0.67 Len 275 nfit 273
See file SUBT.5 for the alignment and transformations
Cluster: 6 ( 1gci 1st3 & 1sbt 1sbc 3vyv 1scj 1mee ) Sc 9.36 RMS 0.69 Len 275 nfit 264
See file SUBT.6 for the alignment and transformations
```

Subtilisins - Essential structural features



Subtilisin BPN (1SBT)

Subtilisins - Active site

SER 221 → forms a covalent bond with the substrate

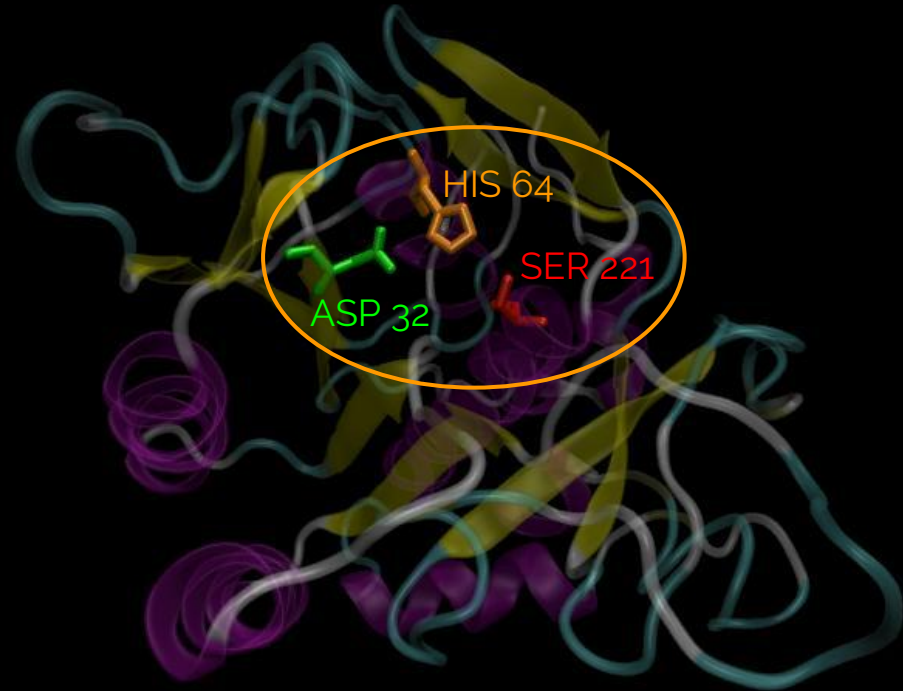
HIS 64 → has a dual role:

- facilitate formation of the covalent bond
- stabilizes the negatively charged transition state

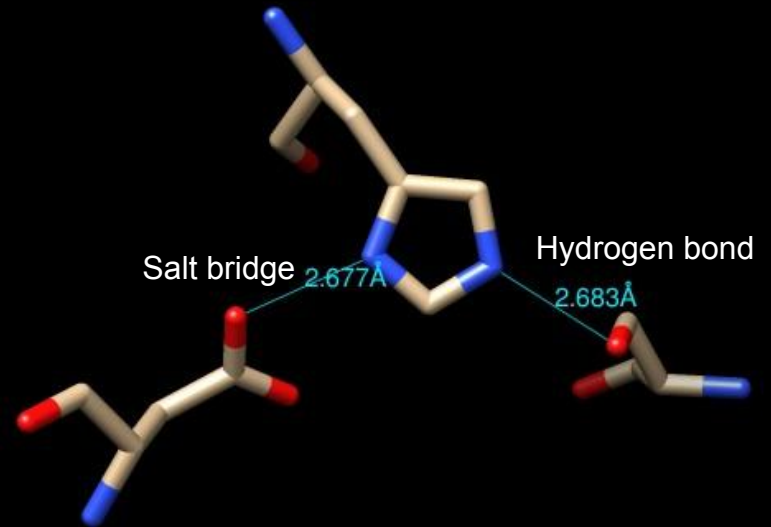
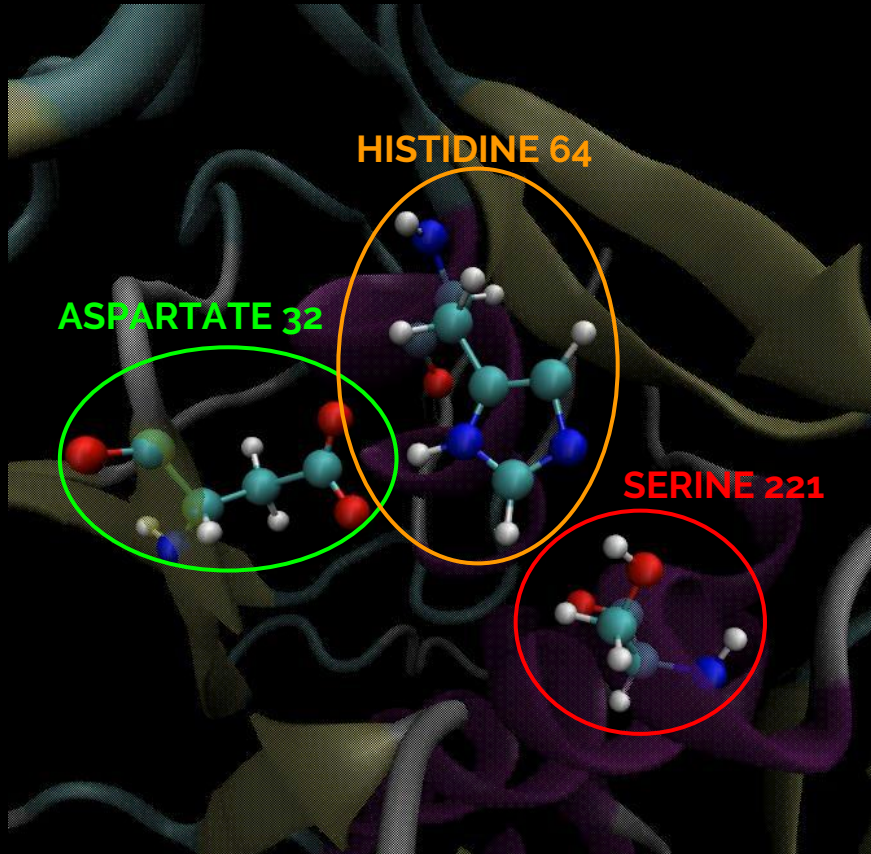
ASP 32 → stabilizes the positive charge of His.



Located far in the sequence
converge in the native structure



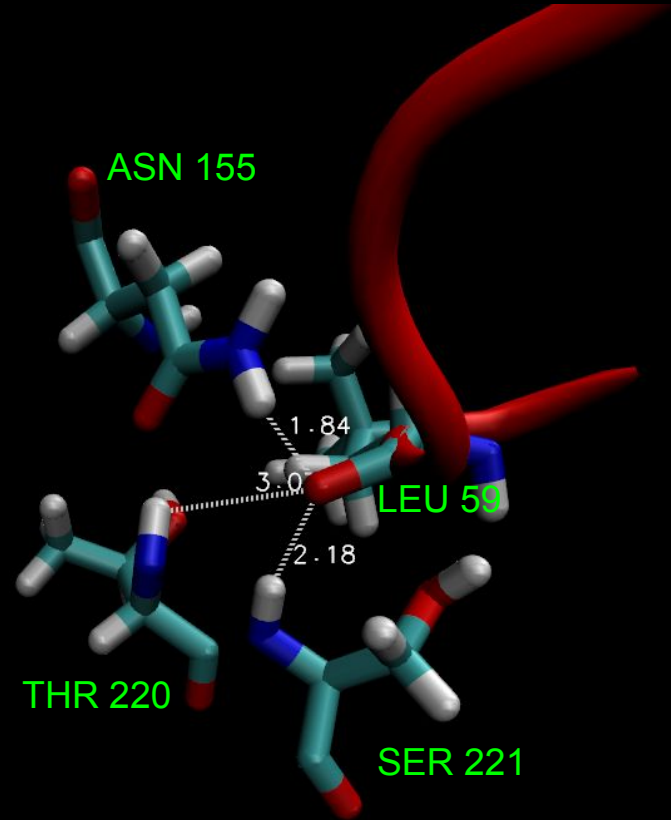
Subtilisins - Catalytic triad



Catalytic triad bonds Subtilisin BPN	Distances (Å)
His (N2) - Ser (OH)	2,683
His (N1) - Asp (O1)	2,677

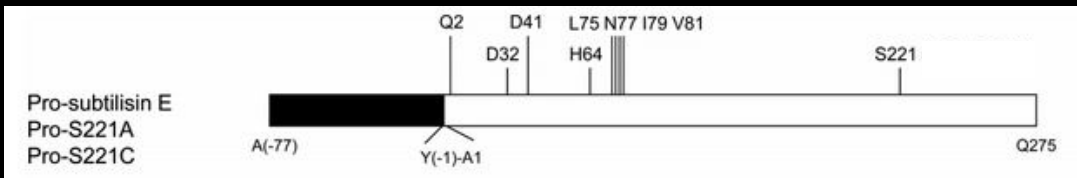
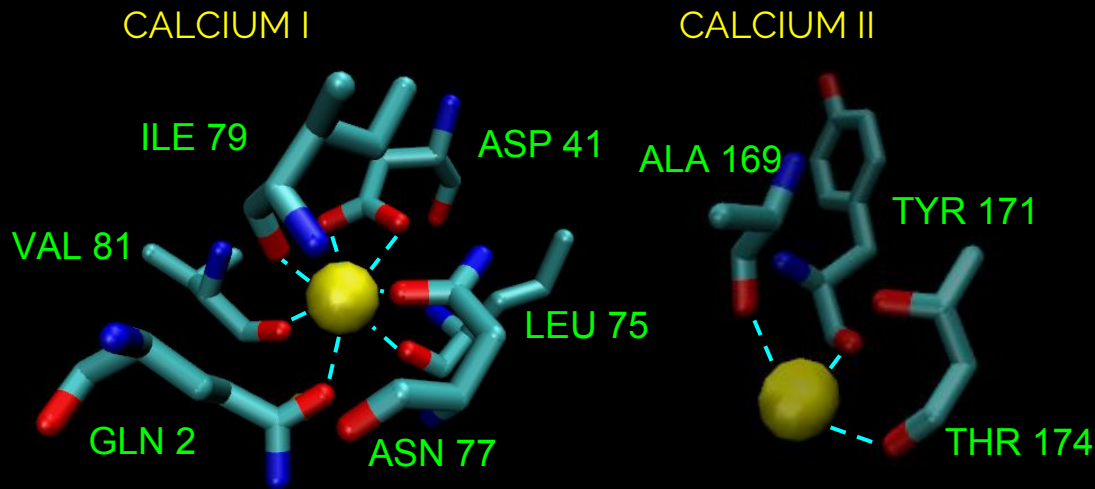
Subtilisins - Oxyanion hole

Oxyanion hole is formed by the side chains of **Asn155** and **Thr220** and the backbone NH of **Ser221** → stabilize the oxyanion of the tetrahedral intermediate



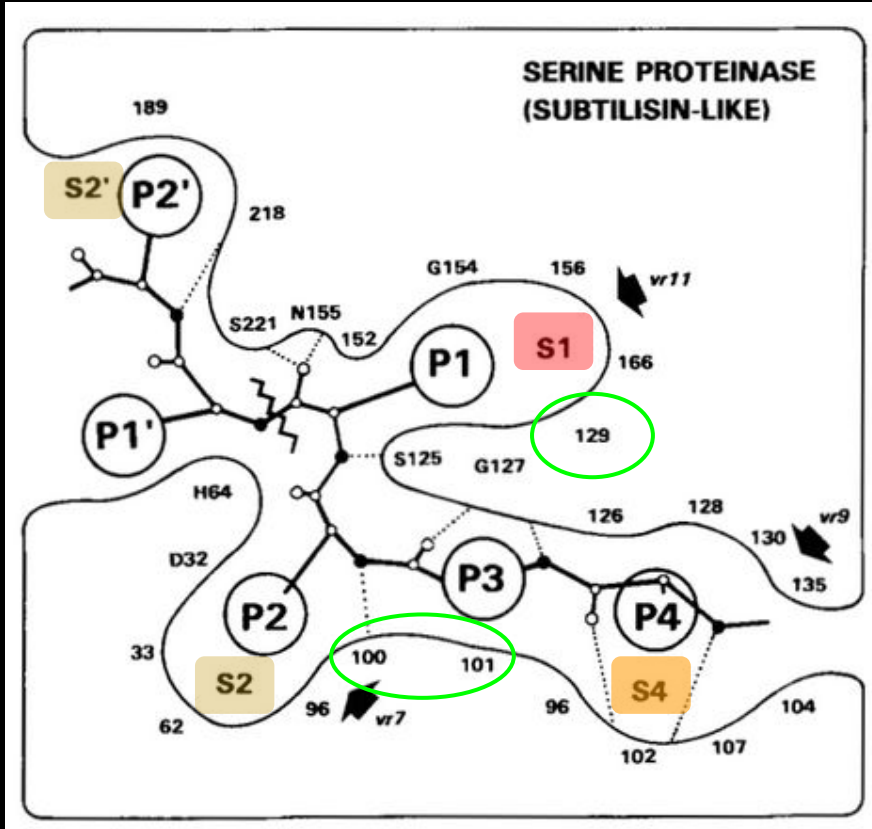
Subtilisin Novo (2SEC) bound to the inhibitor eglin

Subtilisins - Calcium binding domain

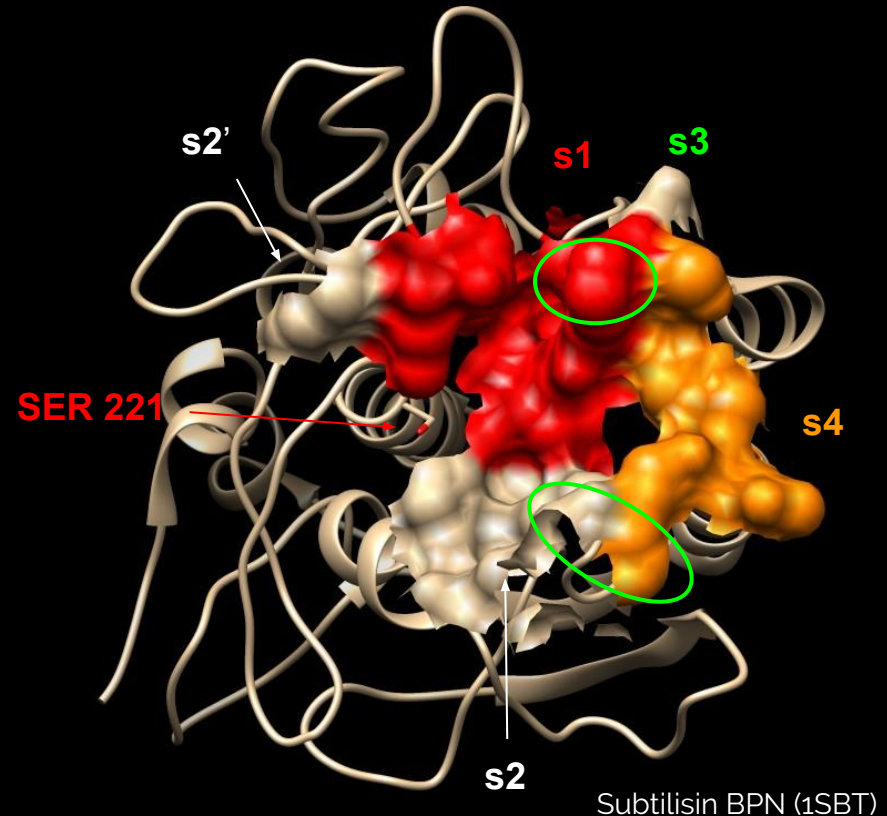


Uehara R, Angkawidjaja C, Koga Y, Kanaya S. Formation of the high-affinity calcium binding site in pro-subtilisin E with the insertion sequence IS1 of pro-Tk-subtilisin. *Biochemistry*. 2013 Dec 17;52(50):9080-8.

Subtilisins - S1-S4 Specificity pockets

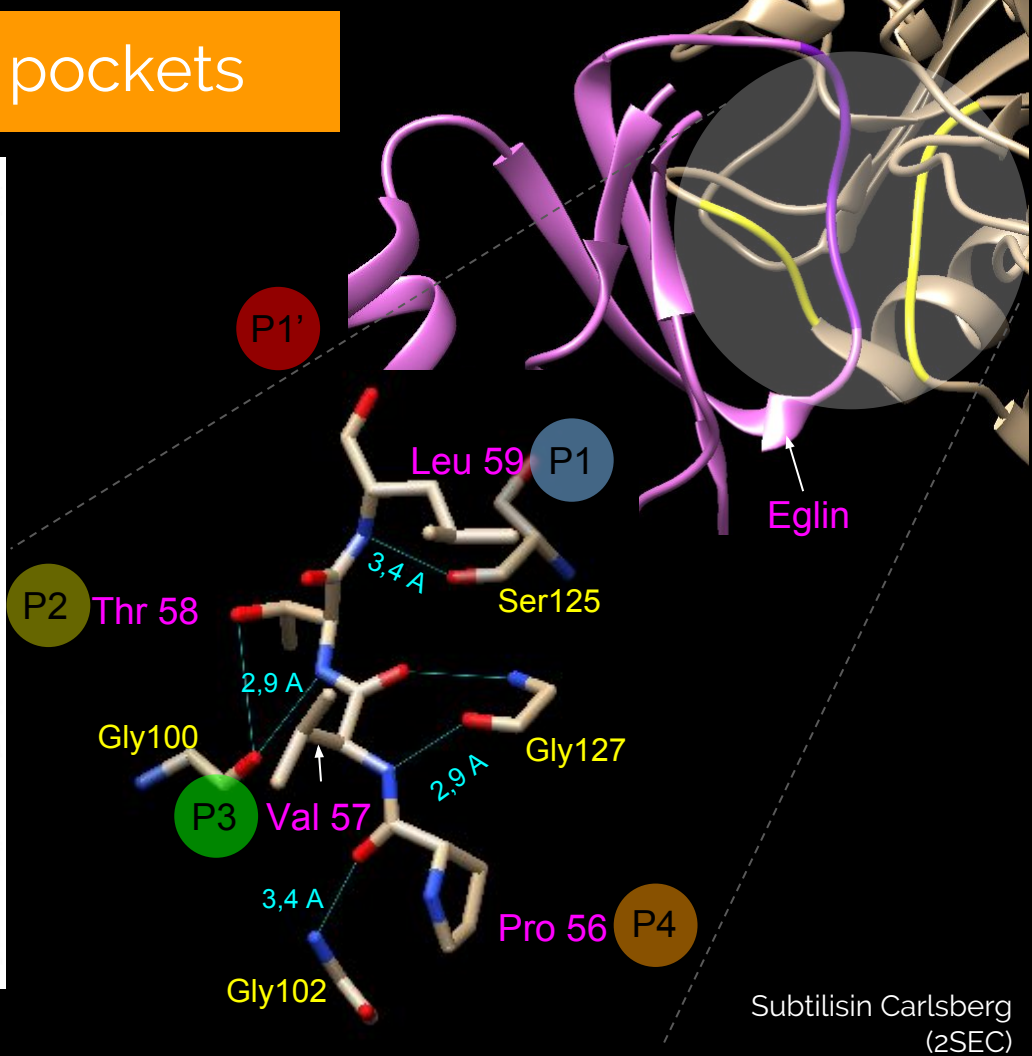
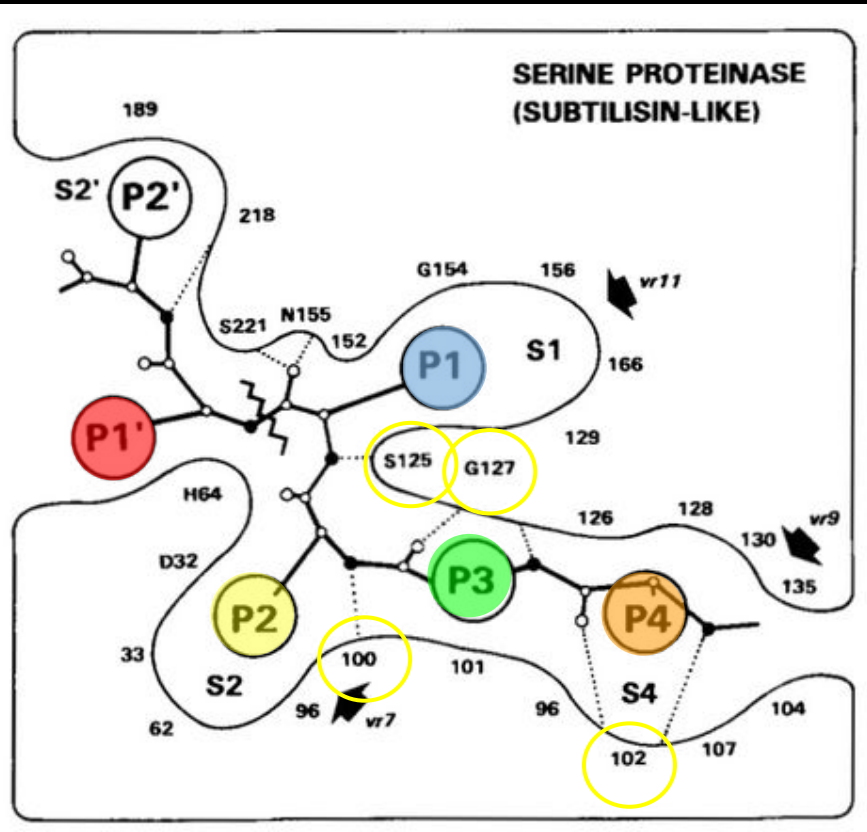


Siezen R, Leunissen J. Subtilases: The superfamily of subtilisin-like serine proteases. *Protein Science*. 2008;6(3):501-523.



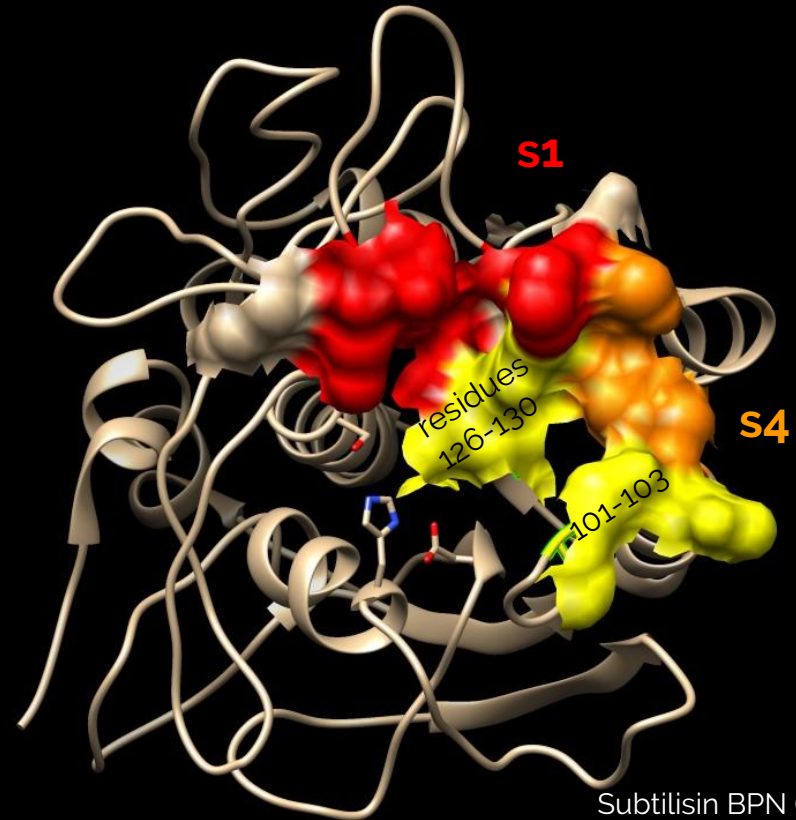
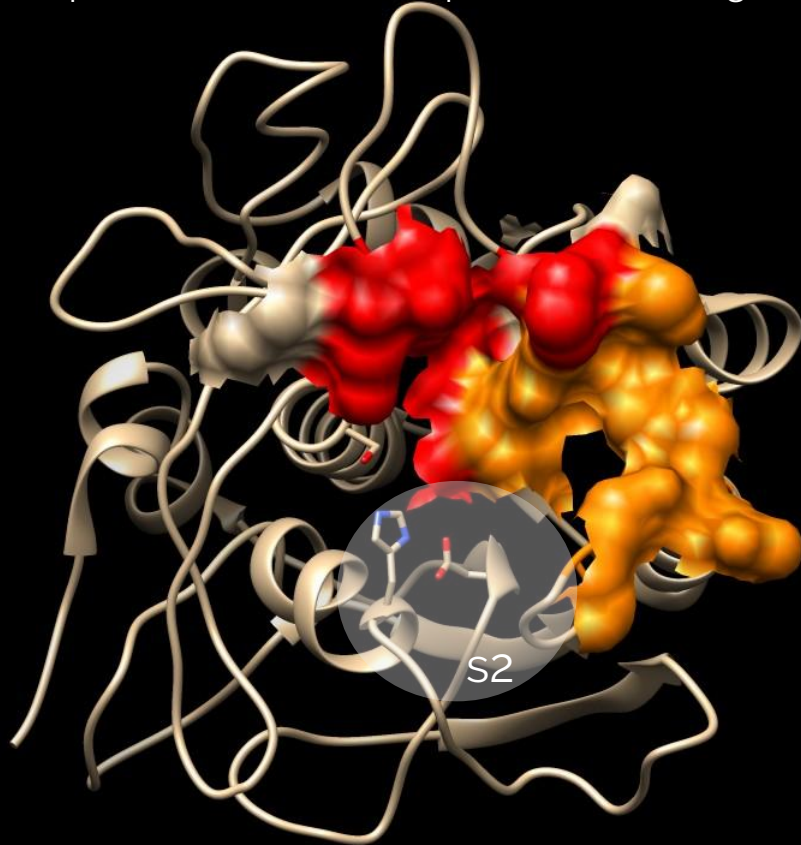
Subtilisin BPN (1SBT)

Subtilisins - S1-S4 Specificity pockets

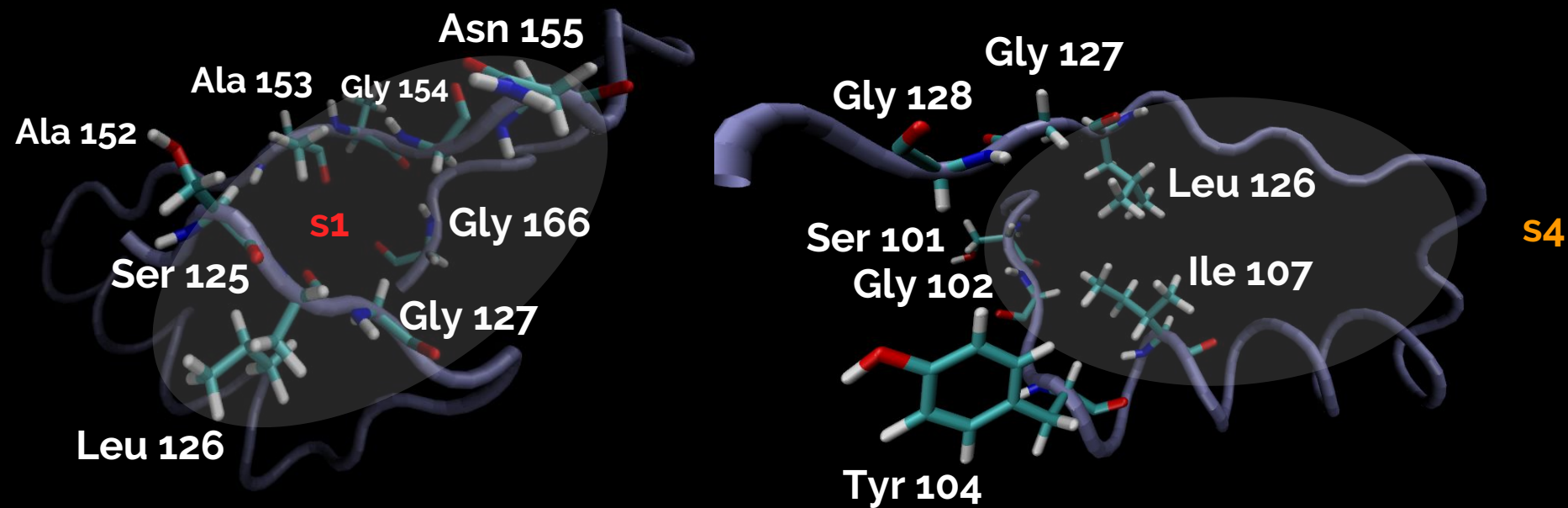


Subtilisins - S1-S4 Specificity pockets

Asp-32 and His-64 form part of S2 binding site



Subtilisins - S1-S4 Specificity pockets



Bacterial subtilisins sequence alignment, HMM

001	sp P29599 SUBB_BACLE	a	q	s	v	p	w	g	i	s	r	v	q	a	p	a	a	h	n	r	g	l	t	G	S	G	V	K	V	A	V	L	D	T	G	I	-	S	T	H	P	D	L	N	-	I	R	G	G	A	S
002	sp P00782 SUBT_BACAM	a	q	s	v	p	y	g	v	s	q	i	k	a	p	a	l	h	s	q	g	y	t	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H	P	D	L	K	V	A	G	G	A	S	M
003	sp P04189 SUBT_BACSU	a	q	s	v	p	y	g	i	s	q	i	k	a	p	a	l	h	s	q	g	y	t	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H	P	D	L	N	V	R	G	G	A	S	F
004	sp P00783 SUBT_BACSA	a	q	s	v	p	y	g	i	s	q	i	k	a	p	a	l	h	s	q	g	y	t	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H	P	D	L	N	V	R	G	G	A	S	F
005	sp P07518 SUBT_BACPU	a	q	s	v	p	y	g	i	s	q	i	k	a	p	a	l	h	s	q	g	y	t	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H	P	D	L	N	V	R	G	G	A	S	F
006	sp P35835 SUBN_BACNA	a	q	s	v	p	y	g	i	s	q	i	k	a	p	a	l	h	s	q	g	y	t	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H	P	D	L	N	V	R	G	G	A	S	F
007	sp P29142 SUBT_GEOSE	a	q	s	v	p	y	g	i	s	q	i	k	a	p	a	l	h	s	q	g	y	t	G	S	N	V	K	V	A	V	I	D	S	G	I	D	S	S	H	P	D	L	N	V	R	G	G	A	S	F
008	sp P29600 SUBS_BACLE	a	q	s	v	p	w	g	i	s	r	v	q	a	p	a	a	h	n	r	g	l	t	G	S	G	V	K	V	A	V	L	D	T	G	I	-	S	T	H	P	D	L	N	-	I	R	G	G	A	S
009	sp P00780 SUBT_BACLI	a	q	t	v	p	y	g	i	p	l	i	k	a	d	k	v	q	a	q	g	f	k	G	A	N	V	K	V	A	V	L	D	T	G	I	Q	A	S	H	P	D	L	N	-	V	V	G	G	A	S

Ca²⁺

Asp 32 Gly 34

Gly 65 Gly 70

001	sp P29599 SUBB_BACLE	F	V	P	G	E	P	S	T	Q	D	G	N	G	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	E	L	Y	A	V	K	V	L	G	A	D	G
002	sp P00782 SUBT_BACAM	V	P	S	E	T	N	P	F	Q	D	N	N	S	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	S	L	Y	A	V	K	V	L	G	A	D	G
003	sp P04189 SUBT_BACSU	V	P	S	E	T	N	P	Y	Q	D	G	S	S	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	S	L	Y	A	V	K	V	L	D	S	T	G
004	sp P00783 SUBT_BACSA	V	P	S	E	T	N	P	Y	Q	D	G	S	S	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	S	L	Y	A	V	K	V	L	D	S	T	G
005	sp P07518 SUBT_BACPU	V	P	S	E	T	N	P	Y	Q	D	G	S	S	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	S	L	Y	A	V	K	V	L	D	S	T	G
006	sp P35835 SUBN_BACNA	V	P	S	E	T	N	P	Y	Q	D	G	S	S	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	S	L	Y	A	V	K	V	L	D	S	T	G
007	sp P29142 SUBT_GEOSE	V	P	S	E	T	N	P	Y	Q	D	G	S	S	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	S	L	Y	A	V	K	V	L	D	S	T	G
008	sp P29600 SUBS_BACLE	F	V	P	G	E	P	S	T	Q	D	G	N	G	H	G	T	H	V	A	G	T	I	A	A	L	N	N	S	I	G	V	L	G	V	A	P	S	A	E	L	Y	A	V	K	V	L	G	A	S	G
009	sp P00780 SUBT_BACLI	F	V	A	G	E	A	Y	N	T	D	G	N	G	H	G	T	H	V	A	G	T	V	A	A	L	D	N	T	T	G	V	L	G	V	A	P	S	V	S	L	Y	A	V	K	V	L	N	S	S	G

His 64

Val 68

Calcium binding site

Leu 90

001	sp P29599 SUBB_BACLE	R	G	A	-	I	S	S	I	A	Q	G	L	E	W	A	G	N	N	G	M	H	V	A	N	L	S	L	G	-	S	P	S	P	S	A	T	L	E	Q	A	V	N	S	A	T	S	R	G	V	L
002	sp P00782 SUBT_BACAM	S	G	Q	Y	-	S	W	I	I	N	G	I	E	W	A	I	A	N	N	M	D	V	I	N	M	S	L	G	G	P	S	G	S	A	A	L	K	A	-	A	V	D	K	A	V	A	S	G	V	V
003	sp P04189 SUBT_BACSU	S	G	Q	Y	-	S	W	I	I	N	G	I	E	W	A	I	S	N	N	M	D	V	I	N	M	S	L	G	G	P	T	G	S	T	A	L	K	T	-	V	V	D	K	A	V	S	S	G	I	V
004	sp P00783 SUBT_BACSA	S	G	Q	Y	-	S	W	I	I	N	G	I	E	W	A	I	S	N	N	M	D	V	I	N	M	S	L	G	G	P	T	G	S	T	A	L	K	T	-	V	V	D	K	A	V	S	S	G	I	V
005	sp P07518 SUBT_BACPU	S	G	Q	Y	-	S	W	I	I	N	G	I	E	W	A	I	S	N	N	M	D	V	I	N	M	S	L	G	G	P	T	G	S	T	A	L	K	T	-	V	V	D	K	A	V	S	S	G	I	V
006	sp P35835 SUBN_BACNA	S	G	Q	Y	-	S	W	I	I	N	G	I	E	W	A	I	S	N	N	M	D	V	I	N	M	S	L	G	G	P	T	G	S	T	A	L	K	T	-	V	V	D	K	A	V	S	S	G	I	V
007	sp P29142 SUBT_GEOSE	S	G	Q	Y	-	S	W	I	I	N	G	I	E	W	A	I	S	N	N	M	D	V	I	N	M	S	L	G	G	P	S	G	S	T	A	L	K	T	-	V	V	D	K	A	V	S	S	G	I	V
008	sp P29600 SUBS_BACLE	S	G	S	V	-	S	S	I	A	Q	G	L	E	W	A	G	N	N	G	M	H	V	A	N	L	S	L	G	-	S	P	S	P	S	A	T	L	E	Q	A	V	N	S	A	T	S	R	G	V	L
009	sp P00780 SUBT_BACLI	S	G	T	Y	-	S	G	I	V	S	G	I	E	W	A	T	T	N	G	M	D	V	I	N	M	S	L	G	G	P	S	G	S	T	A	M	K	Q	-	A	V	D	N	A	Y	A	R	G	V	V

1 2 3 4 5 6 7 8 9
Variable Average Conserved

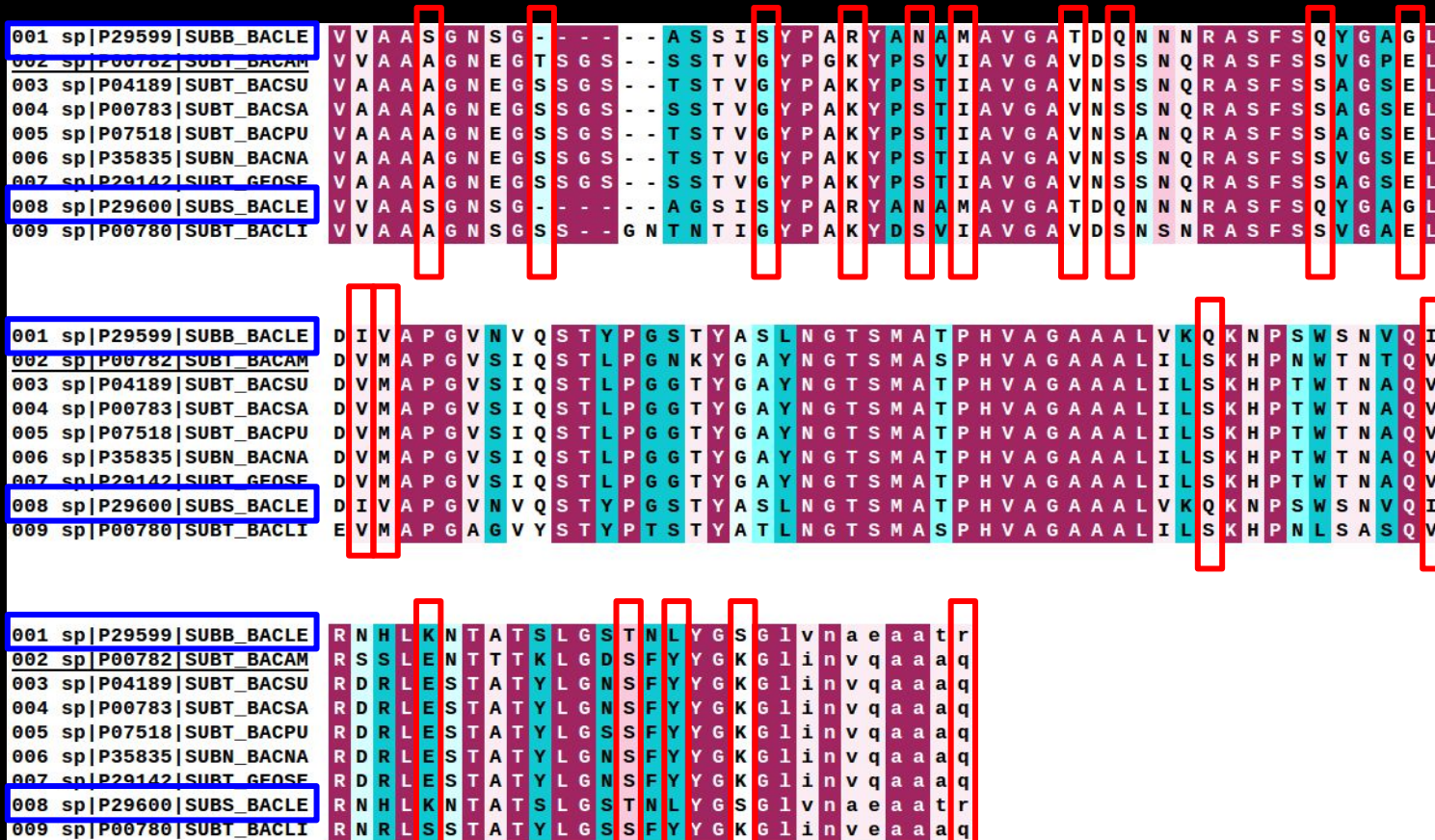
S4

S1/S4

Bacterial subtilisins sequence alignment, HMM



Bacterial subtilisins sequence alignment, HMM

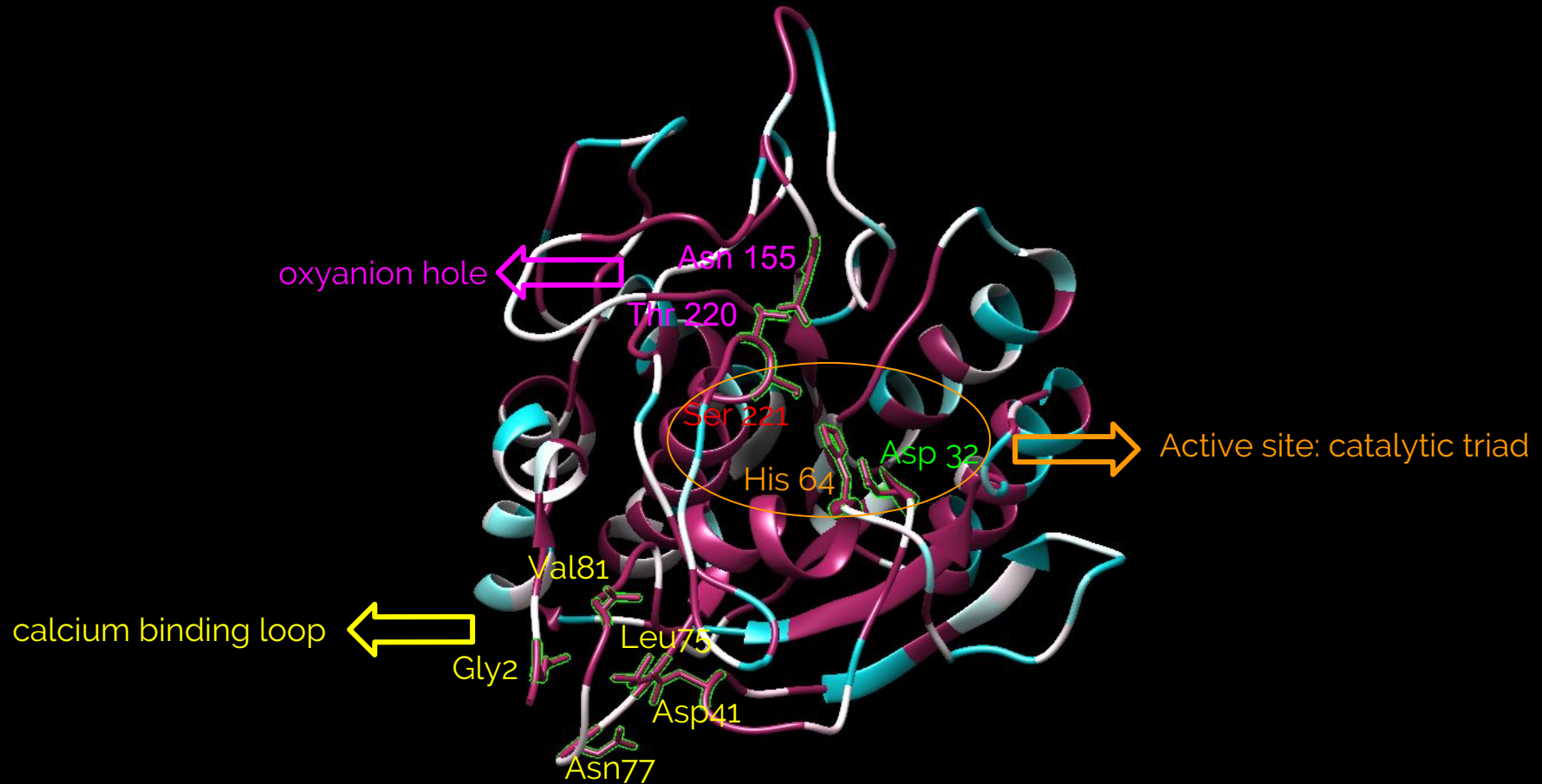


1 2 3 4 5 6 7 8 9
 Variable Average Conserved

Bacterial subtilisins sequence alignment, HMM

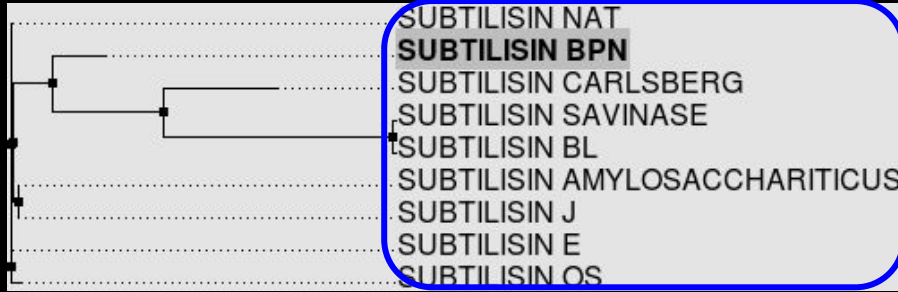
001	sp P29599 SUBB_BACLE	V	V	A	A	S	G	N	S	G	-	-	-	-	A	S	S	I	S	Y	P	A	R	Y	A	N	A	M	A	V	G	A	T	D	Q	N	N	R	A	S	F	S	Q	Y	G	A	G	L					
002	sp P00782 SUBT_BACAM	V	V	A	A	A	G	N	E	G	T	S	G	S	-	S	S	T	V	G	Y	P	G	K	Y	P	S	V	I	A	V	G	A	V	D	S	S	N	Q	R	A	S	F	S	S	V	G	P	E	L			
003	sp P04189 SUBT_BACSU	V	A	A	A	A	G	N	E	G	S	S	G	S	-	T	S	T	V	G	Y	P	A	K	Y	P	S	T	I	A	V	G	A	V	N	S	S	N	Q	R	A	S	F	S	S	A	G	S	E	L			
004	sp P00783 SUBT_BACSA	V	A	A	A	A	G	N	E	G	S	S	G	S	-	S	S	T	V	G	Y	P	A	K	Y	P	S	T	I	A	V	G	A	V	N	S	S	N	Q	R	A	S	F	S	S	A	G	S	E	L			
005	sp P07518 SUBT_BACPU	V	A	A	A	A	G	N	E	G	S	S	G	S	-	T	S	T	V	G	Y	P	A	K	Y	P	S	T	I	A	V	G	A	V	N	S	A	N	Q	R	A	S	F	S	S	A	G	S	E	L			
006	sp P35835 SUBN_BACNA	V	A	A	A	A	A	G	N	E	G	S	S	G	S	-	T	S	T	V	G	Y	P	A	K	Y	P	S	T	I	A	V	G	A	V	N	S	S	N	Q	R	A	S	F	S	S	V	G	S	E	L		
007		No.	Domain1	Domain2	Sc	RMS	Len1	Len2	Align	NFit	Eq.	Secs.	%I	%S	P(m)																																						
00E	Pair	1	1sbt	1gci	8.98	0.94	275	269	275	266	266	0	60.90	100.00	1.76e-91																																						
00E	Pair	2	1sbt	1sbc	9.41	0.73	275	274	275	274	274	0	69.34	100.00	0.00e+00																																						
	Pair	3	1sbt	3vyv	9.40	0.74	275	274	275	274	274	0	85.04	100.00	0.00e+00																																						
	Pair	4	1sbt	1scj	9.45	0.73	275	275	275	275	275	0	84.00	100.00	0.00e+00																																						
	Pair	5	1sbt	1st3	8.99	0.91	275	269	275	266	266	0	60.15	100.00	4.51e-88																																						
001	Pair	6	1sbt	1mee	9.47	0.70	275	275	275	275	275	0	84.36	100.00	0.00e+00																																						
002	Pair	7	1gci	1sbc	9.13	0.88	269	274	274	266	266	0	62.78	100.00	2.70e-97																																						
003	Pair	8	1gci	3vyv	9.17	0.70	269	274	275	266	266	0	61.65	100.00	8.73e-94																																						
004	Pair	9	1gci	1scj	9.21	0.71	269	275	275	266	266	0	61.28	100.00	1.76e-91																																						
00E	Pair	10	1gci	1st3	9.75	0.25	269	269	269	269	269	0	97.77	100.00	0.00e+00																																						
00E	Pair	11	1gci	1mee	9.19	0.79	269	275	275	268	268	0	61.94	100.00	2.28e-95																																						
007	Pair	12	1sbc	3vyv	9.49	0.57	274	274	275	273	273	0	70.33	100.00	0.00e+00																																						
00E	Pair	13	1sbc	1scj	9.55	0.55	274	275	275	274	274	0	69.34	100.00	0.00e+00																																						
00E	Pair	14	1sbc	1st3	9.13	0.87	274	269	274	266	266	0	61.28	100.00	1.76e-91																																						
	Pair	15	1sbc	1mee	9.54	0.58	274	275	275	274	274	0	70.44	100.00	0.00e+00																																						
	Pair	16	3vyv	1scj	9.70	0.25	274	275	275	274	274	0	99.27	100.00	0.00e+00																																						
	Pair	17	3vyv	1st3	9.17	0.67	274	269	275	265	265	0	61.13	100.00	7.66e-92																																						
001	Pair	18	3vyv	1mee	9.66	0.33	274	275	275	274	274	0	98.91	100.00	0.00e+00																																						
002	Pair	19	1scj	1st3	9.23	0.69	275	269	275	266	266	0	60.53	100.00	2.45e-90																																						
003	Pair	20	1scj	1mee	9.72	0.35	275	275	275	275	275	0	98.55	100.00	0.00e+00																																						
004	Pair	21	1st3	1mee	9.20	0.77	269	275	275	268	268	0	61.19	100.00	4.72e-93																																						
005	sp P07518 SUBT_BACPU	R	D	R	L	E	S	T	A	T	Y	L	G	S	S	F	Y	Y	G	K	G	l	i	n	v	q	a	a	a	a	q																						
006	sp P35835 SUBN_BACNA	R	D	R	L	E	S	T	A	T	Y	L	G	N	S	S	F	Y	Y	G	K	G	l	i	n	v	q	a	a	a	q																						
007	sp P29142 SUBT_GEOSE	R	D	R	L	E	S	T	A	T	Y	L	G	N	S	S	F	Y	Y	G	K	G	l	i	n	v	q	a	a	a	q																						
008	sp P29600 SUBS_BACLE	R	N	H	L	K	N	T	A	T	S	L	G	S	T	N	L	Y	G	S	G	l	v	n	a	e	a	a	t	r																							
009	sp P00780 SUBT_BACLI	R	N	R	L	S	S	T	A	T	Y	L	G	S	S	F	Y	Y	G	K	G	l	i	n	v	e	a	a	a	q																							

Subtilisins sequence alignment

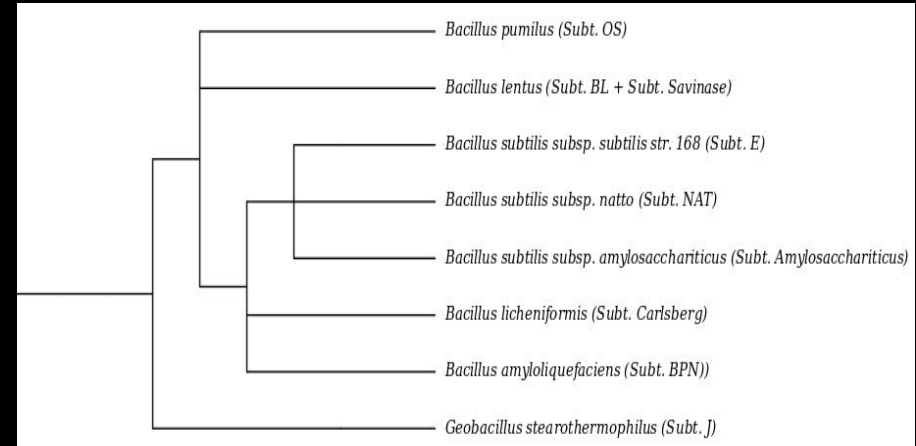


Subtilisins sequence alignment: PHYLOGENY

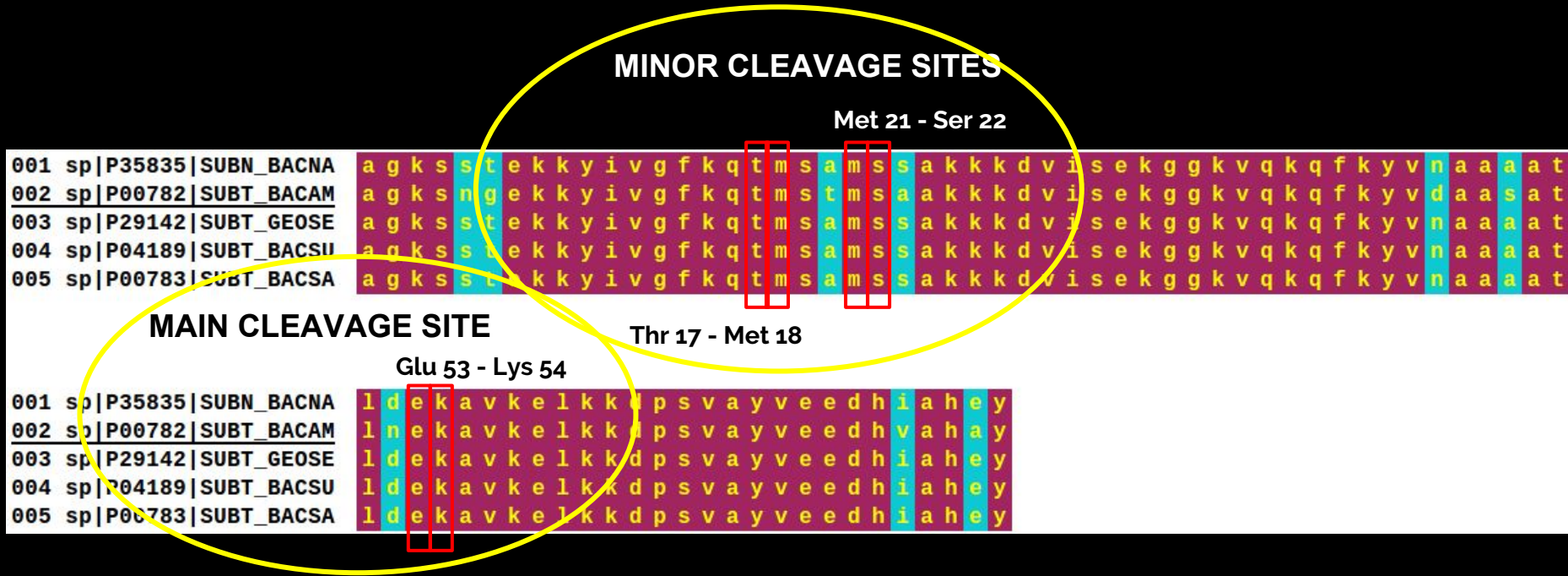
POOR CORRELATION BETWEEN PROTEIN SEQUENCES AND SPECIES PHYLOGENIA



↓
ORTHOLOGS



Subtilisins sequence alignment: PROPEPTIDE CHAIN



Altered flexibility in the substrate-binding site of related native and engineered high-alkaline *Bacillus subtilis*ins.

Mulder FA et al. *J Mol Biol.* 1999.

Requirement of left-handed glycine residue for high stability of the Tk-subtilisin propeptide as revealed by mutational and crystallographic analyse

Pulido MA et al. *J Mol Biol.* 2007.

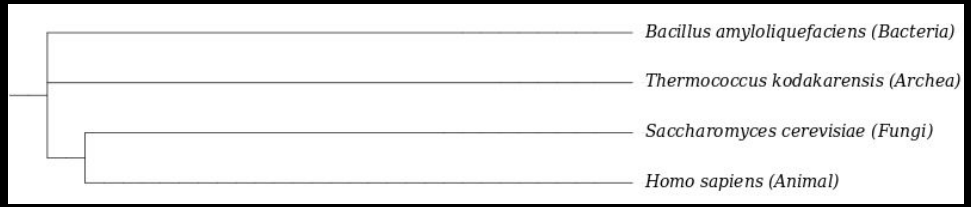
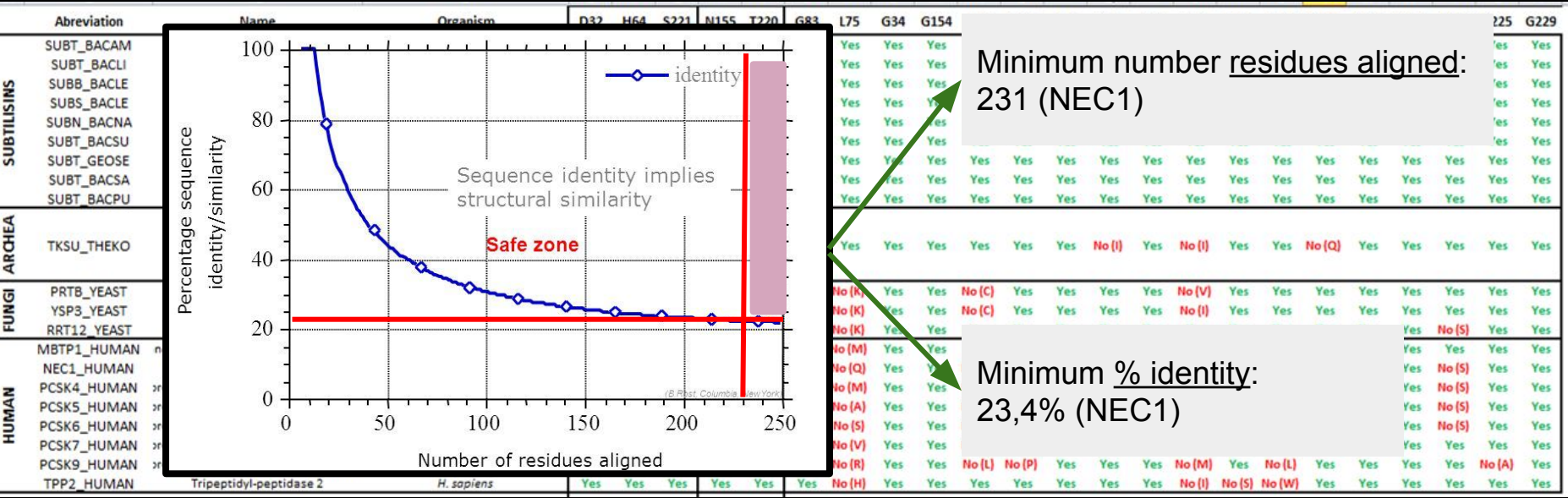
Subtilisin homologues

	Abbreviation	Name	Organism	D32	H64	S221	N155	T220	G83	L75	G34	G154	V68	S125	G65	A69	G70	L90	A152	F189	G193	P201	G219	A223	P225	G229		
SUBTILISINS	SUBT_BACAM	Subtilisin BPN'	<i>B. amyloliquefaciens</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
	SUBT_BACLI	Subtilisin Carlsberg	<i>B. licheniformis</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
	SUBB_BACLE	Subtilisin BL	<i>B. lentus</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
	SUBS_BACLE	Subtilisin Savinase	<i>B. lentus</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
	SUBN_BACNA	Subtilisin NAT	<i>B. subtilis</i> subsp. <i>Natto</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
	SUBT_BACSU	Subtilisin E	<i>B. subtilis</i> (strain 168)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	SUBT_GEOSE	Subtilisin J	<i>Geob. stearothermophilus</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
	SUBT_BACSA	Subtilisin Amylosacchariticus	<i>B. subtilis</i> subsp. <i>Amylosacchariticus</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
SUBT_BACPU	Subtilisin OS	<i>B. pumilus</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
ARCHEA	TKSU_THEKO	Tk-subtilisin	<i>Thermococcus kodakarensis</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No (I)	Yes	No (I)	Yes	Yes	No (Q)	Yes	Yes	Yes	Yes	Yes		
FUNGI	PRTB_YEAST	Cerevisin	<i>S. cerevisiae</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (K)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (V)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
	YSP3_YEAST	Subtilisin-like protease 3	<i>S. cerevisiae</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (K)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (I)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
	RRT12_YEAST	Subtilase-type proteinase RRT12	<i>S. cerevisiae</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (K)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No (S)	Yes	Yes		
HUMAN	MBTP1_HUMAN	ne-bound transcription factor site-1	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (M)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No (Y)	Yes	Yes	Yes		
	NEC1_HUMAN	Neuroendocrine convertase 1	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (Q)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (V)	Yes	No (Y)	No (C)	No (S)	Yes	No (S)	Yes	Yes		
	PCSK4_HUMAN	rotein convertase subtilisin/kexin ty	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (M)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (I)	Yes	No (Y)	No (C)	No (S)	Yes	No (S)	Yes	Yes		
	PCSK5_HUMAN	rotein convertase subtilisin/kexin ty	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (A)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (I)	Yes	No (Y)	No (C)	No (S)	Yes	No (S)	Yes	Yes		
	PCSK6_HUMAN	rotein convertase subtilisin/kexin ty	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (S)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (I)	Yes	No (Y)	No (C)	No (S)	Yes	No (S)	Yes	Yes		
	PCSK7_HUMAN	rotein convertase subtilisin/kexin ty	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (V)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (I)	Yes	No (Y)	No (C)	No (S)	Yes	Yes	Yes	Yes		
	PCSK9_HUMAN	rotein convertase subtilisin/kexin ty	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (R)	Yes	Yes	No (L)	No (P)	Yes	Yes	Yes	No (M)	Yes	No (L)	Yes	Yes	Yes	Yes	Yes	No (A)	Yes	
	TPP2_HUMAN	Tripeptidyl-peptidase 2	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (H)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No (I)	No (S)	No (W)	Yes	Yes	Yes	Yes	Yes	Yes	



Subtilisin homologues

Modified from: <http://www.cmbi.ru.nl/edu/bioinf4/hommod/>



Subtilisin homologues vs non-homologues

Alpha and beta proteins (a/b)	Beta Sheet $\alpha/\beta/\alpha$, Left-Handed Crossover	Subtilisin-like (SB)	Subtilases (subtilisin) Serine-carboxyl proteinase (sedolisin)
-------------------------------	--	-----------------------------	---

	Abbreviation	Name	Organism	D32	H64	S221	N155	T220	G83	L75	G34	G154	V68	S125	G65	A69	G70	L90	A152	F189	G193	P201	G219	A223	P225	G229			
SUBTILISINS	SUBT_BACAM	Subtilisin BPN'	<i>B. amyloliquefaciens</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
	SUBT_BACLI	Subtilisin Carlsberg	<i>B. licheniformis</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
	SUBB_BACLE	Subtilisin BL	<i>B. lentus</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
	SUBS_BACLE	Subtilisin Savinase	<i>B. lentus</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
	SUBN_BACNA	Subtilisin NAT	<i>B. subtilis subsp. Natto</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
	SUBT_BACSU	Subtilisin E	<i>B. subtilis (strain 168)</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	SUBT_GEOSE	Subtilisin J	<i>Geob. stearothermophilus</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
	SUBT_BACSA	Subtilisin Amylosacchariticus	<i>B. subtilis subsp. Amylosacchariticus</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
SUBT_BACPU	Subtilisin OS	<i>B. pumilus</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
ARCHAEA	TKSU_THEKO	Tk-subtilisin	<i>Thermococcus kodakarensis</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No (I)	Yes	No (I)	Yes	Yes	No (Q)	Yes	Yes	Yes	Yes	Yes	Yes		
FUNGI	PRTB_YEAST	Cerevisin	<i>S. cerevisiae</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (K)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (V)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
	YSP3_YEAST	Subtilisin-like protease 3	<i>S. cerevisiae</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (K)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (I)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
	RRT12_YEAST	Subtilase-type proteinase RRT12	<i>S. cerevisiae</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (K)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No (S)	Yes	Yes		
HUMAN	MBTP1_HUMAN	ne-bound transcription factor site-1	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (M)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
	NEC1_HUMAN	Neuroendocrine convertase 1	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (Q)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (V)	Yes	No (Y)	No (C)	No (S)	Yes	No (S)	Yes	Yes	Yes		
	PCSK4_HUMAN	rotein convertase subtilisin/kexin ty	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (M)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (I)	Yes	No (Y)	No (C)	No (S)	Yes	No (S)	Yes	Yes	Yes		
	PCSK5_HUMAN	rotein convertase subtilisin/kexin ty	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (A)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (I)	Yes	No (Y)	No (C)	No (S)	Yes	No (S)	Yes	Yes	Yes		
	PCSK6_HUMAN	rotein convertase subtilisin/kexin ty	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (S)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (I)	Yes	No (Y)	No (C)	No (S)	Yes	No (S)	Yes	Yes	Yes		
	PCSK7_HUMAN	rotein convertase subtilisin/kexin ty	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (V)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (I)	Yes	No (Y)	No (C)	No (S)	Yes	Yes	Yes	Yes	Yes		
	PCSK9_HUMAN	rotein convertase subtilisin/kexin ty	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (R)	Yes	Yes	No (L)	No (P)	Yes	Yes	Yes	Yes	No (M)	Yes	No (L)	Yes	Yes	Yes	Yes	No (A)	Yes		
	TPP2_HUMAN	Tripeptidyl-peptidase 2	<i>H. sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	No (H)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No (I)	No (S)	No (W)	Yes	Yes	Yes	Yes	Yes	Yes		
SEDOLISINS	SED1_ASPFU	Tripeptidyl-peptidase sed1	<i>Aspergillus fumigatus</i>	No (-)	No (-)	Yes	No (D)	Yes	No (-)	No (-)	No (-)	Yes	No (-)	Yes	No (-)	No (-)	No (-)	No (-)	Yes	Yes	Yes	No (I)	Yes	Yes	Yes	No (A)			
	TPP1_HUMAN	Tripeptidyl-peptidase 1	<i>Homo sapiens</i>	No (-)	No (-)	Yes	No (D)	Yes	No (-)	No (-)	No (-)	Yes	No (-)	Yes	No (-)	No (-)	No (-)	No (-)	Yes	Yes	No (S)	No (H)	Yes	No (S)	Yes	No (A)			
	SED3_ARTBC	Tripeptidyl-peptidase SED3	<i>Arthroderma benhamiae</i>	No (-)	No (-)	Yes	No (D)	Yes	No (-)	No (-)	No (-)	Yes	No (-)	Yes	No (-)	No (-)	No (-)	No (-)	No (S)	No (L)	No (Q)	Yes	No (S)	Yes	No (A)				
	SED3_ARTOC	Tripeptidyl-peptidase SED3	<i>Arthroderma otae</i>	No (-)	No (-)	Yes	No (D)	Yes	No (-)	No (-)	No (-)	Yes	No (-)	Yes	No (-)	No (-)	No (-)	No (-)	No (S)	No (L)	Yes	No (Q)	Yes	No (S)	Yes	No (A)			

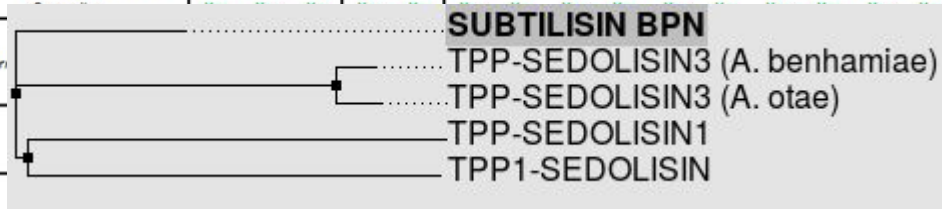
NEGATIVE CONTROL
(same clan: SB)

Subtilisin homologues vs non-homologues

Alpha and beta proteins (a/b)	Beta Sheet $\alpha/\beta/\alpha$, Left-Handed Crossover	Subtilisin-like (SB)	Subtilases (subtilisin) Serine-carboxyl proteinase (sedolisin)
-------------------------------	--	-----------------------------	---

	Abbreviation	Name	Organism	Subtilisin-like (SB)				Subtilases (subtilisin) Serine-carboxyl proteinase (sedolisin)																		
				D32	H64	S221	N155	T220	G83	L75	G34	G154	V68	S125	G65	A69	G70	L90	A152	F189	G193	P201	G219	A223	P225	G229
SUBTILISINS	SUBT_BACAM	Subtilisin BPN'	<i>B. amyloliquefaciens</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	SUBT_BACLI	Subtilisin Carlsberg	<i>B. licheniformis</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	SUBB_BACLE	Subtilisin BL	<i>B. lentus</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	SUBS_BACLE	Subtilisin Savinase	<i>B. lentus</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	SUBN_BACNA	Subtilisin NAT	<i>B. subtilis subsp. Natto</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	SUBT_BACSU	Subtilisin E	<i>B. subtilis (strain 168)</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	SUBT_GEOSU	Subtilisin J	<i>Geob. stearothermophilus</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	SUBT_BACSA	Subtilisin Amylosacchariticus	<i>B. subtilis subsp. Amylosacchariticus</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SUBT_BACPU	Subtilisin OS																									

	Abbreviation	Name	Organism	Subtilisin-like (SB)				Subtilases (subtilisin) Serine-carboxyl proteinase (sedolisin)																			
				D32	H64	S221	N155	T220	G83	L75	G34	G154	V68	S125	G65	A69	G70	L90	A152	F189	G193	P201	G219	A223	P225	G229	
FUNGI	TKSU_THEKO	Tk-subtilisin	<i>Thermoplasma</i>																								
	PRTB_YEAST	Cerevisin	<i>Saccharomyces cerevisiae</i>																								
FUNGUS	YSP3_YEAST	Subtilisin-like protease 3	<i>Yarrowia lipolytica</i>																								
	RRT12_YEAST	Subtilase-type proteinase RRT12	<i>Rhizoglyphus nigellii</i>																								
HUMAN	MBTP1_HUMAN	ne-bound transcription factor site-1	<i>Homo sapiens</i>																								
	NEC1_HUMAN	Neuroendocrine convertase 1	<i>Homo sapiens</i>																								
HUMAN	PCSK4_HUMAN	protein convertase subtilisin/kexin type 4	<i>Homo sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No (M)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (I)	Yes	No (Y)	No (C)	No (S)	Yes	No (S)	Yes	Yes
	PCSK5_HUMAN	protein convertase subtilisin/kexin type 5	<i>Homo sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No (S)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (I)	Yes	No (Y)	No (C)	No (S)	Yes	No (S)	Yes	Yes
HUMAN	PCSK6_HUMAN	protein convertase subtilisin/kexin type 6	<i>Homo sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No (S)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (I)	Yes	No (Y)	No (C)	No (S)	Yes	No (S)	Yes	Yes
	PCSK7_HUMAN	protein convertase subtilisin/kexin type 7	<i>Homo sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No (V)	Yes	Yes	No (C)	Yes	Yes	Yes	Yes	No (I)	Yes	No (Y)	No (C)	No (S)	Yes	Yes	Yes	Yes
HUMAN	PCSK9_HUMAN	protein convertase subtilisin/kexin type 9	<i>Homo sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No (R)	Yes	Yes	No (L)	No (P)	Yes	Yes	Yes	No (M)	Yes	No (L)	Yes	Yes	Yes	Yes	No (A)	Yes
	TPP2_HUMAN	Tripeptidyl-peptidase 2	<i>Homo sapiens</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No (H)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No (I)	No (S)	No (W)	Yes	Yes	Yes	Yes	Yes	Yes

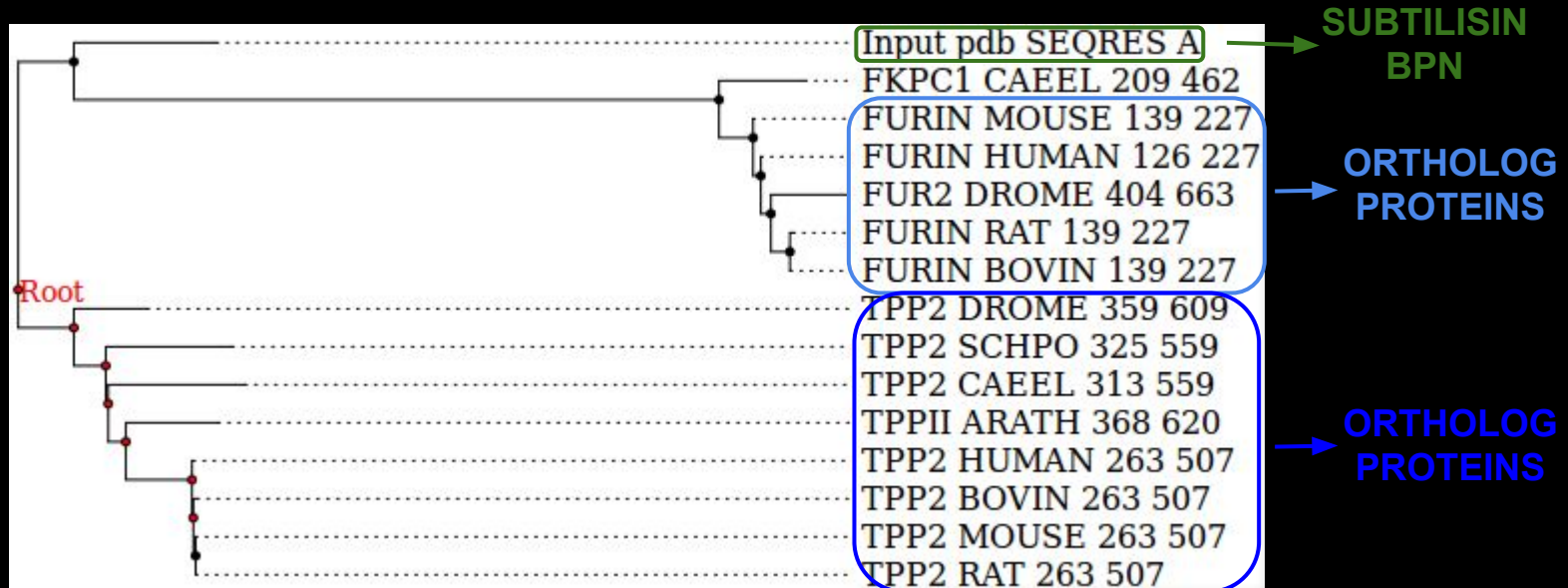


	Abbreviation	Name	Organism	Subtilisin-like (SB)				Subtilases (subtilisin) Serine-carboxyl proteinase (sedolisin)																			
				D32	H64	S221	N155	T220	G83	L75	G34	G154	V68	S125	G65	A69	G70	L90	A152	F189	G193	P201	G219	A223	P225	G229	
SEDOLISINS	SED1_ASPFU	Tripeptidyl-peptidase sed1	<i>Aspergillus fumigatus</i>	No (-)	No (-)	Yes	No (D)	Yes	No (-)	No (-)	No (-)	Yes	No (-)	Yes	No (-)	No (-)	No (-)	No (-)	Yes	Yes	Yes	No (I)	Yes	Yes	Yes	No (A)	
	TPP1_HUMAN	Tripeptidyl-peptidase 1	<i>Homo sapiens</i>	No (-)	No (-)	Yes	No (D)	Yes	No (-)	No (-)	No (-)	Yes	No (-)	Yes	No (-)	No (-)	No (-)	No (-)	Yes	Yes	No (S)	No (H)	Yes	No (S)	Yes	Yes	
	SED3_ARTBC	Tripeptidyl-peptidase SED3	<i>Arthroderma benhamiae</i>	No (-)	No (-)	Yes	No (D)	Yes	No (-)	No (-)	No (-)	Yes	No (-)	Yes	No (-)	No (-)	No (-)	No (-)	No (S)	No (L)	No (Q)	Yes	No (Q)	Yes	No (S)	Yes	No (A)
	SED3_ARTOC	Tripeptidyl-peptidase SED3	<i>Arthroderma otae</i>	No (-)	No (-)	Yes	No (D)	Yes	No (-)	No (-)	No (-)	Yes	No (-)	Yes	No (-)	No (-)	No (-)	No (-)	No (S)	No (L)	Yes	No (Q)	Yes	No (S)	Yes	No (A)	

NEGATIVE CONTROL
(same clan: SB)

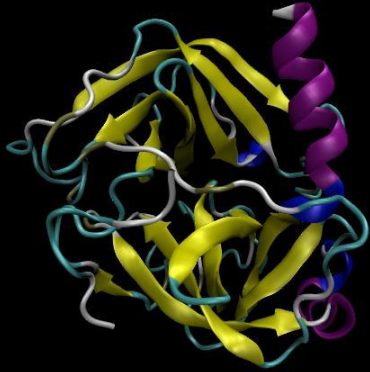
Subtilisin homologues: ORTHOLOGUES

ORTHOLOG PROTEINS: homologous proteins that retain the same function along the evolution course and are present in different species.

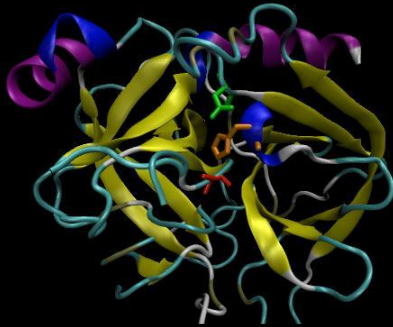


Subtilisins vs Chymotrypsins

CHYMOTRYPSIN (1AB9)



His 57
Asp 102
Ser 195



SUBTILISIN (1SBT)



Active site groups



- Identical (different order)
- Almost identical in the three-dimensional structure

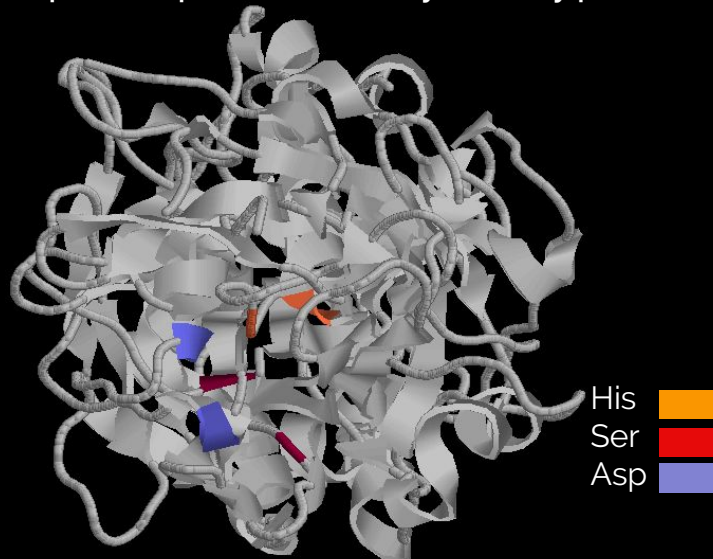
Asp 32
His 64
Ser 221



An example of **CONVERGENT EVOLUTION**

Subtilisins vs Chymotrypsins

Superimposition chymotrypsin and subtilisin with STAMP (USING ALIGNFIT)



STAMP
(Rough)
Score: 0.42
RMS: 2.52



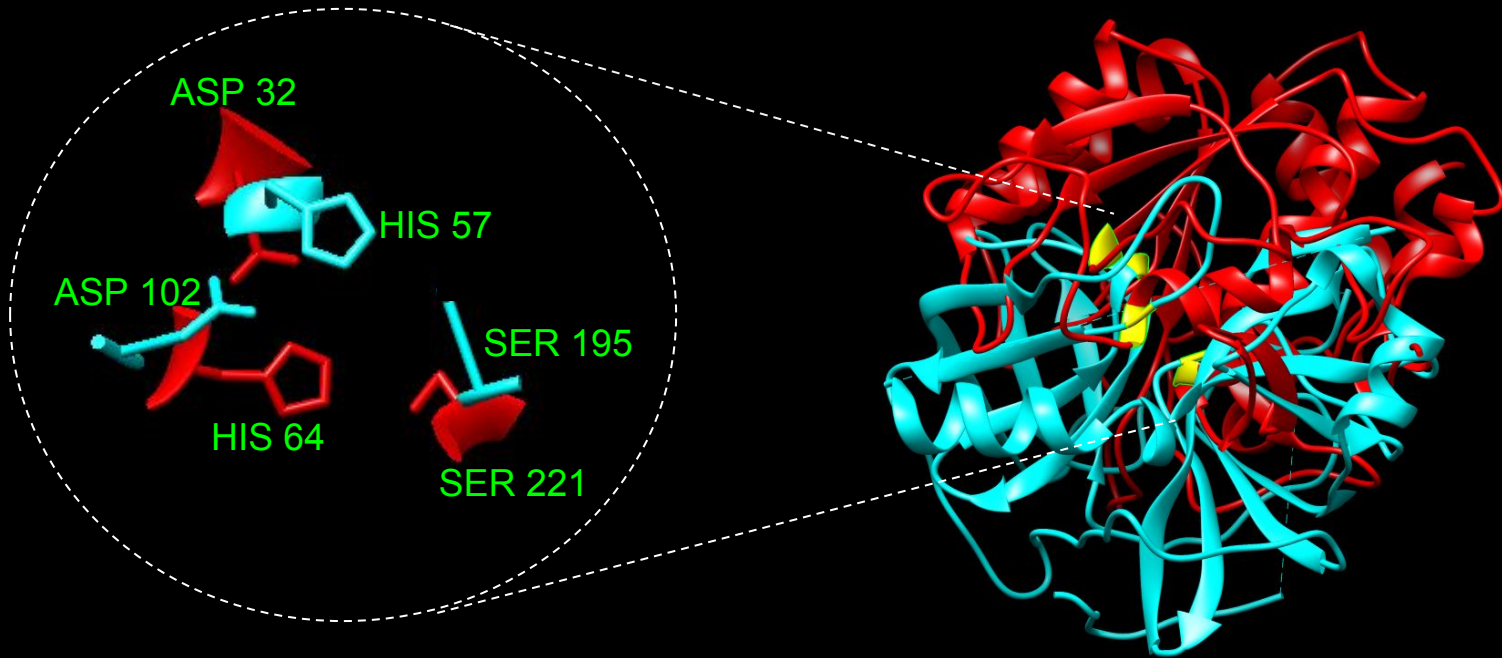
STAMP
(alignfit)
Score: 0.54
RMS: 2.26



```
Cluster: 1 ( 1gci & 1st3 ) Sc 9.75 RMS 0.25 Len 269 nfit 269
See file SUBTCHYM.1 for the alignment and transformations
Cluster: 2 ( 1scj & 1mee ) Sc 9.71 RMS 0.35 Len 275 nfit 275
See file SUBTCHYM.2 for the alignment and transformations
Cluster: 3 ( 3vyv & 1scj 1mee ) Sc 9.58 RMS 0.24 Len 275 nfit 274
See file SUBTCHYM.3 for the alignment and transformations
Cluster: 4 ( 1sbc & 3vyv 1scj 1mee ) Sc 9.53 RMS 0.54 Len 275 nfit 273
See file SUBTCHYM.4 for the alignment and transformations
Cluster: 5 ( 1sbt & 1sbc 3vyv 1scj 1mee ) Sc 9.52 RMS 0.67 Len 275 nfit 273
See file SUBTCHYM.5 for the alignment and transformations
Cluster: 6 ( 1gci 1st3 & 1sbt 1sbc 3vyv 1scj 1mee ) Sc 9.36 RMS 0.69 Len 275 nfit 264
See file SUBTCHYM.6 for the alignment and transformations
Cluster: 7 ( 1ab9 & 1gci 1st3 1sbt 1sbc 3vyv 1scj 1mee ) Sc 1.72 RMS 3.58 Len 372 nfit 65 LOW SCORE
```


Subtilisins vs Chymotrypsins

Superimposition catalytic triad with XAM



CONCLUSIONS

- Subtilisins contain specific regions highly conserved, which allows these proteins to carry out the peptidase function
- The propeptide chain regulates the catalytic activity of subtilisins to assure a proper function in the organism. Its sequence is conserved among the different subtilisins analysed. Not all subtilisins have a propeptide chain
- In this project, we described 9 subtilisins with a high homology both in sequence and structure. All them are found in bacillus species
- We have found many other subtilisin homologues in archaea, fungi and animal kingdoms
- The homology decreases as we go further away in the evolution course
- Most of the residues of these homologues involved in subtilisins function remain conserved, although some other differ
- TPP2 and Furin protein families have been described as ortholog proteins in some specific species
- Despite the different structure, the active site of subtilisins and chymotrypsin can be superimposed. This suggests a convergent evolution process

QUESTIONS

1. What are the main residues involved in the catalytic activity of subtilisins?
 - a) **Asp, His, Ser**
 - b) Asn, Phe, Val
 - c) Ala, Thr, Gly
 - d) All the residues in the active site are equally important
 - e) None of the answers is correct
2. What is the correct answer related to subtilisin's autoprocessing?
 - a) It requires the folding of the subtilisin domain
 - b) Another activated subtilisin is necessary to complete the process
 - c) Both answers above are correct
 - d) The propeptide chain regulates the catalytic activity
 - e) **All the answers are correct**
3. Related with subtilisin homologues:
 - a) No homologues are found in any animal species
 - b) Structure of homologues in plants are the most conserved
 - c) The only kingdom with homologues, a part from bacteria, is archea
 - d) **Some human proteins can be considered subtilisin homologues**
 - e) None of the answers is correct

QUESTIONS

4. When comparing subtilisins and chymotrypsins:
- a) They have a high structural homology, but low sequence homology
 - b) They have a low structural homology, but a high sequence homology
 - c) Both sequence and structural homology are high
 - d) They both use the same catalytic mechanism**
 - e) Chymotrypsins cannot be crystallized, so we can't obtain any structural parameter
5. Related to subtilisin's structure:
- a) Subtilisin BPN has two Calcium binding domains**
 - b) The active site is formed by three essential residues: Ala, Leu, Gly
 - c) There are three oxyanion holes
 - d) All subtilisins have a propeptide chain
 - e) None of the answers is correct
6. In subtilisin BPN, the propeptide:
- a) Acts as a chaperone
 - b) Is a temporal inhibitor
 - c) Both answers above are correct
 - d) Contains 2 cleavage sites
 - e) All the answers are correct**

QUESTIONS

7. Which of the following is not an essential feature in subtilisin BPN:

- a) The catalytic triad
- b) The oxyanion hole
- c) The specificity pockets
- d) The potassium binding loop**
- e) The calcium binding domain

8. Related to subtilisin's classification (SCOP):

- a) Subtilisins belong to trypsin-like superfamily
- b) Trypsin and chymotrypsin belong to subtilisin-like superfamily
- c) Both answers above are correct
- d) Subtilisins belong to subtilisin-like superfamily whilst trypsin and chymotrypsin belong to trypsin-like superfamily**
- e) None of the answers is correct

QUESTIONS

9. Chymotrypsin and subtilisin:

- a) Have identical structures but different functions
- b) Have a similar structure and different functions
- c) Are an example of convergent evolution**
- d) Are an example of divergent evolution
- e) None of the options is correct

10. In the catalytic mechanism:

- a) His acts as a nucleophile
- b) Ser acts as a nucleophile and is involved in the tetrahedral transition state**
- c) Both answers above are correct
- d) Water is not necessary to allow the start of another catalytic cycle
- e) All the answers are correct

REFERENCES

Branden C, Tooze J. Introduction to protein Structure. 2nd ed. New York: Garland Publishing; 1998.

Braxton S, Wells JA. The importance of a distal hydrogen bonding group in stabilizing the transition state in subtilisin BPN'. *J Biol Chem*. 1991;266(18):11797-800.

Di Cera E. Serin Proteases. *IUMB Life*. 2009;61(5):510-5.

Gallagher T, Gilliland G, Wang L, Bryan P. The prosegment-subtilisin BPN' complex: crystal structure of a specific 'foldase'. *Structure*. 1995;15:3(9):907-14.

Jain SC, Shinde U, Li Y, Inouye M, Berman HM. The crystal structure of an autoprocessed Ser221Cys-subtilisin E-propeptide complex at 2.0 Å resolution. *J. Mol. Biol*. 1998;285:137-44.

Mulder FA, Schipper D, Bott R, Boelens R. Altered flexibility in the substrate-binding site of related native and engineered high-alkaline Bacillus subtilisins. *J Mol Biol*. 1999;292(1):111-23.

Osmulski PA, Gaczynska M. A new large proteolytic complex distinct from the proteasome is present in the cytosol of fission yeast. *Curr Biol*. 1998;8(18):1023-6.

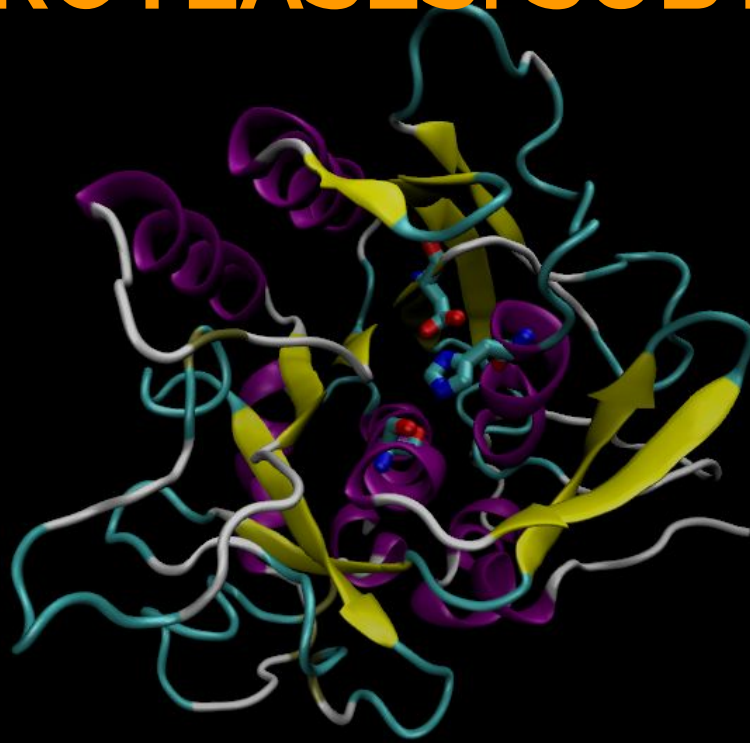
Pulido MA, Tanaka S, Sringiew C, You DJ, Matsumura H, Koga Y, Takano K, Kanaya S. Requirement of left-handed glycine residue for high stability of the Tk-subtilisin propeptide as revealed by mutational and crystallographic analyses. *J Mol Biol*. 2007;374(5):1359-73.

REFERENCES

- Siezen RJ, Leunissen JA. Subtilases: the superfamily of subtilisin-like serine proteases. *Protein Science: A Publication of the Protein Society*. 1997;6(3):501-23.
- Uehara R, Angkawidjaja C, Koga Y, Kanaya S. Formation of the high-affinity calcium binding site in pro-subtilisin E with the insertion sequence IS1 of pro-Tk-subtilisin. *Biochemistry-US*. 2013;52(50):9080-8.
- Takeuchi Y, Satow Y, Nakamura KT, Mitsui Y. Refined crystal structure of the complex of subtilisin BPN' and Streptomyces subtilisin inhibitor at 1.8 Å resolution. *J Mol Biol*. 1991;221(1):309-25.

Thank you for your attention!

SERIN-PROTEASES: SUBTILISINS



Júlia Garcia, Gemma Núñez, Marina Solà, Marc Ten
24th February 2017