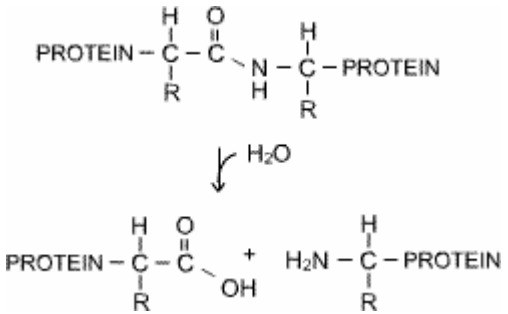
Protease

- Branden & Tooze, Chapter 11
- An enzyme that hydrolyzes the peptide bond
  - works without consuming energy because peptide bond hydrolysis is exothermic
  - releases ~ 2 kcal/mol when the bond breaks
  - resonance structure that gives it partial double bond characteristics makes the bond kinetically stable although thermodynamically unstable
  - study the mechanism of serine protease
  - determinants of specificity



- Role of protease in disease development
  - implicated in numerous hereditary diseases
  - normal developmental process and lifecycle of pathogens (virus and parasite) both depend on protease activity
  - cancer needs proteases to break loose and metastasize
  - structure-based design of protease inhibitor has a potential of regulating disease propagation

Peptide bond hydrolysis

- Proteases accelerate peptide bond hydrolysis
- Often form an acyl-enzyme intermediate which is subsequently resolved
- Tetrahedral transition state is formed at two different time points
- End result is an addition of water (i.e. hydrolysis) but water does not attack the main chain directly

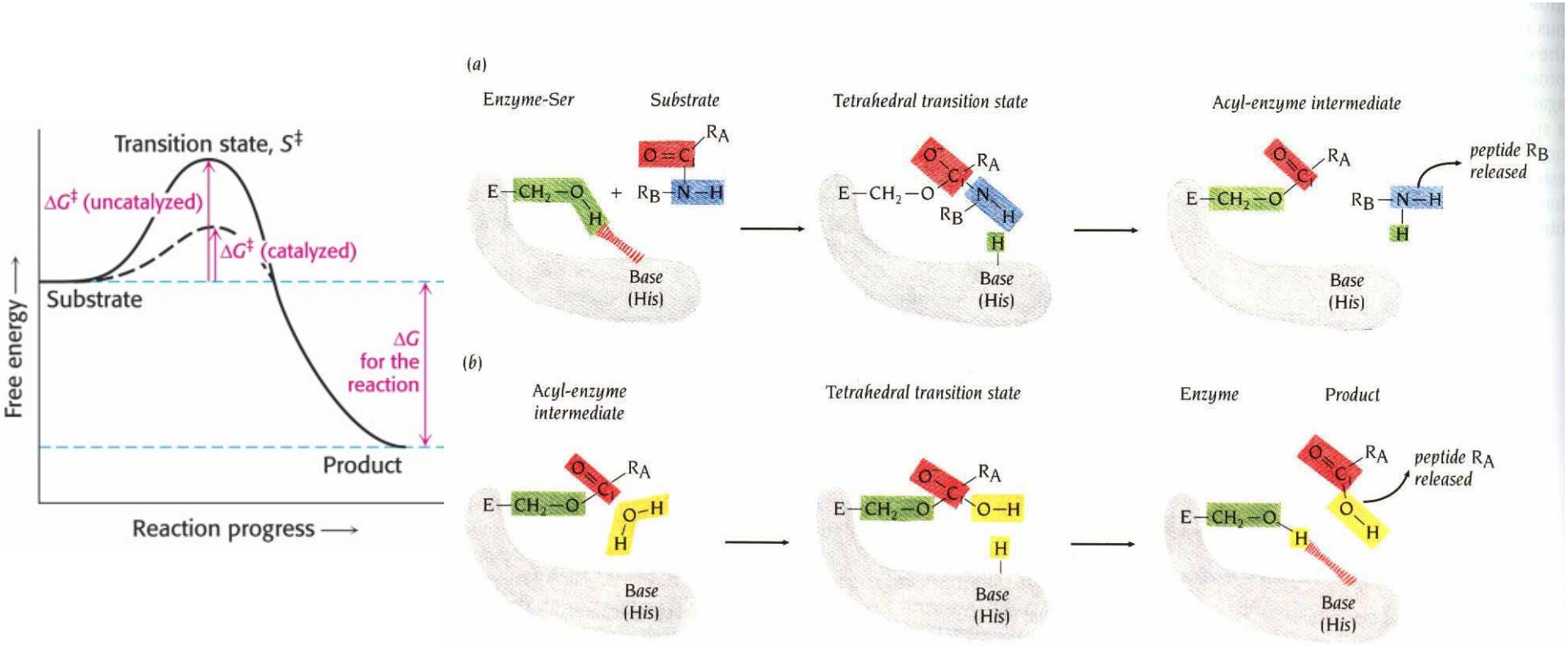


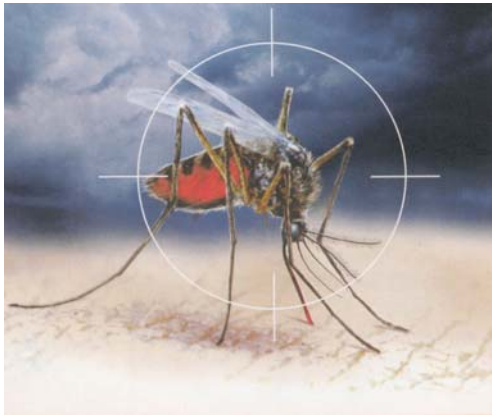
Table 1

Peptidases used in structure-based drug design.

Peptidase	Biological function	Disease
Cysteine peptidases		
Cathepsin B	Antigen processing	Acute pancreatitis, cancer
Cathepsins L and S	Lysosomal proteolysis	Inflammation
FP-2	Haemoglobin degradation	Malaria
Caspase-1	Maturation of interleukin 1-β	Ameliorate inflammation, endotoxic shock
Caspase-3 and -7	Executioner caspases in apoptosis	Neuronal and cardiac ischemic injuries
Calpain-1 and -2	Degradation of cytoskeletal proteins	Stroke, neural injuries
Picomain cysteine peptidases	Processing of viral pro-protein	Virus infection
Serine peptidases		
Thrombin	Proteolysis of fibrinogen	Thrombosis
Factor Xa	Conversion of prothrombin to thrombin	Thrombosis
Factor VIIa	Activation of factors IX and X	Thrombosis
Urokinase	Activation of plasminogen	Cancer
Flavivirus peptidases	Processing of polyprotein	Viral infection
DPP-4	Processing of hormone precursors	Type 2 diabetes mellitus
20S proteasome	Ubiquitin-dependent protein degradation	Cancer
Aspartic peptidases		
HIV peptidase	Processing of viral pro-protein	HIV infection
Renin	Processing of angiotensinogen	Blood pressure
Memapsin-2	β-Secretase activity	Alzheimer's disease
Plasmepsin	Haemoglobin degradation	Malaria
Metallo peptidases		
Angiotensin-converting enzyme	Conversion of angiotensin	Hypertension
Botulinium neurotoxin	Cleavage of SNAP proteins	*Clostridium* and tetanus infection
Anthrax lethal factor	Cleavage of MAPKK	*B. anthracis* infection
Matrix metallopeptidase-1	Degradation of connective tissue	Tissue damage in tumour invasion
FtsH	Elimination of misfolded proteins	Neurological diseases
Carboxypeptidases B and U	Cleavage of tissue plasminogen activator	Blood coagulation
PSMA	Liberates glutamate from Ac-Asp-Glu in brain	Marker for prostate cancer

Mittl and Grutter, COSB 16, 769 (2006)

Cysteine protease



Papain

- falcipain 2—helps malaria parasite P. falciparum degrade hemoglobin
- cathepsin—lysosomal protease, implicated in musculoskeletal disease and emphysema
- cathepsin K inhibitor is a potential therapeutic against osteoporosis

Caspase

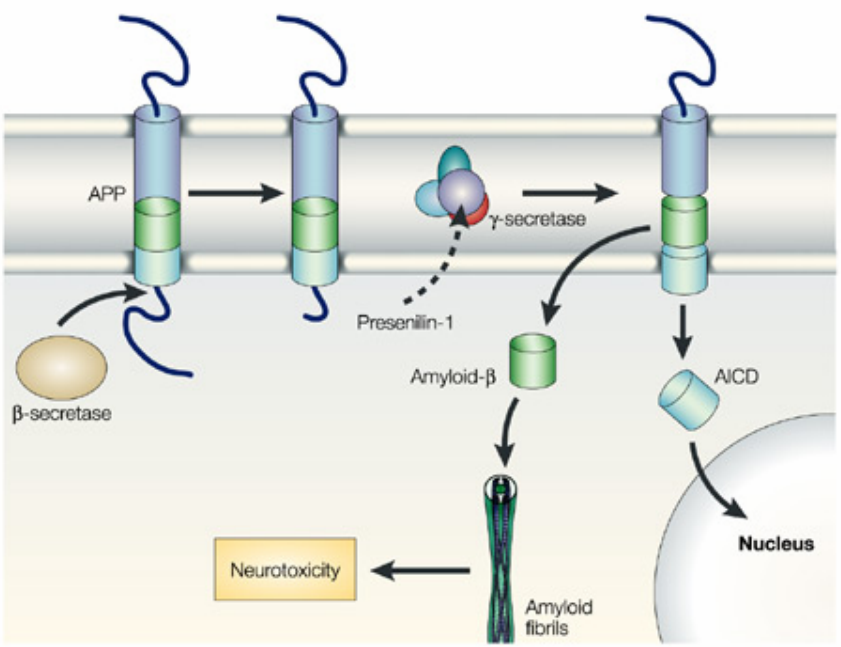
- cuts proteins after asp
- involved in programmed cell death (cancer and AD) and inflammation
- usually dimeric although pro-enzyme (zymogen) may be monomeric
- autocatalysis induces dimerization and activation

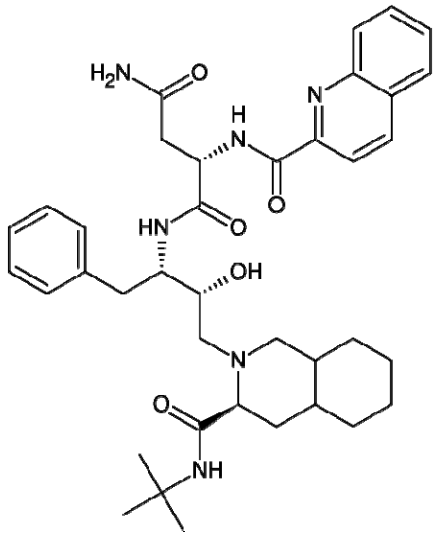
Picornain

- structurally similar to serine protease chymotrypsin but active site ser has been mutated to cys
- important for the spread of hepatitis A

Aspartic protease

- Smallest class of human proteases with only 15 members
- Play important roles in physiological and pathological processes
- beta-site amyloid precursor protein cleaving enzyme (BACE, β-secretase)
  - Alzheimer's disease (AD) is triggered by proteolytic degradation
  - potential target for treatment of AD



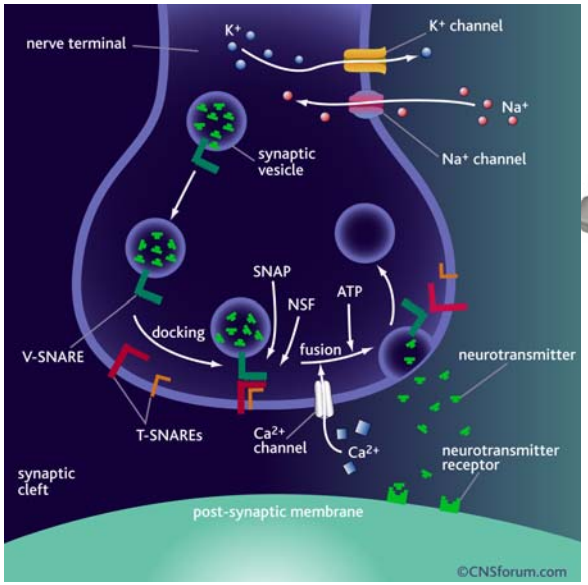


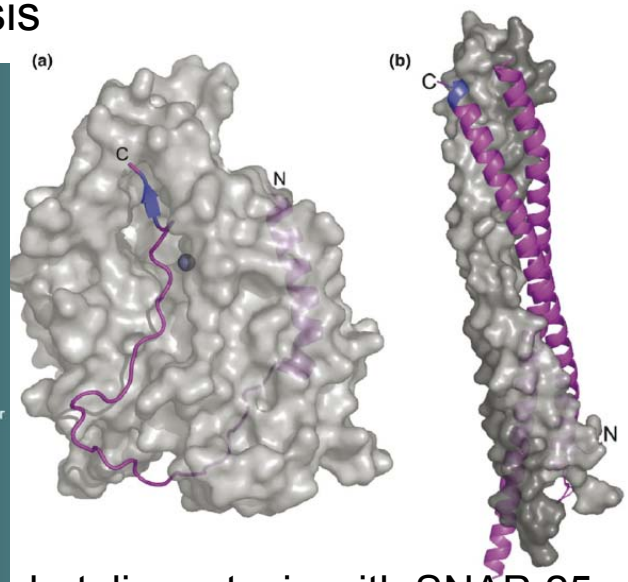
Saquinavir: HIV protease inhibitor

- HIV-1 protease—active site at the homodimeric interface
- Plasmepsin—breaks down hemoglobin and thus helps spread malaria

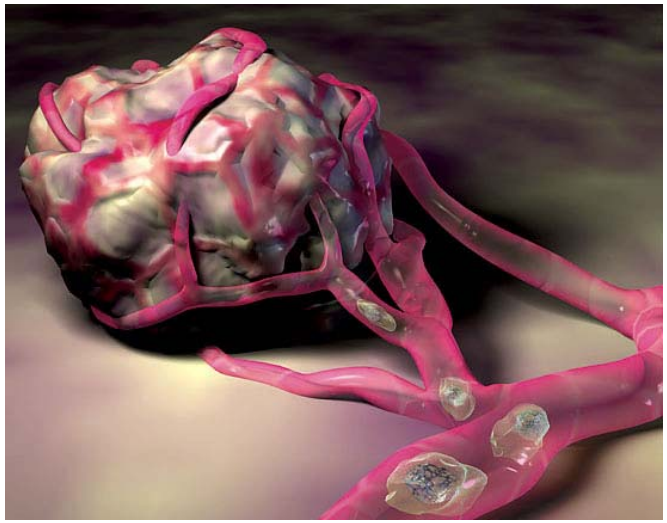
Metalloprotease

- Uses activated metal to hydrolyze a peptide bond
- Botulinum and tetanus toxins are both zinc metalloproteases
  - cleaves the SNARE complex and thus prevents neurotransmission, causing paralysis





botulinum toxin with SNAP-25 peptide

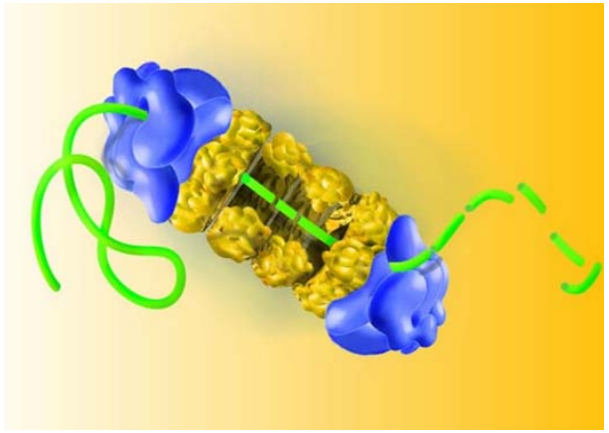


metastasis

- Anthrax lethal factor is a zinc protease and cleaves signaling proteins in the MAP kinase pathway
- Metastasizing cancer requires matrix metalloprotease (MMP) to migrate to distant parts of the body

Serine protease

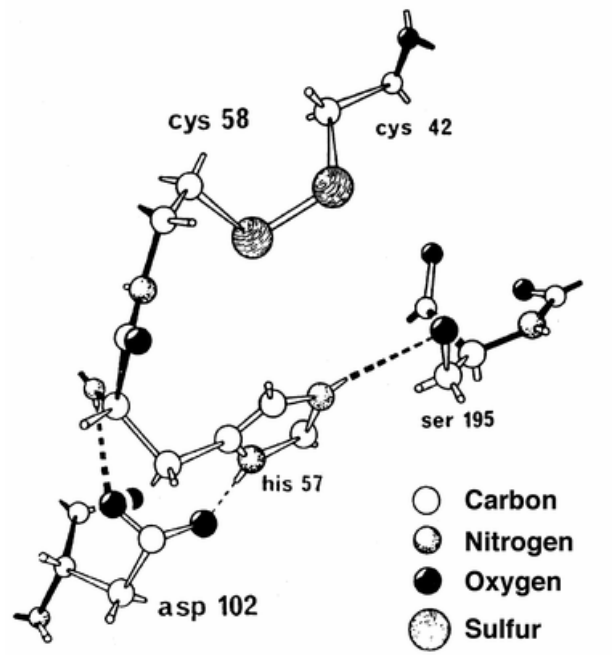
- First class of proteases to be determined to high resolution
- Digestive enzymes—trypsin, chymotrypsin, elastase
- Factor Xa
  - initiates blood coagulation by cleaving prothrombin
  - used in biotechnology to cleave recombinant protein from a fusion
- Protein C—anti-coagulant enzyme—mutation causes thrombosis
- S20 subunit of the proteasome
  - inhibition of S20 leads to accumulation of protein in the cytosol, inducing apoptosis
  - potential cancer drug
- Dipeptidyl peptidase IV (DPP-4)
  - inactivates glucagon-like peptide 1 (GLP-1) and glucose-dependent insulinotropic polypeptide (GIP)
  - inhibition of DPP-4 increases GLP-1 and GIP and decreases blood glucose
  - potential target for treating type 2 diabetes mellitus



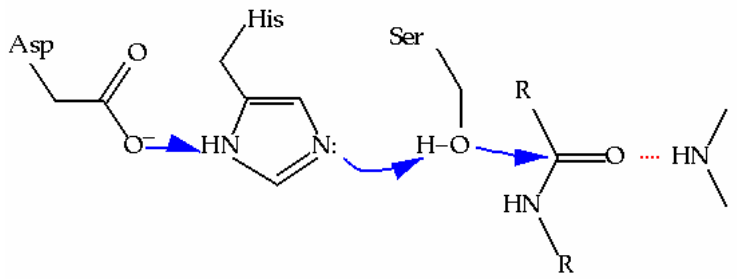
Catalytic triad

The active site of a serine protease consists of three catalytic residues

- **Ser, His, and Asp**
- although they are sequentially far apart, they are structurally close
- charge relay (proton shuttle) set up among these side chains make serine hydroxyl a good nucleophile for attack on the amide bond



active site of chymotrypsin
Blow et al, Nature 221, 337 (1969)



The catalytic triads may appear in evolutionarily unrelated proteins, which have independently evolved the catalytic mechanism

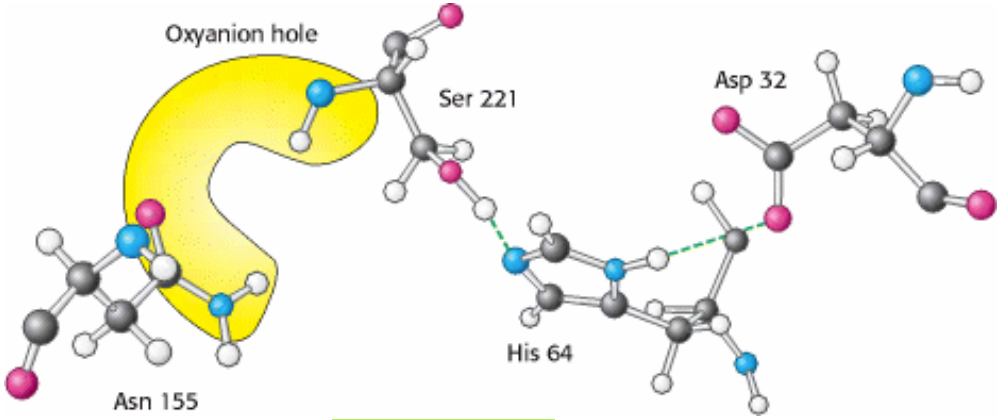
- in some enzymes asp may be substituted with glu and ser with thr
- convergent evolution

Oxyanion hole

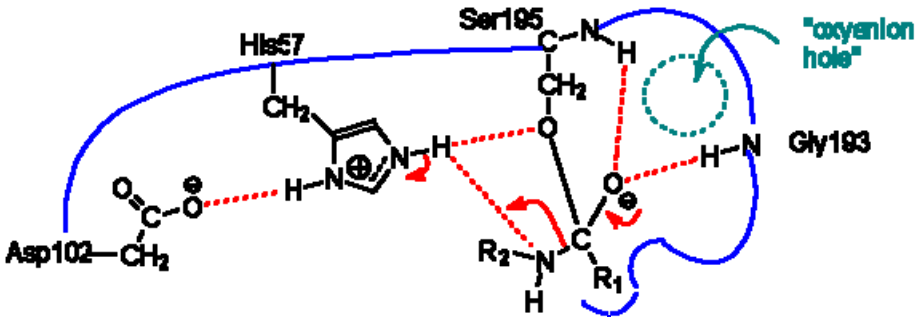
As the amide bond is attacked by the catalytic serine residue, the carbonyl oxygen develops a negative charge

Oxyanion hole is lined with potential hydrogen bond donors to neutralize this excess net negative charge and thus stabilize the transition state

- main chain and side chain NH of N155 of subtilisin
- S195 NH and G193 NH of chymotrypsin



subtilisin



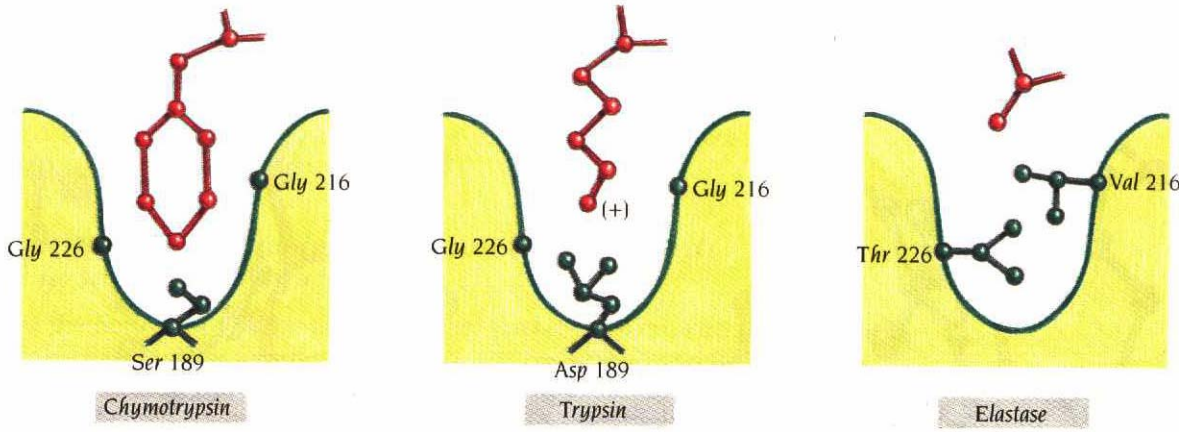
chymotrypsin

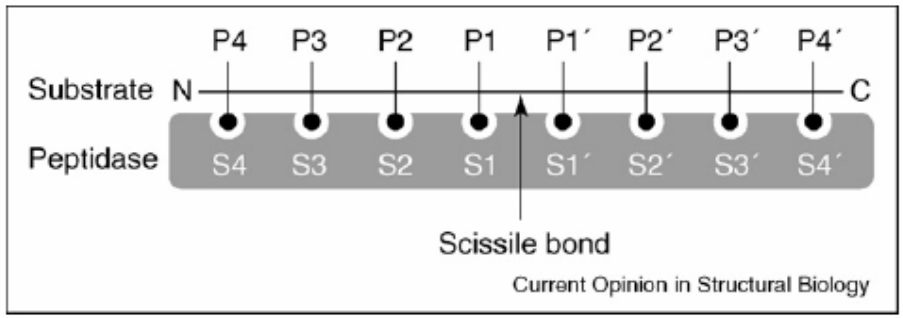
Specificity pocket

Proteases recognize specific amino acid sequences

- **trypsin** cuts after positively charged residues—K, R
- **chymotrypsin/subtilisin** cuts after large hydrophobic residues—F, Y, W, L
- elastase cuts after small hydrophobic residues—A, V, G

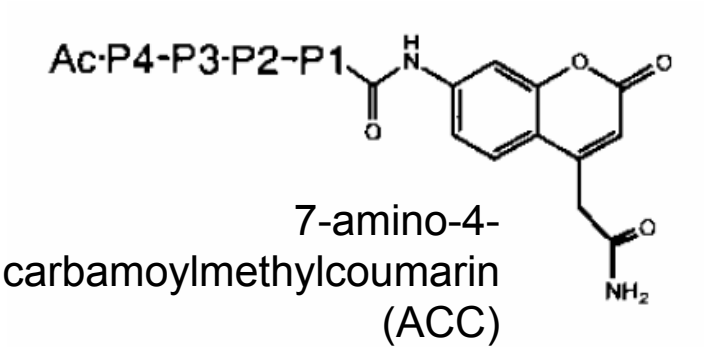
Side chain of specificity determining residue fits in the specificity pocket





Determining specificity

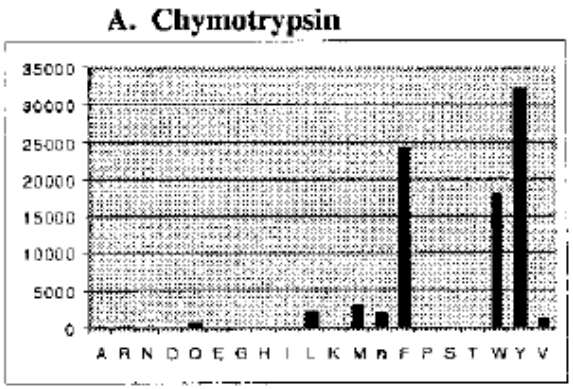
- Many proteases have not been characterized
- Proteolytic specificity may be examined by constructing a fluoregenic peptide library and determining the kinetics in high throughput

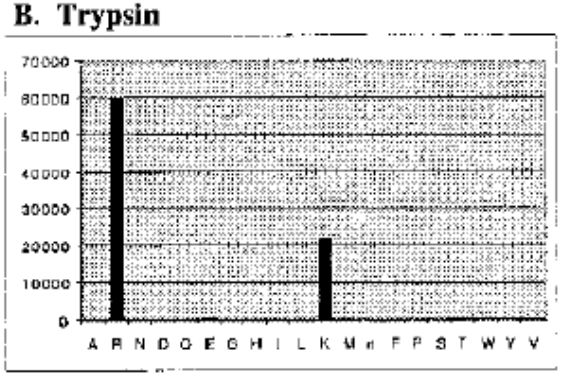


7-amino-4-carbamoylmethylcoumarin (ACC)

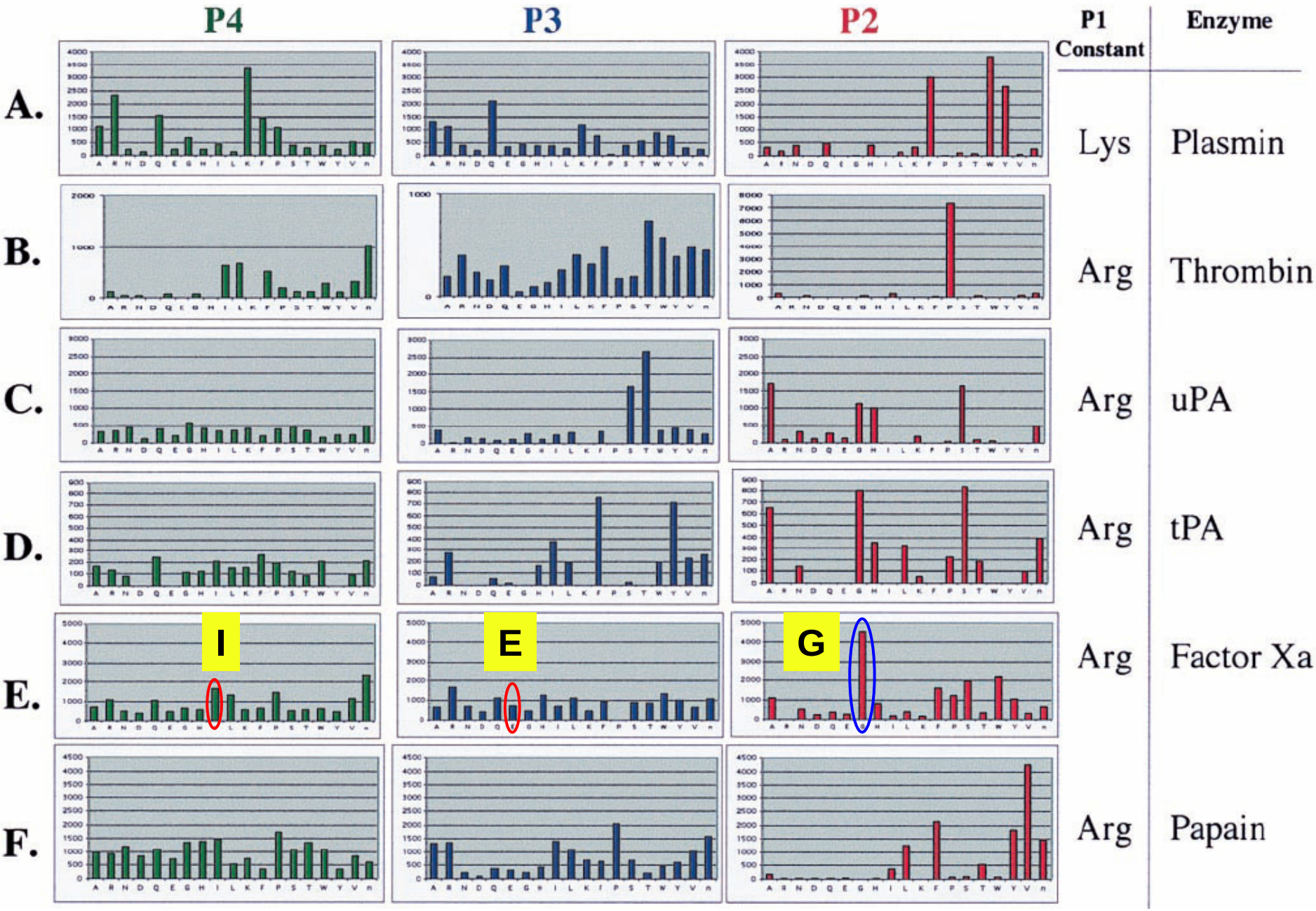
Compound	λ_{max}, nm ex	em
ACC	350	450
7-Nle-Thr-Pro-Lys-ACC	325	400

- total of 137,180 peptide sequences





Harris et al, PNAS 97, 7754 (2000)

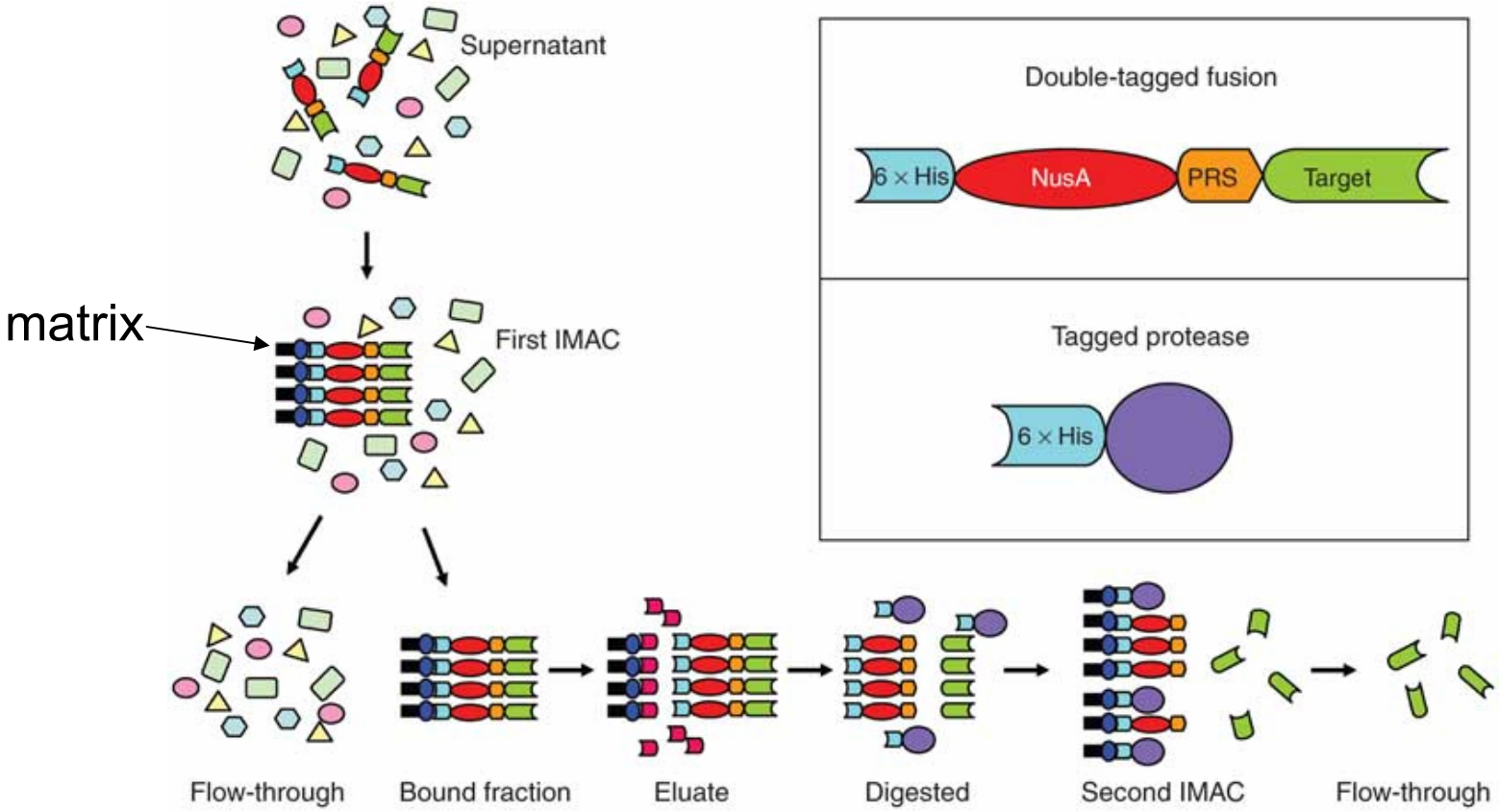

P4
P3
P2
P1 Constant
Enzyme
A.
Lys
Plasmin
B.
Arg
Thrombin
C.
Arg
uPA
D.
Arg
tPA
E.
I
E
G
Arg
Factor Xa
F.
Arg
Papain

Biotechnological application

Some proteases recognize short peptide sequences

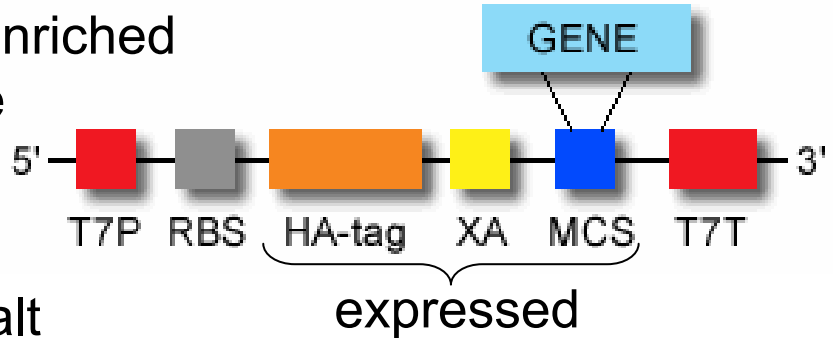
- factor Xa: Ile-Glu-Gly-Arg (IEGR)
- enterokinase: Asp-Asp-Asp-Asp-Lys (DDDDK)
- tobacco etch virus (TEV): Glu-Asn-Leu-Tyr-Phe-Gln ↓ Gly (ExxYxQ/S)(G)

Protein purification is often performed using fusion proteins



Affinity purification

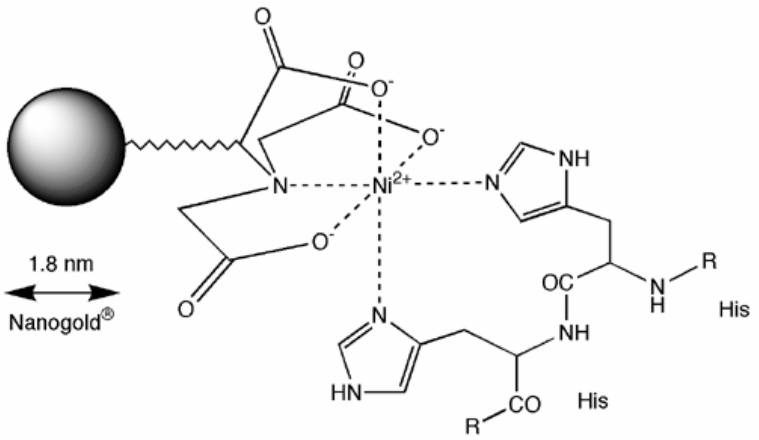
One step purification that yields a highly enriched sample ready for biochemical and some structural studies (NMR and CD)

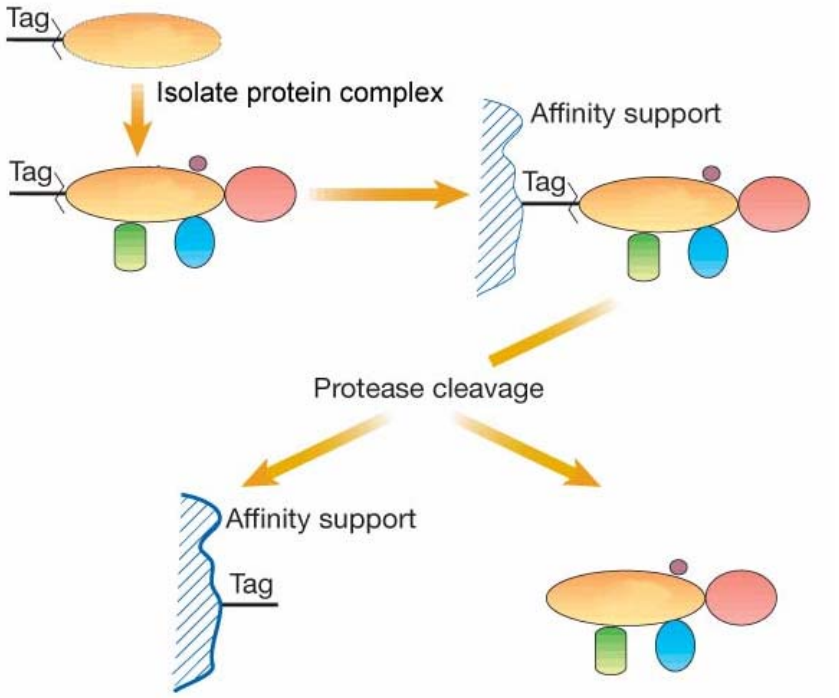


Hexahis (6xHis) : immobilized nickle, cobalt

Affinity tags (e.g. hemagglutinin tag) : affinity column made of monoclonal antibody

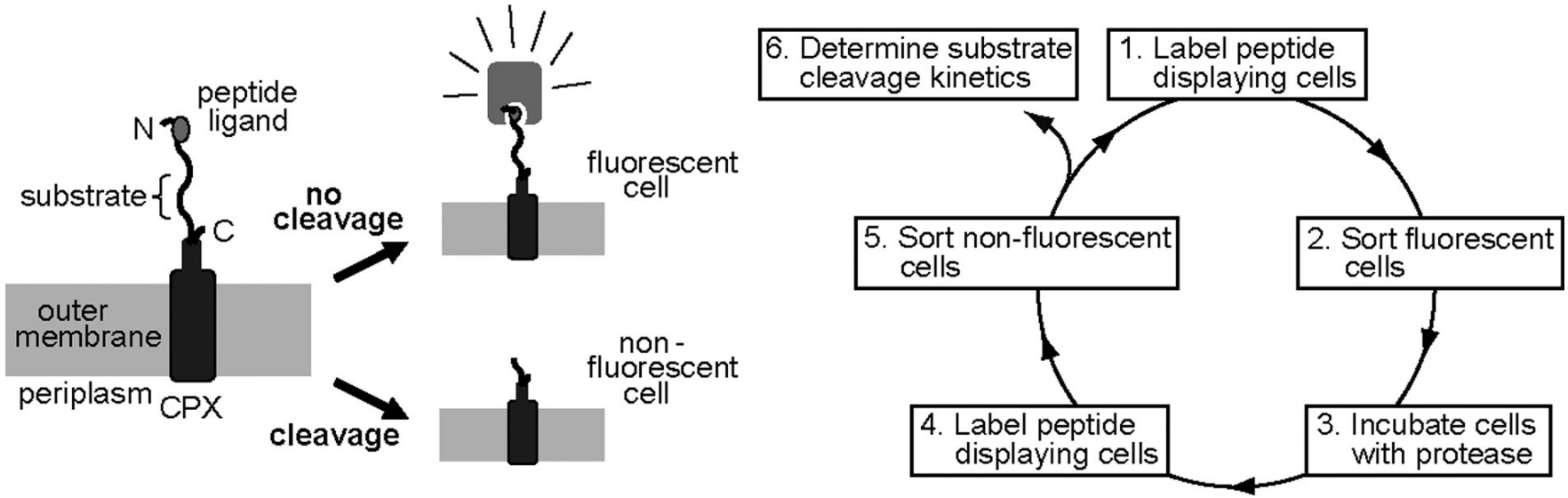
Biotinin : streptavidin, avidin





High throughput specificity determination

- Identify optimal protease substrates by using kinetic analysis of a **random peptide library** expressed on the outer surface of bacteria
- Each bacteria harbors a unique peptide sequence
- Genomic information (i.e. DNA) and phenotypic information (i.e. peptide) are physically linked to each other—once the desired peptides have been identified, their sequences can be easily decoded by sequencing the DNA

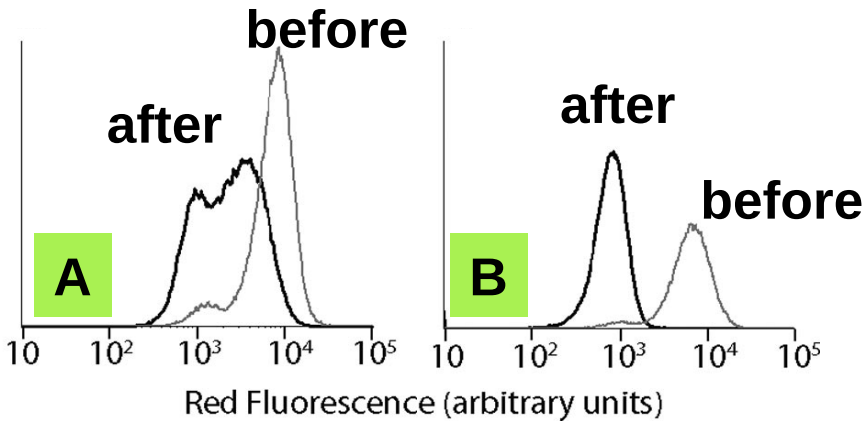


Boulware and Daugherty, PNAS 103, 7583 (2006)

Control substrates for EK

A: negative control: GGQSGQ

B: positive control: DDDDK



A peptide library with six randomized positions—a theoretical diversity of 6.4 x 10^7

Screen against caspase-3 (DEVD) and enteropeptidase (DDDDK)

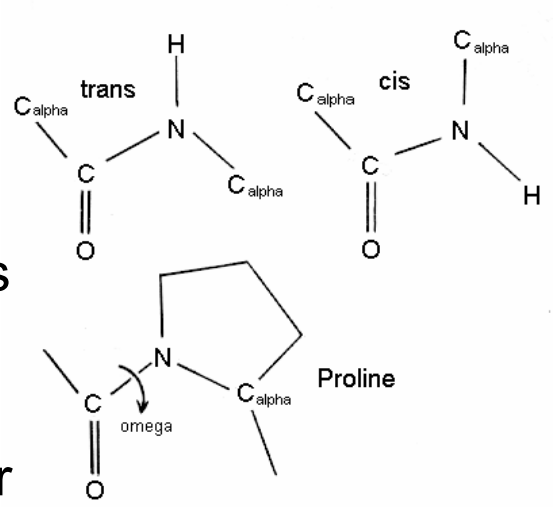
Table 1. Caspase-3 substrates identified by using CLiPS with a 6-mer library

Substrate	P5	P4	P3	P2	P1	P1′	P2′	P3′	P4′	Conversion
Canonical	q	D	E	V	D	G	g	q	s	0.95 ± 0.03
CS 2.7	s	D	G	V	D	G	W	g	g	0.95
CS 2.14 (4)*	s	D	V	V	D	G	W	g	g	0.94 ± 0.03
CS 2.23	s	D	G	V	D	G	V	g	g	0.93
CS 2.11	g	g	s	L	D	T	W	T	A	0.81
CS 2.20 (2)	L	D	T	V	D	R	g	g	q	0.79 ± 0.01

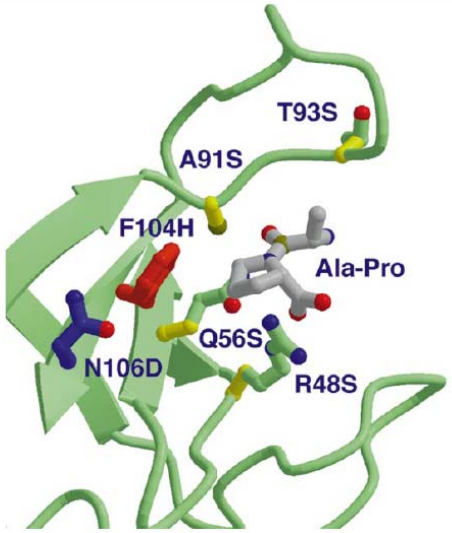
Table 2. Enteropeptidase substrates identified by using CLiPS with a 6-mer library

Substrate	P5	P4	P3	P2	P1	P1′	P2′	P3′	P4′	Conversion
Canonical	D	D	D	D	K	g	g	q	s	0.15 ± 0.08
EP 4.3	s	S	G	D	R	M	W	g	g	0.97 ± 0.01
EP 4.6	s	S	G	E	R	M	M	G	g	0.93 ± 0.03
EP 4.7	g	s	D	D	R	R	A	G	g	0.91 ± 0.03
EP 4.8	V	R	D	Y	R	M	g	g	q	0.87 ± 0.04
EP 3.6	g	s	s	D	R	A	R	V	W	0.86 ± 0.05

Designing new protease



- Cyclophilin is a prolyl peptidyl isomerase—facilitates the interconversion between the trans and cis conformation of proline
- R48, Q56, A91, T93 were individually mutated to ser
- Mutate three amino acids close to the peptide binding cleft to form a catalytic triad similar to that found in serine proteases
  - A91S, F104H, N106D
- Hydrolyzes Ala-Pro on the amino side of pro



Quemeneur et al, Nature 391, 301 (1998)

Table 1 Kinetic parameters for ECypP variants

	Efficiency k_{cat}/K_m ($s^{-1}M^{-1}$)	K_m ($\times 10^{-3}$ M)	k_{cat} (s^{-1})	Rate enhancement k_{cat}/k_{uncat}
WT	$<10^{-3}$	ND		
R48S	$<10^{-3}$	ND		
Q56S	$5.8 \pm 0.8 \times 10^{-3}$	ND		
A91S	73.1 ± 2.2	0.6 ± 0.15	0.044 ± 0.012	0.9×10^{7}
T93S	$<10^{-3}$	ND		
A91S-F104H	81.8 ± 3.3	1.9 ± 0.10	0.155 ± 0.015	3.23×10^{7}
A91S-F104H-N106D	1675 ± 12	2.4 ± 0.09	4.0 ± 0.2	8.33×10^{8}