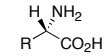
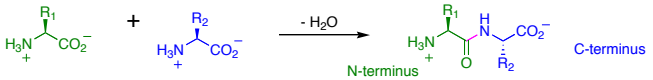
Chapter 25: Amino Acids, Peptides, and Proteins.
monomer unit: α-amino acids

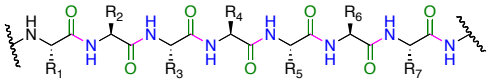


R = sidechain

α- Amino Acid

Biopolymer: the monomeric amino acids are linked through an amide bond (the carboxylic acids of one AA with the α-amino group of a second)



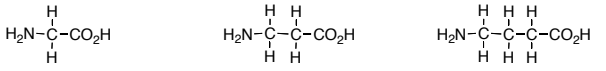


Peptide or protein (polypeptide)

peptide (< 50 amino acids)
protein (> 50 amino acids)

323

25.1: Classification of Amino Acids. AA's are classified according to the location of the amino group.



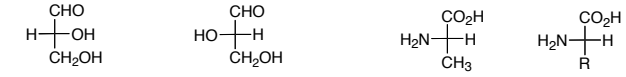
α-amino acid (2-amino carboxylic acid)

β-amino acid (3-amino carboxylic acid)

γ-amino acid (4-amino carboxylic acid)

There are 20 genetically encoded α-amino acids found in peptides and proteins (Table 20.1, p. 1032-3)

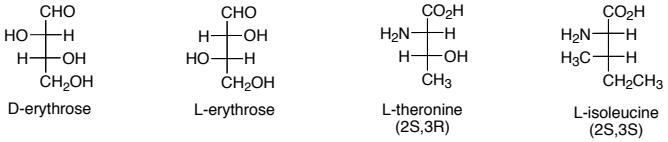
19 are primary amines, 1 (proline) is a secondary amine
19 are "chiral", 1 (glycine) is achiral; the natural configuration of the α-carbon is L. (Ch. 25.2)



D-glyceraldehyde

L-glyceraldehyde

L-alanine



D-erythrose

L-erythrose

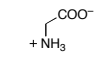
L-theronine (2S,3R)

L-isoleucine (2S,3S)

324

α-Amino acids are classified by the properties of their sidechains.

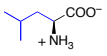
Nonpolar (hydrophobic):



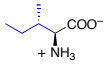
Glycine *(Gly, G)*

(S)-(+)-Alanine *(Ala, A)*

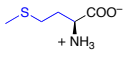
(S)-(+)-Valine *(Val, V)*



(S)-(–)-Leucine *(Leu, L)*

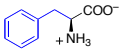


(2S,3S)-(+)-Isoleucine *(Ile, I)*

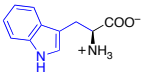


(S)-(–)-Methionine *(Met, M)*

(S)-(–)-Proline *(Pro, P)*

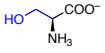


(S)-(–)-Phenylalanine *(Phe, F)*



(S)-(–)-Tryptophan *(Trp, W)*

Polar but non-ionizable:

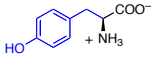


(S)-(–)-Serine *(Ser, S)*

pKa ~ 13

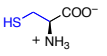
(2S,3R)-(–)-Threonine *(Thr, T)*

pKa ~ 13



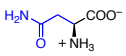
(S)-(–)-Tyrosine *(Tyr, Y)*

pKa ~ 10.1

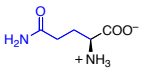


(R)-(–)-Cysteine *(Cys, C)*

pKa ~ 8.2



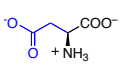
(S)-(–)-Asparagine *(Asn, N)*



(S)-(+)-Glutamine *(Gln, Q)*

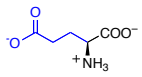
325

Acidic:



(S)-(+)-Aspartic Acid *(Asp, D)*

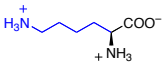
pKa ~ 3.6



(S)-(+)-Glutamic Acid *(Glu, E)*

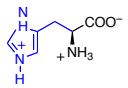
pKa ~ 4.2

Basic:



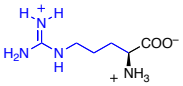
(S)-(+)-Lysine *(Lys, K)*

pKa ~ 10.5



(S)-(–)-Histidine *(His, H)*

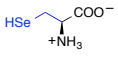
pKa ~ 6.0



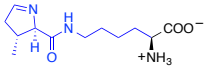
(S)-(+)-Arginine *(Arg, R)*

pKa ~ 12.5

Non-standard (but genetically encoded):



(*R*)-Selenocysteine (*Sec, U*)

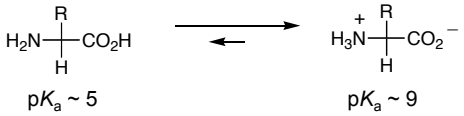


Pyrrolysine (*Pyl, O*)

25.2: Stereochemistry of Amino Acids: The natural configuration of the α-carbon is L. D-Amino acids are found in the cell walls of bacteria. The D-amino acids are not genetically encoded, but derived from the epimerization of L-isomers (Ch. 25.6).

326

25.3: Acid-Base Behavior of Amino Acids. Amino acids exist as a zwitterion: a dipolar ion having both a formal positive and formal negative charge (overall charge neutral).



Amino acids are *amphoteric*: they can react as either an acid or a base. Ammonium ion acts as an acid, the carboxylate as a base.

Isoelectric point (pI): The pH at which the amino acid exists largely in a neutral, zwitterionic form (influenced by the nature of the sidechain)

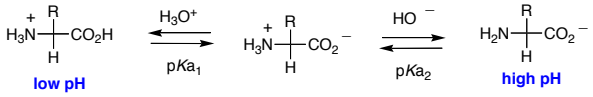


Table 25.2 (p. 1037) & 25.3 (p. 1038)

327

Bronsted Acid: proton donor (H^+)

weak acids (and bases) do not fully dissociate

$$H\text{-}A + H_2O \rightleftharpoons H_3O^+ + A^-$$

$$K_a = \frac{[H_3O^+]\,[A^-]}{[H\text{-}A]} \quad \textit{acid dissociation constant}$$

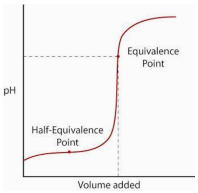
$$pK_a = -\log K_a$$

$$pH = -\log [H^+]$$

Henderson-Hasselbalch Equation: Relates pK_a with pH

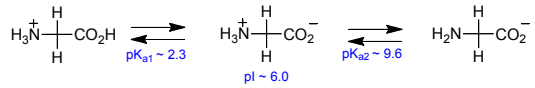
$$pH = pK_a + \log \frac{[A^-]}{[H\text{-}A]}$$

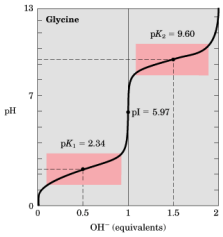
when $[A^-] = [H\text{-}A]$, the pH = pK_a

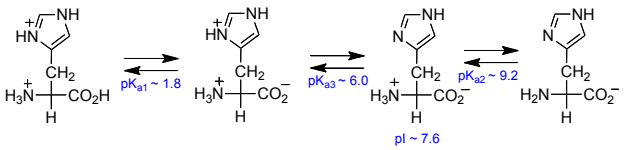


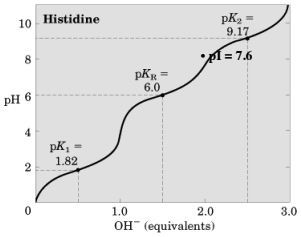
For a review see: http://themedicalbiochemistrypage.org/

328



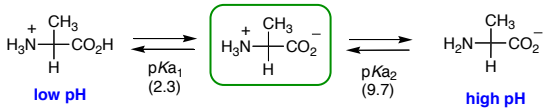






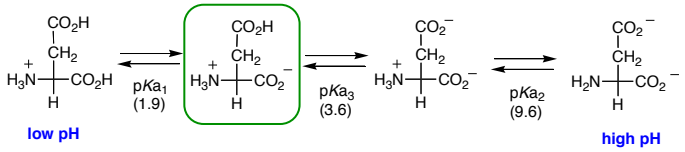
329

$$pI = \frac{pKa_x + pKa_y}{2}$$



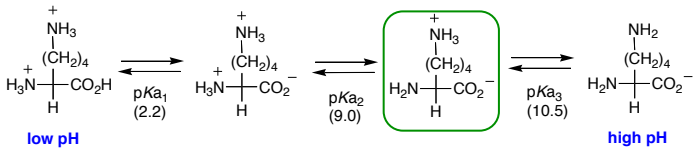
$$pI = \frac{pKa_1 + pKa_2}{2}$$

$$pI = 6.0$$



$$pI = \frac{pKa_1 + pKa_3}{2}$$

$$pI = 2.7$$

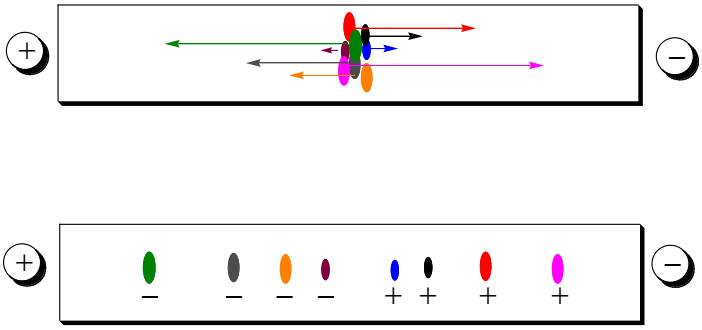


$$pI = \frac{pKa_2 + pKa_3}{2}$$

$$pI = 9.7$$

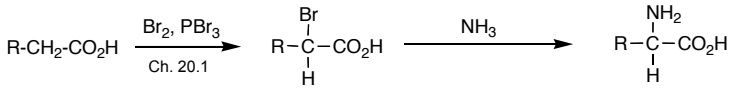
330

Electrophoresis: separation of polar compounds based on their mobility through a solid support. The separation is based on charge (pI) or molecular mass. (Fig. 25.2, p. 1039)

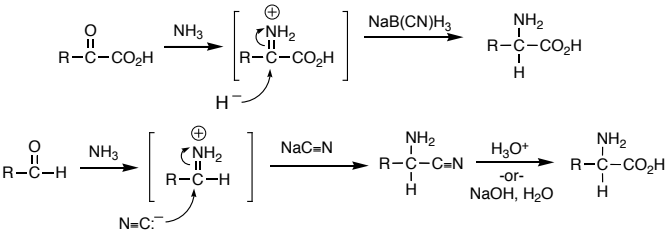


331

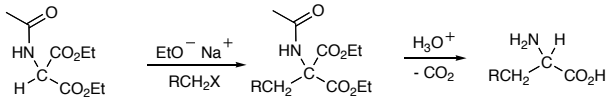
25.4: Synthesis of Amino Acids:



Strecker Synthesis: recall reductive amination (Ch. 21.10) and cyanohydrin formation (Ch. 17.7)

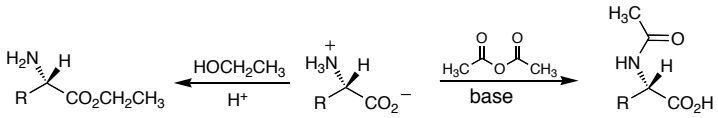


Amidomalonate Synthesis: recall the malonic acid synthesis (Ch. 20.6)

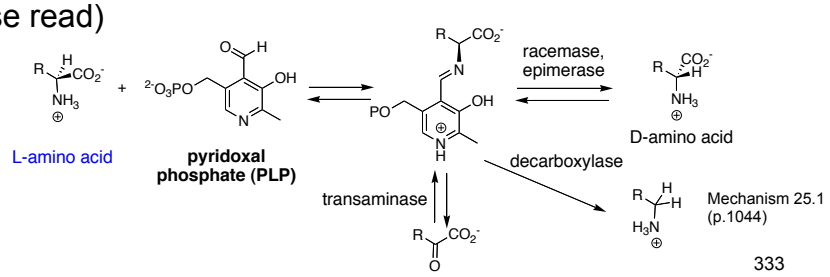


332

25.5: Reactions of Amino Acids. Amino acids will undergo reactions characteristic of the amino (amide formation) and carboxylic acid (ester formation) groups.



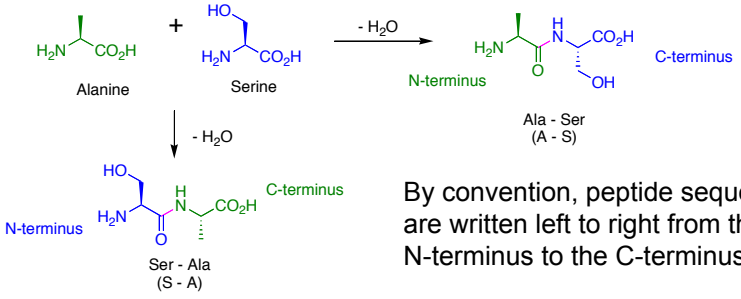
25.6: Some Biochemical Reactions of Amino Acids. Many enzymes involved in amino acid biosynthesis, metabolism and catabolism are pyridoxal phosphate (vitamin B_6, PLP) dependent. (please read)



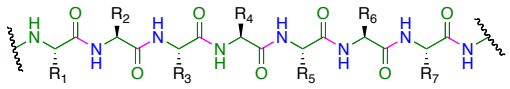
Mechanism 25.2 (p.1047-8)

333

25.7: Peptides. Proteins and peptides are polymers made up of amino acid units (residues) that are linked together through the formation of amide bonds (peptide bonds) from the amino group of one residue and the carboxylate of a second residue



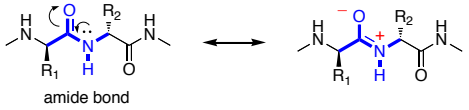
By convention, peptide sequences are written left to right from the N-terminus to the C-terminus



backbone

334

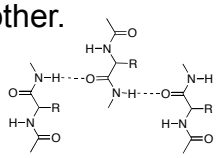
The amide (peptide) bond has C=N double bond character due to resonance resulting in a planar geometry



amide bond

restricts rotations
resistant to hydrolysis

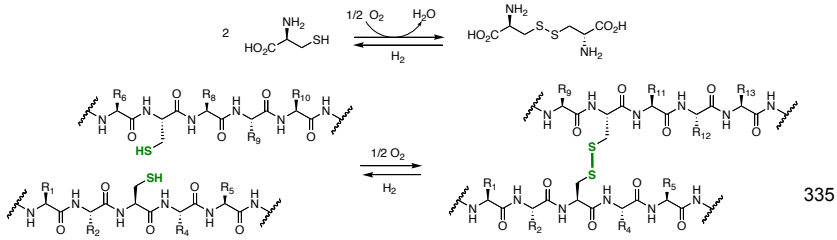
The N-H bond of one amide linkage can form a hydrogen bond with the C=O of another.



N-O distance 2.85 - 3.20 Å

optimal N-H-O angle is 180 °

Disulfide bonds: the thiol groups of cysteine can be oxidized to form disulfides (Cys-S-S-Cys) (Ch. 15.2)



335

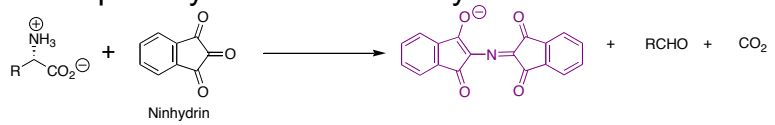
25.8: Introduction to Peptide Structure Determination.
Protein Structure:
primary (1°) structure: the amino acid sequence
secondary (2°): frequently occurring substructures or folds
tertiary (3°): three-dimensional arrangement of all atoms in a single polypeptide chain
quaternary (4°): overall organization of non-covalently linked subunits of a functional protein.

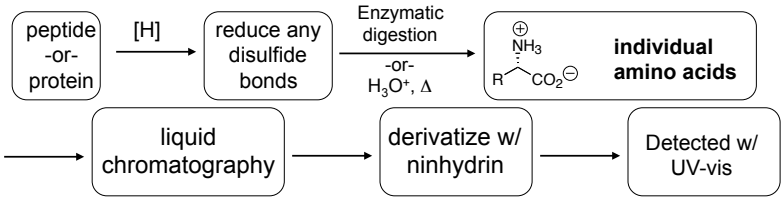
1. Determine the amino acids present and their relative ratios
2. Cleave the peptide or protein into smaller peptide fragments and determine their sequences
3. Cleave the peptide or protein by another method and determine their sequences. Align the sequences of the peptide fragments from the two methods

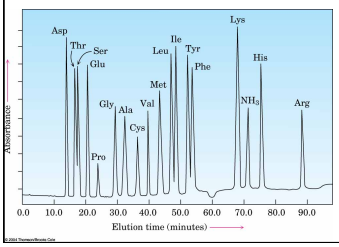
336

25.9: Amino Acid Analysis. automated method to determine the amino acid content of a peptide or protein

Reaction of primary amines with ninhydrin







Different amino acids have different chromatographic mobilities (retention times)

1972 Nobel Prize in Chemistry
William Stein
Stanford Moore

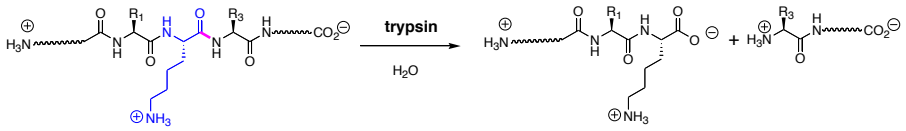
337

25.10: Partial Hydrolysis and End Group Analysis. Acidic hydrolysis of peptides cleave the amide bonds indiscriminately.

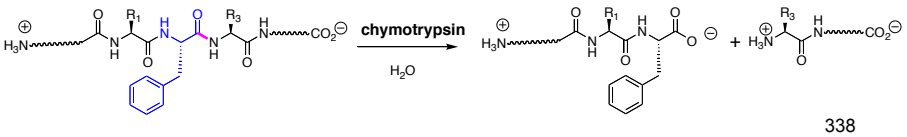
Proteases (peptidases): Enzymes that catalyzed the hydrolysis of the amide bonds of peptides and proteins.

Enzymatic cleavage of peptides and proteins at defined sites:

- *trypsin*: cleaves at the C-terminal side of basic residues,
  Arg, Lys but not His



- *chymotrypsin*: cleaves at the C-terminal side of aromatic residues
  Phe, Tyr, Trp



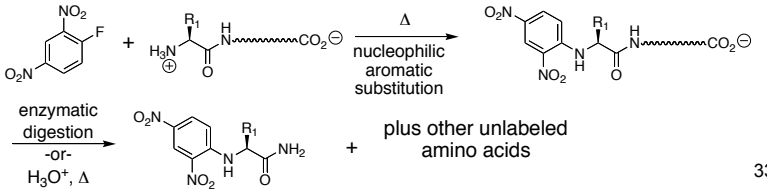
338

Trypsin and chymotrypsin are endopeptidases

Carboxypeptidase: Cleaves the amide bond of the C-terminal amino acid (exopeptidase)

End Group Analysis. The C-terminal AA is identified by treating with peptide with carboxypeptidase, then analyzing by liquid chormatography (AA Analysis).

N-labeling: The peptide is first treated with 1-fluoro-2,4-dinitro-benzene (Sanger's reagent), which selectively reacts with the N-terminal amino group. The peptide is then hydrolyzed to their amino acids and the N-terminal amino acid identified as its N-(2,4-dinitrophenyl) derivative (DNP).



339

Trypsin:

E-A-Y-L-V-C-G-E-R
F-V-N-Q-H-L-F-S-H-L-K
G-C-F-L-P-K
L-G-A

Chymotrypsin:

F-V-N-Q-H-L-F
S-H-L-K-E-A-Y
L-V-C-G-E-R-G-C-F
L-P-K-L-G-A

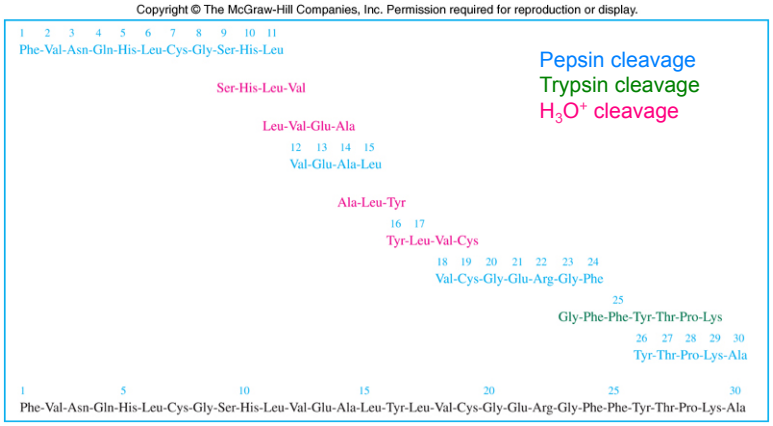
F-V-N-Q-H-L-F
*F-V-N-Q-H-L-F-*S-H-L-K
S-H-L-K-E-A-Y
E-A-Y-L-V-C-G-E-R
L-V-C-G-E-R-G-C-F
G-C-F-L-P-K
L-P-K-L-G-A
L-G-A

F-V-N-Q-H-L-F-S-H-L-K-E-A-Y-L-V-C-G-E-R-G-C-F-L-P-K-L-G-A

340

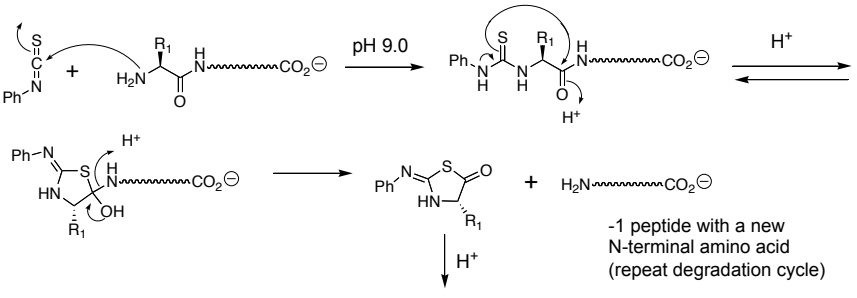
25.11: Insulin. Insulin has two peptide chains (the A chain has 21 amino acids and the B chain has 30 amino acids) held together by two disulfide linkages. (please read)

Pepsin: cleaves at the C-terminal side of Phe, Tyr, Leu; but not at Val or Ala.



341

25.12: The Edman Degradation and Automated Peptide Sequencing. Chemical method for the sequential cleavage and identification of the amino acids of a peptide, one at a time starting from the N-terminus. (Mechanism 25.3, p. 1057)
Reagent: Ph-N=C=S, phenylisothiocyanate (***PITC***)



N-phenylthiohydantoin:
separated by liquid chromatography (based of the R group) and detected by UV-vis

342

Peptide sequencing by Edman degradation:

- Cycle the pH to control the cleavage of the N-terminal amino acid by PITC.
- Monitor the appearance of the new *N*-phenylthiohydantoin for each cycle.
- Good for peptides up to ~ 25 amino acids long.
- Longer peptides and proteins must be cut into smaller fragments before Edman sequencing.

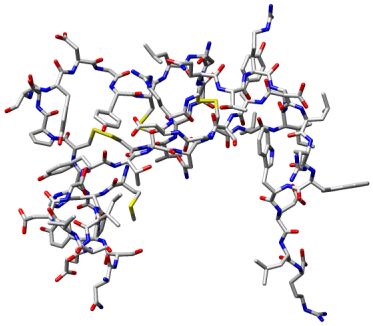
Tandem mass spectrometry has largely replaced Edman degradation for peptide sequencing (Fig. 25.9, p. 1058)

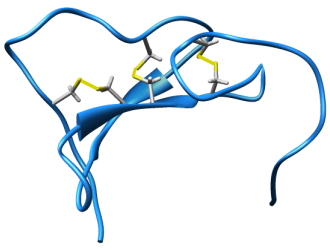
343

Epidermal Growth Factor (EGF): the miracle of mother's spit
53 amino acid, 3 disulfide linkages

H_2N–ASN1•SER2•TYR3•PRO4•GLY5•CYS6•PRO7•SER8•SER9•**TYR10**•
ASP11•GLY12•**TYR13**•CYS14•LEU15•ASN16•GLY17•GLY18•VAL19•
CYS20•**MET21**•HIS22•ILE23•GLU24•SER25•LEU26•ASP27•SER28•
TYR29•THR30•CYS31•ASN32•CYS33•VAL34•ILE35•GLY36•**TYR37**•
SER38•GLY39•ASP40•**ARG41**•CYS42•GLN43•THR44•**ARG45**•ASP46•
LEU47•**ARG48**•TRP49•TRP50•GLU51•LEU52•ARG53–CO_2H

Trypsin **Chymotrypsin** **Cyanogen Bromide**



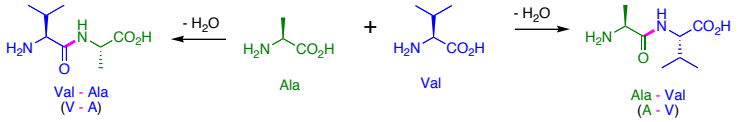


1986 Nobel Prize in Medicine or Physiology :
Stanley Cohen & Rita Levi-Montalcini

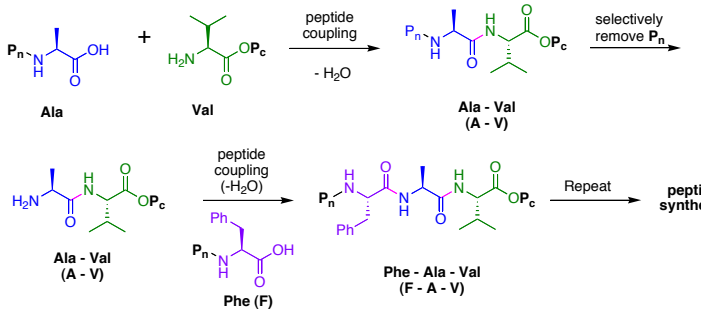
344

25.13: The Strategy for Peptide Synthesis:

Chemical synthesis of peptide:
1. Solution phase synthesis
2. Solid-phase synthesis



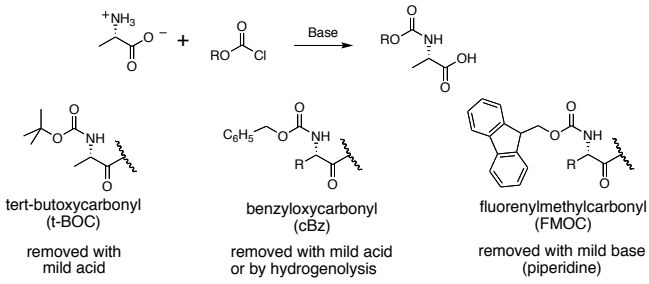
The need for protecting groups



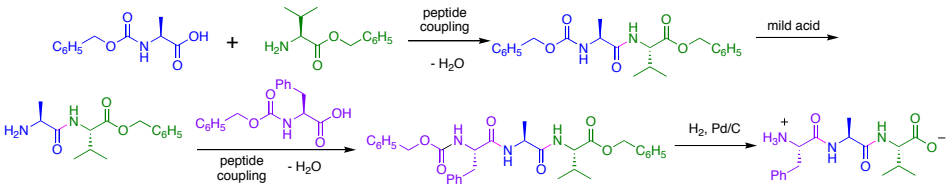
Orthogonal protecting group strategy: the carboxylate protecting group must be stable to the reaction conditions for the removal of the α-amino protecting group and (*vice versa*)

345

25.14: Amino and Carboxyl Group Protection and Deprotection. The α-amino group is protected as a carbamate.

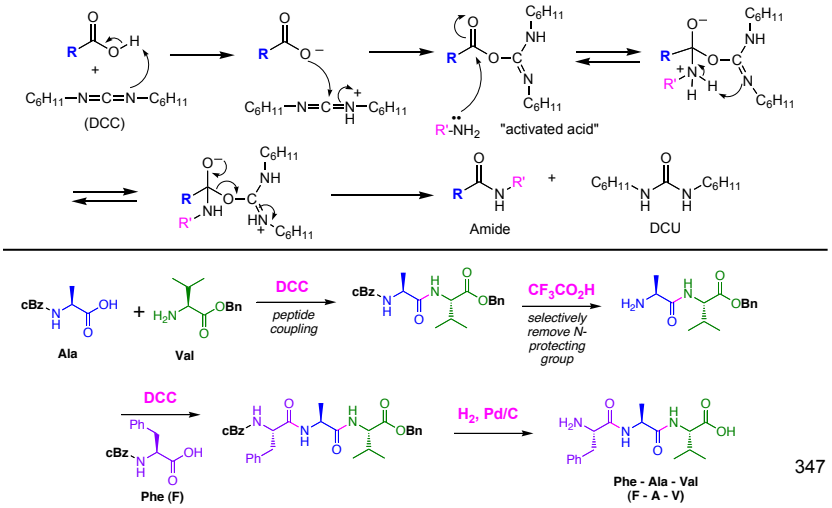


Protected as a benzyl ester; removed by hydrogenolysis



346

25.15: Peptide Bond Formation. Amide formation from the reaction of an amine with a carboxylic acid is slow. Amide bond formation (peptide coupling) can be accelerated if the carboxylic acid is activated. *Reagent: dicyclohexylcarbodiimide (**DCC**)* (Mechanism 25.4, p. 1063)



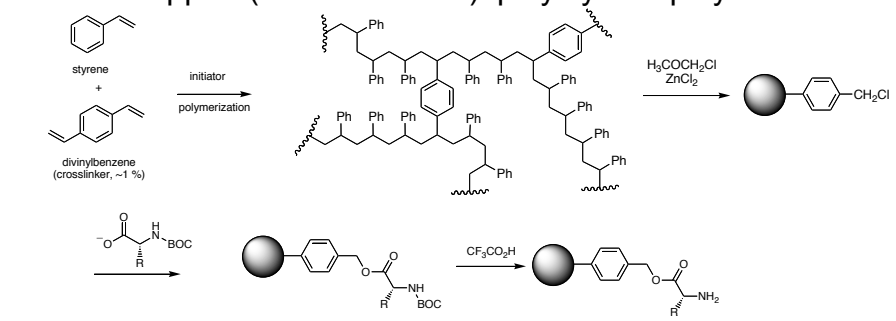
347

- In order to practically synthesize peptides and proteins, time consuming purifications steps must be avoided until the very end of the synthesis.
- Large excesses of reagents are used to drive reactions forward and accelerate the rate of reactions.
- How are the excess reagents and by-products from the reaction, which will interfere with subsequent coupling steps, removed without a purification step?

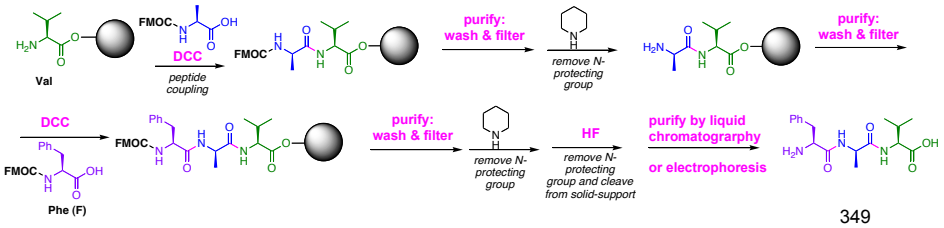
25.16: Solid-Phase Peptide Synthesis: The Merrifield Method. Peptides and proteins up to ~ 100 residues long are synthesized on a solid, insoluble, polymer support. Purification is conveniently accomplished after each step by a simple wash and filtration.

348

The solid support (Merrifield resin): polystyrene polymer



Solid-phase peptide synthesis

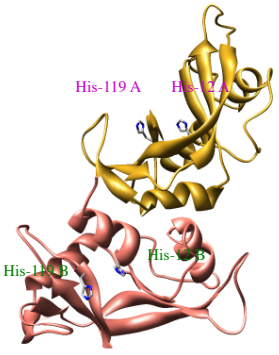


349

Ribonuclease A- 124 amino acids, catalyzes the hydrolysis of RNA
Solid-phase synthesis of RNase A:

Synthetic RNase A: 78 % activity
0.4 mg was synthesized
2.9 % overall yield
average yield ~ 97% per coupling step

```
LYS GLU THR ALA ALA ALA LYS PHE GLU ARG
GLN HIS MET ASP SER SER THR SER ALA ALA
SER SER SER ASN TYR CYS ASN GLN MET MET
LYS SER ARG ASN LEU THR LYS ASP ARG CYS
LYS PRO VAL ASN THR PHE VAL HIS GLU SER
LEU ALA ASP VAL GLN ALA VAL CYS SER GLN
LYS ASN VAL ALA CYS LYS ASN GLY GLN THR
ASN CYS TYR GLN SER TYR SER THR MET SER
ILE THR ASP CYS ARG GLU THR GLY SER SER
LYS TYR PRO ASN CYS ALA TYR LYS THR THR
GLN ALA ASN LYS HIS ILE ILE VAL ALA CYS
GLU GLY ASN PRO TYR VAL PRO VAL HIS PHE
ASP ALA SER VAL
```

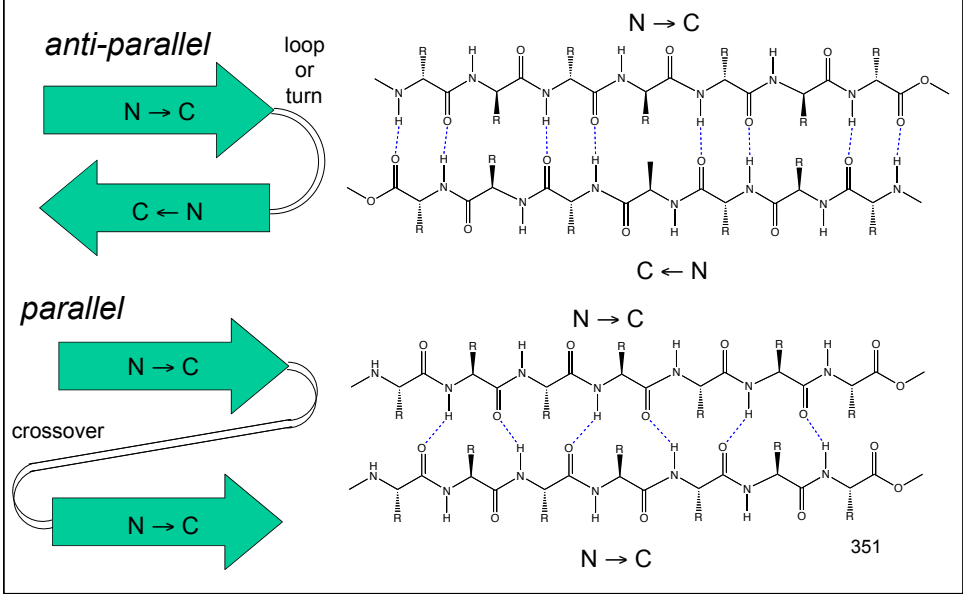


pdb code: 1AFL

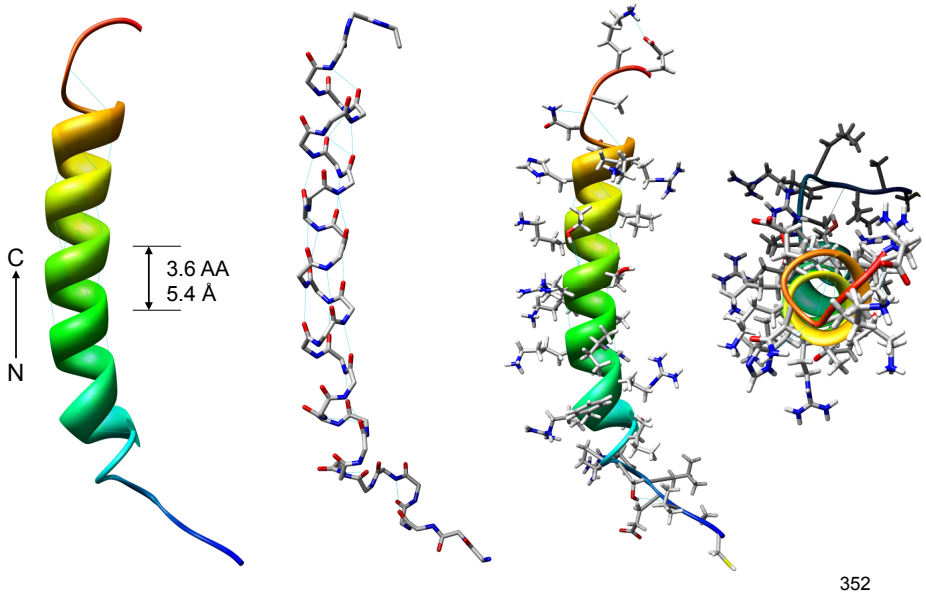
R. Bruce Merrifield, Rockefeller University, 1984 Nobel Prize in Chemistry:
"for his development of methodology for chemical synthesis on a solid matrix."

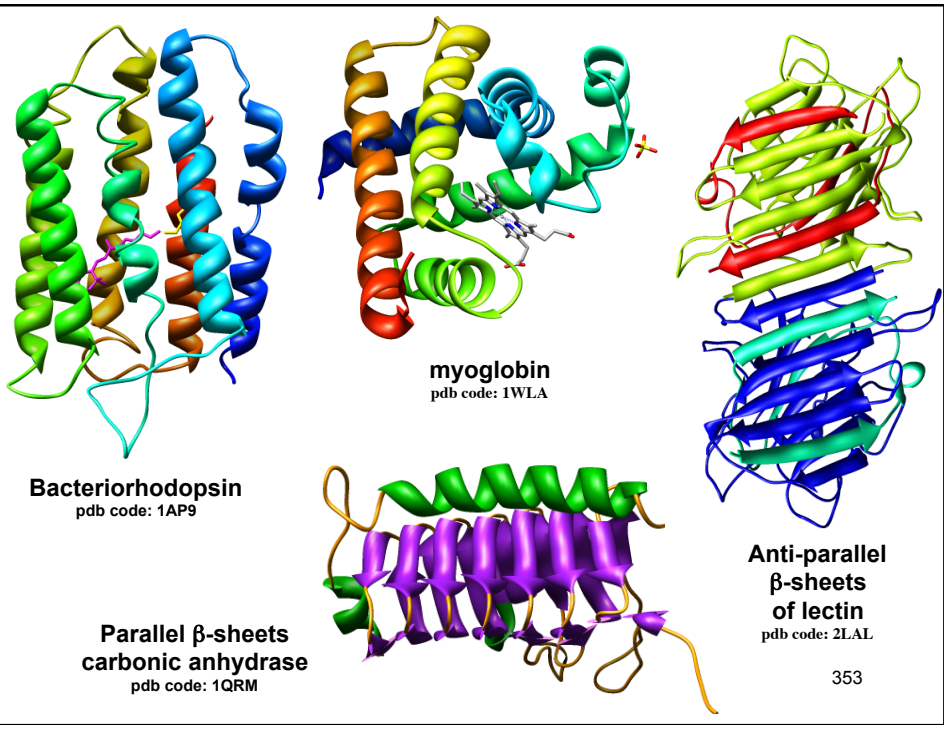
350

25.17: Secondary Structures of Peptides and Proteins.
β-sheet: Two or more extended peptide chain, in which the amide backbones are associated by hydrogen bonded



α-helix: 3.6 amino acids per coil, 5.4 Å





353

25.18: Tertiary Structure of Polypeptides and Proteins.

Fibrous. Polypeptides strands that "bundle" to form elongated fibrous assemblies; insoluble.

Globular. Proteins that fold into a "spherical" conformation.

Factors that contribute to 3° structure (Table 25.4, p. 1070)

Disulfide bonds between cysteine

Hydrogen bonds

Salt bridges (ion pairs)

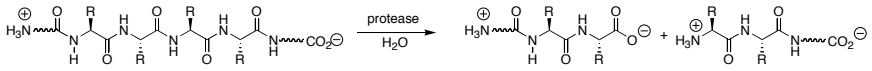
Van der Waal forces (hydrophobic effect). Proteins will fold so that *hydrophobic* amino acids are on the inside (shielded from water) and *hydrophilic* amino acids are on the outside (exposed to water)

354

Enzymes: proteins that catalyze biochemical reactions.

- by bringing the reactive atoms together in the optimal geometry for the reaction.
- lowering the activation energy ($\Delta G^{\ddagger}$) by stabilizing the transition state and/or high energy intermediate.
- many enzymes use the functional groups of the amino acid sidechain to carry out the reactions

Proteases (peptidases): catalyzes the hydrolysis of peptide bonds

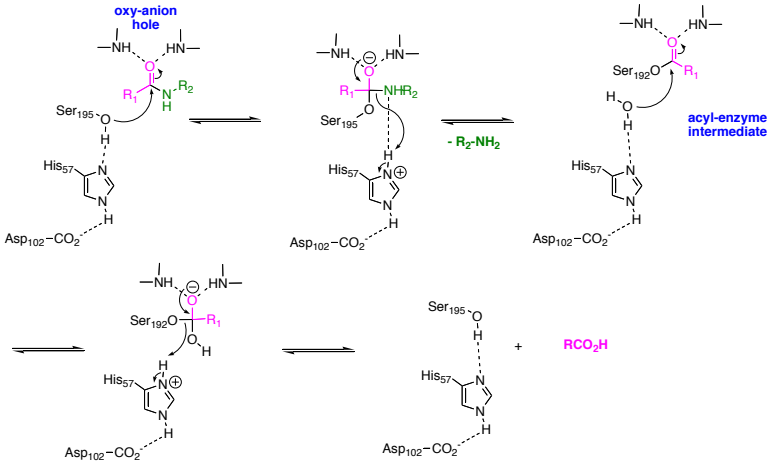


Four classes of proteases:

Serine (trypsin): aspartate-histidine-serine
Aspartyl (HIV protease, renin): two aspartates
Cysteine (papain, caspase): histidine-cysteine
Metallo (Zn^{2+}) (carboxypeptidase, ACE): glutamate

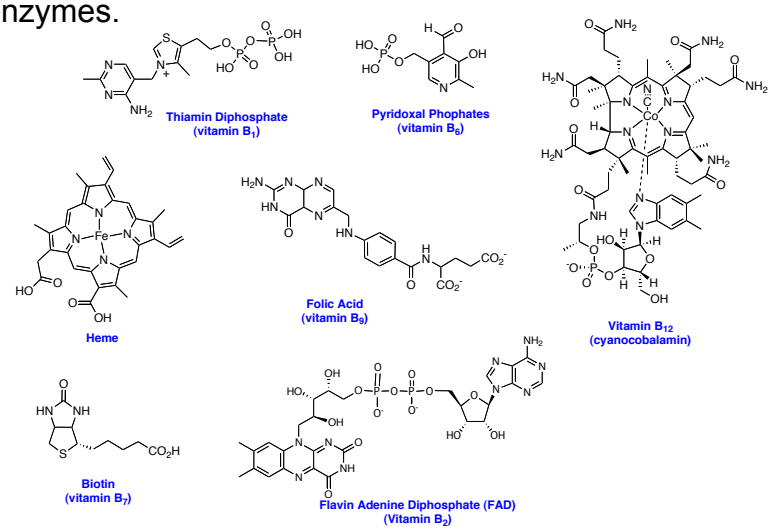
355

Mechanism of carboxylpeptidase, metalloprotease
(Mechanism 25.5, p. 1072)
Mechanism of a serine protease (trypsin, chymotrypsin):



356

25.19: Coenzymes. Some reactions require additional organic molecules or metal ions. These are referred to as cofactors or coenzymes.



25.20: Protein Quaternary Structure: Hemoglobin. (please read)
25.21: G-Coupled Protein Receptors. (please read)

357

Chapter 26: Nucleosides, Nucleotides, and Nucleic Acids. Nucleic acids are the third class of biopolymers (polysaccharides and proteins being the others).

Two major classes of nucleic acids

deoxyribonucleic acid (DNA): carrier of genetic information

ribonucleic acid (RNA): an intermediate in the expression of genetic information and other diverse roles

The Central Dogma (F. Crick):

DNA ⟶ mRNA ⟶ Protein
(genome) (transcriptome) (proteome)

The monomeric units for nucleic acids are nucleotides. Nucleotides are made up of three structural subunits

1. Sugar: ribose in RNA, 2-deoxyribose in DNA
2. Heterocyclic base
3. Phosphodiester

358