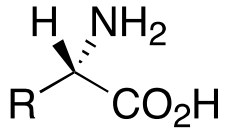
CYI101
Common CHEMISTRY(Organic)

Macromolecules: Introduction to peptides and proteins

21st February 2022/Sec G & H

Amino Acids: *Structure & Diversity*



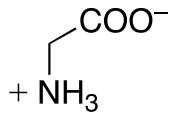
α- Amino Acid

R = sidechain

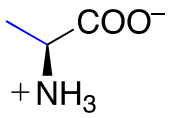
Peptides, and Proteins

monomer unit: α-amino acids

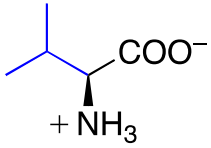
Nonpolar



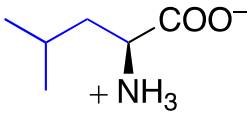
Glycine *(Gly, G)*



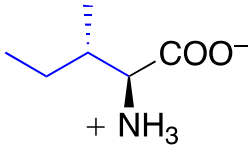
(S)-(+)-Alanine *(Ala, A)*



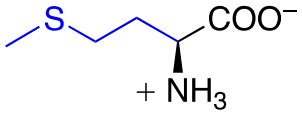
(S)-(+)-Valine *(Val, V)*



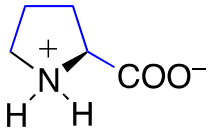
(S)-(–)-Leucine *(Leu, L)*



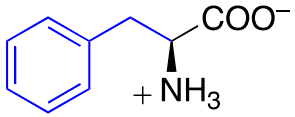
(2S,3S)-(+)-Isoleucine *(Ile, I)*



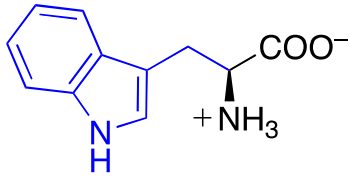
(S)-(–)-Methionine *(Met, M)*



(S)-(–)-Proline *(Pro, P)*

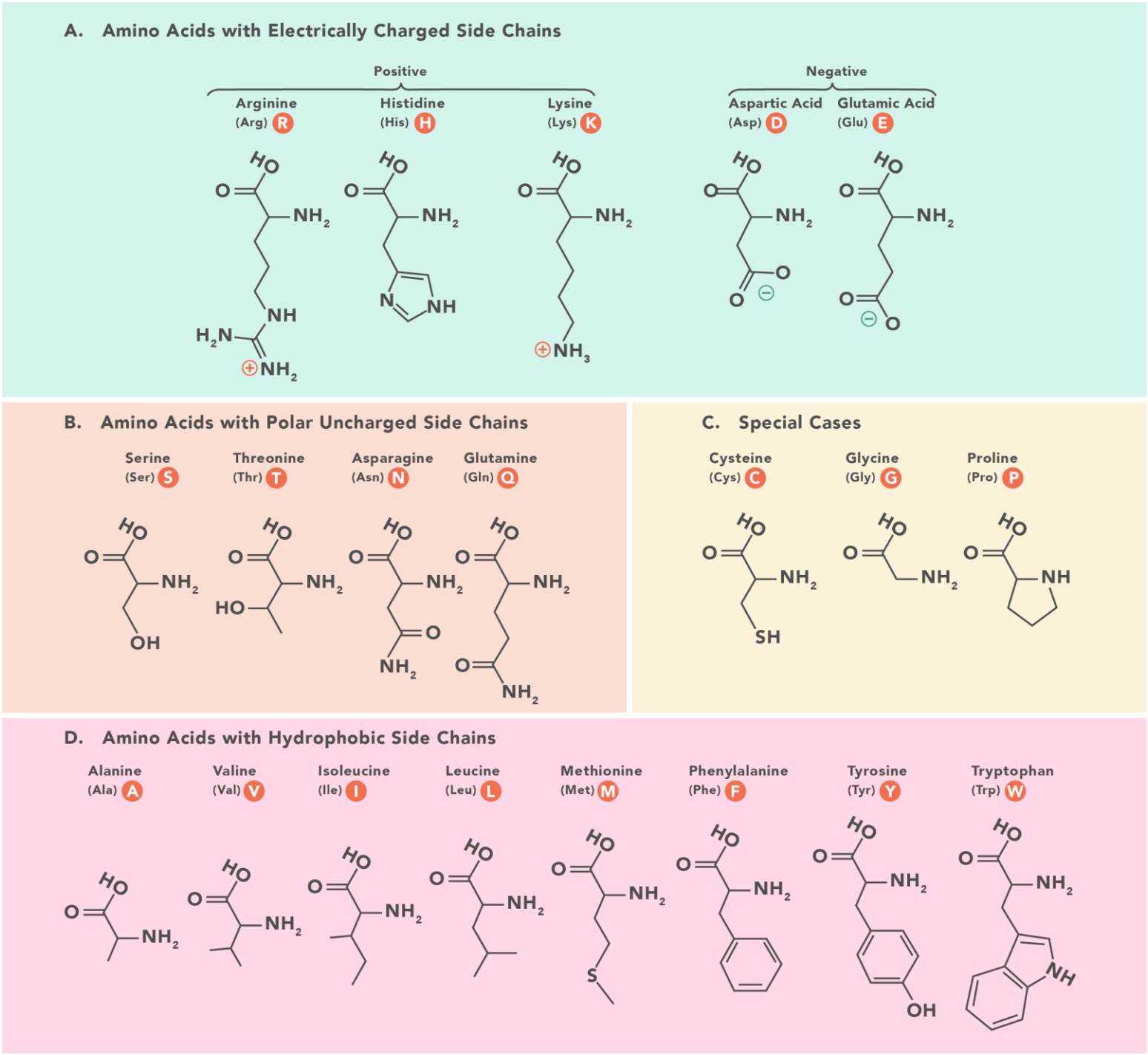


(S)-(–)-Phenylalanine *(Phe, F)*

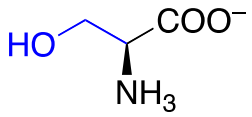


(S)-(–)-Tryptophan *(Trp, W)*

Amino Acids: Structure & Diversity

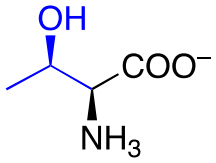


Polar



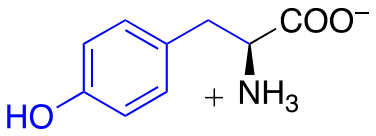
(S)-(–)-Serine *(Ser, S)*

pKa ~ 13



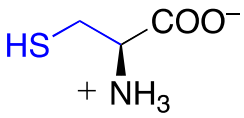
(2S,3R)-(–)-Threonine *(Thr, T)*

pKa ~ 13



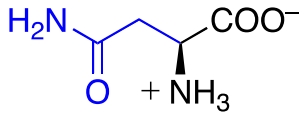
(S)-(–)-Tyrosine *(Tyr, Y)*

pKa ~ 10.1

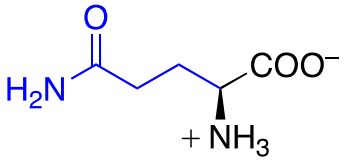


(R)-(–)-Cysteine *(Cys, C)*

pKa ~ 8.2

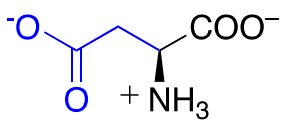


(S)-(–)-Asparagine *(Asn, N)*



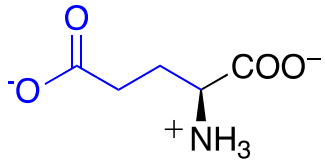
(S)-(+)-Glutamine *(Gln, Q)*

Acidic:



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

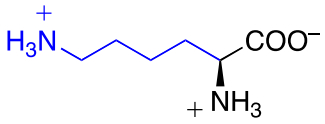
pKa ~ 3.6



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

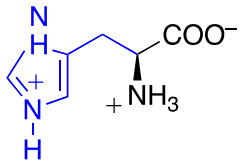
pKa ~ 4.2

Basic:



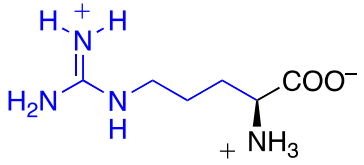
(S)-(+)-Lysine *(Lys, K)*

pKa ~ 10.5



(S)-(–)-Histidine *(His, H)*

pKa ~ 6.0

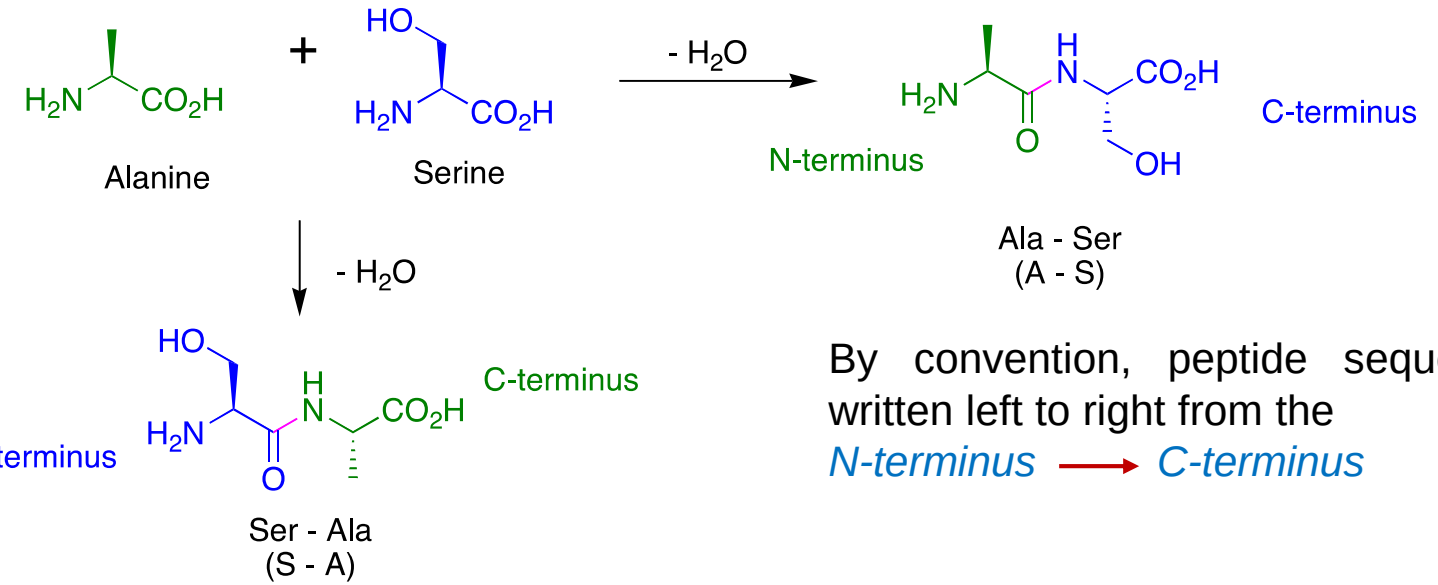


(S)-(+)-Arginine *(Arg, R)*

pKa ~ 12.5

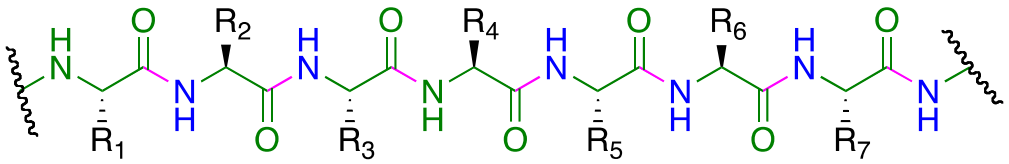
Peptide & Protein: *Sequences*

Proteins and peptides are Biopolymers made up of amino acid units (residues) that are linked together through the formation of amide bonds (peptide bonds) from the aminogroup of one residue and the carboxylate of a second residue



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

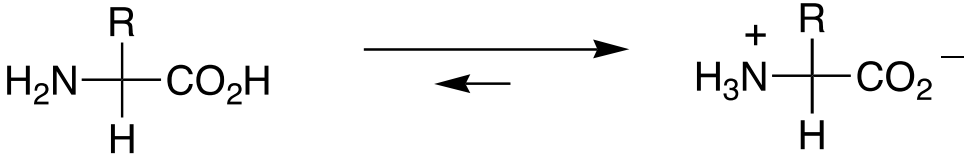
peptide (< 50 amino acids)
protein (> 50 amino acids)



Proteins are large, complex molecules that serve diverse functional and structural roles within cells.

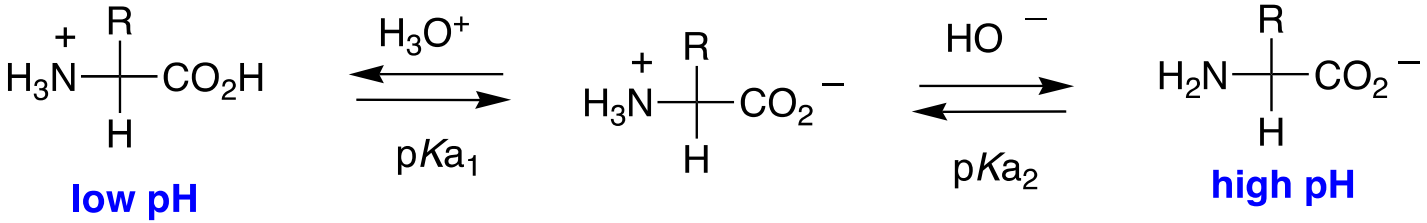
Amino Acids: *Zwitterionic Behaviour*

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).

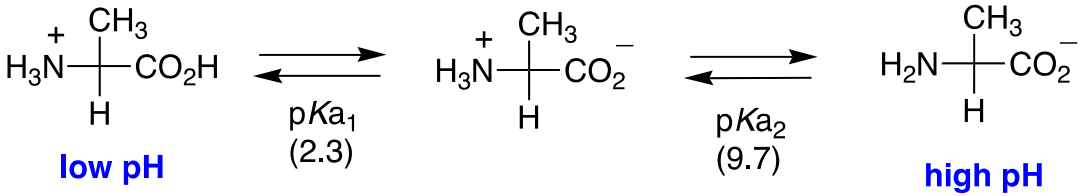


Amino acids are amphoteric: they can react as either an acid or a base. Ammonium ion acts as an acid, the carboxylate as a base.

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



Calculating the Isoelectric Point of Alanine:

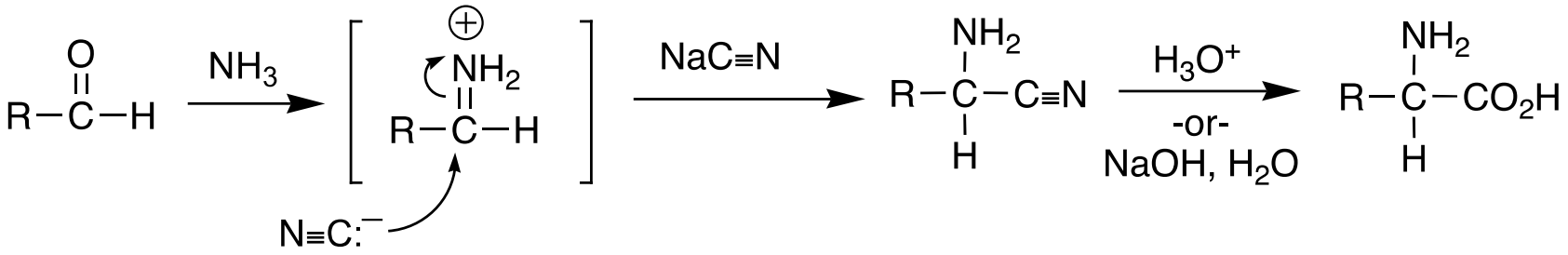


$$pI = \frac{pKa_1 + pKa_2}{2}$$

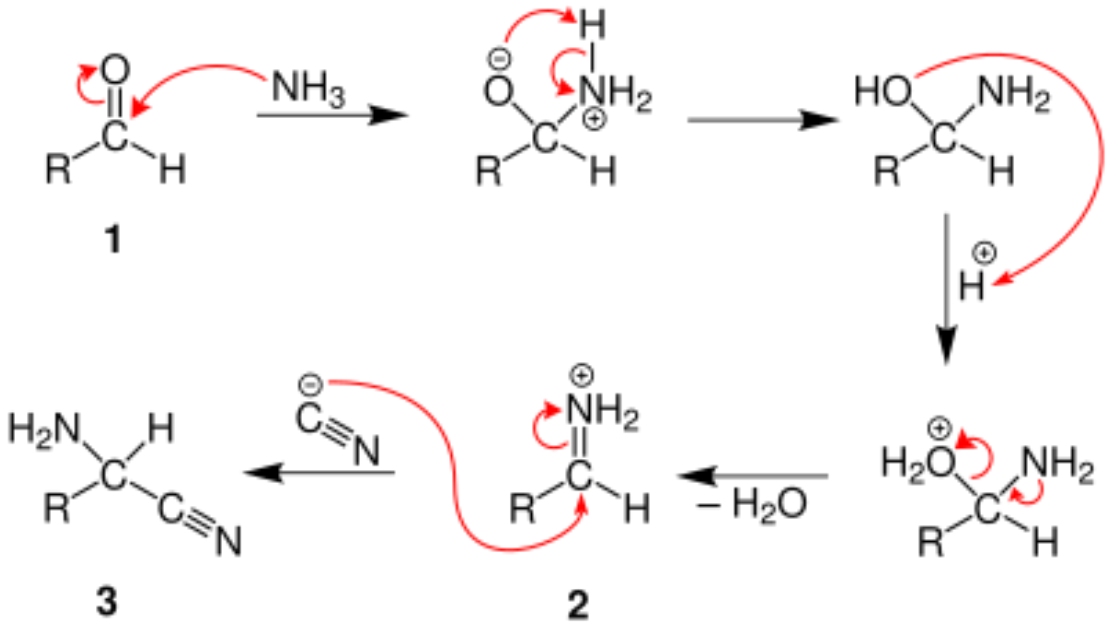
pI = 6.0

Amino Acids: *Synthesis*

Strecker Synthesis: The Strecker amino acid synthesis, is a method for the synthesis of amino acids by the reaction of an aldehyde with ammonium chloride in the presence of Sodium cyanide.

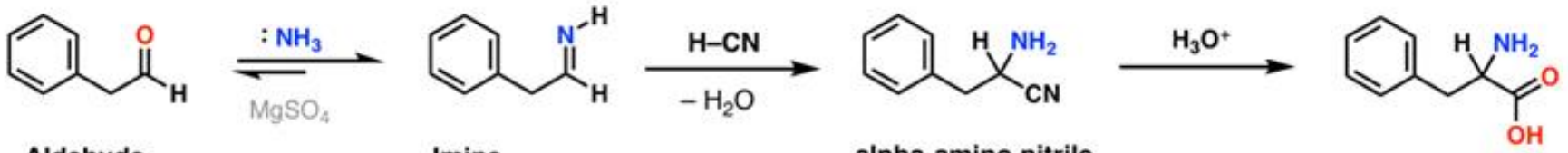


Mechanism



Amino Acids: *Synthesis*

The Strecker Synthesis of Amino Acids From Aldehydes

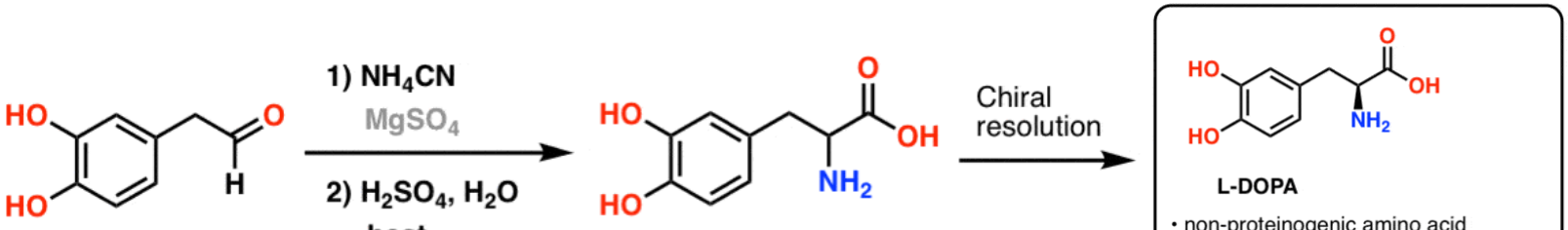


Aldehyde | **Imine** | **alpha-amino nitrile** (racemic mixture) | **(+/–) Phenylalanine** *(one of 20 "proteinogenic" amino acids)*

Notes:

- although "NH_4CN" may sometimes be written, NH_4CN itself is not a stable solid. However NH_4CN is made in solution by combining NH_4Cl and KCN, with KCl as a byproduct
- $MgSO_4$ is sometimes used to assist in imine formation, as it absorbs H_2O and drives the equilibrium towards the imine
- (greatly preferred to using HCN gas!)

An application of the Strecker: synthesis of important non-proteinogenic amino acids



(3,4-dihydroxyphenyl)acetaldehyde

DL-DOPA

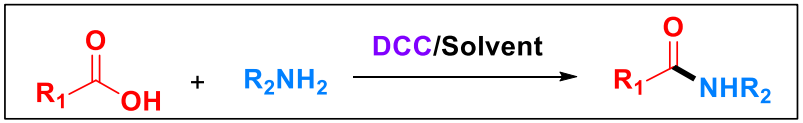
(racemic mixture)

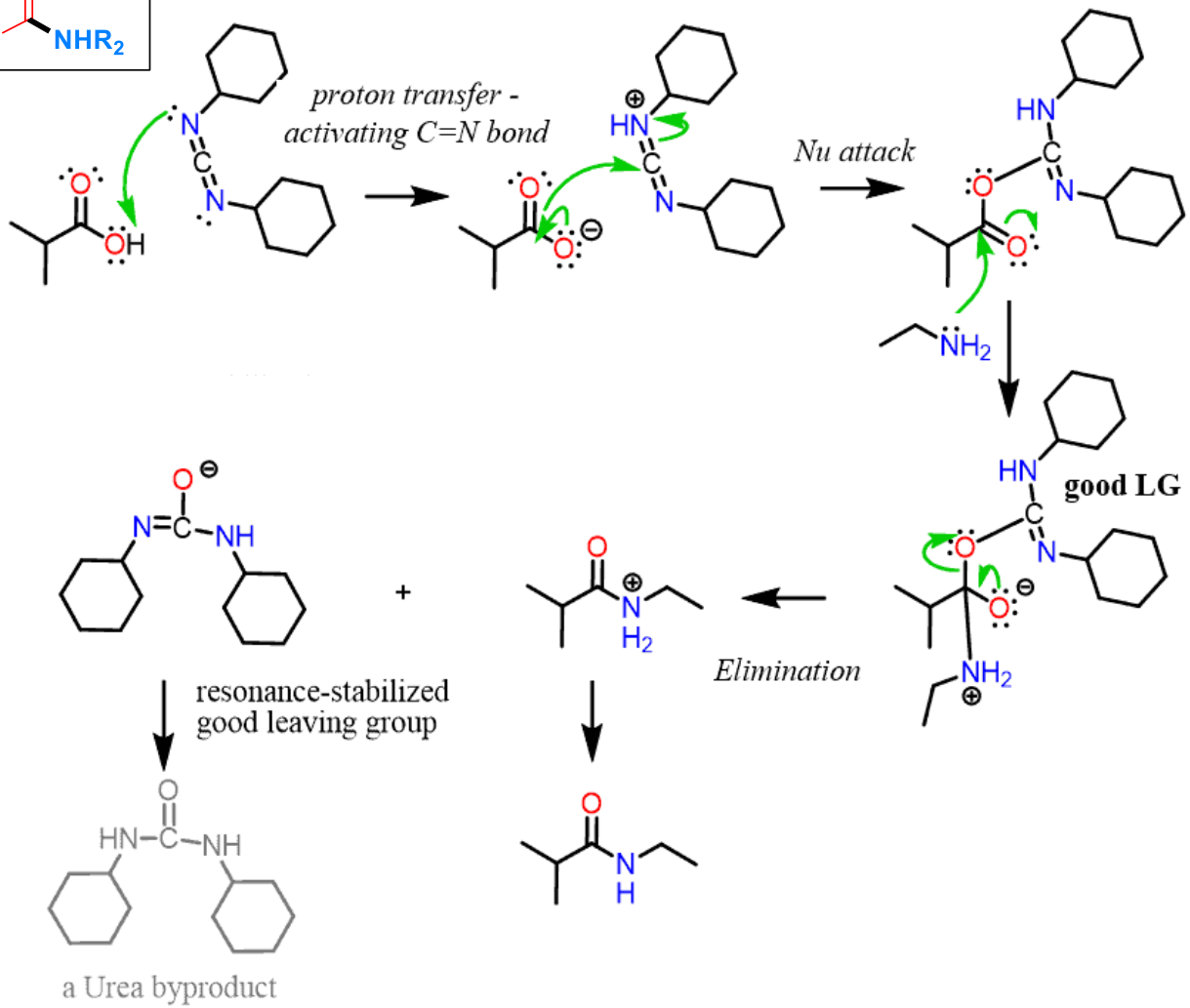
L-DOPA

- non-proteinogenic amino acid
- Treatment for Parkinson's disease

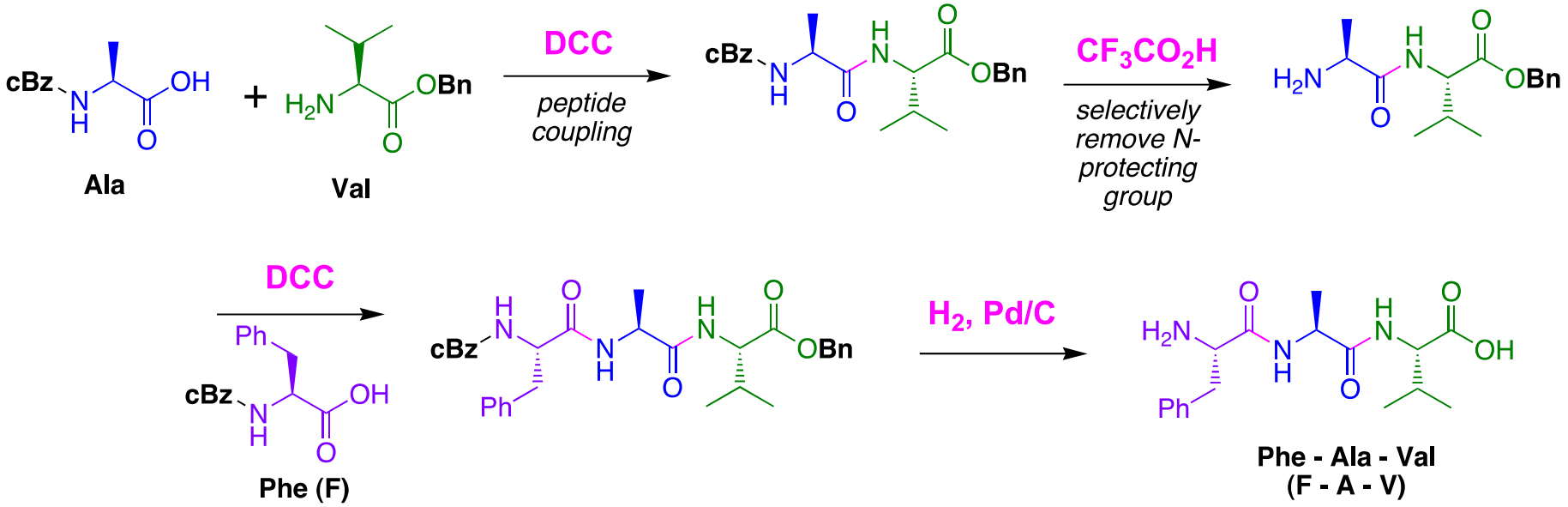
Peptide Synthesis: *Amide Bond*

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)*

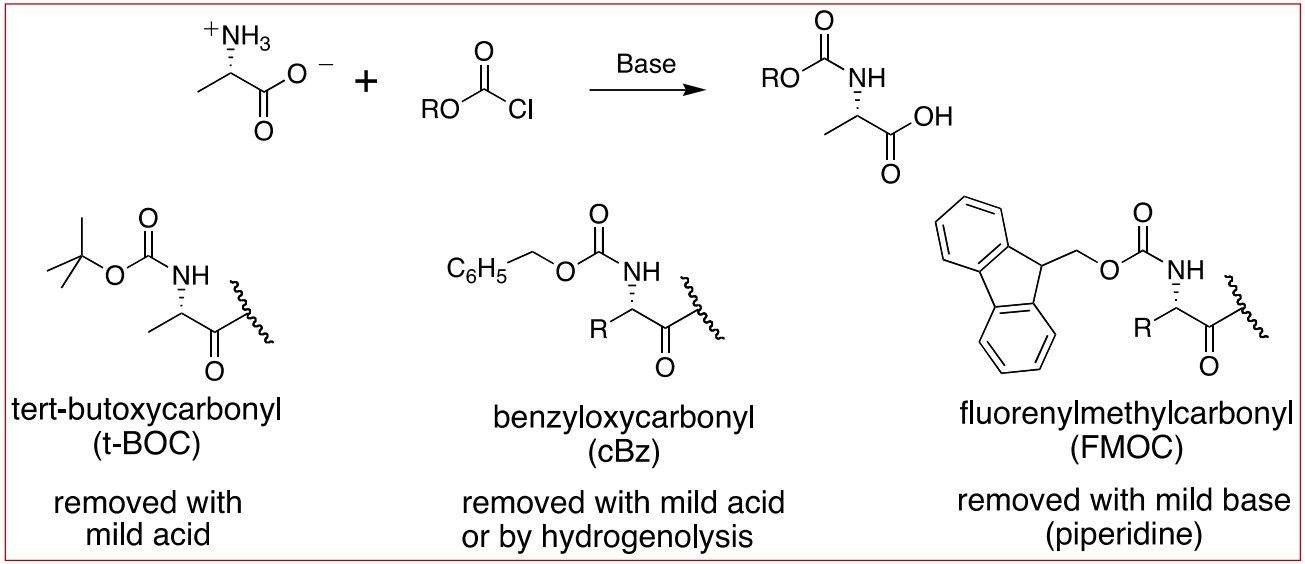




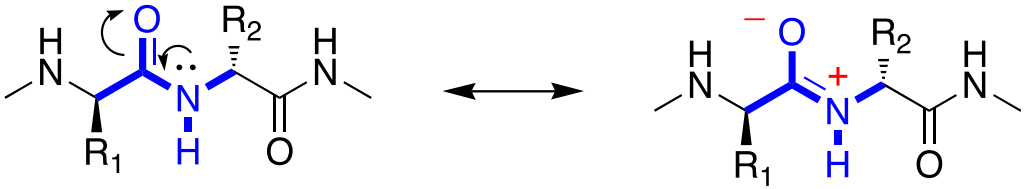
Peptide Synthesis: *Phe-Ala-Val*



Common Protecting group in Peptide Synthesis

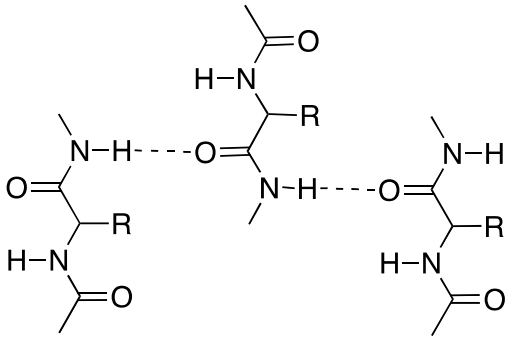


The amide (peptide) bond has C=N double bond character due to resonance resulting in a planar geometry

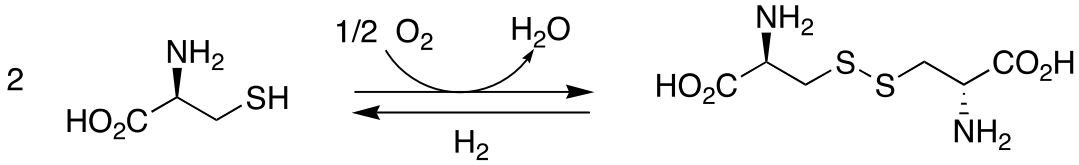


- *restricts rotations*
- *resistant to hydrolysis*

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



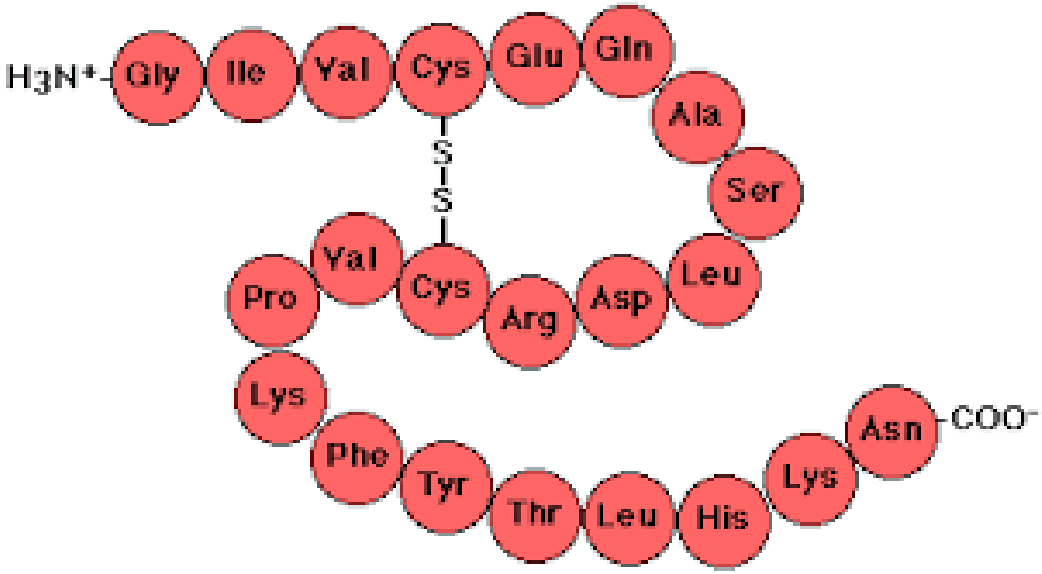
Disulfide bonds: the thiol groups of cysteine can be oxidized to form disulfides (Cys S-S-Cys)



Protein: *Primary structure*

The primary structure is the sequence of amino acids.

Protein primary structure is the linear sequence of *amino acids in a peptide* or protein. By convention, the primary structure of a protein is reported starting from the amino-terminal (N) end to the carboxyl-terminal (C) end. The amino acids in the primary structure are held together by covalent bonds, which are made during the process of protein synthesis.

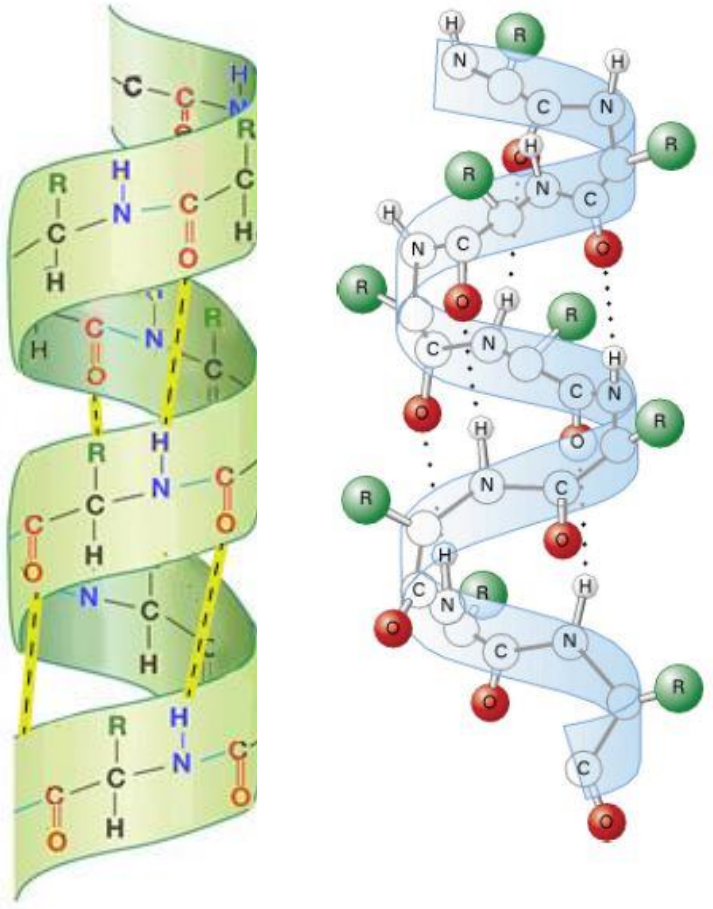


The secondary structure is primarily composed of alpha helices and beta-sheets.

Protein: *Secondary structure*

Secondary structure of proteins is formally defined by the *pattern of hydrogen bonds* between the *amino hydrogen* and *carboxyl oxygen* atoms in the peptide backbone. The most common secondary structures are *alpha helices* and *beta sheets*.

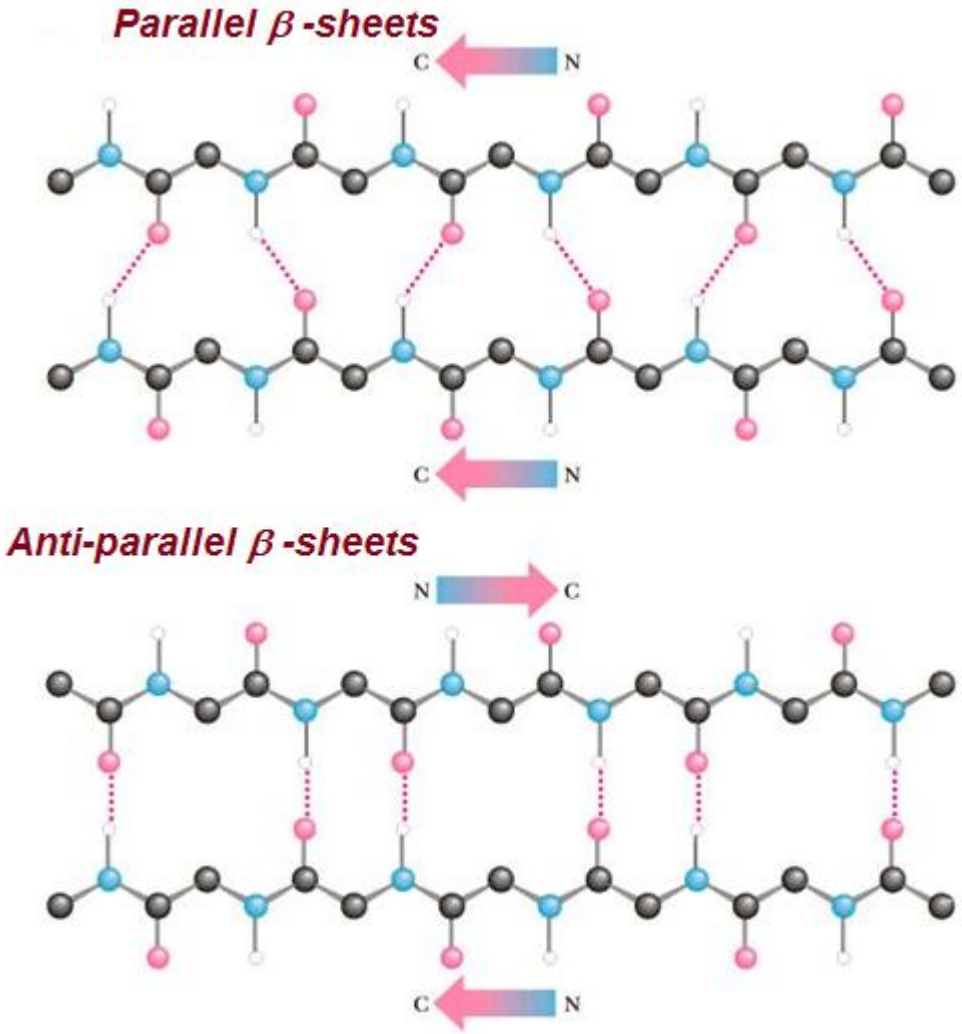
The alpha helix (α-helix) is a common motif in the secondary structure of proteins and is a right hand-helix conformation in which every *backbone N−H group* hydrogen bonds to the *backbone C=O group* of the *amino acid located three or four residues earlier along the protein sequence.*



Alpha helix (α-helix)

Protein: *Secondary structure*

The β-sheet is a common motif of regular secondary structure in proteins. Beta sheets consist of so β-strand connected laterally by at least two or three backbone hydrogen bonds, forming a generally twisted, pleated sheet.



- β-sheets formed from multiple side-by-side beta-strands.
- Can be parallel or anti-parallel configuration.
- Anti-parallel beta-sheets are more stable.

The supramolecular association of β-sheets has been implicated in formation of the protein aggregates and fibrils observed in many human diseases, notably the amyloidosis such as Alzheimer's disease.