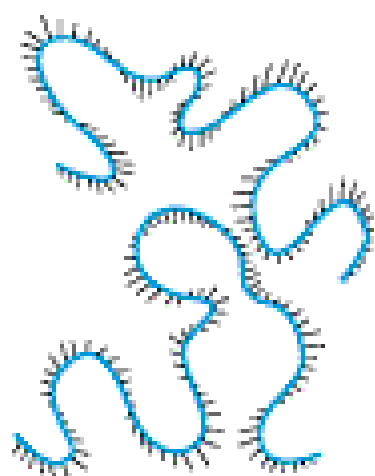


Native DNA



Denatured DNA

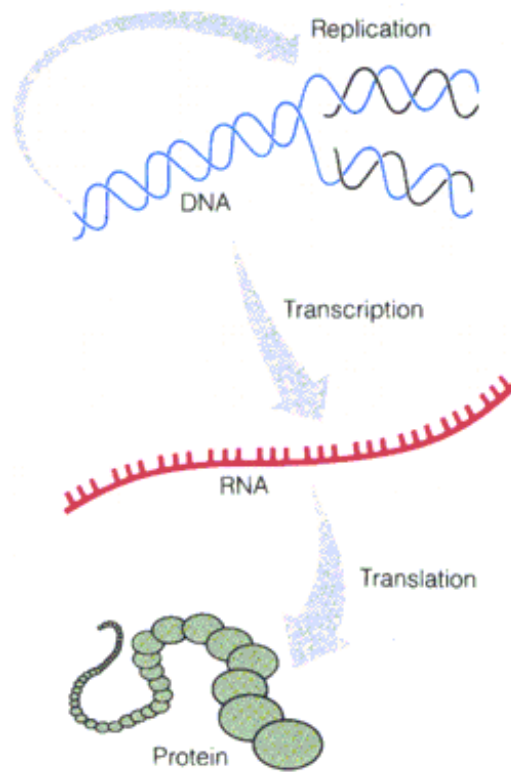
)

Central Dogma

DNA

RNA

Protein



Base Pairing is important at each step

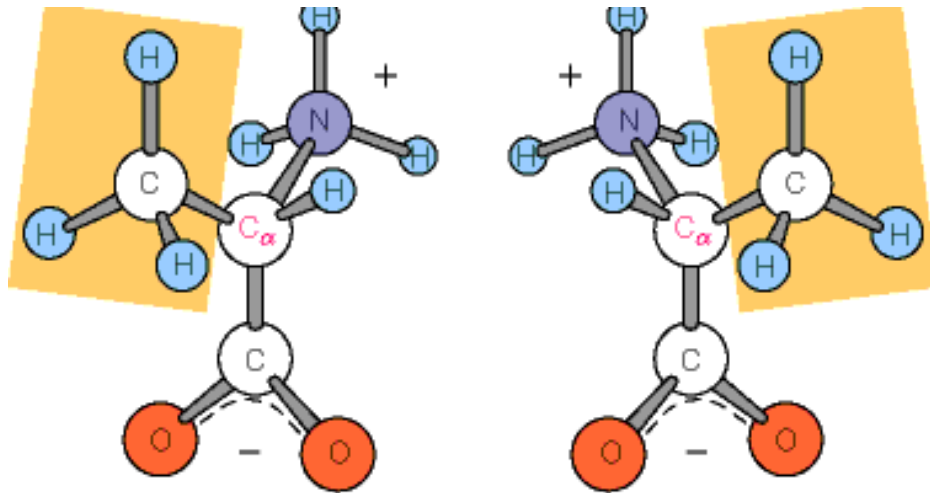
Chapter 5--Proteins

- Building Blocks --chemical functionality
- Primary Structure
 - Peptide bond
 - Metastability and hydrolysis
 - Determining and Using primary sequences

Amino Acids--20

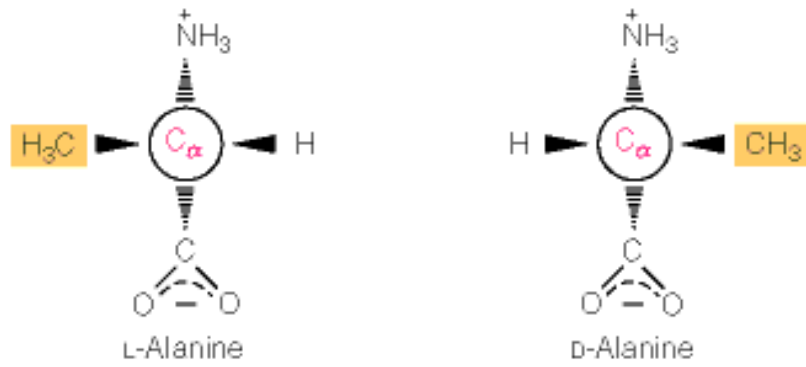
Why just L form??

- Hydrophobic
- Aliphatic
- Aromatic
- Hydrophilic
- Acidic
- Basic
- Polar



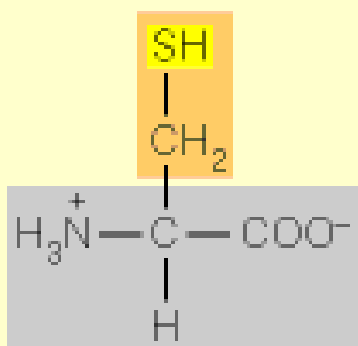
(a) L-Alanine

D-Alanine



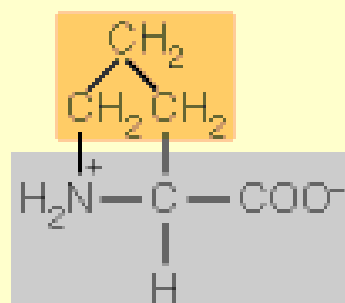
(b) L-Alanine

D-Alanine



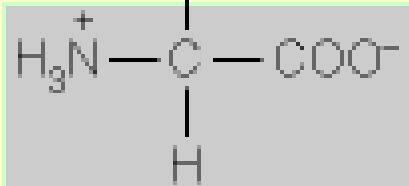
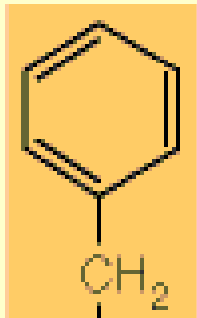
Cysteine (Cys) C

CYCLIC AMINO ACID

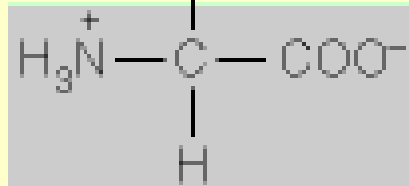
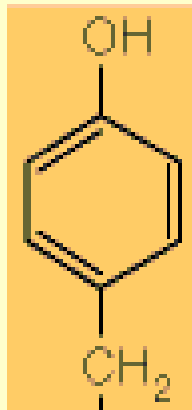


Proline (Pro) P

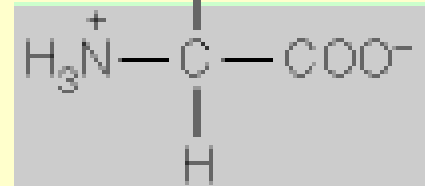
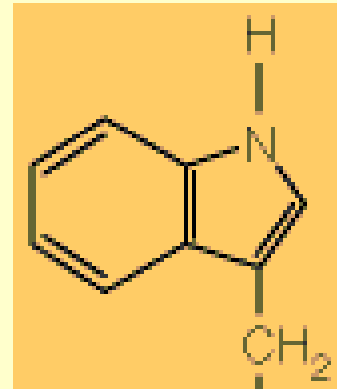
AROMATIC AMINO ACIDS



Phenylalanine (Phe) F

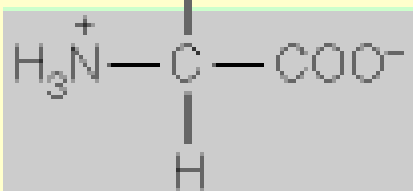
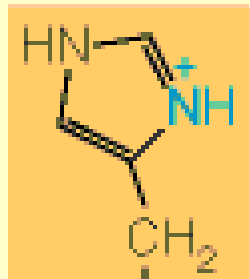


Tyrosine (Tyr) Y

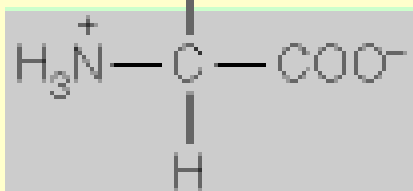
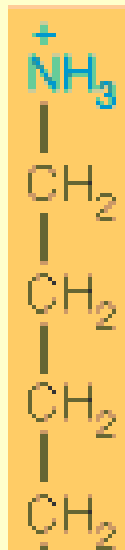


Tryptophan (Trp) W

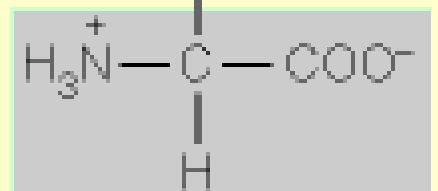
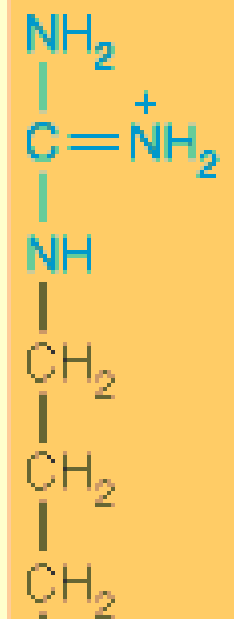
BASIC AMINO ACIDS



Histidine (His) H

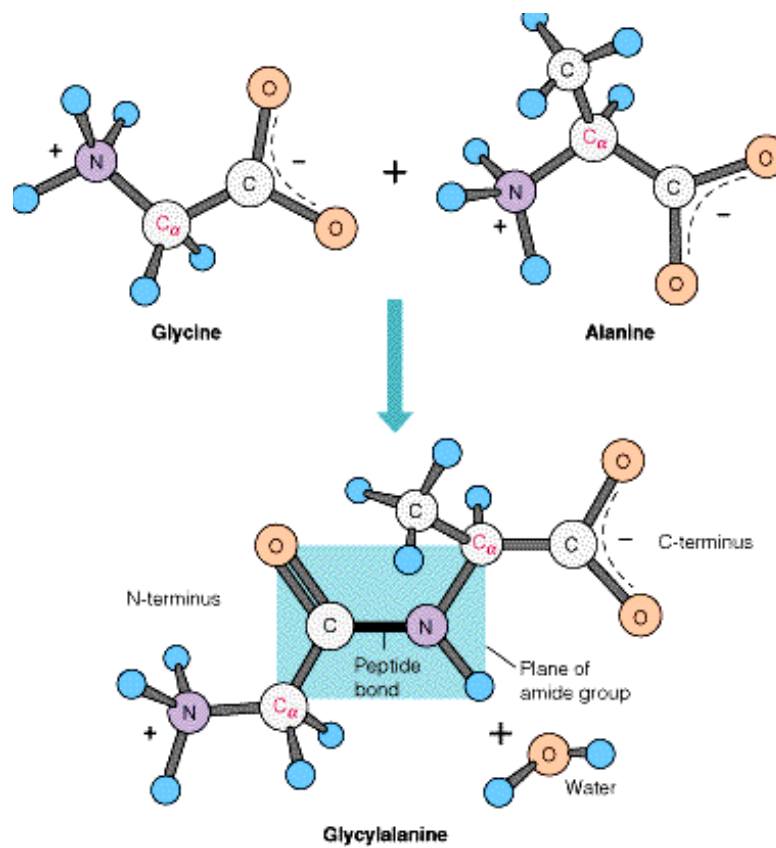


Lysine (Lys) K

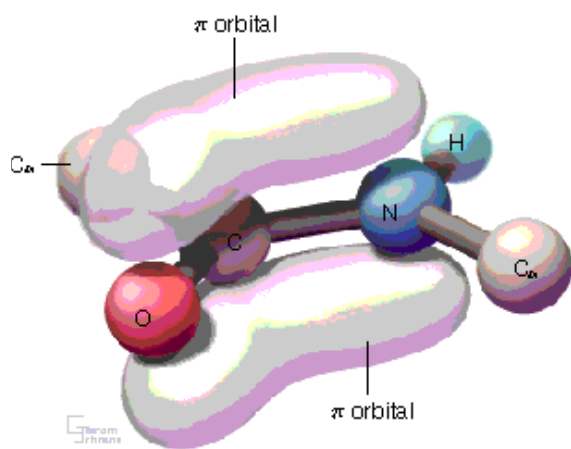


Arginine (Arg) R

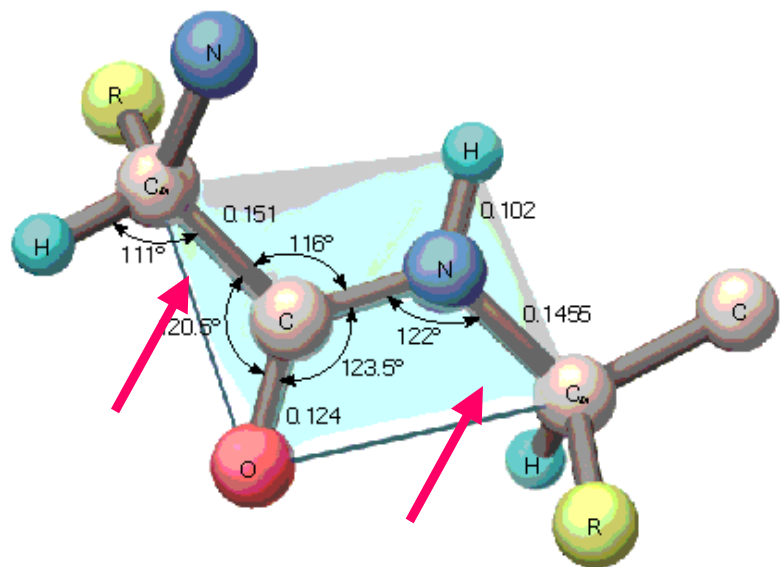
This reaction is energetically unfavorable



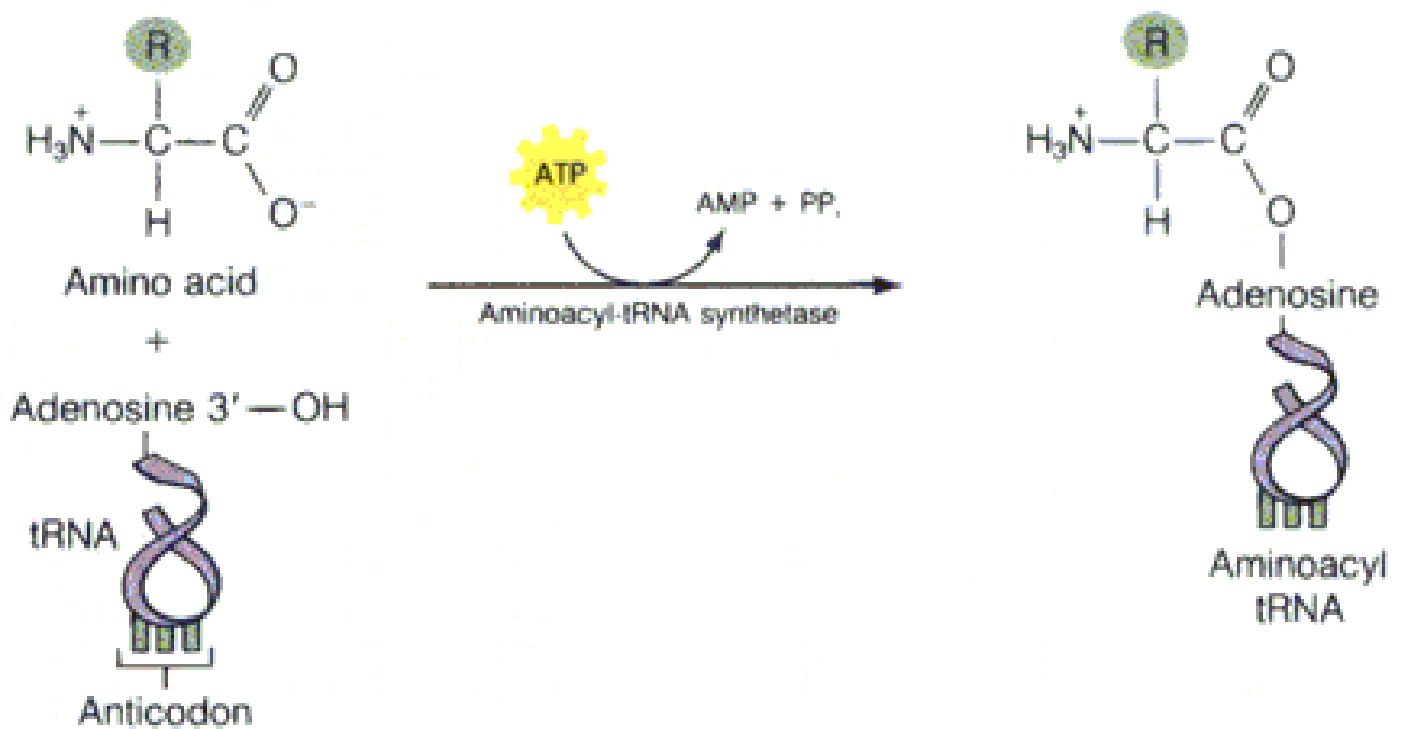
The peptide bond is planar. The range of allowable Phi and Psi angles depend on steric hindrance.



a) Partial double-bond character of peptide bond



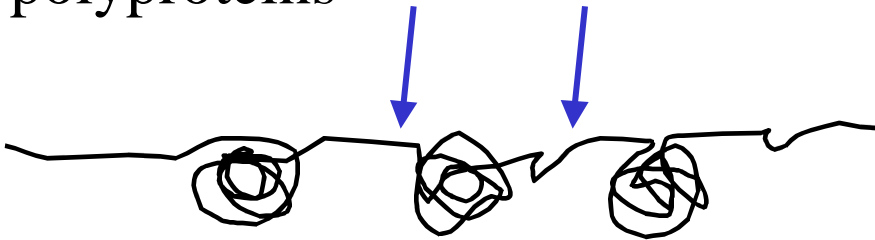
(b) Bond angles and lengths



To make peptide polymers ($G > 0$), ATP is hydrolyzed ($G < 0$)

HIV Protease

- HIV Proteins are produced as “polyproteins”

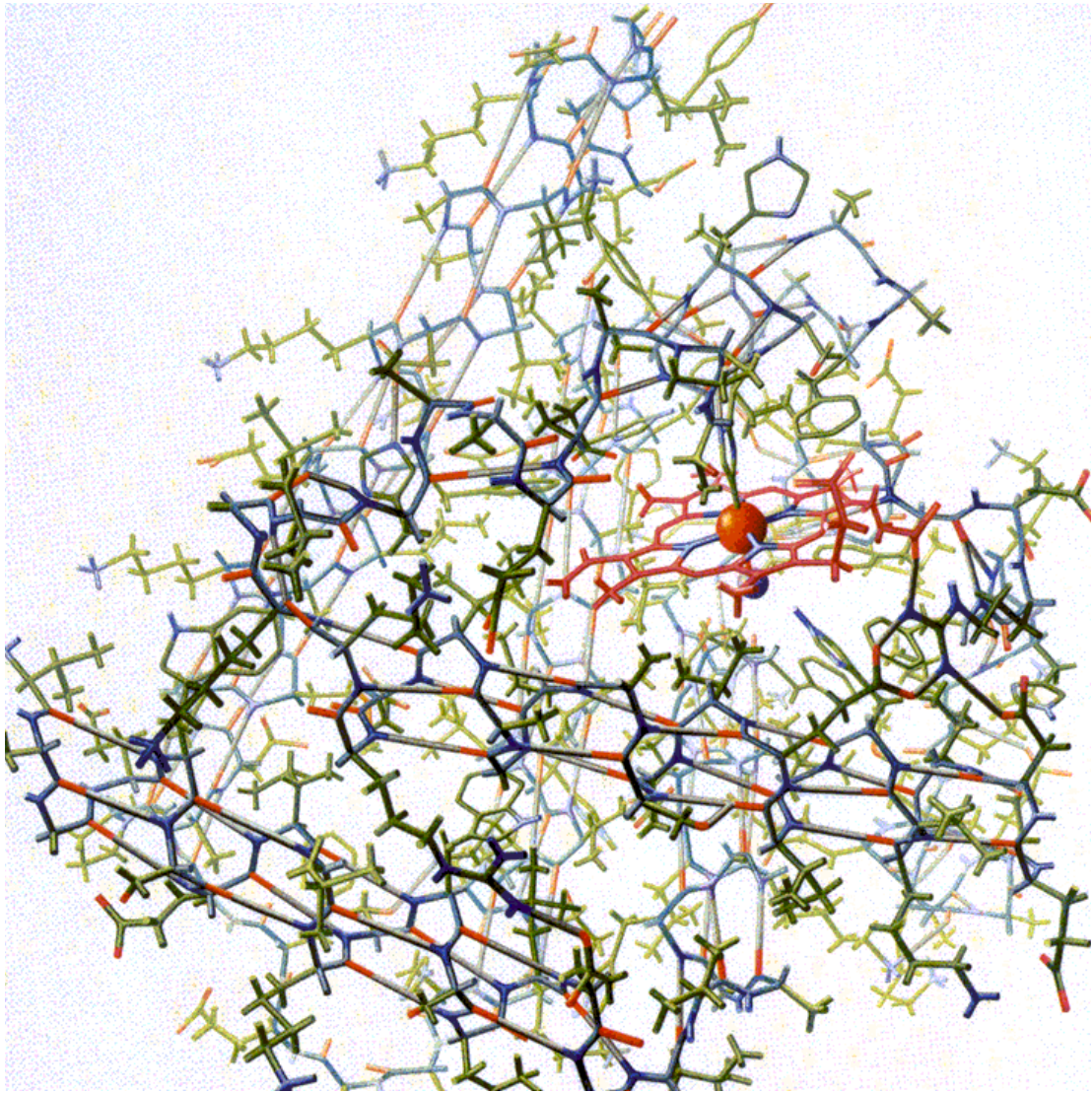


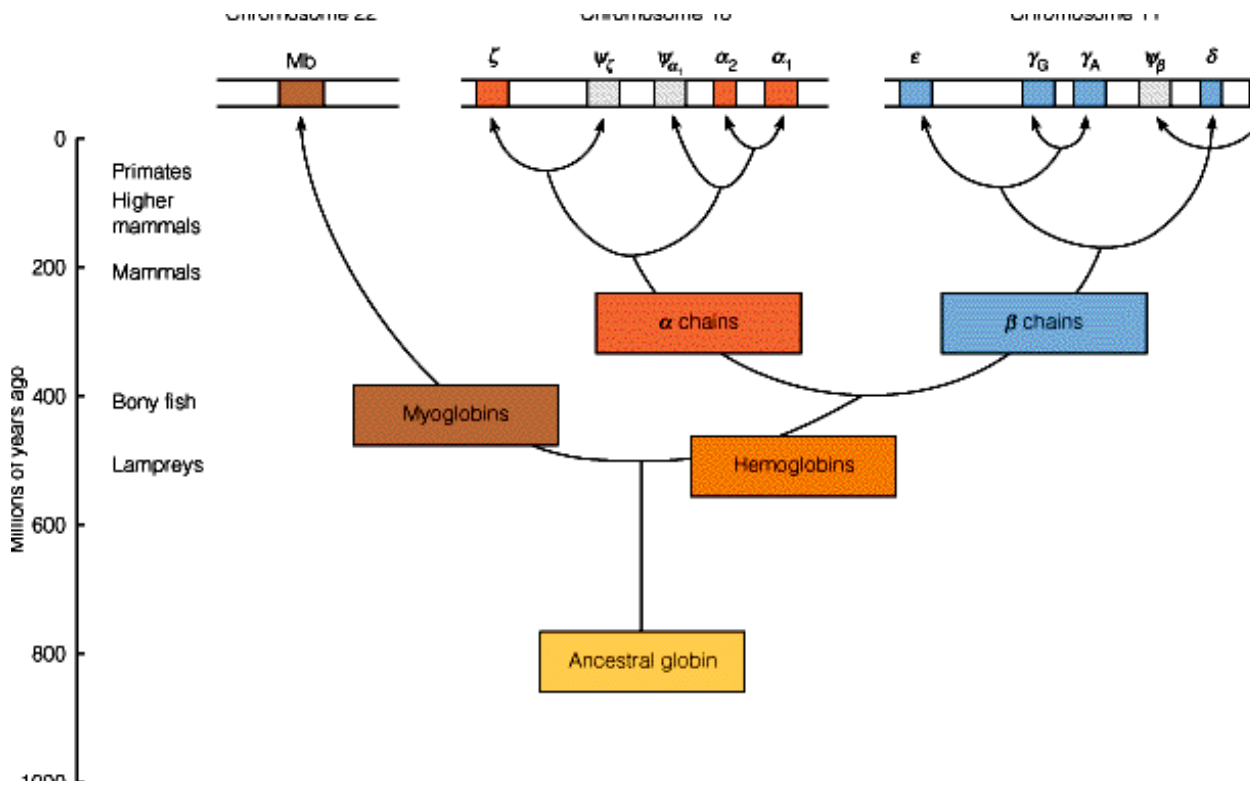
Protease Inhibitors may mimic tetrahedral transition State.

Key:

- Identical amino acids
- Conservative substitutions
- Nonconservative substitutions

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Human	G	L	S	D	G	E	W	Q	L	V	L	N	V	W	G
Whale	V	L	S	E	G	E	W	Q	L	V	L	H	V	W	A
Number	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Human	K	V	E	A	D	I	P	G	H	G	Q	E	V	L	I
Whale	K	V	E	A	D	V	A	G	H	G	Q	D	I	L	I
Number	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
Human	R	L	F	K	G	H	P	E	T	L	E	K	F	D	K
Whale	R	L	F	K	S	H	P	E	T	L	E	K	F	D	R
Number	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Human	F	K	H	L	K	S	E	D	E	M	K	A	S	E	D
Whale	F	K	H	L	K	T	E	A	E	M	K	A	S	E	D



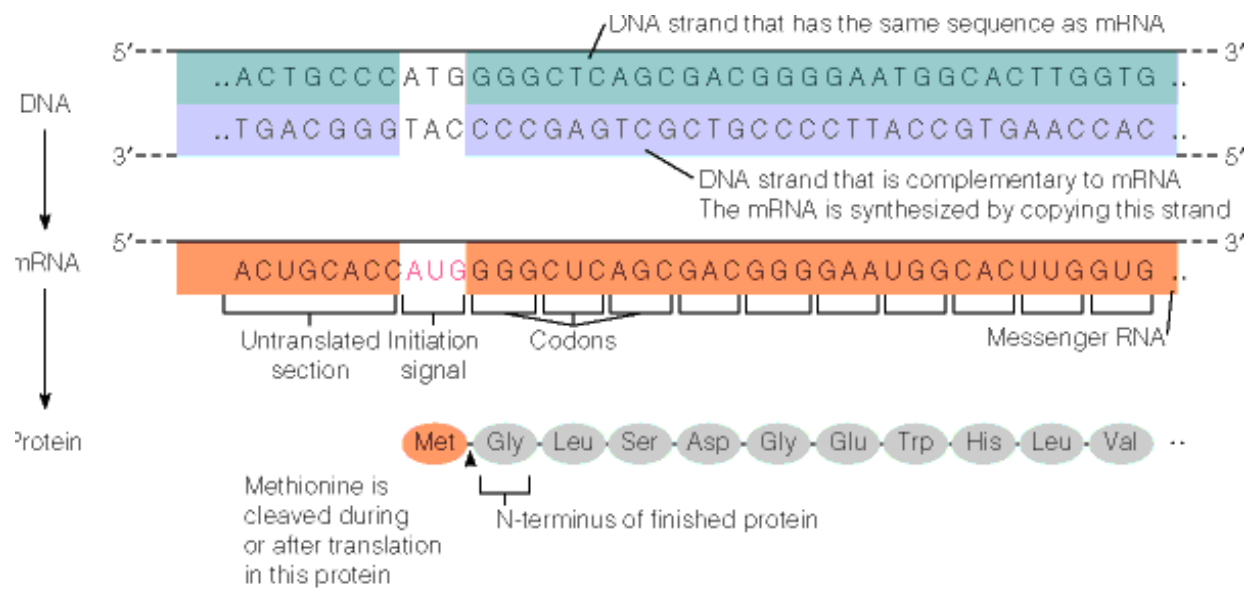


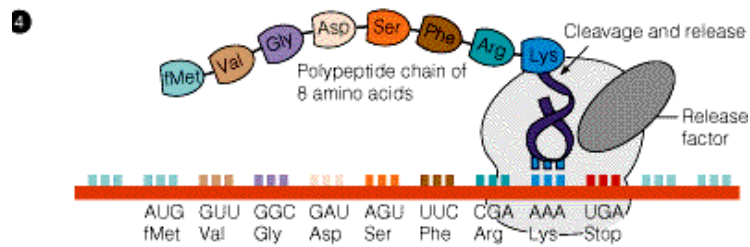
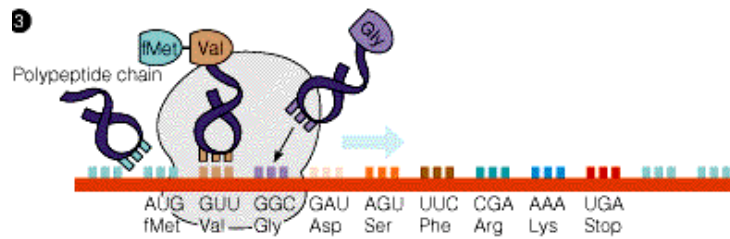
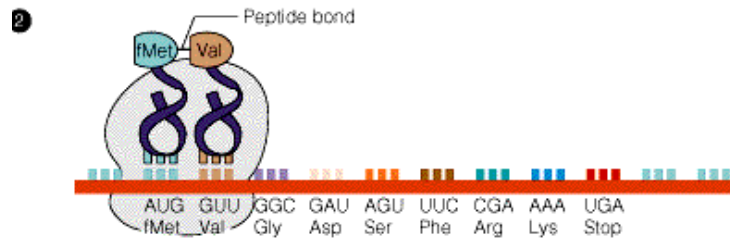
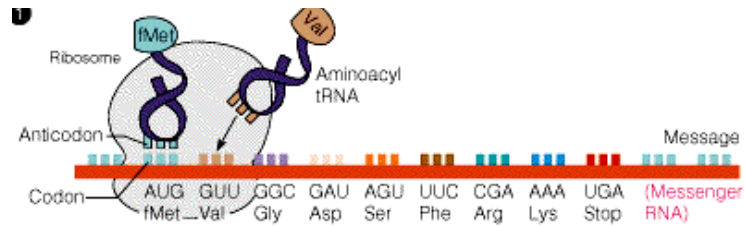
Primary Sequences can be used to
Infer evolutionary relationships among

Contemporary organisms

Individuals of the same species

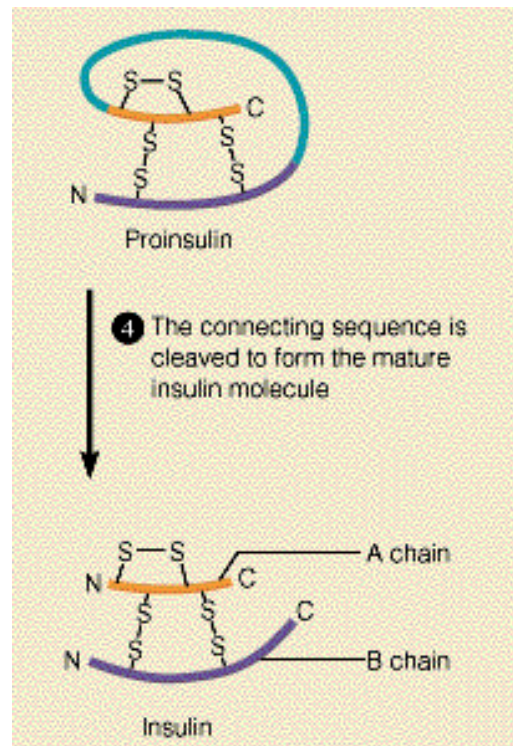
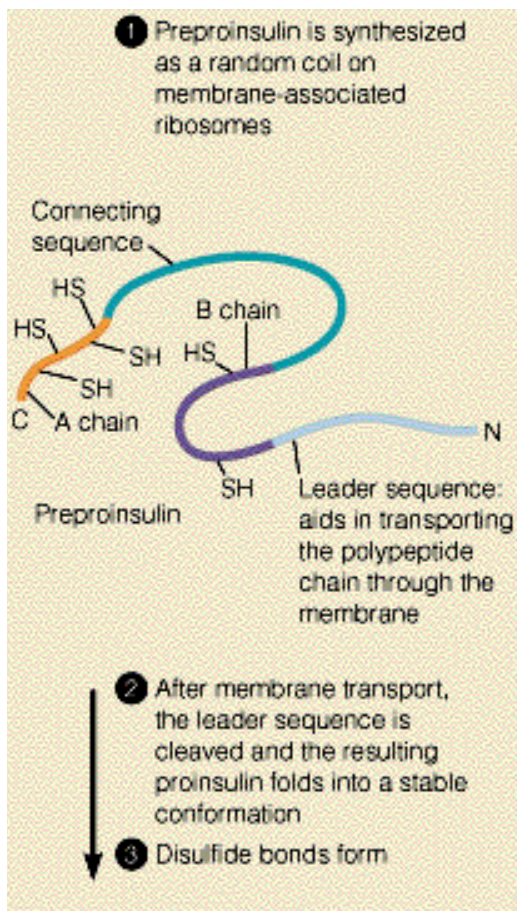
Related genes





		Second position				
		U	C	A	G	
U	UUU } Phe	UCU } Ser	UAU } Tyr	UGU } Cys	U	
	UUC } Phe		UAC } Tyr	UGC } Cys	C	
	UUA } Leu		UCA } Ser	UAA Stop	UGA Stop	A
	UUG } Leu		UCG } Ser	UAG Stop	UGG Trp	G
C	CUU } Leu	CCU } Pro	CAU } His	CGU } Arg	U	
	CUC } Leu		CCC } Pro	CAC } His	CGC } Arg	C
	CUA } Leu		CCA } Pro	CAA } Gln	CGA } Arg	A
	CUG } Leu		CCG } Pro	CAG } Gln	CGG } Arg	G
A	AUU } Ile	ACU } Thr	AAU } Asn	AGU } Ser	U	
	AUC } Ile		ACC } Thr	AAC } Asn	AGC } Ser	C
	AUA } Ile		ACA } Thr	AAA } Lys	AGA } Arg	A
	AUG Met/start		ACG } Thr	AAG } Lys	AGG } Arg	G
G	GUU } Val	GCU } Ala	GAU } Asp	GGU } Gly	U	
	GUC } Val		GCC } Ala	GAC } Asp	GGC } Gly	C
	GUA } Val		GCA } Ala	GAA } Glu	GGA } Gly	A
	GUG } Val		GCG } Ala	GAG } Glu	GGG } Gly	G

Processing and targeting of Eukaryotic protein

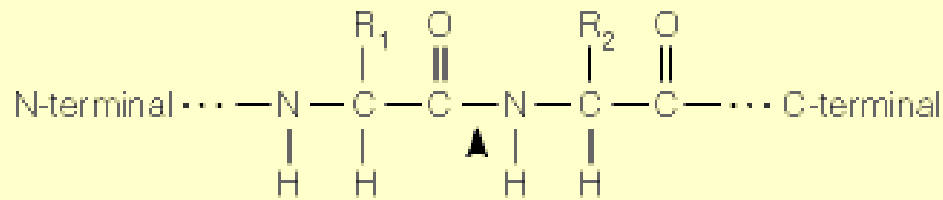


Sequencing proteins (traditional)

1. Cleave into small peptides
2. Use Edman Degradation to modify/cleave N ter amino acid.
3. Identify by HPLC vs known aa's
4. Repeat (robots do this)

Modern

1. Base pair DNA oligomer to start of sequence.
2. Use NTPs, DNA polymerase, and ddNTPs
To produce "truncated" DNAs.
3. Read sequence by gel electrophoresis.



Enzyme	Preferred Site ^a	Source
Trypsin	R ₁ = Lys, Arg	From digestive systems of animals, many other sources
Chymotrypsin	R ₁ = Tyr, Trp, Phe, Leu	Same as trypsin
Thrombin	R ₁ = Arg	From blood; involved in coagulation
V-8 protease	R ₁ = Asp, Glu	From <i>Staphylococcus aureus</i>
Prolyl endopeptidase	R ₁ = Pro	Lamb kidney, other tissues
Subtilisin	Very little specificity	From various bacilli
Carboxypeptidase A	R ₂ = C-terminal amino acid	From digestive systems of animals
Thermolysin	R ₂ = Leu, Val, Ile, Met	From <i>Bacillus thermoproteolyticus</i>

^aThe residues indicated are those next to which cleavage is most likely. Note that in some cases preference is determined by the residue on the N-terminal side of the cleaved bond (R₁) and sometimes by the residue to the C-terminal side (R₂). Generally, proteases do not cleave where proline is on the other side of the bond. Even prolyl endopeptidase will not cleave if R₂ = Pro.

Sequence and Structural Information

Available in Public Databases

*****ENTREZ/BLAST/
National Library of Medicine**

*****Protein Database (“pdb lite”)
Rutgers Center for Structural Biology**