2.1. Introduction - Overview of Chemical Biology

2.1.1. What Is Chemical Biology?

- Although the The term “Chemical Biology originates in 1945 by Linus Pauling, it is a relatively new scientific area related to Molecular biology, Structural chemistry, Bioinformatics, Proteomics, Organic chemistry, Pharmacology, Medicinal chemistry and many others.

- The field is primarily concerned with research at the interface of Chemistry and Biology. But in practical sense it is largely focused on small molecules as research tools and potential therapeutics. The crucial part of the work in Chemical Biology is the design and synthesis of small molecules and screening of these with bioassays by utilizing instrumentation of chemistry and biology. It takes the helps of all established field in sciences and biosciences including efficient methods for the synthesis of large numbers of compounds, characterization tools, computational structure-activity relationship, and high-throughput screening in bioassays. Thus the chemical biology is important approach for combating disease such as Cancer, Neurodegenerative diseases, Inflammation and many others.

- Thus, Chemical Biology-related research is strongly interdisciplinary and contains elements from many scientific disciplines such as Medicinal chemistry, Molecular biology, Pharmacology, Biophysical chemistry, Biochemistry, Organic chemistry, Structural chemistry, Bioinformatics, Proteomics, Genetics and more (Figure 2.1 and 2.2).

- Chemical biology may also be defined as the application of synthetic chemical techniques and tools to the study and manipulation of biological systems.

- **Chemical biology** is a scientific discipline spanning the fields of chemistry and biology that involves the application of chemical techniques and tools, often compounds produced through synthetic chemistry, to the study and manipulation of biological systems.
- This is slightly different from biochemistry, which is classically defined as the study of the chemistry of biomolecules. While, biochemists study biomolecules, their three-dimensional structure, dynamics, functions and the inhibition/ activation of enzymes/receptors with small organic molecules, Chemical biologists attempt to utilize chemical principles to modulate systems to either investigate the underlying biology or create new function. Thus, Chemical Biology is often closer related to that of cell biology than biochemistry. In short, biochemists deal with the chemistry *of* biology, chemical biologists deal with chemistry *applied to* biology.

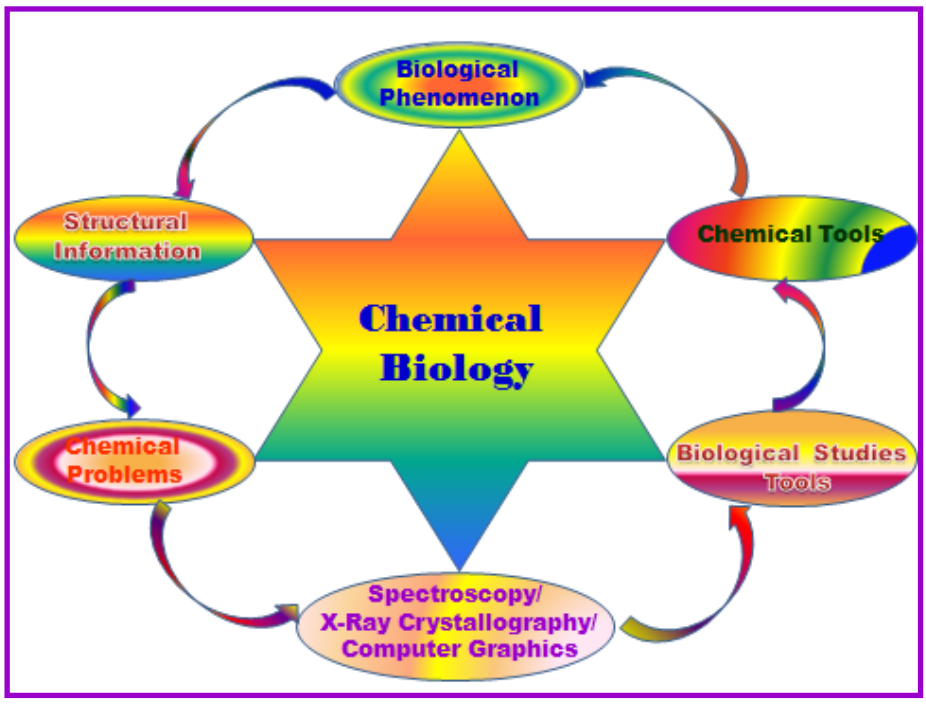


Figure 2.1: Chemical biology takes the help of various research tools from other branches of sciences.

- Some forms of chemical biology attempt to answer biological questions by directly probing living systems at the chemical level. In contrast to research using biochemistry, genetics, or molecular biology, where mutagenesis can provide a new version of the organism or cell of interest, chemical biology studies sometime probe systems *in vitro* and *in vivo* with small molecules that have been designed for a specific purpose or identified on the basis of biochemical or cell-based screening.

- **According to Schreiber's Definition:** Chemical biologists make both small and large 'small molecules'. They make them in tubes and cells, on the surfaces of glass, in monolayers, and even on phage viruses, and they use them to illuminate the principles that underlie life.

- **To my understanding, Chemical Biology is the scientifically designed chemical-molecular approach to investigate cellular principles at molecular levels with the help of biochemical research tools.**

- **Systems of Interest of Chemical Biology**:

Chemical Biology is driven by the fast expansion of scientific knowledge and methods that has occurred during the last decades in areas such as Molecular biology. Most importantly, recent mapping of the Human genome has made it possible to open up the field of Proteomics. However, the fast development in other areas including Bioinformatics and Structural biology

has also had a great impact. In addition, the technical development in for instance imaging, high-throughput screening (HTS) and Synthetic chemistry has been important.

Research in Chemical Biology is made in many ways and includes studies in for instance:

- Proteomics
- Glycobiology
- Combinatorial chemistry
- Molecular sensing
- siRNA-A tool in chemical biology
- small molecules as probes of biology and medicine

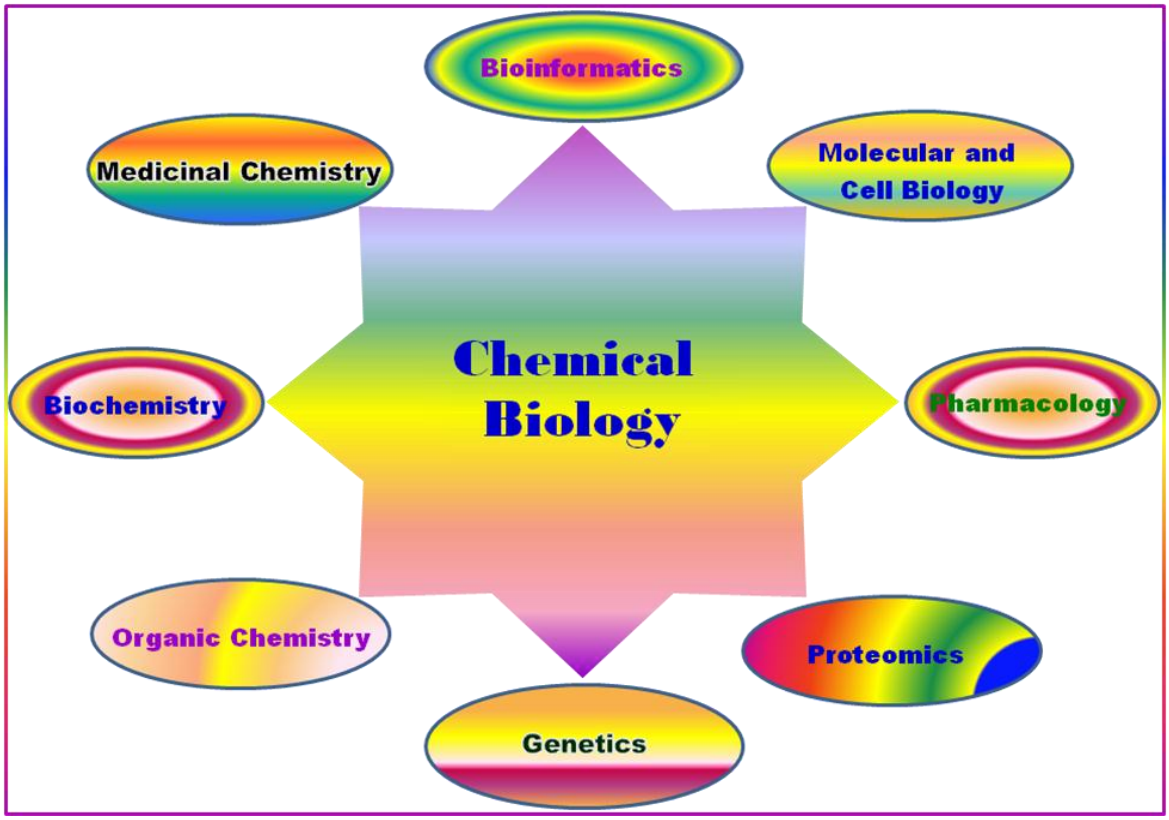


Figure 2.2: Relation between Chemical biology and other branches of sciences.

2.1.2. What is not Chemical Biology?

- **Biochemistry**
  - chemistry of biological systems
  - relatively few classes of reactions mediated by enzymes

- **Bio-organic Chemistry**
  - extension of organic chemistry toward biology
  - e.g. synthesis of natural products

2.1.3. Some Basic Principles

- **Chemoselectivity**
  - Biologically inherent (thiols)
  - Exogenous via promiscuous systems (azido sugars, azidohomoalanine)
- **Molecular Recognition**
  - Stereocomplementarity
  - H-bonding
  - Hydrophobic interactions
- **Ultimately Everything Is Aqueous**
- **Probes, Diagnostics, and Drugs**
  - Tumor cell labeling
  - Enzyme and metabolite tracking
  - Catalysis
  - Modern medicinal chemistry (inhibitor design) is essentially chemical biology
- **Different Drugs**
  - Antibody therapeutics
  - Interference RNAs
- **21st Century Medicine**
  - Proteomics, and Glycobiology, Small molecular approach.

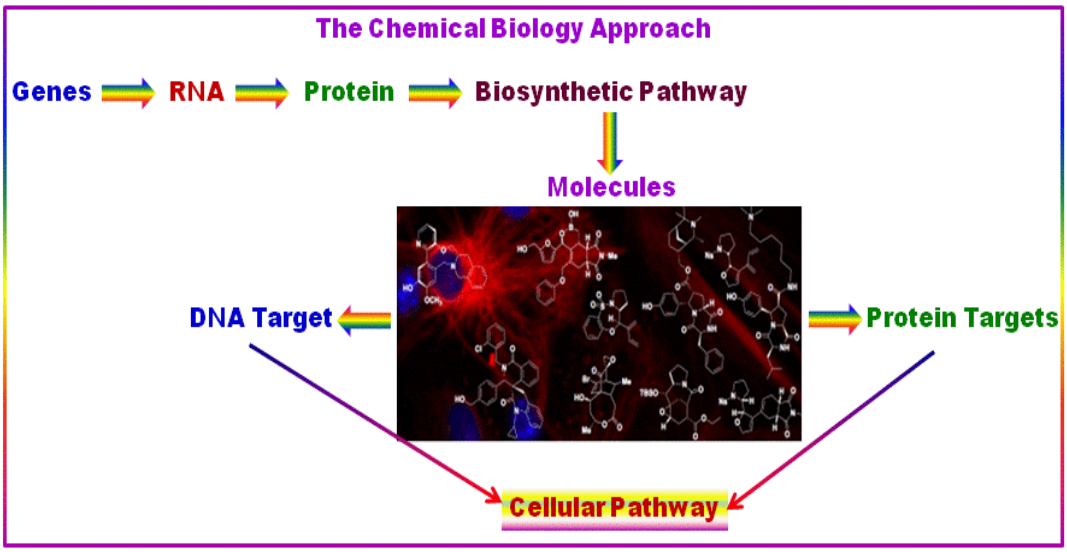


Figure 2.3: The approach of Chemical biology.

2.1.4. Proteomics and Glycobiology

Proteomics

- Proteomics investigates the proteome, the set of expressed proteins at a given time under defined conditions. Proteomics deals with rapid protein identification and has developed into a biological assay for quantitative analysis of complex protein samples by comparing protein changes in differently perturbed systems. Current goals in proteomics include determining protein sequences, abundance and any post-translational modifications. Other interests are study of protein-protein interactions, cellular distribution of proteins and understanding protein activity. Another important aspect of proteomics is the advancement of technology to achieve these goals.

- Within Chemical Biology-related Proteomics, both the structure and function of the proteins are of importance as a protein may have fundamentally different functions dependent on its three dimensional structure. An example of when this can give clinical effects is the Mad Cow Disease which is caused by a specific type of misfolded protein called a prion.

- Chemical biologists are poised to impact proteomics through the development of techniques, probes and assays with synthetic chemistry for the characterization of protein samples of high complexity. These approaches include the development of enrichment strategies, chemical affinity tags and probes.

Glycobiology

- **Glycobiology** is the study of the structure, biological functions, and biosynthesis, of saccharides presents in all living cells and are widely distributed in nature with the help of organic chemistry, molecular and cellular biology, enzymology and related domains.

- Sugars or saccharides are essential components of all living things and aspects of the various different roles they play in biology are researched in various different medical, biochemical and biotechnological fields.

- Sugars control and influence almost every aspect of the cellular processes ranging from cell-cell interactions, energy intermediates, adhesion mechanisms, growth factor signaling, blood clotting, receptor binding, regulating the activity of hormones in the blood, directing embryonic development, and serving structural roles. In this new light, the role of sugars is being reconsidered with a new zeal to understand the working of the intricate processes of life.

- Thus, glycobiology is an area of dense research for chemical biologists. For instance, live cells can be supplied with synthetic variants of natural sugars in order to probe the function of the sugars in vivo. Carolyn Bertozzi at University of California, Berkeley has developed a method for site-specifically reacting molecules the surface of cells that have been labeled with synthetic sugars.

- Glycomics, analogous to Genomics and Proteomics, is the systematic study of all glycan structures of a given cell type or organism and is a subset of glycobiology.

- While DNA, RNA and Proteins are encoded at the genetic level, sugars or glycans or glycoconjugates (Sugars linked with other types of biological molecule, like: Glycoproteins; Proteoglycans; and Glycolipids) in the cell are not encoded directly at any level.

- It is well established fact that simple sugars have their own complex language, perhaps more complex than those of proteins and DNA. The molecules continuously steer and guide many activities relevant to the proper functioning of a cell. In fact, many proteins undergo post-translational modifications to form conjugated molecules with sugar (Glycoconjugates) to function properly. However, research in this field is not growing at a fast space compared to those of other major macromolecules, mostly because of the structural complexity of the sugars. They are structurally so complex, can rival the size and complexity of DNA and proteins. Furthermore, their complexity goes on increasing as they come together to form a range of homo and hetero polymeric compounds inside the cell, controlling and influencing a wide array of functional mechanics.

- Fast Atom Bombardment (FAB) mass spectrometry is a powerful tool for characterizing the complex carbohydrates. This technique can be coupled with sensitive array detector technology to elucidate the structures of glycans.

- **The Scope of Glycomics Research:** Growing understanding, knowledge, and investigations reveal beyond doubt that the sugar molecules are not mere decorative elements serving simply structural and energy requirements in a cell. It is clear that their involvement in the intricate design of life is far more crucial than that understood a few years ago. Glycans are at the center of many disorders and diseases sparking the possibility of exploiting them for therapeutic and diagnostic purposes. There are many biochemical pathways and diseases in which glycans are intricately involved. There are many ways in which sugars may affect physiological conditions. What is yet unclear is how many of the possible structures can be predicted by permutation and combination which

actually exist in nature. Thus, investigating the structure is a subject of great deal of research.

2.1.5. Goal of Chemical Biology

- In the postgenomic era it is major challenge to determine how proteins function together in complex molecular systems to drive a diverse set of cellular processes. Molecularly focused technologies that can probe and manipulate protein function in a controlled manner can manage this great job. Chemical Biology at its position at the interface between chemistry and biology can provide a unique platform for the development of such a molecular toolkit.
- Chemical Biology research is diverse. Main goal of Chemical Biology is to synthesize molecules that can be used as tools to selectively and reversibly modulate proteins.
- The synthesis of molecules to study extra- and intracellular signaling is another aim of Chemical Biology as well as some aspects of Glycobiology which is concerned with the study of different types of sugar molecules in the cell. In Chemical Biology studies, synthetic variations of sugar molecules can be used as tools for research.
- Another important branch of Chemical Biology uses endogenous biomolecules to develop chemical processes and/or materials.
- Research in the field is often concerned with the understanding of biological functions in the healthy individual as well as of the pathological mechanisms related to disease conditions including cancer, neurodegenerative disorders, renal and pulmonary dysfunctions and metabolic disorders.
- The field is therefore important for the generation of knowledge and tools for basic science as well as for the study of disease mechanisms. For many diseases, it is also important for the production of countermeasures and preventive actions.

- **Small Molecule Approach:** Building of library of bioactive small molecules is an example of Chemical Biology research. Recent research of small molecule approach to Chemical Biology includes:

    - The study of the different subtypes of Adenosine receptors reported by Jacobson et al., and their agonists and antagonists (activators and blockers).
    - Adenosine receptors are known or suspected to be involved in multiple diseases and conditions including Inflammation, Endocrine disorders, Cancer, Vision disorders, Renal disorders, Pulmonary disorders, Dementia, Anxiety, Pain, Parkinson’s disease, Sleep disorders and Ischaemia. It is therefore clear that molecular tools to reversibly and selectively manipulate the functions of different

subtypes of receptors involved in the conditions may have the potential to become useful tools in the research and also in some cases useful therapeutics.

- Research also emphasized on the manipulation of receptors and reengineer the binding sites on the receptors is the target of chemical biologists. This type of approach could lead to insights into the accuracy of G-protein coupled receptor modeling, signaling pathways, and not least, the design of small molecules to be able to rescue disease-related mutations and do small-molecule directed gene therapy.
- The combination of tailoring of small molecules and their protein targets is therefore of special interest.
- Naturally occurring small molecules called natural products or their derivatives encampus a substantial fraction of the current pharmacopeia. Understanding the relationship of protein targets of natural products to heritable disease genes by comparing the biological functional connections between genes and gene products, e.g., protein/protein interactions (network connectivities) of these targets and genes will lead to the drug discovery for disease treatment at protein level. By determining whether natural products are intrinsically suited for targeting disease genes and whether their enrichment among current drugs reflects a historical focus or special properties intrinsic to these molecules. Prof. Stuart L. Schreiber is working in a data-driven way, to discover whether or not the propensity of natural products for interaction with biological targets is an advantage for probe or drug discovery directed at the genes determined to be causal for human disease.
- They are successful in identifying natural product targets and evaluating a database of natural products and targets from GVKBio. They have standardized 5581 target names and species to human proteins, as either direct natural product targets or orthologous human target proteins, and mapped these targets to 946 human proteins with connections in STRING. For human disease genes, they combined 3655 genes contained in the OMIM Morbid Map with 1580 genes from a genome-wide association study SNP database and mapped these to 2681 human proteins with connections in STRING.

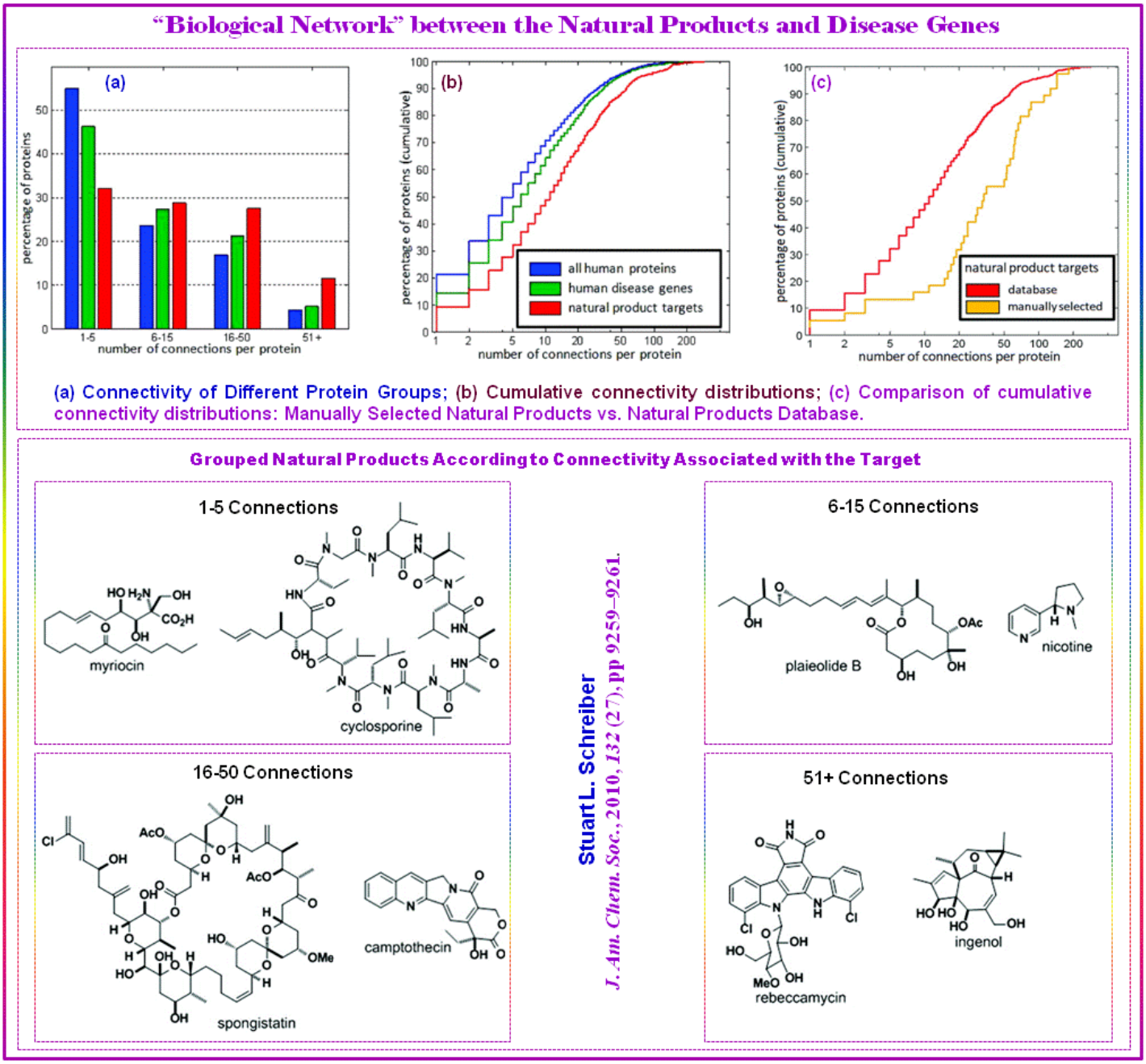


Figure 2.4: The small molecular approach of Chemical biology.

- Schreiber's results indicate that targets of natural products are highly connected, more than genes connected to human disease. This finding indicates that targeting at protein level is much more beneficial than at gene level for diagnostic.
- Many natural products act as basic defense mechanisms against invaders in the absence of tissue specialization or an advanced immune response leading to the death of the invading organism.

Therefore, the highly connected proteins can be targeted by natural products which will interrupt the activities of essential protein of the invader. Therefore the ultimate goal of chemical biology remains in the development of library of small natural molecules to target the root cause of the diseases and thus the treatment of all the heritable diseases can better way be done.

2.1.6. Manipulation of Biological Activity with Small Molecule:

In the postgenomic era it is major challenge to determine how proteins function together in complex molecular systems to drive a diverse set of cellular processes. Molecularly focused technologies that can probe and manipulate protein function in a controlled manner can manage this great job. Chemical Biology at its position at the interface between chemistry and biology can provide a unique platform for the development of such a molecular toolkit.

Due to the fundamental importance of transcriptional regulation in biological systems, small-molecule mimetic that promote exogenous control over transcription represent extremely valuable tools. Such study was reported by Jin Zhang *et al.,* and Helen E. Blackwell et al.

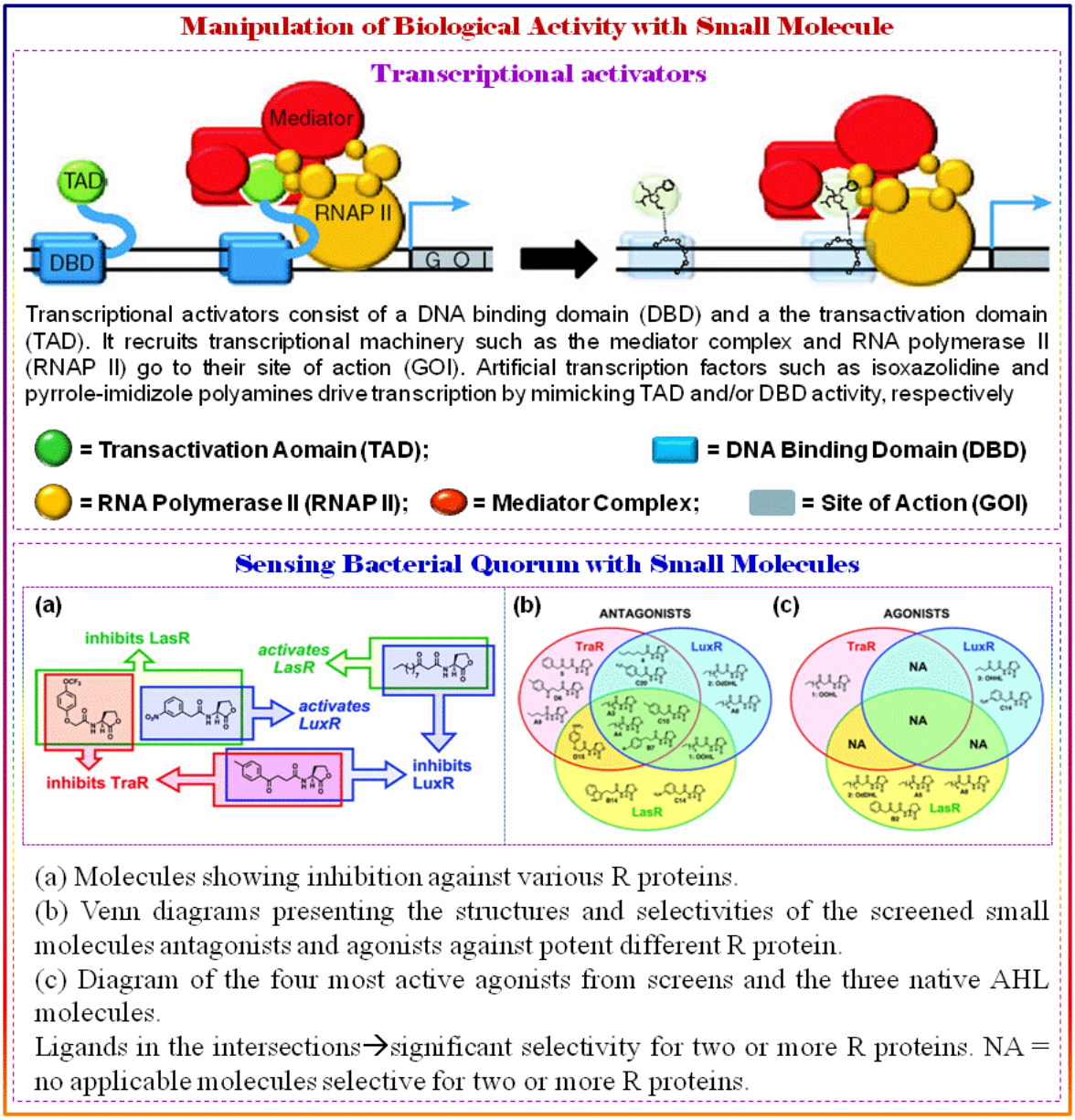


Figure 2.5: Manipulation of biological activity with small molecules.

2.1.7. Recent Advancement and Future Prospect of Chemical Biology:

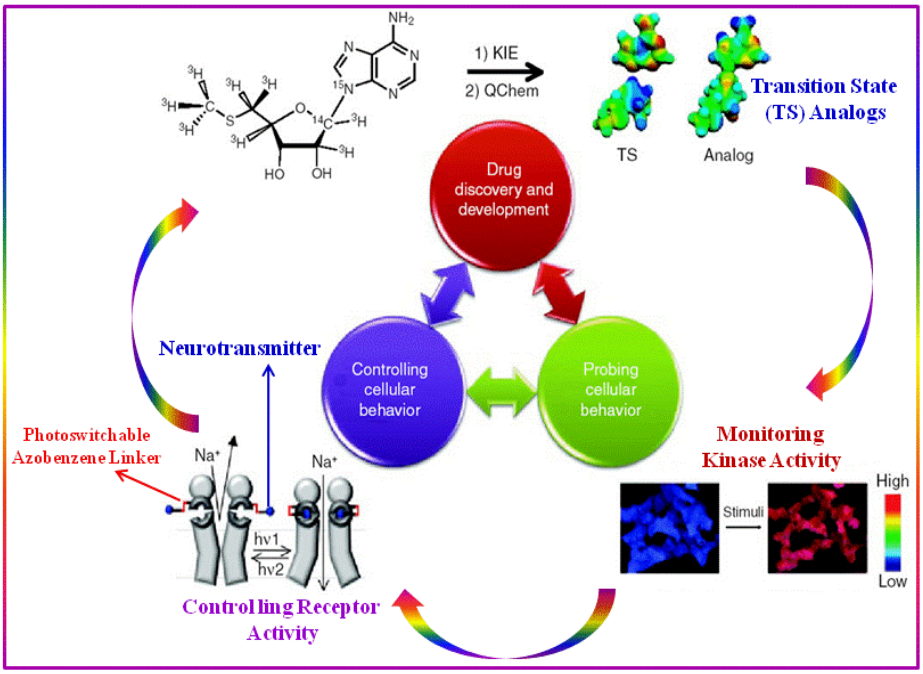
Recent Advancement of Chemical Biology through Several Approaches

1. **Directed Development of Transition State (TS) Analogs:**

- Geometric and electrostatic potential mapping of enzymatic transition states by the use of several possible techniques of Chemistry, biology and computaion will allow us design transition state analogues molecules which will provide a wealth of enezyme inbitors drugs for arresting the enezyme related with diseases. Several of such inhibitors are in clinical trials or in preclinical development methods. Thus designing small molecules inhibitors is the fruitful approach for the drug development.

2. **A FRET-Based Strategy to Monitor Kinases Activity:**

- The activity assays on kinases based on FRET reporters as was reported by Allen, M. D. may find immediate application in high-throughput screening of small molecules drugs. It will also help in probing of multiple physiological and pharmacological events at subcellular locations in living cells in chemical and functional genomics studies. Thus, bridging of high-throughput technology with dynamic live-cell activity measurement can laid a establishment of mechanistic studies and versatile drug discovery processes targeting protein kinases.



3. **Photoswitchable Gate: Remote Control of Receptor Activity:**

- Neurons have voltage-gated, ligands-gated, and temperature-gated ion channels but not light-gated ion channel. Recently a structure-based design of a new chemical gate was developed by Isacoff *et al.* that present light sensitivity to an ion channel. The gate includes a functional group for selective binding of a Shaker K^+ ion channel, a pore blocker and a photoisomerizable azobenzene linker. Light induced *Trans-Cis* isomerization of azobenbenzene linker allowed the photoswitchable gate to switch potential on and off. This strategy was used to control the activity of the Shaker potassium ion channel and ionotropic neurotransmitter receptors in a variety of systems, including zebrafish. The Advancement of such photoswitchable ion channel gate will allow allow rapid, precise and reversible control over neuronal firing. These will find potential applications for controlling the activity of such receptors and designing the potent molecules for future drug design targeting receptor.

Figure 2.6: Recent advancement of chemical biology.

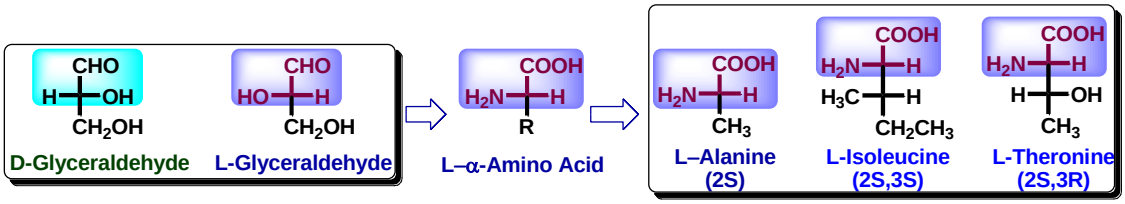
The studies reported till the date accentuate the great impact of chemical biology to the study and modulate the biology and future scope of research at the interface of chemistry and biology.

Exploring the many existing small biological compounds and designing several enzyme inhibitors will also provide essential mechanistic probes and new candidates for drug discovery. Thus, many of such molecules showed and will continue to show the capability of targeting to an entire proteome. Also, the development of robust chemical/bio-chemical tools to probe and to investigate cellular principles at molecular levels will be of great achievement toward the success of conceptual approach, the Chemical Biology and thus, will facilitate the development of new therapeutics for treatment of diseases.

2.2. Amino Acids and their Asymmetric Synthesis

2.2.1. Introduction

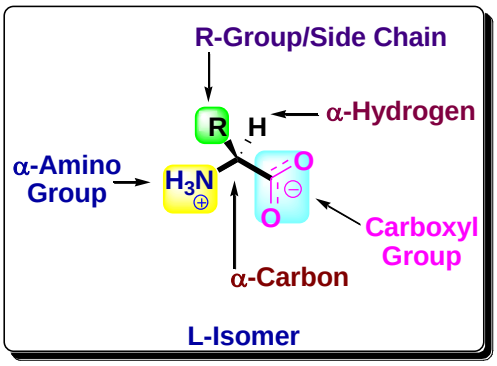
- Amino acids are building blocks of proteins.
- Proteins are composed of 20 different amino acid (encoded by standard genetic code, construct proteins in all species).
- Their molecules containing both amino and carboxyl groups attached to the same a-carbon
- 19 are 1°-amines, 1 (proline) is a 2°-amine
- 19 amino acids are “chiral” and 1 (glycine) is achiral (R=H)
- The configuration of the “natural” amino acids is L (L-a-amino acids).



- Their chemical structure influences three dimensional structures of proteins.
- They are important intermediates in metabolism (porphyrins, purines, pyrimidines, creatin, urea etc).
- They can have hormonal and catalytic function.
- Several genetic disorders are cause in amino acid metabolism errors (aminoaciduria - presence of amino acids in urine)

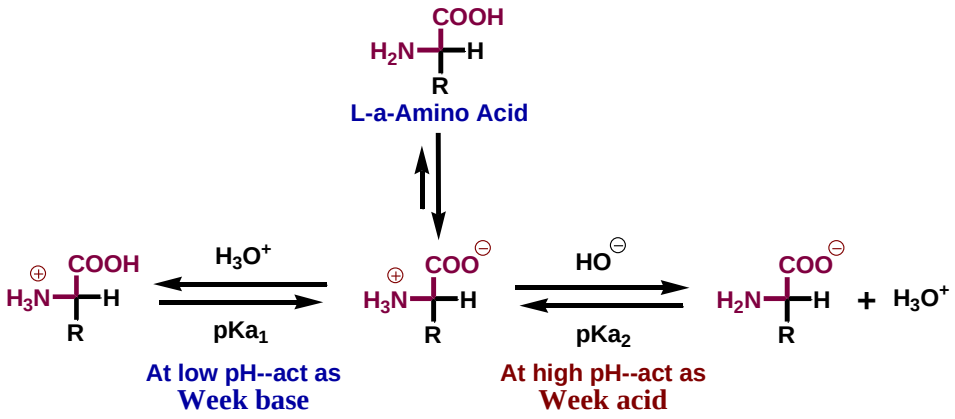
2.2.2. The Basic Structure of Amino Acids

- It differs only in the structure of the side chain (R-group).
- L-isomer is normally found in proteins.



2.2.3. Nonionic and Zwitterionic Forms (Dipolar Structure) of Amino Acids

- Zwitterion = in German for hybrid ion
- The zwitterion predominates at neutral pH
- Isoelectric point (pI): pH at which the amino acid exists in a neutral, zwitterionic form (influenced by the nature of the side chain).
- At acidic pH, the carboxyl group is protonated and the amino acid is in the cationic form
- At neutral pH, the carboxyl group is deprotonated but the amino group is protonated. The net charge is zero; such ions are called Zwitterions
- At alkaline pH, the amino group is neutral $–NH_2$ and the amino acids are in the anionic form.



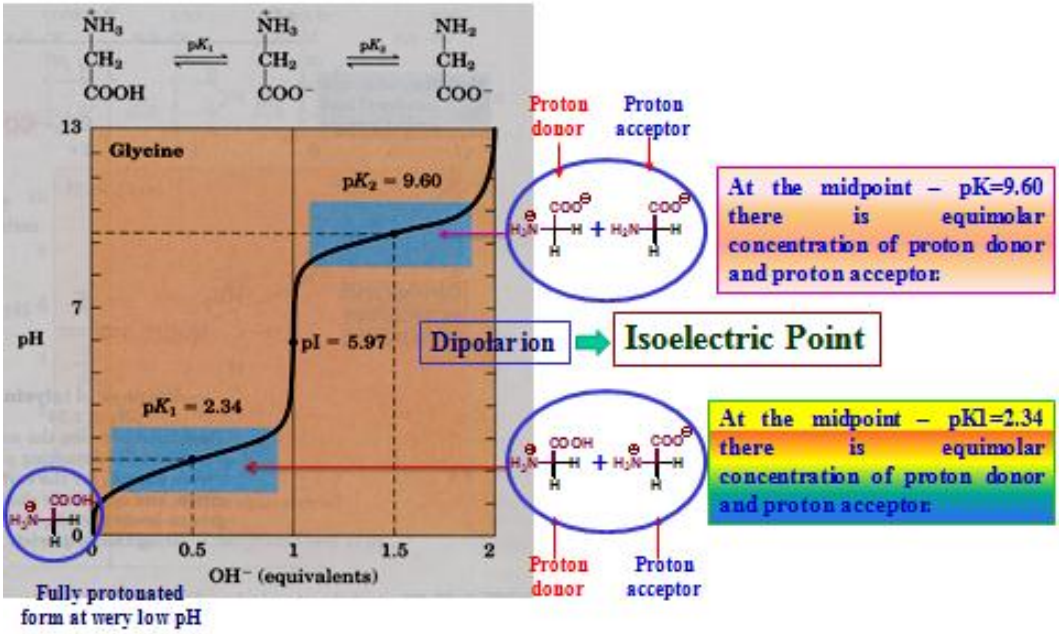
- Amino Acids Carry a Net Charge of Zero at a Specific pH

- **Zwitterions predominate at pH values between the pK_a values of amino and carboxyl group**
- **For amino acid without ionizable side chains, the Isoelectric Point (equivalence point, pI) is:**

$$pI = \frac{pK_1 + pK_2}{2}$$

- **At this point, the net charge is zero**
  - **AA is least soluble in water**
  - **AA does not migrate in electric field**

- **Amino acids have characteristic titration curves:**



- Henderson/Hasselbach equation and pK_a

Protonated form Unprotonated form (conjugate base)

$$HA \rightleftharpoons H^+ + A^-$$

$$K_a = \frac{[H^+]\,[A^-]}{[HA]}$$

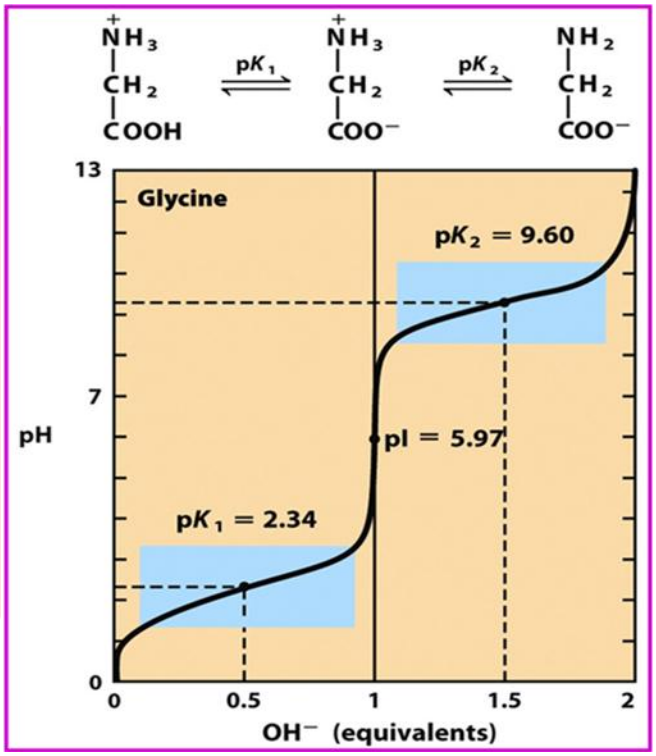
$$[H^+] = K_a \times \frac{[HA]}{[A^-]}$$

$$-\log [H^+] = -\log K_a - \log \frac{[HA]}{[A^-]}$$

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

- **Amino Acids Can Act as Buffers**

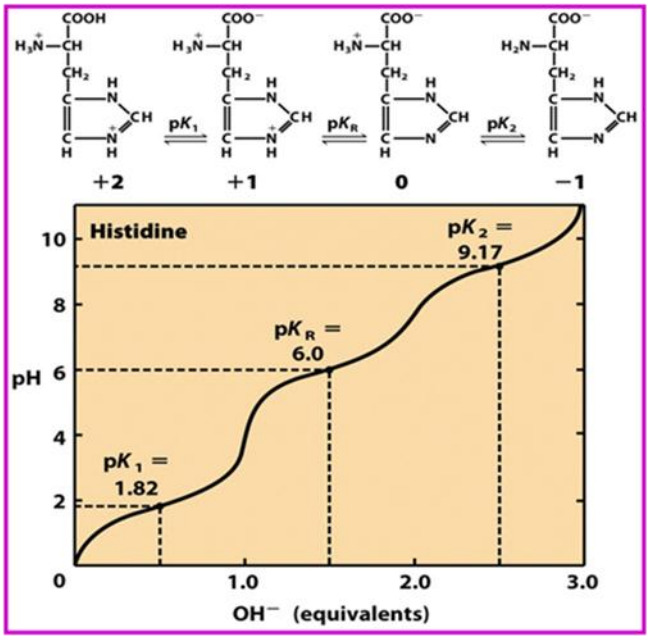
1. **Amino acids with uncharged side-chains, such as glycine, have two pK_a values:**
   - **The pK_a of the α-carboxyl group is 2.34**
   - **The pK_a of the α-amino group is 9.6**
2. **Thus, it can act as a buffer in two pH regimes.**



- **Side-chain Ionization and The pI**

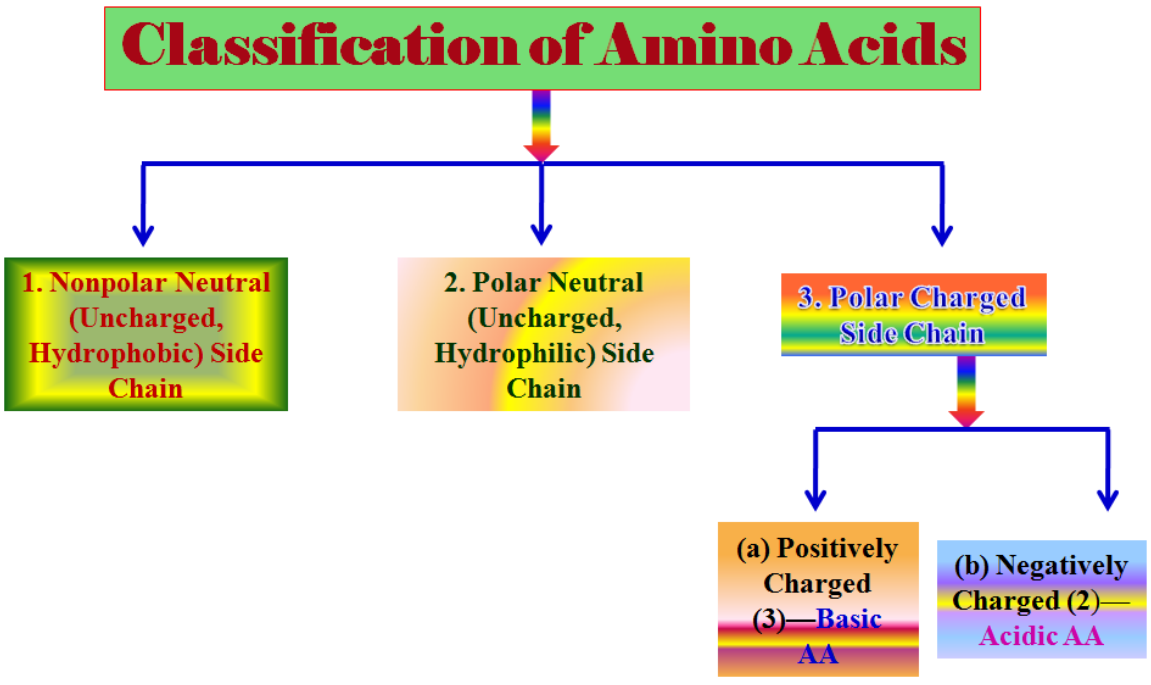
How to calculate the pI when the side-chain is ionizable?

- Identify species that carries a net zero charge
- Identify pK_a value that defines the acid strength of this zwitterion: (pK_2)
- Identify pK_a value that defines the base strength of this zwitterion: (pK_R)
- Take the average of these two pK_a values



2.2.4. Classification of Amino Acids

Amino acids are generally divided into groups on the basis of their side chains (R groups). The most helpful start-point is to separate amino acids into:



2.2.4.A. Amino Acids with Nonpolar Neutral (Uncharged, Hydrophobic) Side Chain:

- Only carbon and hydrogen in their side chains.
- Generally *unreactive* but *hydrophobic.*
- Determining the 3-D structure of proteins (they tend to cluster on the inside of the molecule).
- The simplest amino acid is **Glycine**, which has a single hydrogen atom as its side chain.
- **Alanine, Valine, Leucine and Isoleucine** have saturated hydrocarbon R groups (i.e. they only have hydrogen and carbon linked by single covalent bonds). Leucine and Isoleucine are isomers of each other.
- The side chain of **Methionine** includes a sulfur atom but remains hydrophobic in nature.
- **Phenylalanine** is Alanine with an extra benzene (sometimes called a Phenyl) group on the end. Phenylalanine is highly hydrophobic and is found buried within globular proteins.
- **Tryptophan** is highly hydrophobic and tends to be found immersed inside globular proteins. Structurally related to Alanine, but with a two ring (bicyclic) indole group added in place of the single aromatic ring found in Phenylalanine. The presence of the nitrogen group makes Tryptophan a little less hydrophobic than Phenylalanine.
- **Proline** is unique amongst the amino acids – its side chain is bonded to the backbone nitrogen as well as to the a-carbon. Because of this proline is technically an *imino* rather than an amino acid. The ring is not reactive, but it does restrict the geometry of the backbone chain in any protein where it is present.

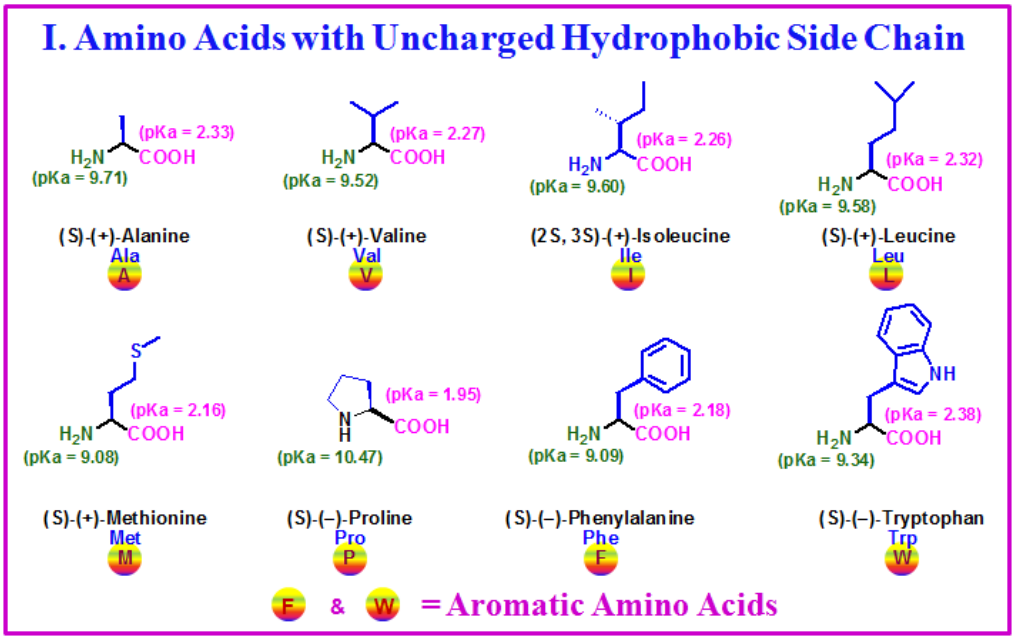


Figure 2.7: Structures of amino acids with uncharged hydrophobic side chain.

2.2.4.B. Amino Acids with Polar Neutral (Uncharged, Hydrophilic) Side Chain:

- **Tyrosine** is Phenylalanine with an extra hydroxyl (-OH) group attached. It is polar and very weakly acidic. Tyrosine can play an important catalytic role in the active site of some enzymes. Reversible phosphorylation of –OH group in some enzymes is important in the regulation of metabolic pathways.
- **Serine and Threonine** play important role in enzymes which regulate phosphorylation and energy metabolism.
- **Cysteine** has sulfur-containing side group.The group has the potential to be more reactive.It is not very polar.Cysteine is most important for its ability to link to another cysteine via the sulfur atoms to form a covalent disulfide bridge, important in the formation and maintenance of the tertiary (folded) structure in many proteins.
- **Asparagine and Glutamine** are the amide derivatives of Aspartate (Aspartic acid) and Glutamate (Glutamic acid) - see below. They cannot be ionised and are therefore uncharged.

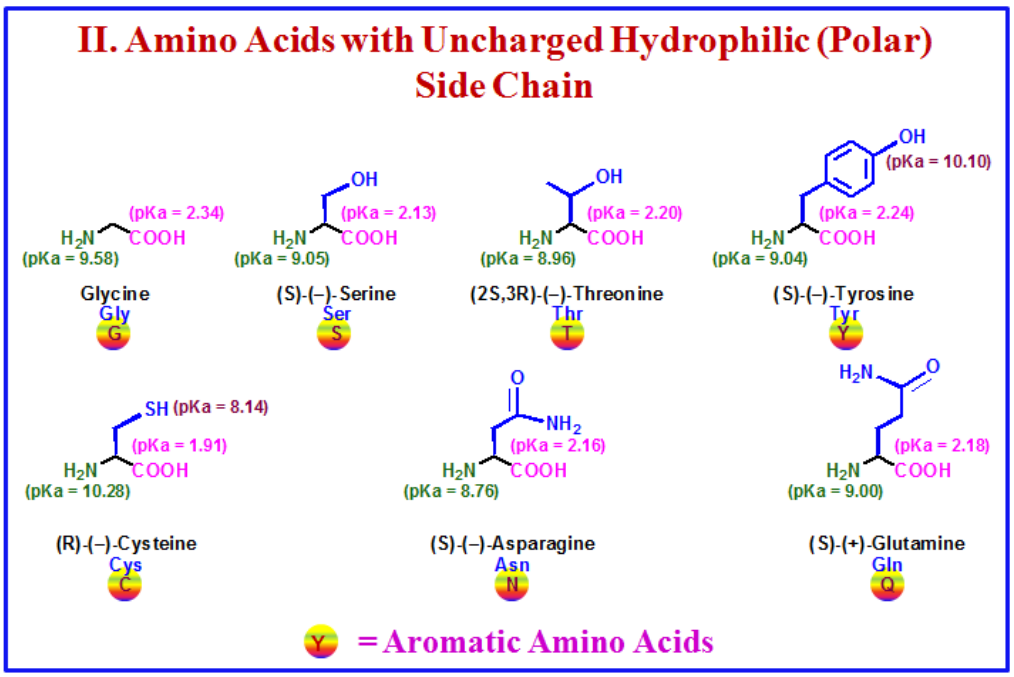


Figure 2.8: Structures of amino acids with uncharged hydrophilic side chain.

2.2.4.C. Amino Acids with Charged Side Chain:

(a) Positively Charged R Groups

- Lysine and Arginine both have pKs around 10.0 and are therefore always positively charged at neutral pH.
- With a pK of 6.5, Histidine can be uncharged or positively charged depending upon its local environment.
- Histidine has an important role in the catalytic mechanism of enzymes and explains why it is often found in the active site.

(b) Negatively (Nonpolar) Charged R Groups:

- Two amino acids with negatively charged (i.e. acidic) side chains - Aspartate (Aspartic acid) and Glutamate (Glutamic acid).
- These amino acids confer a negative charge on the proteins of which they are part.

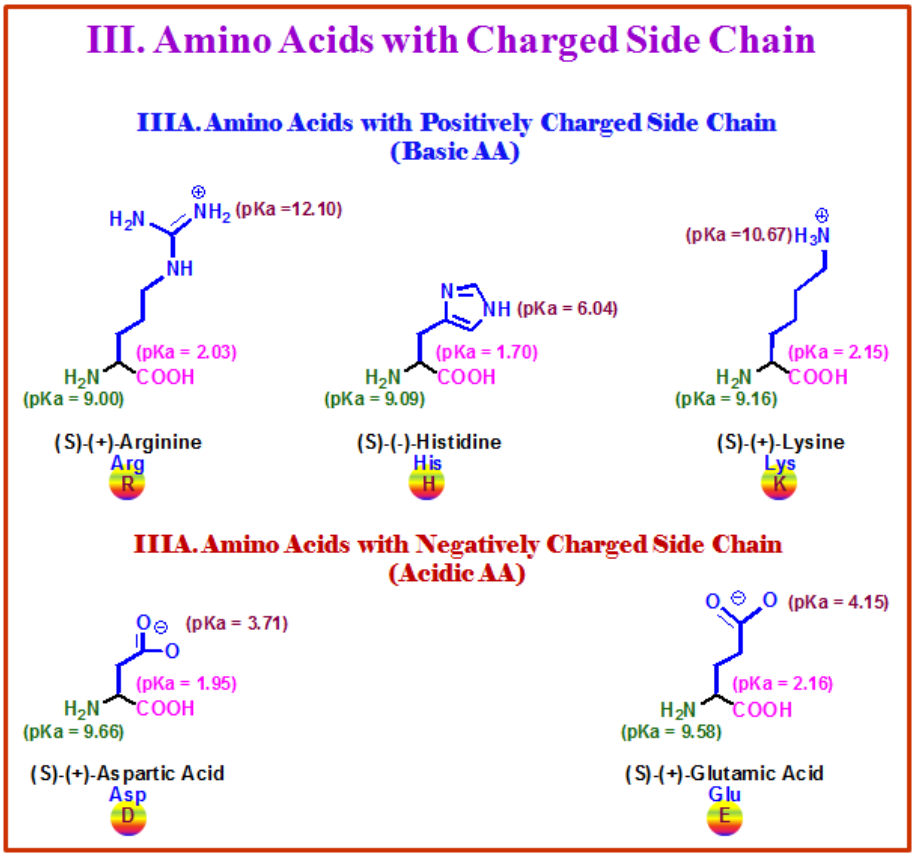


Figure 2.9: Structures of amino acids with charged side chain.

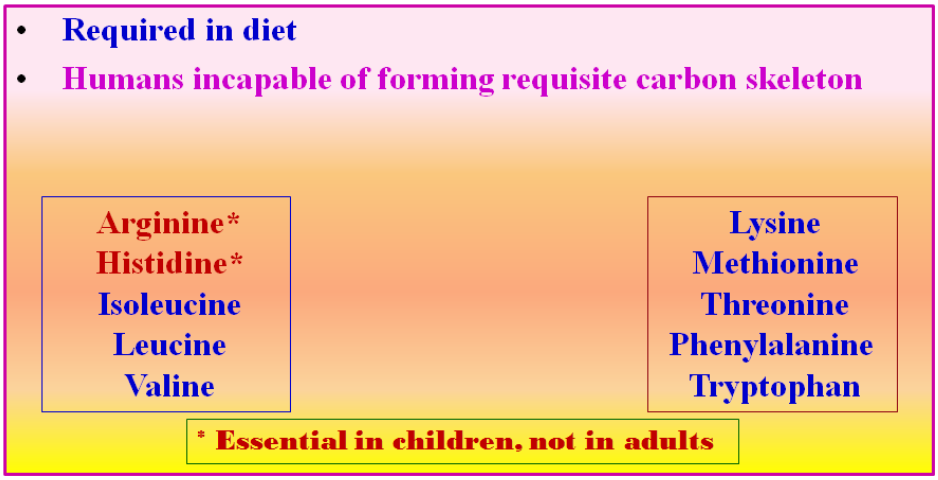
2.2.4.D. "Special" 21st and 22nd Amino Acids:

Figure 2.10: Structures of 21st and 22nd natural amino acids

2.2.4.E. Classification Based on Chemical Constitution

1. **Small amino acids:–** Glycine, Alanine
2. **Branched amino acids:–** Valine, Leucine, Isoleucine
3. **Hydroxy amino acids (-OH group):–** Serine, Threonine
4. **Sulfur amino acids:–** Cysteine, Methionine
5. **Aromatic amino acids:–** Phenylalanine, Tyrosine, Tryptophan
6. **Acidic amino acids and their derivatives:–** Aspartate, Asparagine, Glutamate, Glutamine
7. **Basic amino acids:–** Lysine, Arginine, Histidine
8. **Imino acid:-** Proline
9. **Special/Newly Added amino acids:-** Pyrrolysine, Selenocysteine

2.2.5. Essential Amino Acids in Humans



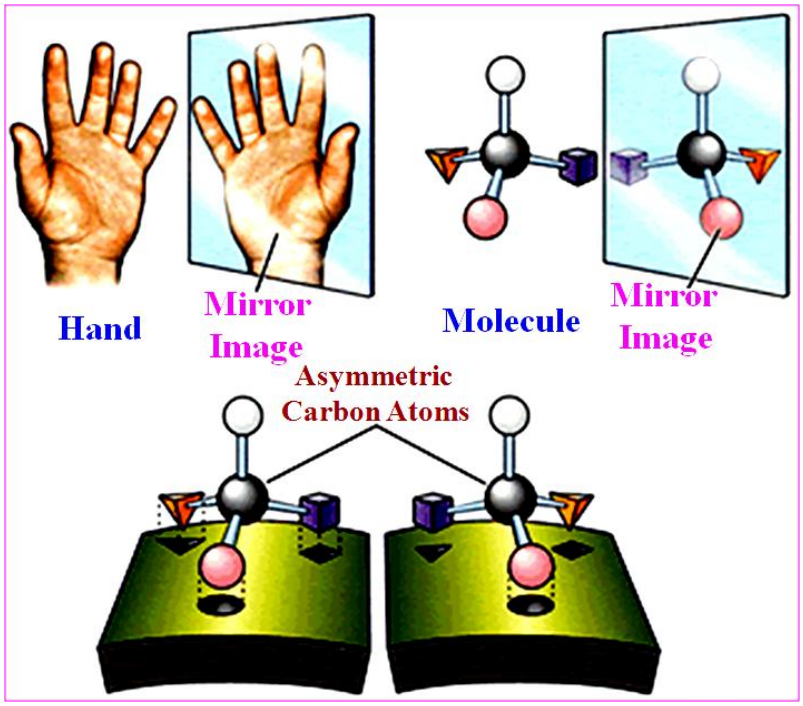
2.2.6. Non-Essential Amino Acids in Humans

- Not required in diet
- Can be formed from α-keto acids by transamination and subsequent reactions

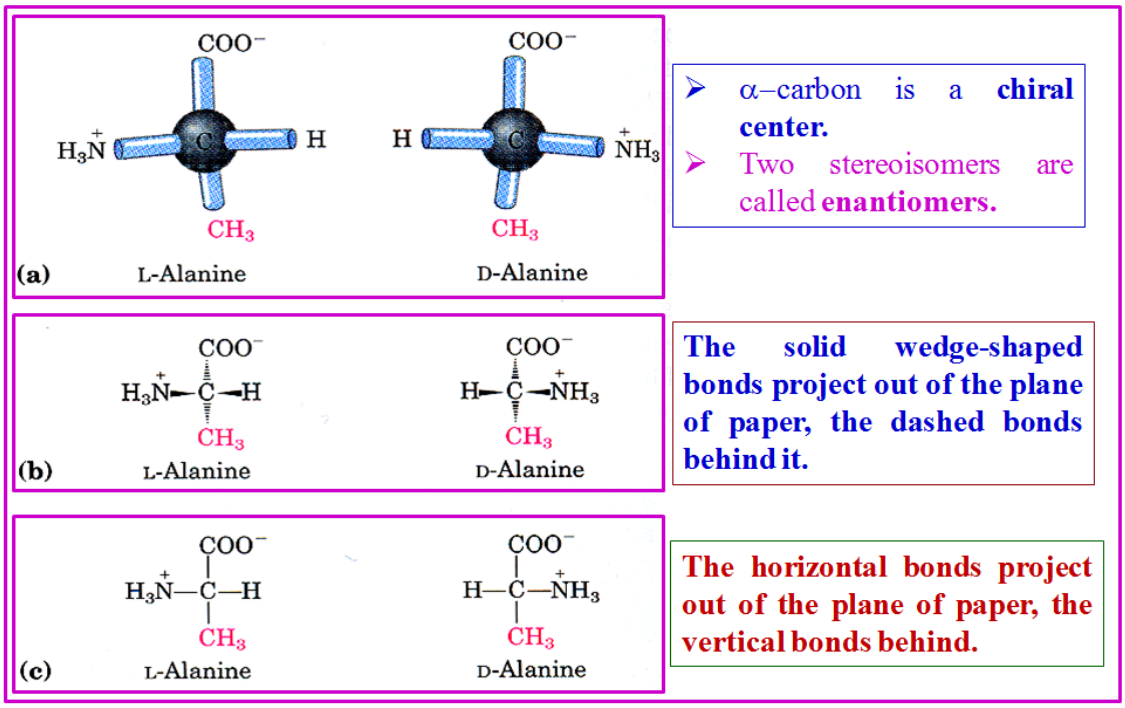
Alanine	Glycine
Asparagine	Proline
Aspartate	Serine
Glutamate	Cysteine (from Met*)
Glutamine	Tyrosine (from Phe*)

* Essential amino acids

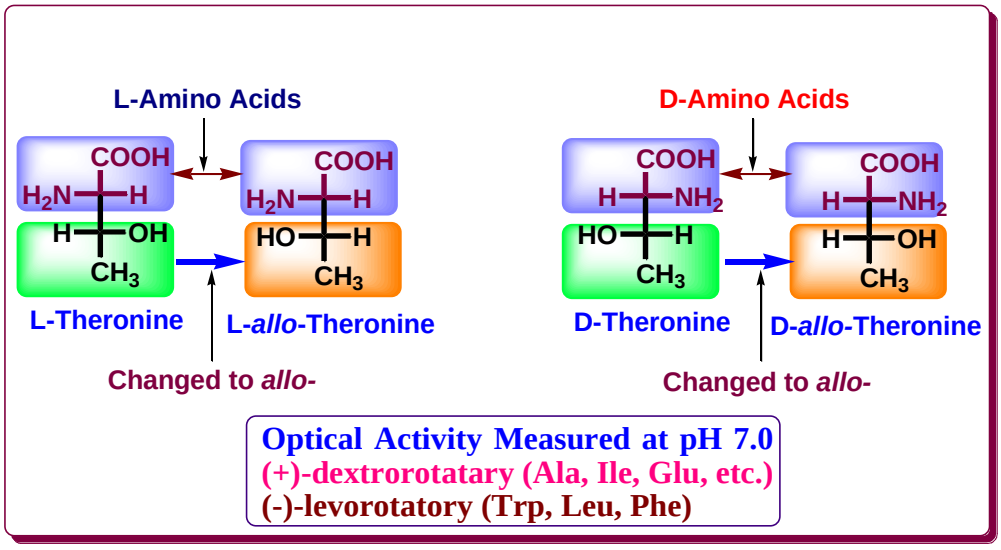
2.2.7. The Stereochemistry of Amino Acids



2.2.7.A. The Two Stereoisomers of Alanine



2.2.7.B. Stereoisomers of Threonine



2.2.8. Uncommon Amino Acids Found in Proteins

- Not incorporated by ribosomes
- Arise by post-translational modifications of proteins
- Reversible modifications, especially phosphorylation is important in regulation and signaling

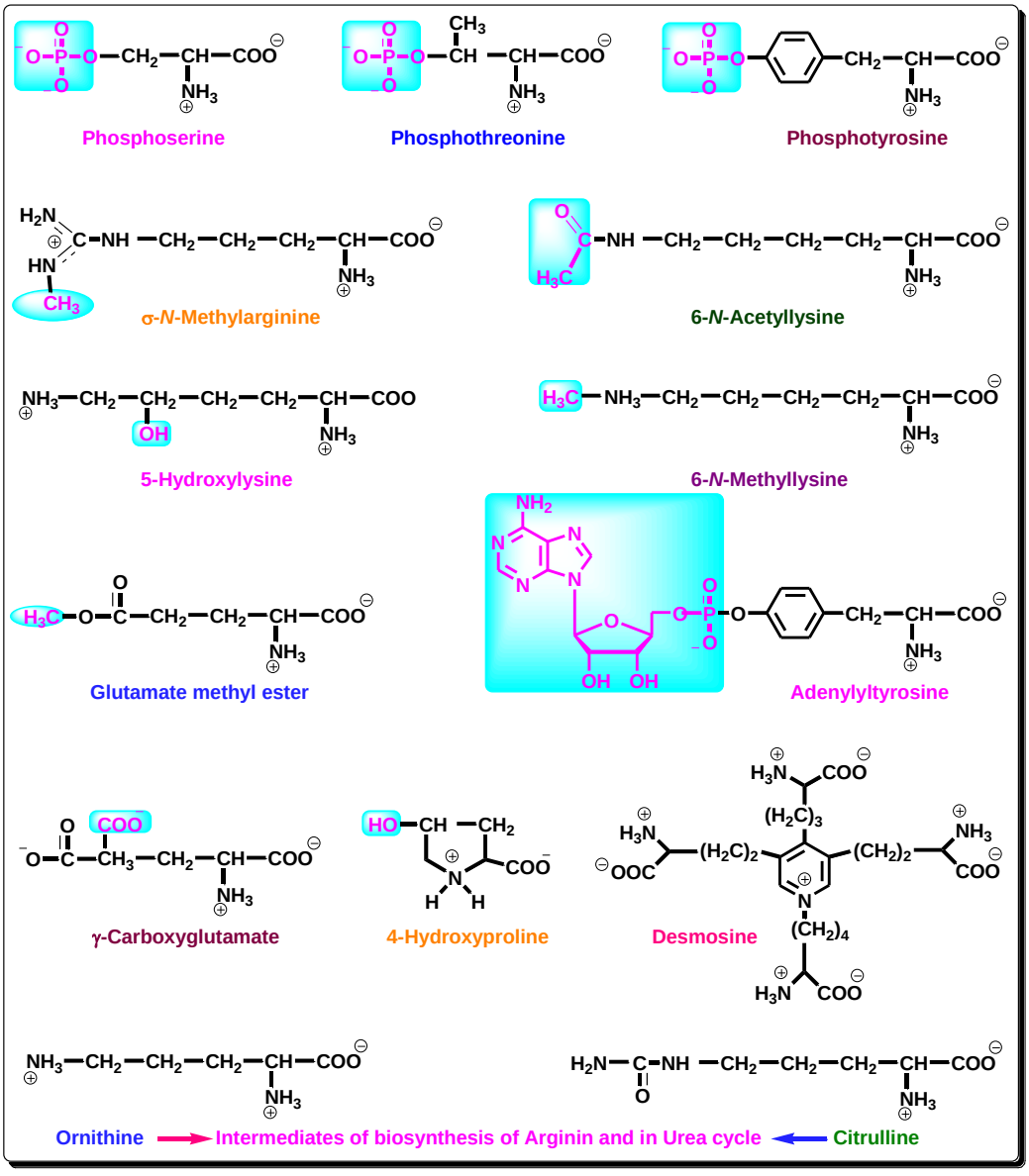


Figure 2.11: Structures of uncommon amino acids (post translational modification).

2.2.9. The Genetic Code: DNA-Amino Acid Dictionary

The Genetic Code: DNA-Amino Acid Dictionary

nonpolar	polar	basic	acidic	(stop codon)

		2nd base							
		T		C		A		G	
1st base	T	TTT	(Phe/F) Phenylalanine	TCT	(Ser/S) Serine	TAT	(Tyr/Y) Tyrosine	TGT	(Cys/C) Cysteine
		TTC	(Phe/F) Phenylalanine	TCC	(Ser/S) Serine	TAC	(Tyr/Y) Tyrosine	TGC	(Cys/C) Cysteine
		TTA	(Leu/L) Leucine	TCA	(Ser/S) Serine	TAA	Ochre (*Stop*)	TGA	Opal (*Stop*)
		TTG	(Leu/L) Leucine	TCG	(Ser/S) Serine	TAG	Amber (*Stop*)	TGG	(Trp/W) Tryptophan
	C	CTT	(Leu/L) Leucine	CCT	(Pro/P) Proline	CAT	(His/H) Histidine	CGT	(Arg/R) Arginine
		CTC	(Leu/L) Leucine	CCC	(Pro/P) Proline	CAC	(His/H) Histidine	CGC	(Arg/R) Arginine
		CTA	(Leu/L) Leucine	CCA	(Pro/P) Proline	CAA	(Gln/Q) Glutamine	CGA	(Arg/R) Arginine
		CTG	(Leu/L) Leucine	CCG	(Pro/P) Proline	CAG	(Gln/Q) Glutamine	CGG	(Arg/R) Arginine
	A	ATT	(Ile/I) Isoleucine	ACT	(Thr/T) Threonine	AAT	(Asn/N) Asparagine	AGT	(Ser/S) Serine
		ATC	(Ile/I) Isoleucine	ACC	(Thr/T) Threonine	AAC	(Asn/N) Asparagine	AGC	(Ser/S) Serine
		ATA	(Ile/I) Isoleucine	ACA	(Thr/T) Threonine	AAA	(Lys/K) Lysine	AGA	(Arg/R) Arginine
		ATG[A]	(Met/M) Methionine	ACG	(Thr/T) Threonine	AAG	(Lys/K) Lysine	AGG	(Arg/R) Arginine
	G	GTT	(Val/V) Valine	GCT	(Ala/A) Alanine	GAT	(Asp/D) Aspartic acid	GGT	(Gly/G) Glycine
		GTC	(Val/V) Valine	GCC	(Ala/A) Alanine	GAC	(Asp/D) Aspartic acid	GGC	(Gly/G) Glycine
		GTA	(Val/V) Valine	GCA	(Ala/A) Alanine	GAA	(Glu/E) Glutamic acid	GGA	(Gly/G) Glycine
		GTG	(Val/V) Valine	GCG	(Ala/A) Alanine	GAG	(Glu/E) Glutamic acid	GGG	(Gly/G) Glycine

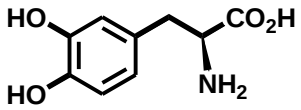
[A] **The codon ATG both codes for methionine and serves as an initiation site: the first ATG in an DNA's coding region is where translation into protein begins.** 21st AA, Selenocysteine, codes "UGA" opal (or umber) stop codon and 22nd AA, Pyrrolysine, Codes "GUA" amber stop codon

.

2.2.10. Synthesis of α-Amino Acids

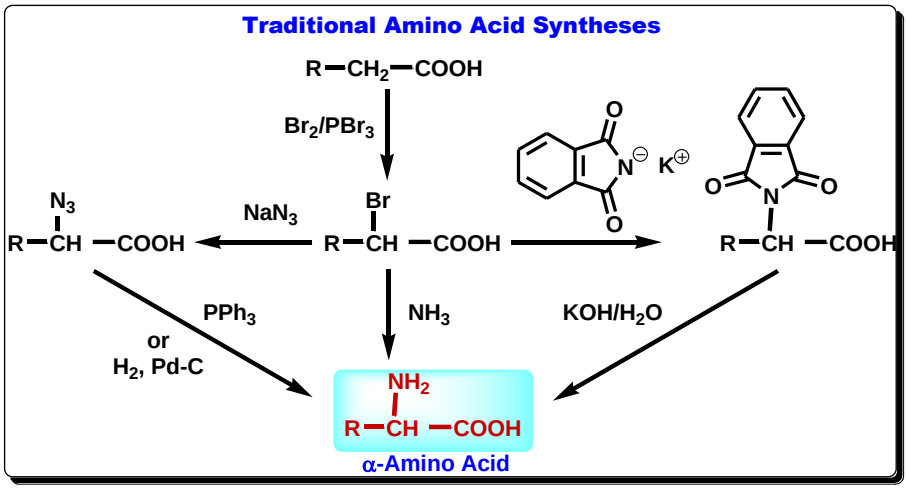
New unnatural amino acids with altered properties

- new therapeutics (lead compounds)
- mechanistic probes
-

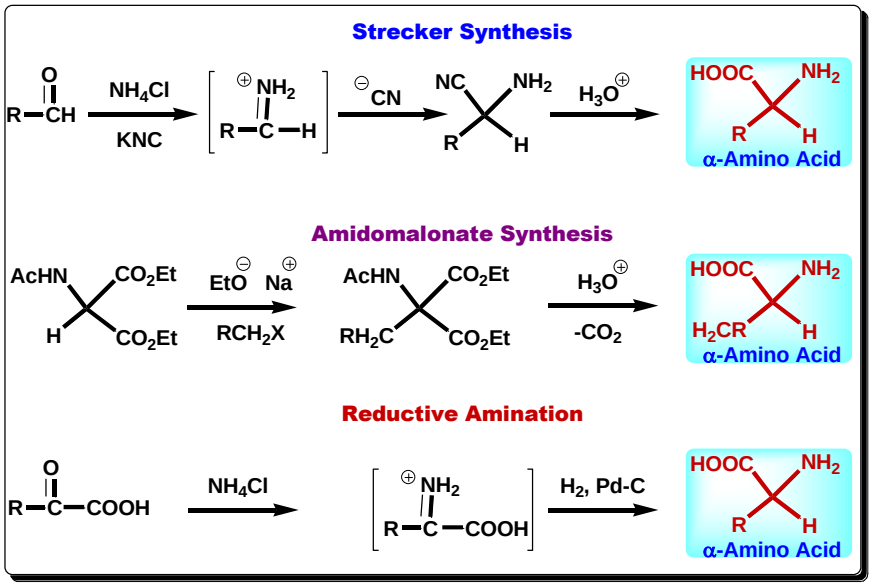


L-DOPA

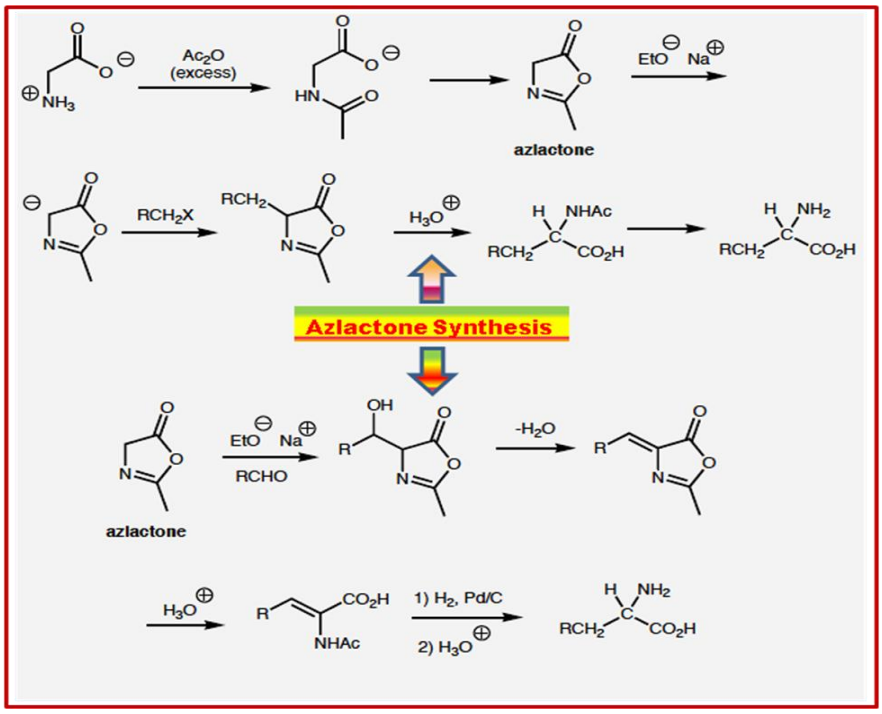
2.2.10.A. Traditional Synthesis of α-Amino Acids



2.2.10.B. Strecker and Other Methods of Synthesis of α-Amino Acids



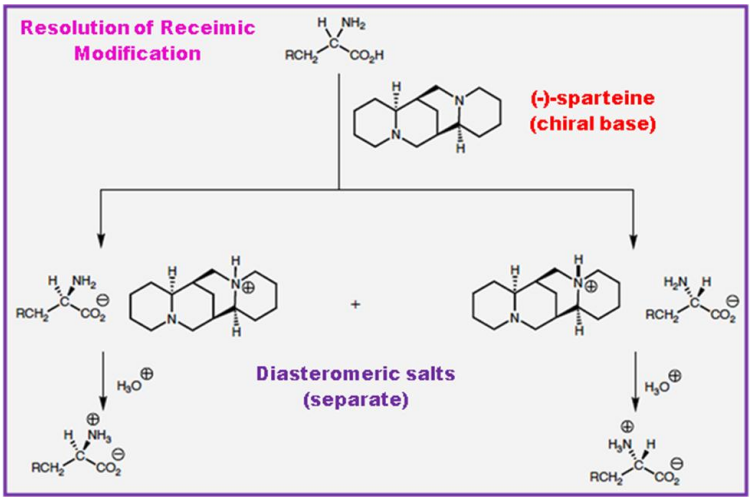
2.2.10.C. Azalactone Synthesis of α-Amino Acids



These all the above methods of amino acid synthesis are racemic syntheses!!!

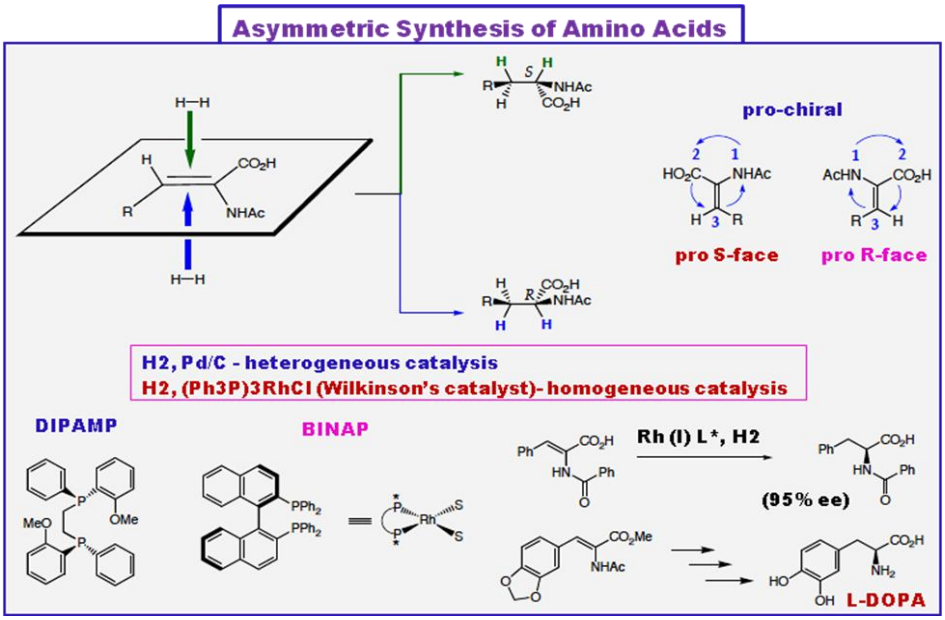
- **So, the two enantiomers need to be separated.**
- **Resolution: separation of enantiomers.**

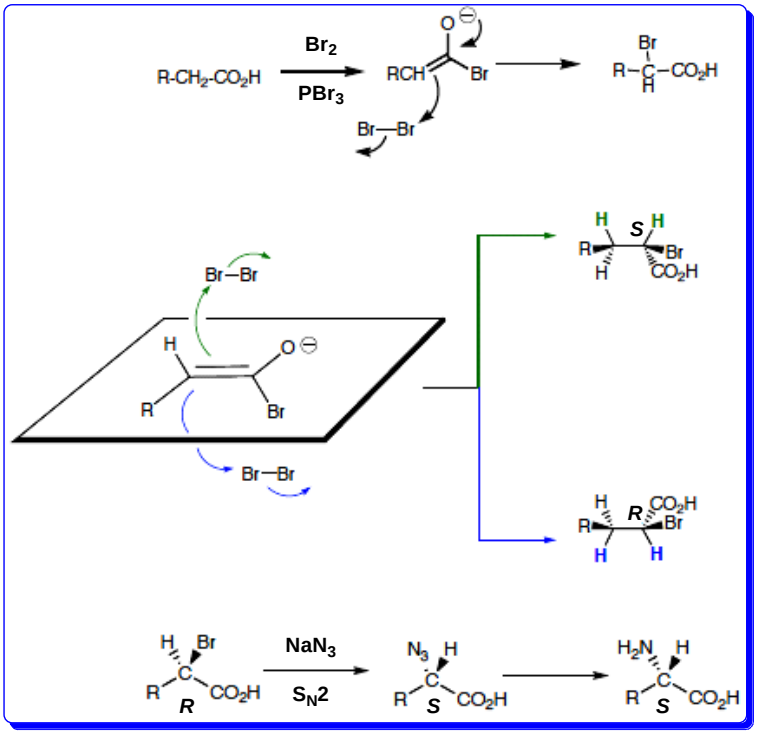
2.2.10.D. Resolution Method to the Synthesis of Asymmetric α-Amino Acids



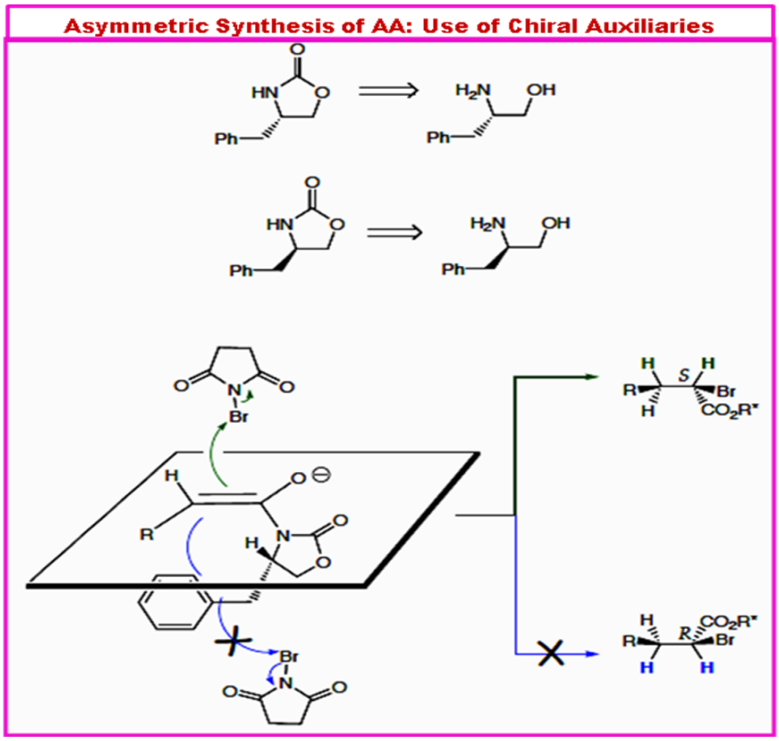
2.2.11. Asymmetric Synthesis of Amino Acids

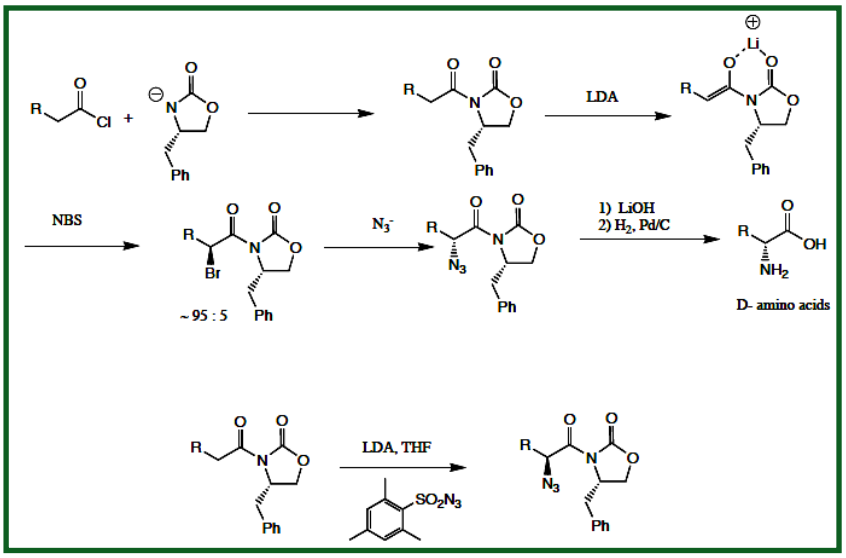
2.2.11.A. Asymmetric Synthesis of Amino Acids: Use of Rh-Catalyst





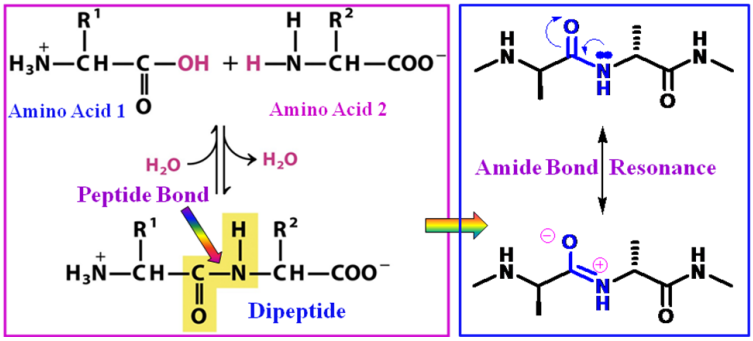
2.2.11.B. Asymmetric Synthesis of Amino Acids: Use of Chiral Auxiliaries





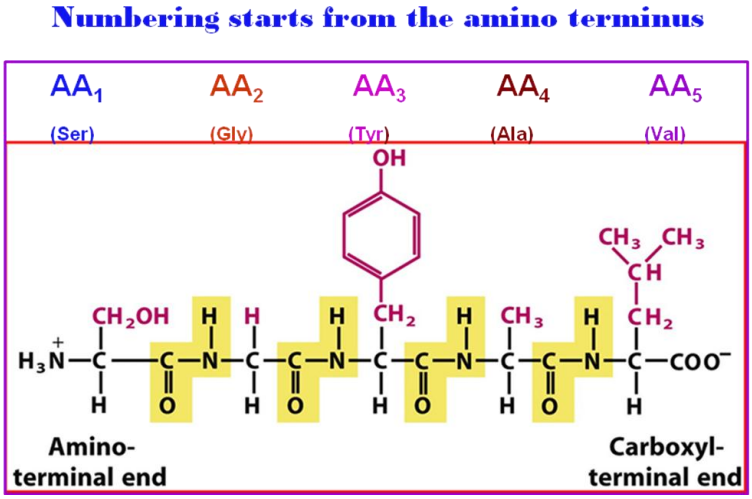
2.3. Chemistry of Peptide Bonds

- **"Peptides" are small condensation products of amino acids**
- **They are "small" compared to proteins (di, tri, tetra… oligo-).**
- C=N double bond character due to the resonance structure
- Restricted rotations about C=N resists hydrolysis



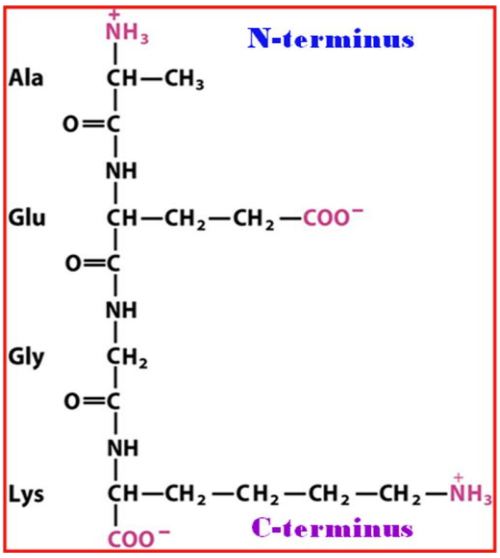
2.3.1. Reading the AA in a Peptide Chain

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



2.3.2. Naming a Peptide Sequence

1. **Naming starts from the N-terminus**
2. **Sequence is written as: Ala-Glu-Gly-Lys**
3. **Sometimes the one-letter code is used: AEGK**



2.3.3. Peptides: A Variety of Functions

• **Hormones and pheromones** ➢ insulin (think sugar) ➢ oxytocin (think childbirth) ➢ sex-peptide (think fruit fly mating)	• **Neuropeptides** ➢ substance P (pain mediator)
• **Antibiotics** ➢ polymyxin B [for Gram (-) bacteria] ➢ bacitracin [for Gram (+) bacteria]	• **Protection, e.g. toxins** ➢ amanitin (mushrooms) ➢ conotoxin (cone snails) ➢ chlorotoxin (scorpions)

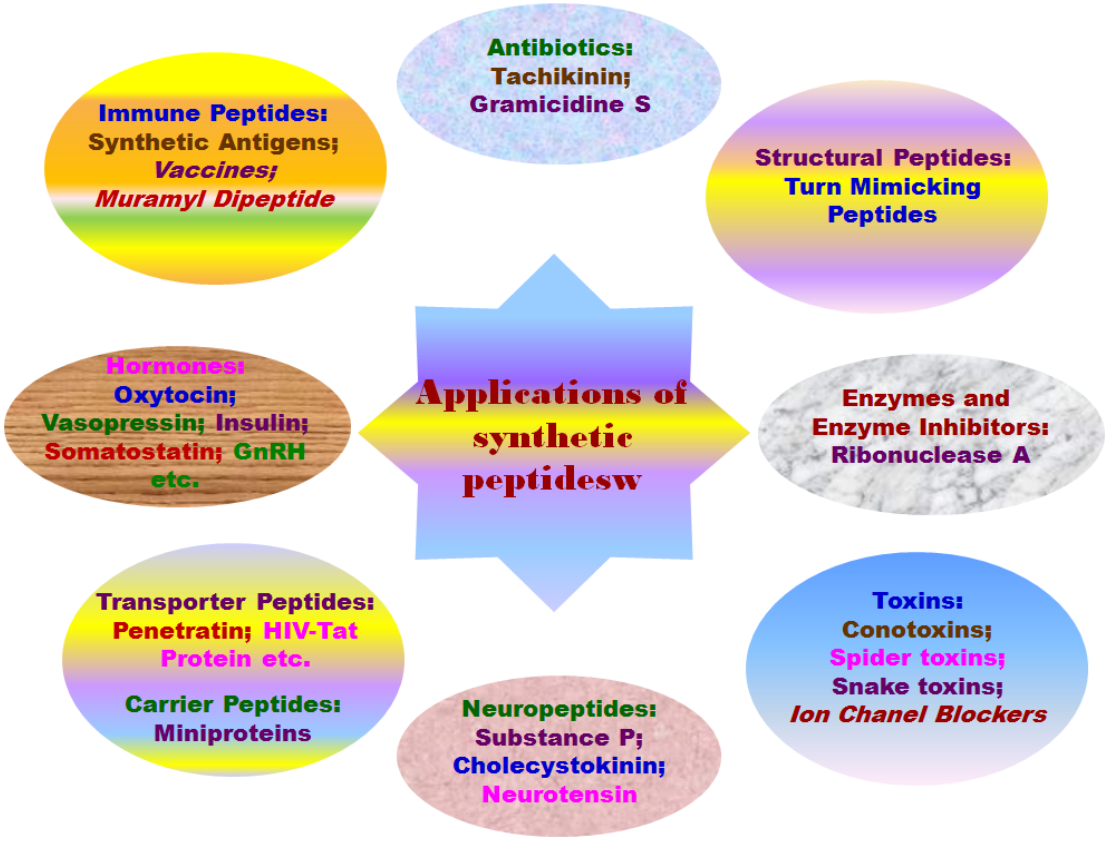
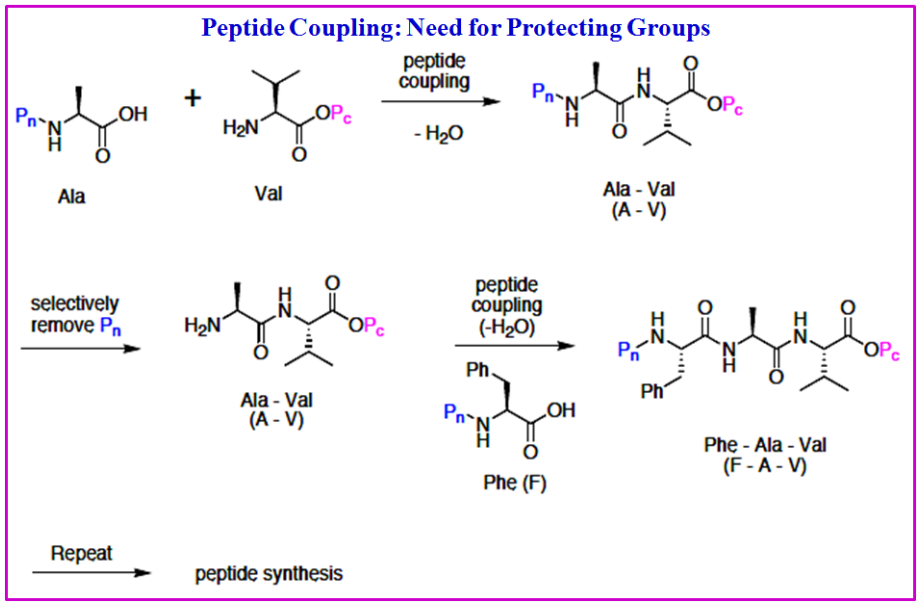
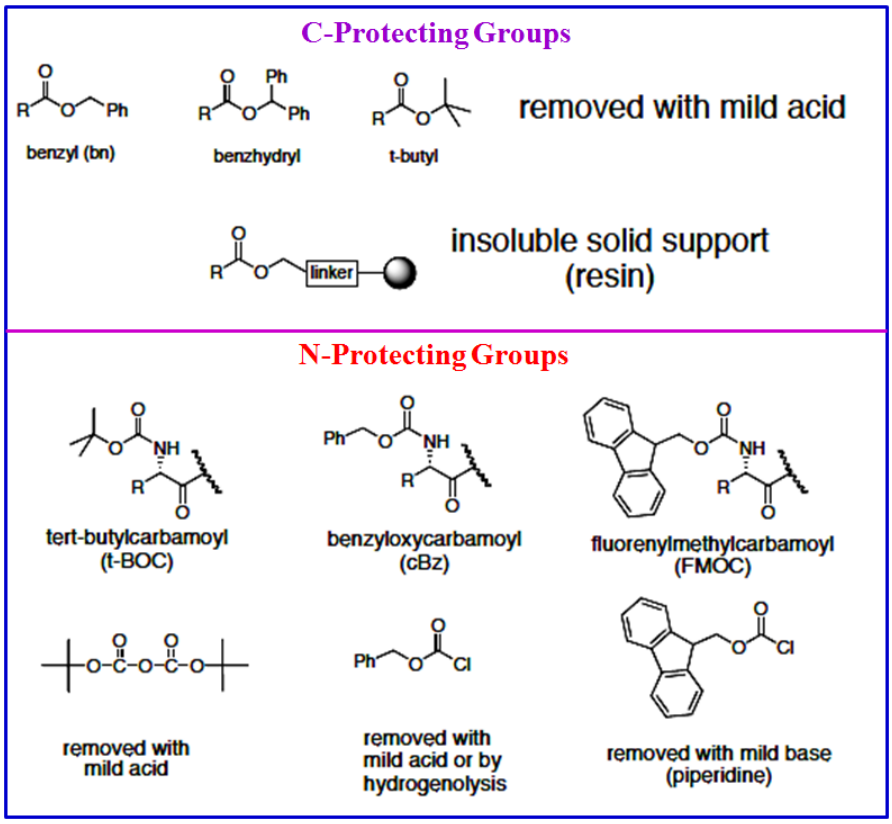


Figure 2.12: Schematic presentation to show a variety of applications of peptides.

2.3.4. Peptide Coupling: Need for Protecting Groups



2.3.5. The Protecting Groups



2.3.6. Solid-Phase Peptide Synthesis (SPPS)

- Peptides up to ~ 100 amino acids can be synthesized in a laboratory
- Laboratory synthesis is from the C-terminus to the N-terminus
- Nature synthesizes peptides from N to C.

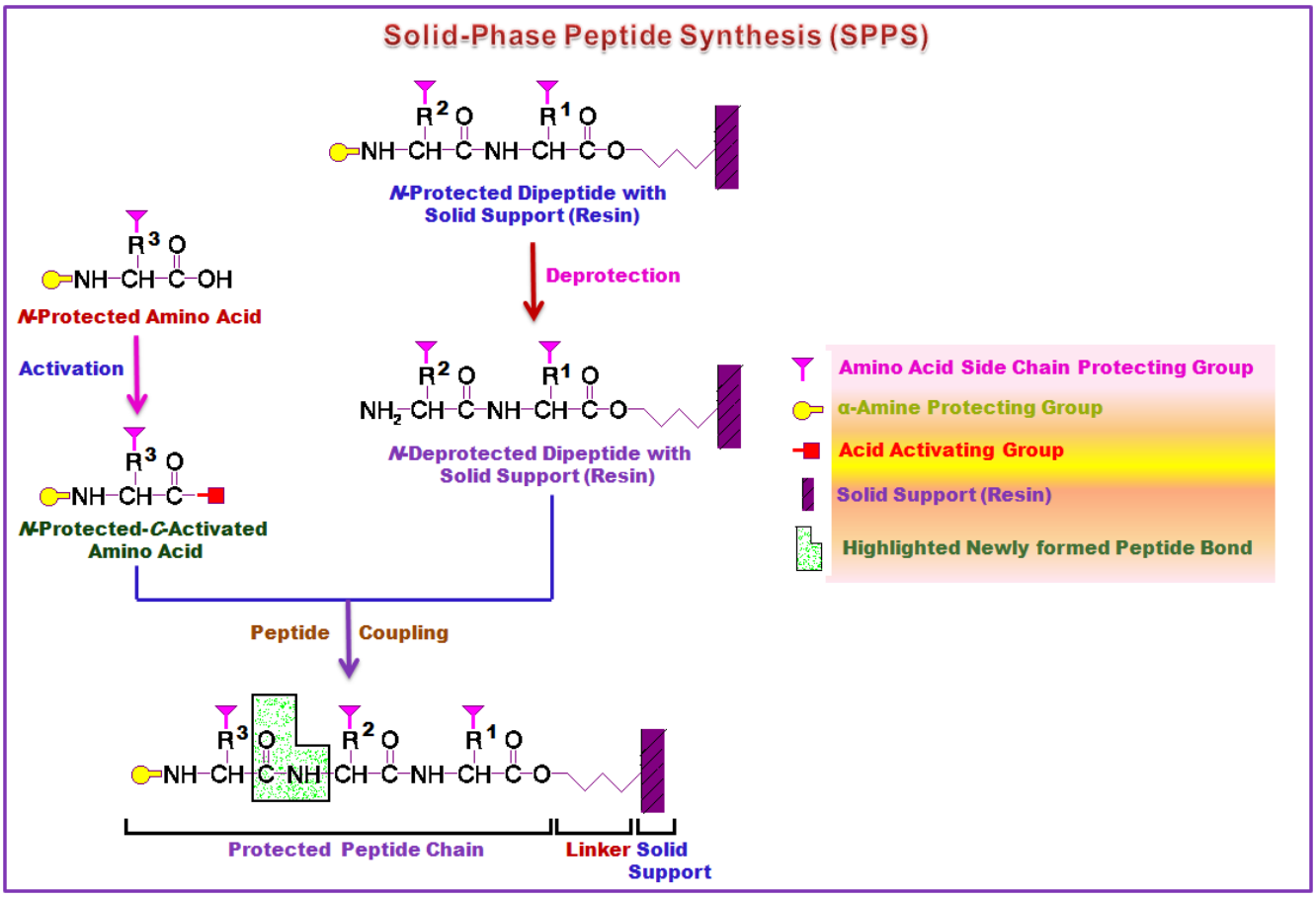
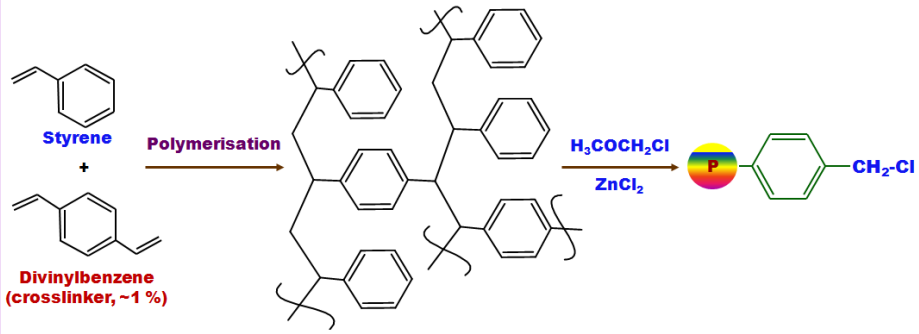


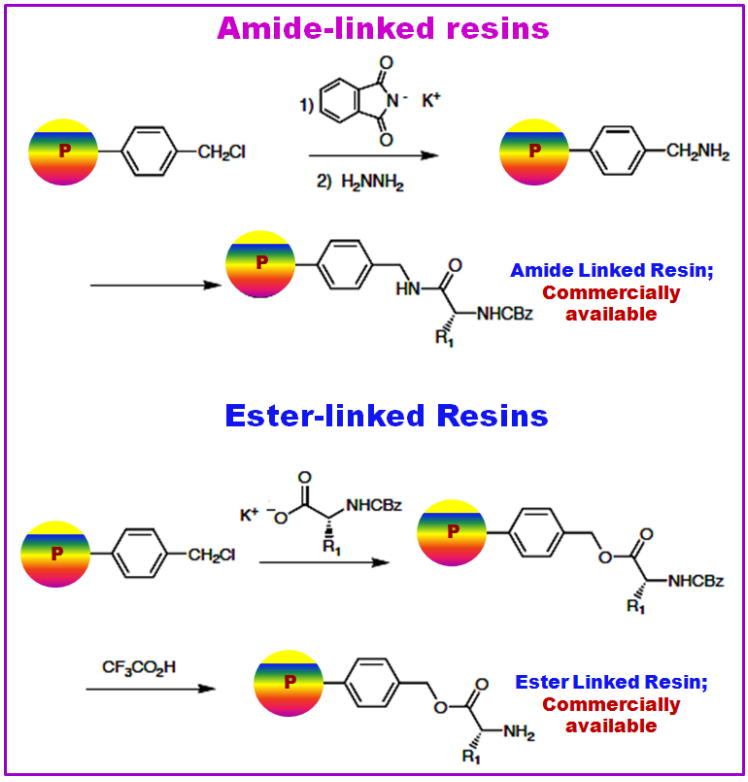
Figure 2.13: Solid phase peptide synthesis protocol

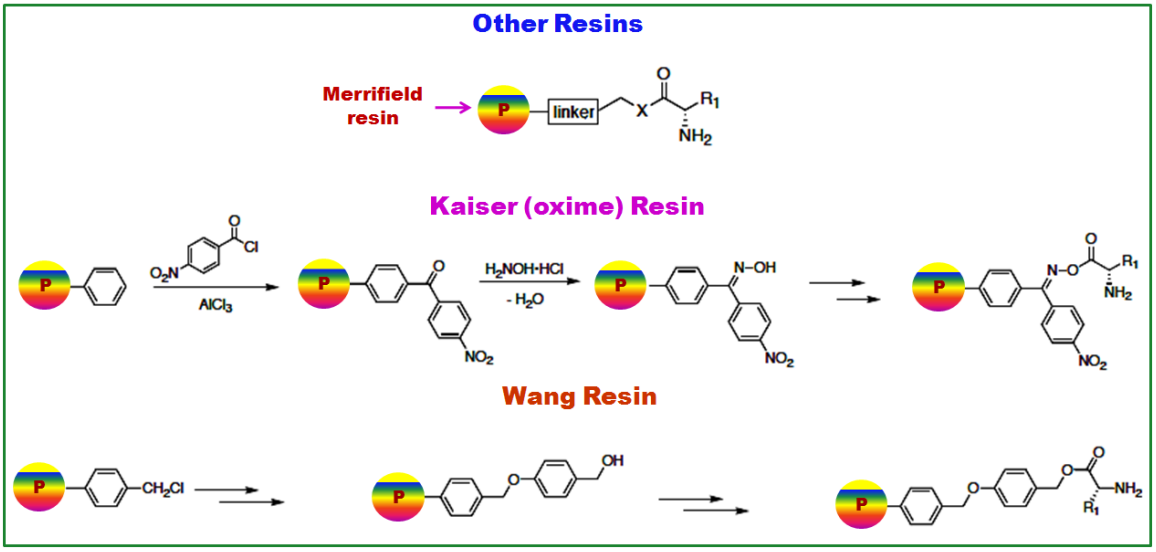
2.3.6.A. Solid-Phase Peptide Synthesis: The solid Support

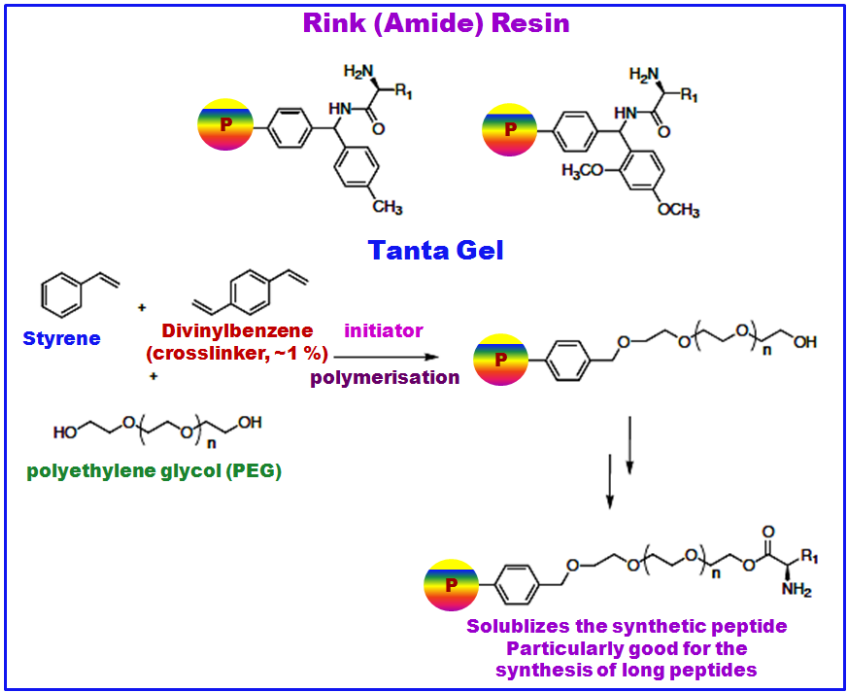
- Can be functionalised;
- Chemical stability (it must be inert to all applied chemicals);
- Mechanical stability (it shouldn't brake under stirring);
- It must swell extensively in the solvents used for the synthesis;
- Peptide-resin bond should be stable during the synthesis;
- Peptide-resin bond can be cleaved effectively at the end of the synthesis;
- The basic of the most common used resins is polystyrene-1,4-divinylbenzene (1-2%) copolymer

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



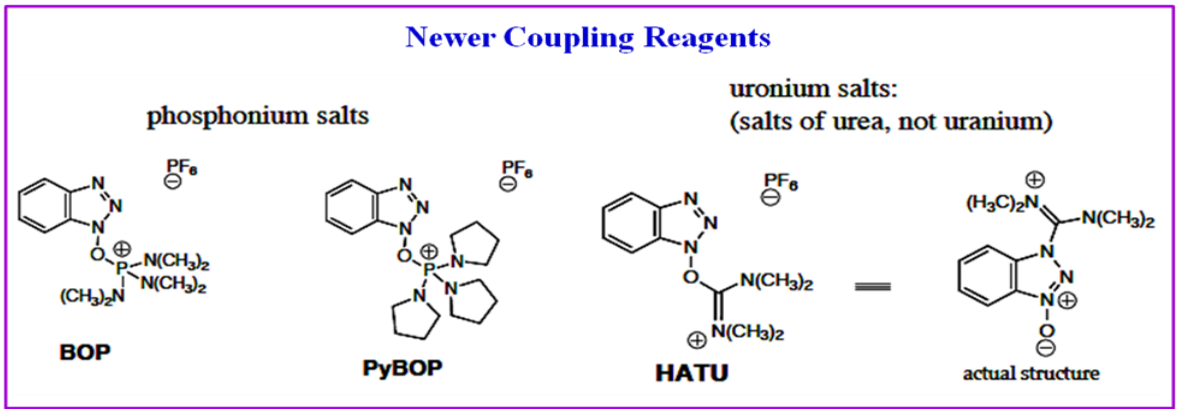




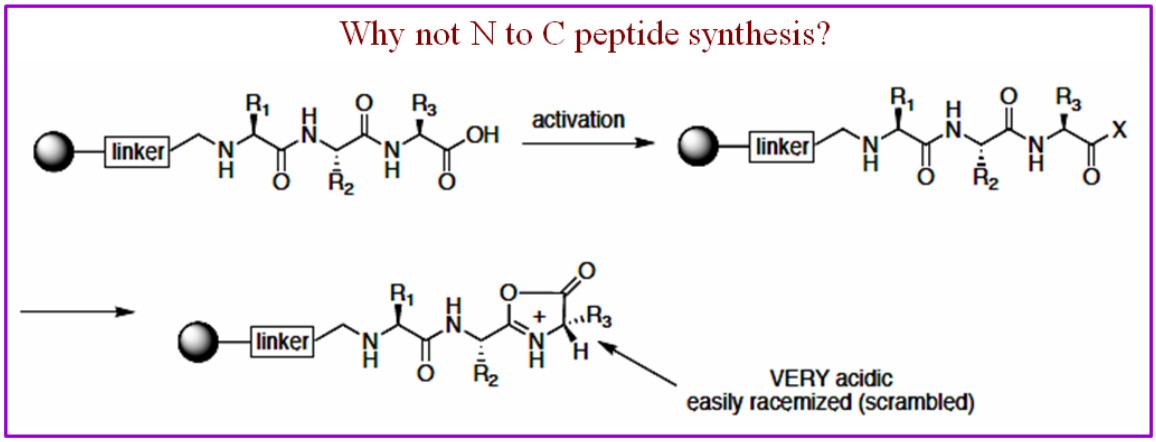


2.3.6.B. Peptide Coupling Reagent:

- **N-protected carboxylic acid,**
- **C-protected amine**
- **DCC, HOBT**

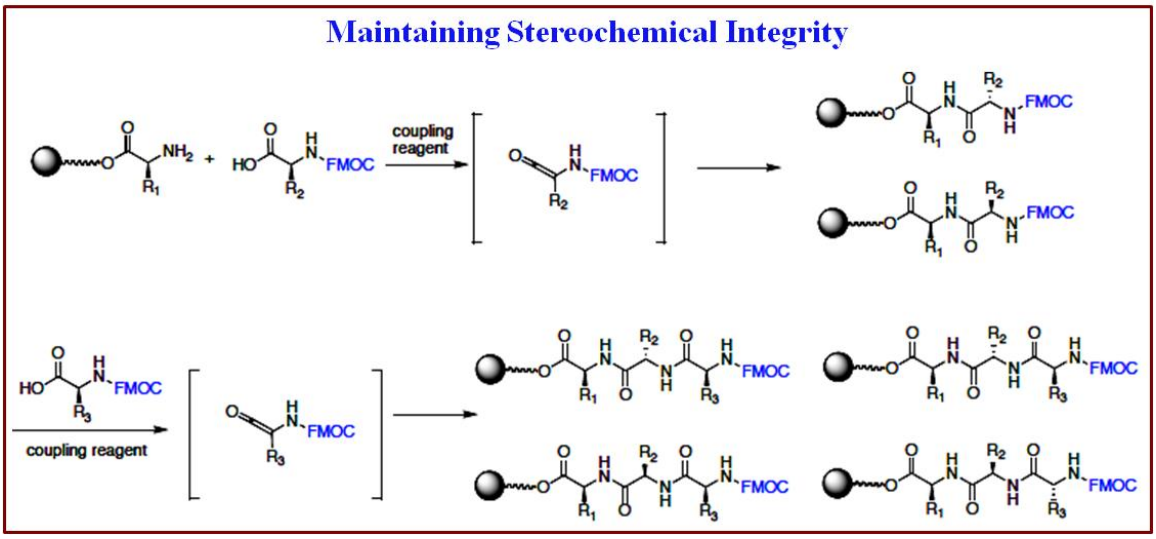


2.3.6.C. Why not N to C peptide synthesis?



2.3.6.D. Importance of Maintaining Stereochemical Integrity During the Coupling Step

- Number of possible stereoisomers = 2n where n= no. of chiral centers
- A peptide w/ 10 AA residues has 210 possible stereoisomers.



2.3.6.D. Proteins are:-

- **Polypeptides (covalently linked α-amino acids)**

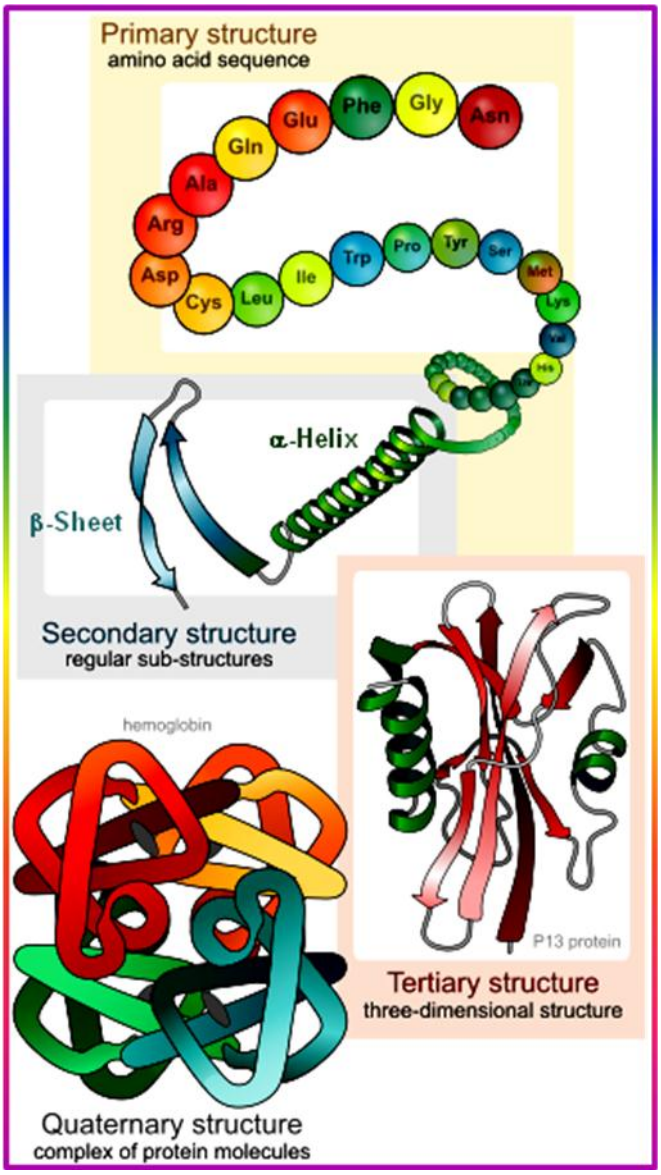
+ possibly –

- cofactors,
- coenzymes,
- prosthetic groups,
- other modifications

- **Cofactor** is a general term for functional non-amino acid component
  - – Metal ions or organic molecules
- **Coenzyme** is used to designate an organic cofactors
  - – NAD^+ in lactate dehydrogenase
- **Prosthetic groups** are covalently attached cofactors
  - – Heme in myoglobin

2.4. Peptide Secondary Structures and Tools for Stabilization

- **Primary (1°) Structure:** It tells about the amino acid sequence. The primary structure is held together by covalent or peptide bonds. The two ends of the polypeptide chain are referred to N-terminus and C-terminus end in a protein. Edman degradation or mass spectrometry can help in sequence sequence determination of a protein. However, it can be read directly from the sequence of the gene using the genetic code.



- **Secondary (2°) Structure:** Frequently occurring substructures or folds in a polypeptide is the concern of Secondary structure. In water a polypeptide chain will not stay in an elongated form, but fold up according to the polarity of the side chains it contains and the rotation of peptide backbone determined by van der Waals radii of side chains.

  To understand protein structures we can measure two *torsion angles/*dihedral angles in the backbone which define the tilt between two neighboring amide planes (the plane of the peptide bond) with the Cα at the center of rotation:

  Φ Cα – N Ψ Cα - C

  Thus we can calculate the backbone conformations of a peptide through interplay of rotation around the bonds defined by the torsion angles Phi (Φ) and Psi (Ψ) and the steric hindrance of side groups determined by their Van der Waals radii. The resulting conformational map is called *Ramachandran plot*. A free rotation because of thermal motion around a C-C bond is possible in the absence of any steric *i.e.,* Van der Waals radii constraints.

- **Tertiary (3°) Structure:** three-dimensional arrangement of all atoms in a single polypeptide chain is the main feature of tertiary structure. The tertiary structure is the 3-dimensional, native structure of a single polypeptide or protein. A protein normally is folded into a compact structure. The secondary structures are stabilized by the final, native fold. The native fold is defined as the *active conformation which is mostly stabilized by* H-bonding, Electrostatic or Ionic, Hydrophobic and Van der Waals forces of interaction and some time by Covalent interaction.

  The location of an amino acid in a protein fold correlates with the energy of hydration of each individual amino acid residue and the entropy of the side chain, polypeptide backbone, and solvent molecules. As mentioned earlier, polar and charged amino acids are likely to be hydrated, whereas the non-polar residues stick to each other and often form the core of a protein, forming a usually hydrophobic core that stabilizes the fold of water soluble, globular proteins. For membrane proteins exhibiting both hydrophobic and hydrophilic surfaces, the amino acid distribution is different from globular proteins but the same rules of the hydrophobic effect apply.

- **Quaternary (4°) structure:** It is the overall organization of non-covalently linked subunits of a functional protein. Single polypeptides can associate with each other to form larger ***protein complexes*** of geometrically specific arrangements, called quaternary structures. Individual polypeptides in protein complexes are referred to as ***subunits***. Most enzymes are complexes of proteins and the symmetry and stoichiometry of the composition of the complexes is crucial for their activity.

  One can distinguish two different compositions, the homomeric and heteromeric complexes. Heteromeric composition of most protein complexes gives the cells an additional level of variability and complexity it can use for its activity. Often, heteromeric compositions of protein complexes are tissue specific or developmental specific and multiple genes can control the activity of a single heteromeric protein complex.

- **Primary Structural Motifs:**
  - **α-Helix:** a right handed helical structure with average torsion angles Φ = -57 and Ψ = -47
  - **β-Sheet:** parallel (Φ = -119 and Ψ = 113) or anti-parallel pleated sheet structures
  - **β-Turn:** minimal loop structures of 3 to 4 amino acids with defined torsion angles
  - **Disulfide bonds:** Peptide or proteins may form covalent linkage between two cysteine amino acids.

- **Forces Stabilizing Peptide Structures:**
  - H-bonding
  - Electrostatic or Ionic
  - Hydrophobic and Van der Waals
  - Covalent

2.4.1. The α-Helix:

- Polypeptides chain folds along with their backbone into regular conformations which would similar to the α-keratin fiber. The most simple and elegant arrangement is a right-handed spiral conformation known as the 'α-helix'.
- Amino acids wound into a helical structure 3.6 amino acids per coil, 5.4 Å.
- **Properties of the α -helix.:** The structure repeats itself every 5.4 Å along the helix axis, *i.e.* the α-helix has a pitch of 5.4 Å. α-helices have 3.6 amino acid residues per turn, *i.e.* a helix of 36 amino acids long would form 10 turns. The separation of residues along the helix axis is 5.4/3.6 or 1.5 Å, *i.e.* the α-helix has a rise per residue of 1.5 Å.
  - Every main chain C=O and N-H group is hydrogen-bonded to a peptide bond 4 residues away (*i.e.* O_i to N_{i+4}). This gives a very regular, stable arrangement.
  - The peptide planes are roughly parallel with the helix axis and the dipoles within the helix are aligned, *i.e.* all C=O groups point in the same direction and all N-H groups point the other way. Side chains point outward from helix axis and are generally oriented towards its N-terminal end.

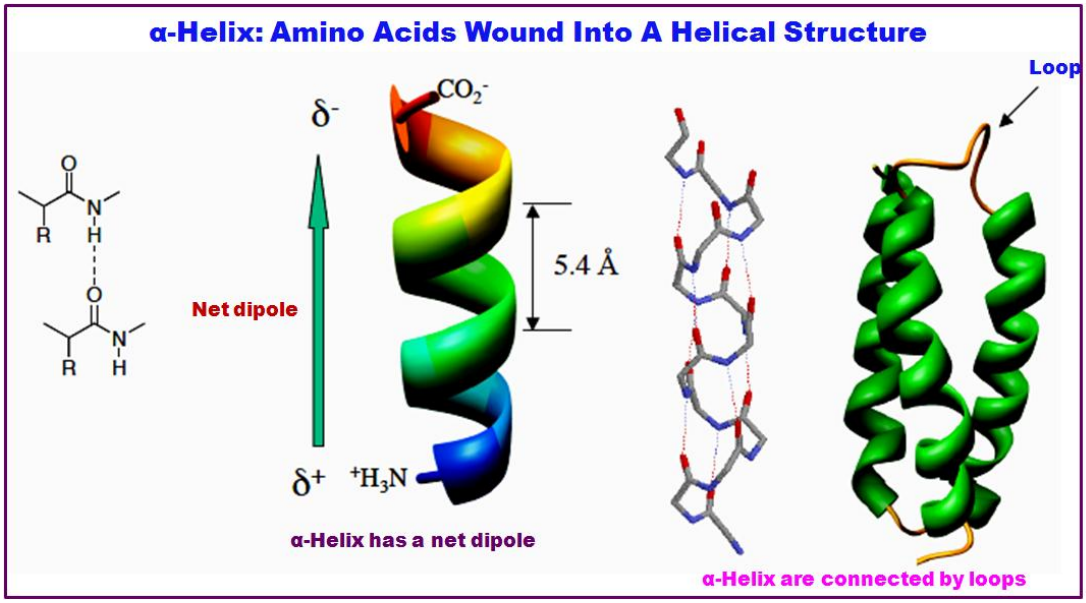


Figure 2.14: The structure of a α-helix.

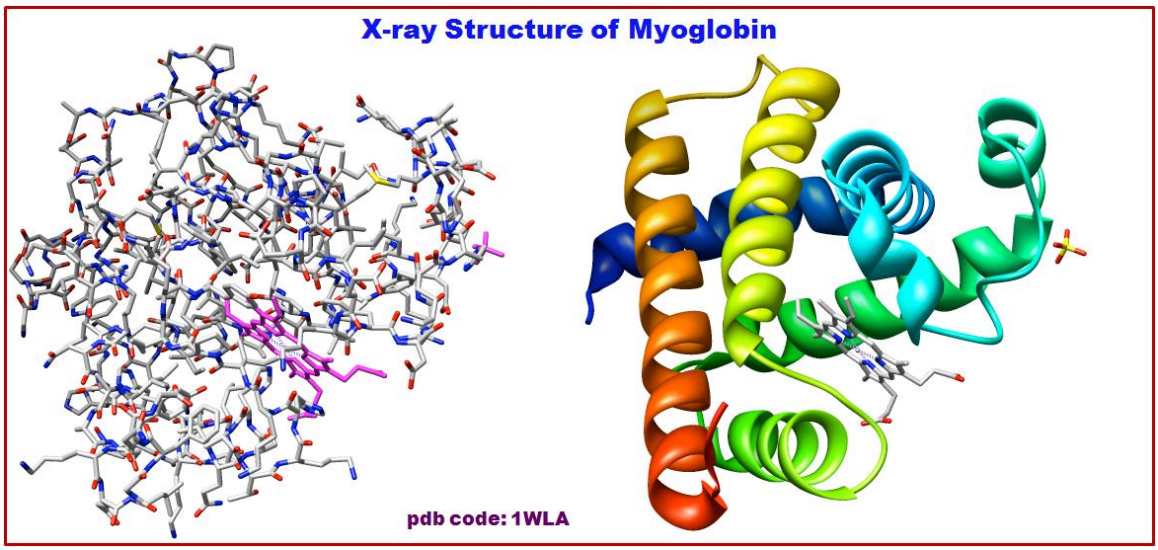


Figure 2.15: The structure of α-helix of Myoglobin.

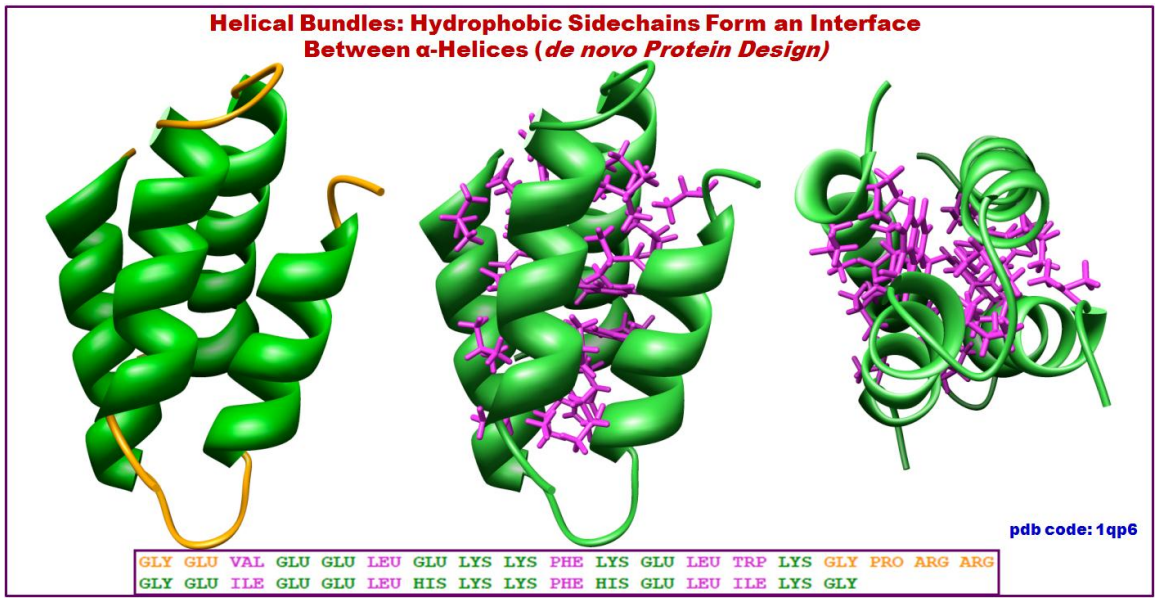


Figure 2.16: The structure of helix bundles.

2.4.2. β-Sheets

The **β-sheet** or **β-pleated sheet** is the second form of regular secondary structure in proteins. It is only somewhat less common than alpha helix. Beta sheets consist of **beta strands** which are connected laterally by at least two or three backbone hydrogen bonds. Thereby it forms a twisted, pleated sheet. A **β-strand** is a stretch of polypeptide chain with 3- to10-amino acids long with backbone in an almost fully extended conformation. The higher-level association of β-sheets has been implicated in formation of the protein aggregates and fibrils responsible for generating many human diseases, notably the amyloidoses such as Alzheimer's disease.

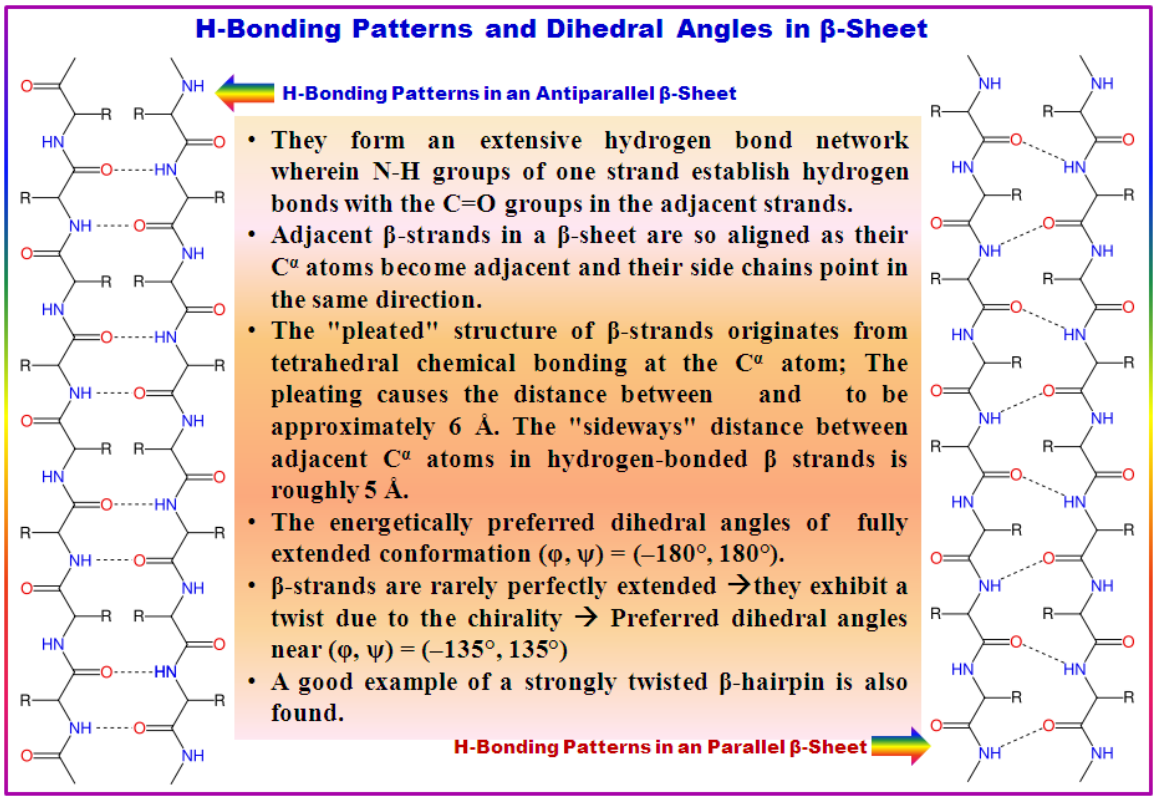


Figure 2.17: The structure of β-sheets-parallel/anti-parallel.

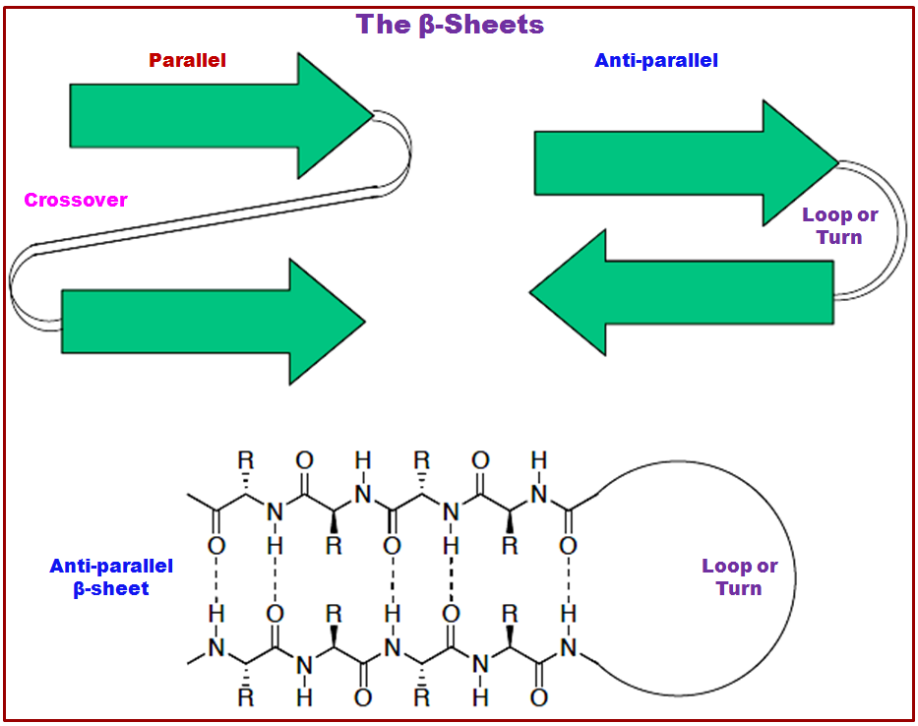


Figure 2.18: The structure of β-sheets-parallel/anti-parallel.

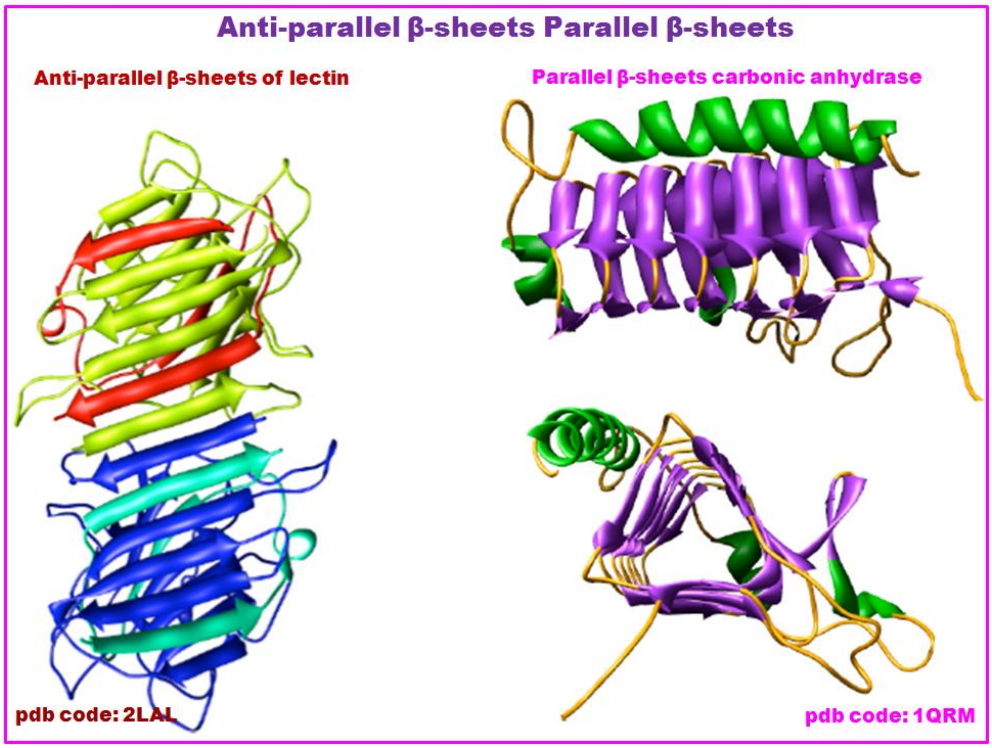


Figure 2.19: The structure of anti-parallel β-sheets of *Lectin* and parallel b-sheet of *Carbonic anhydrase*.

2.4.3. β-Turn

A region of the protein involving four consecutive residues where the polypeptide chain folds back on itself by nearly 180°. This chain reversal gives proteins a globular rather than linear structure. Thus a turn may be defined as is a structural motif where the C^{α} atoms of two residues separated by usually 1 to 5 peptide bonds are in closer in < 7 Å distant while the corresponding residues do not form a regular secondary structure element such as an alpha helix or beta sheet. The backbone dihedral angles are not constant for all the residues in the turn which in contrary to helices.

A turn can be converted into its **inverse turn** (almost *mirror-image turn*) by changing the sign on *all* of its dihedral angles. Thus, the γ-turn has two forms, a classical form with (φ, ψ) dihedral angles of roughly (75°, -65°) and an inverse form with dihedral angles (-75°, 65°). At least eight forms of the β-turn have been identified, varying mainly in whether a *cis* isomer of a peptide bond is involved and on the dihedral angles of the central two residues. Types VIa1, VIa2 and VIb turns are subject to the additional condition that residue (*i* + 2)(*) must be a *cis*-proline. Several turns are listed below in the table 1.

Table 1: Ideal Angles for Different Types of β-turn.

Type	φ_{i+1}	ψ_{i+1}	φ_{i+2}	ψ_{i+2}
I	-60	-30	-90	0
II	-60	120	80	0
VIII	-60	-30	-120	120
I'	60	30	90	0
II'	60	-120	-80	0
VIa1	-60	120	-90	0*
VIa2	-120	120	-60	0*
VIb	-135	135	-75	160*
IV	turns excluded from all the above categories			

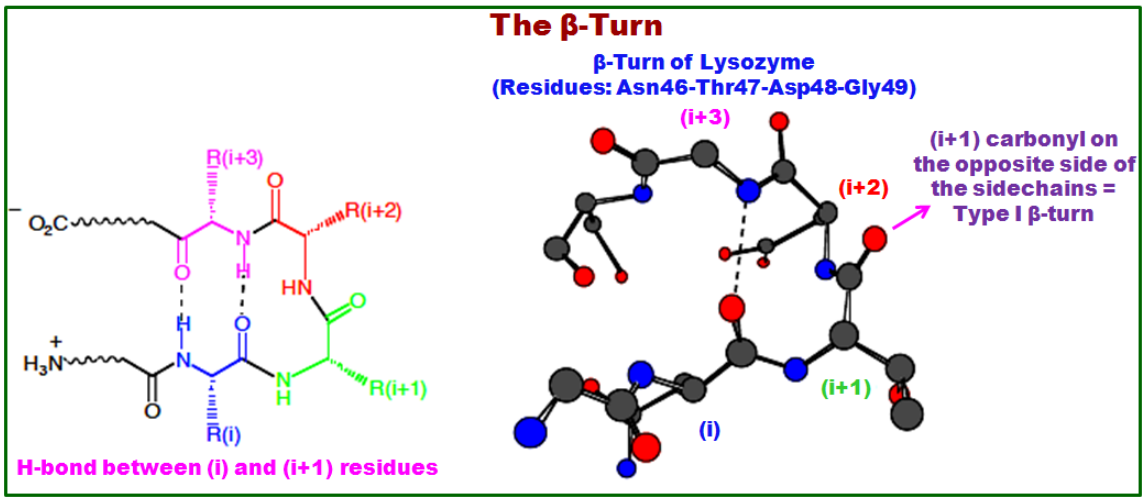


Figure 2.20: The H-bonding in β-turn structure and the β-turn of *Lysozyme*.

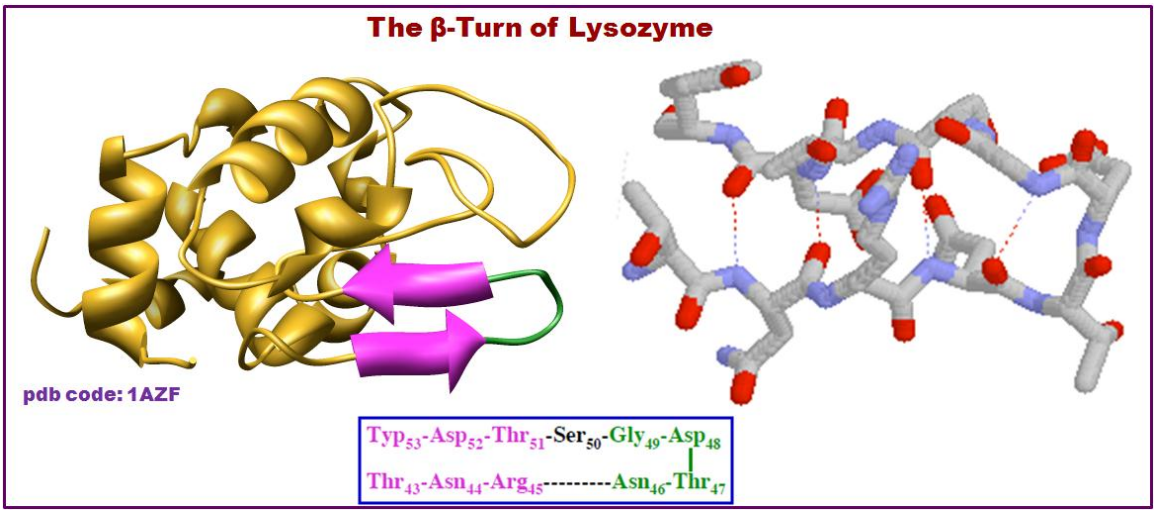


Figure 2.21: The H-bonding in the β-turn of *Lysozyme*.

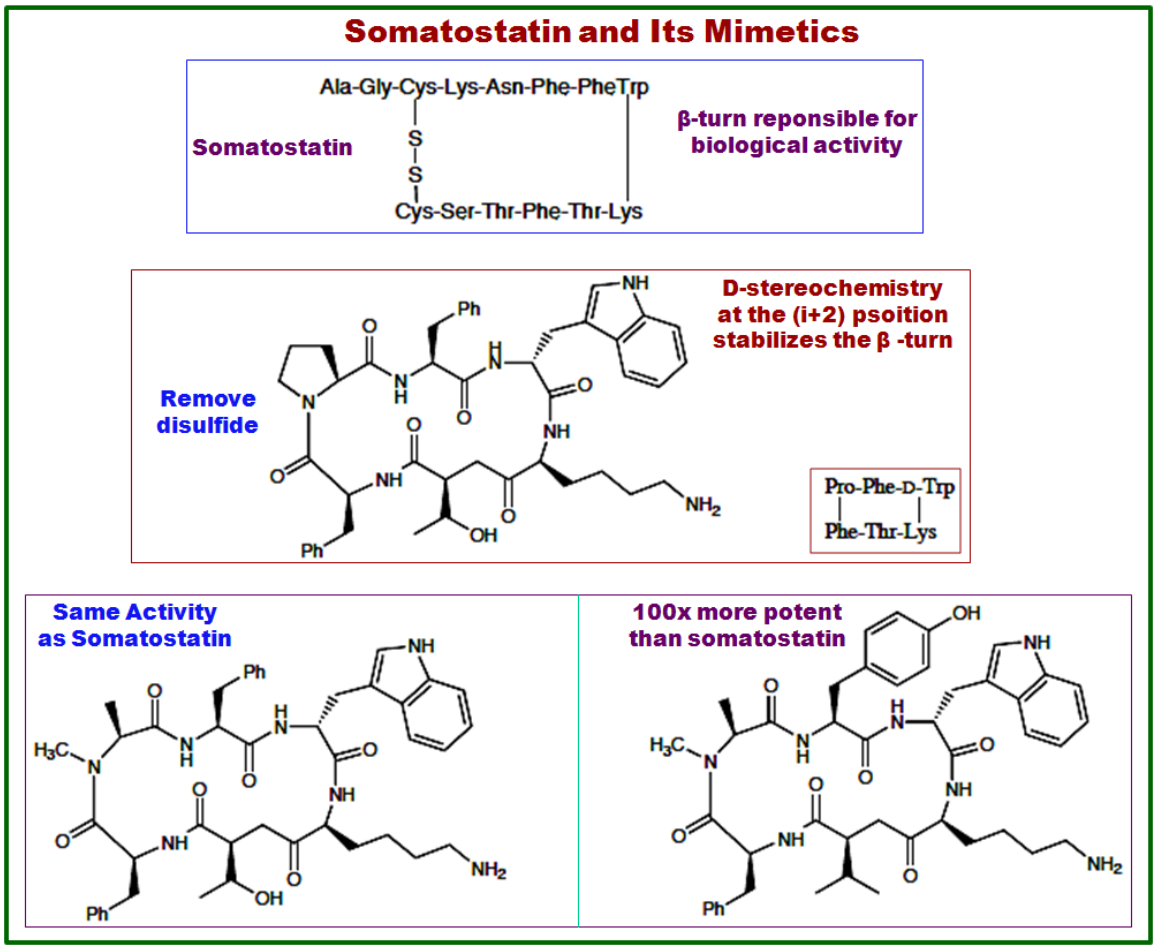


Figure 2.22: Structures of some β -turn mimetics.

2.4.4. Disulfide Bonds: Covalent Structural Scaffolds, Redox Active, Reversible

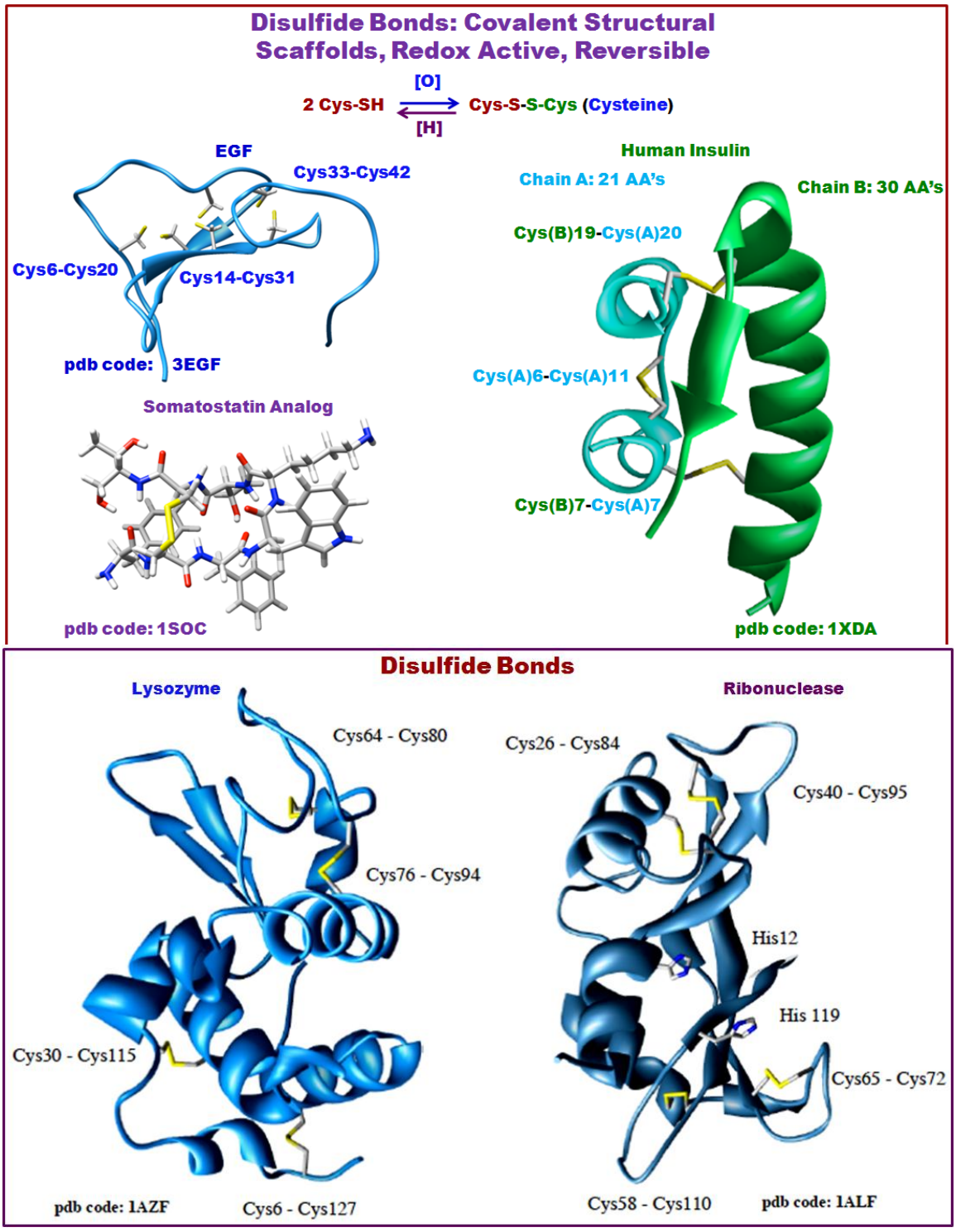


Figure 2.23: The covalent disulphide bonds in cysteine containing proteins.

2.5. Natural β-amino Acids and β-peptides

2.5.1. Introduction

β-peptides consist of β amino acids, which have their amino group bonded to the β- carbon rather than the α-carbon as in the 20 standard biological amino acids. The only commonly naturally occurring β-amino acid is β-alanine which is used as a component of larger bioactive molecules. However, β-peptides in general do not appear in nature. For this reason β-peptide-based antibiotics are being explored as ways of evading antibiotic resistance. Studies in this field were explored first in 1996 by the group of Dieter Seebach and that of Samuel Gellman.

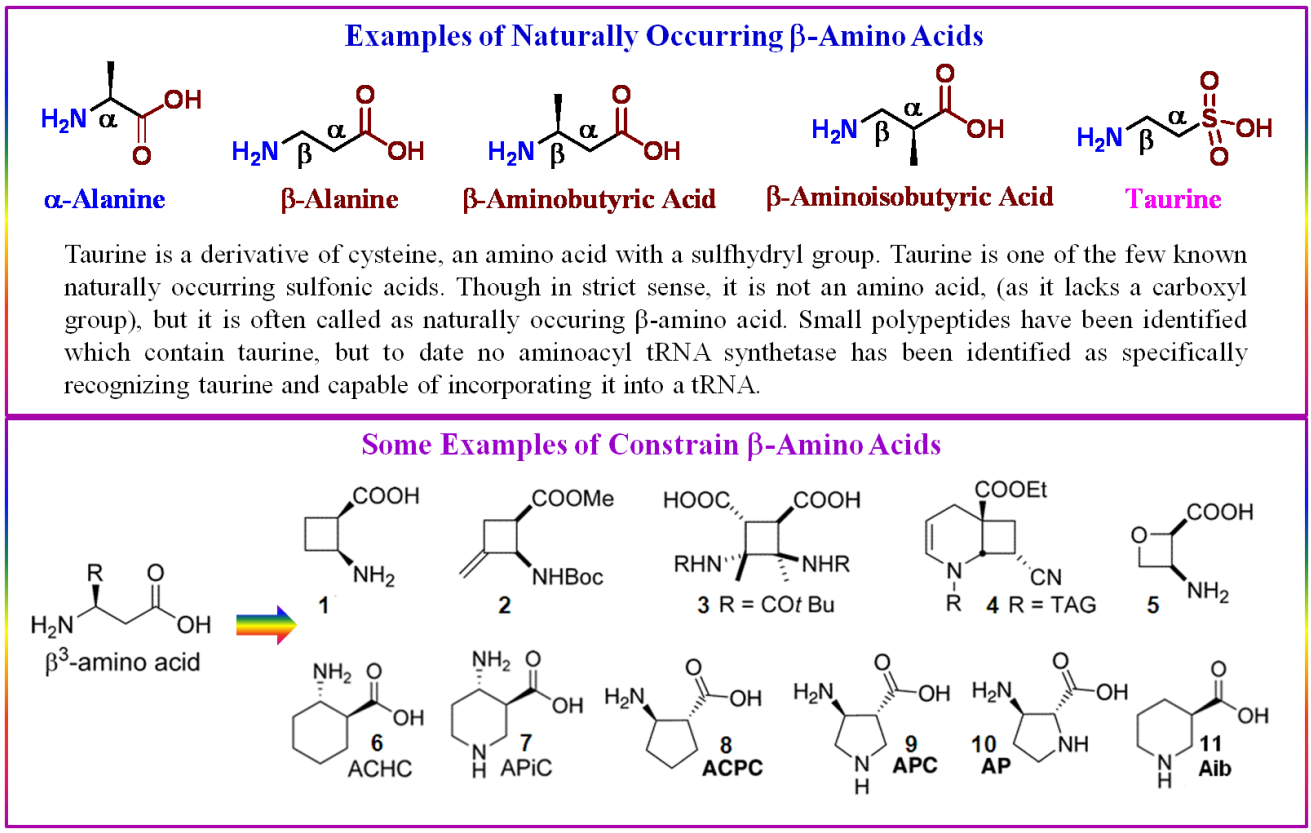


Figure 2.24: The chemical structures of naturally occurring and few unnatural β- amino acids.

The folded structures in natural polypeptides containing α-amino acids are conveniently defined using backbone torsion angles Φ and ψ. Each α-amino acid residue possesses two degrees of torsional freedom about the N−Cα (Φ) and Cα-CO (ψ) bonds, with the peptide bond restricted to a trans, planar (ω ≈ 180°) conformation. The Ramachandran map provides a convenient means of analyzing and representing the observed backbone conformations of α-amino acid residues in peptides and proteins. A very large numbers of natural and unnatural amino acids have been utilized to generate mimic many biologically

and structurally interesting compounds peptides and proteins. Gellman and co-workers have studied several model peptides containing covalently constrained β-amino acids. Their study led to the realization that new classes of folded unnatural polypeptide structures could be generated. The insertion of additional atoms in between the flanking peptide units enhances the number of degrees of torsional freedom, resulting in an expansion of energetically accessible conformational space. As for example, in the β-amino acid residue, local backbone conformations are determined by values of three torsional variables (Φ, θ, and ψ) while for the γ-residue, the number of torsional variables is four (Φ, θ1, θ2, and ψ) (Figure 2.25).

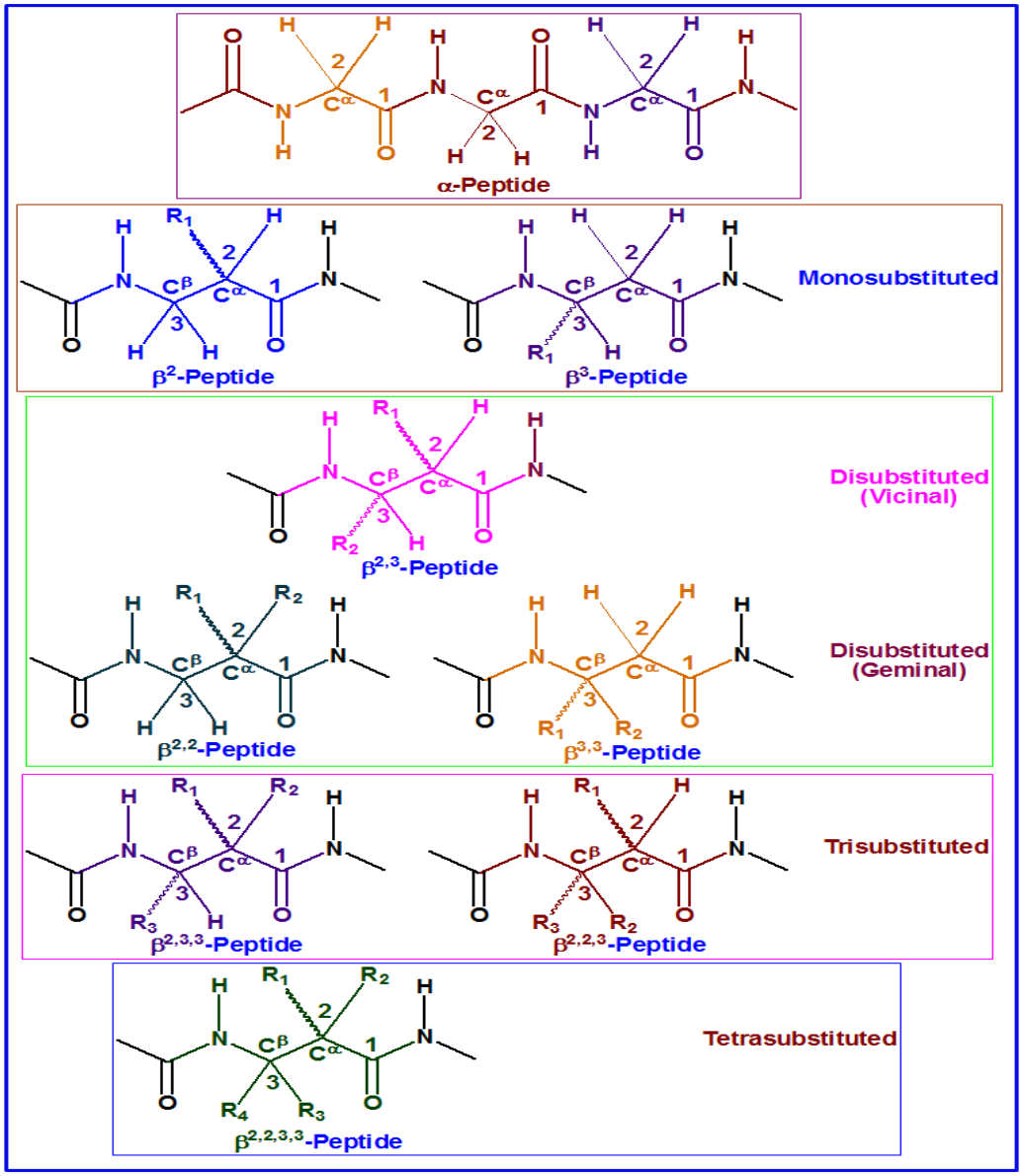


Figure 2.25: Substitution patterns for a β-amino acid residue.

A poly-β-amino acid helix introduced by Kovac et al. in 1965 has 3.4 residues/turn and an axial translation of 1.58 Å/ residue. It is necessarily right-handed for the β-amino acids with the L-configuration at the α-carbon. Recent interest in the chemistry and biology of peptides containing backbone expanded amino acid residues stems from the studies reported in the mid 1990s that novel polypeptide helices could be formed in oligo β-peptides and the characterization of hybrid structures that demonstrated that the β-and γ-residues can be accommodated into canonical helical folds, with an expansion of the intramolecular hydrogen bonded rings.

2.5.2. Chemical structure and synthesis

Figure 2.25 illustrates the substitution pattern for a β-amino acid residue. The most commonly used is monosubstituted (β^3-) residues. These are derived by Arndt-Eistert homologation of the readily available α-amino acid residues. In α-amino acids both the carboxylic acid group and the amino group are bonded to the same carbon center, the α-carbon (C^{α}). In β amino acids, the amino group is bonded to the β carbon (C^{β}). Only glycine lacks a β carbon, which means that β-glycine is not possible.

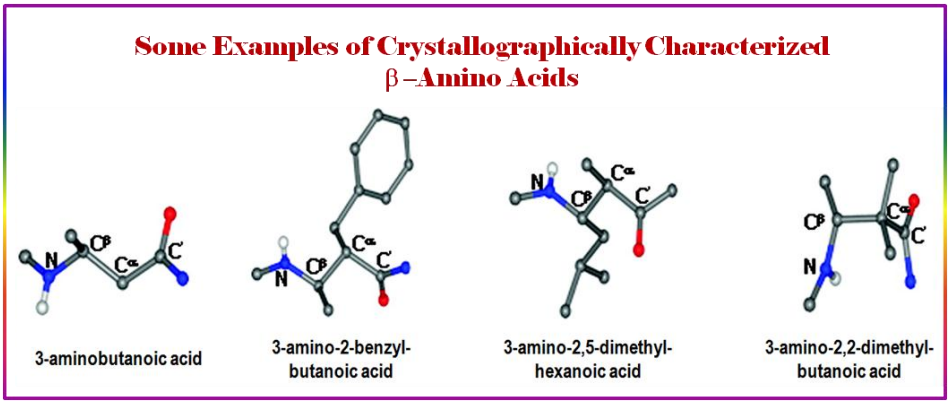


Figure 2.26: Some examples of crystallographically characterized substituted β-residues.

The chemical synthesis of β-amino acids can be challenging, especially given the diversity of functional groups bonded to the β-carbon and the necessity of maintaining chirality (figure 2.27). In the alanine molecule shown, the β-carbon is achiral; however, most larger amino acids have a chiral C^{β} atom. A number of synthesis mechanisms have been introduced to efficiently form β-amino acids and their derivatives notably those based on the Arndt-Eistert synthesis. Two main types of β-peptides exist: those with the organic residue (R) next to the amine are called β^3-peptides and those with position next to the carbonyl group are called β^2-peptides.

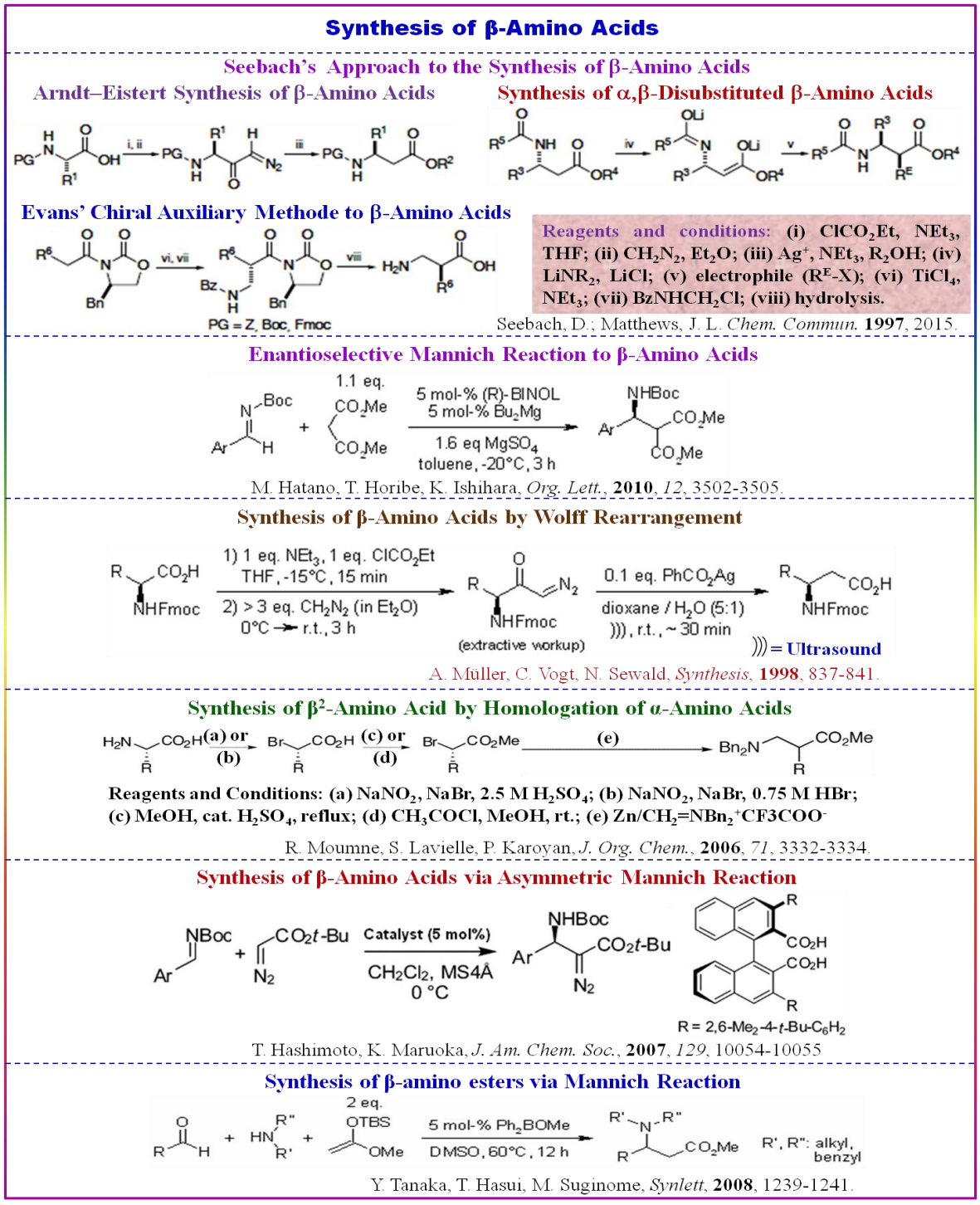


Figure 2.27: Various methods for the synthesis of β-amino acids.

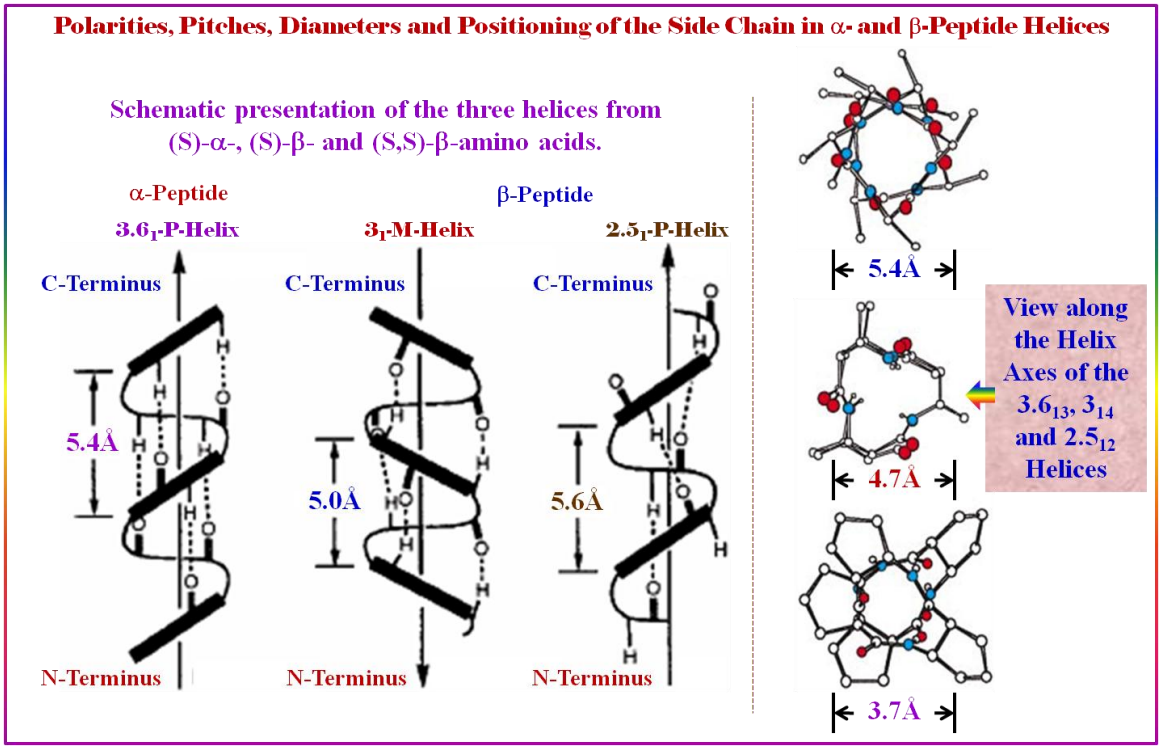
2.5.3. Secondary structure

Because the backbones of β-peptides are longer than those of α-peptides, β-peptides form different secondary structures. The alkyl substituents at both the α- and β-positions in a β-amino acid favor a gauche conformation about the bond between the α-carbon and β-carbon. This also affects the thermodynamic stability of the structure.

2.5.3.1. The Helices of β-Peptides

Many types of helix structures consisting of β-peptides have been reported. These conformation types are distinguished by the number of atoms in the hydrogen-bonded ring that is formed in solution; 8-helix, 10-helix, 12-helix, 14-helix, and 10/12-helix have been reported. In general, β-peptides form a more stable helix than α-peptides.

A comparative structural analysis of an α-peptide 3.6_1 helix with the 3_1 and 2.5_1 helices of β-peptides shows the following informations: (a) the helices have different polarities with respect to their C and N-termini; (b) their shapes and sizes drastically are different; (c) the 3_1 β-peptide helix can be produced with the homochiral residues or with the residues having 2,3-*l* relative configuration; (d) the 3_1 β -peptide helix can be produced with geminal disubstitution, however, the α-helix is stabilized by incorporation of geminally disubstituted α-amino acids; (e) the β-peptide helices are stabilized by hydrophobic interactions between side chains. The hydrophobic interactions contribute a lot to make a difference between β^2- and β^3-peptides (3_1 helix) and the mixed β^2-/β^3-peptides.



igure 2.28: A comparative structures of a α-peptide helix with β-peptides helices.

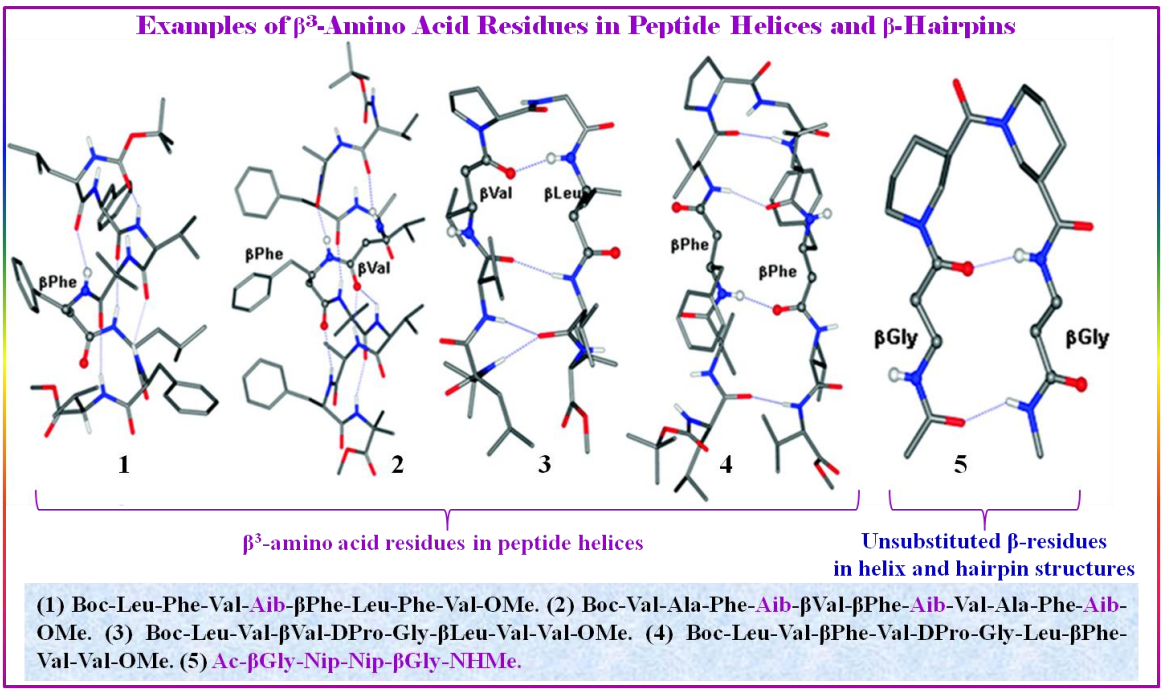


Figure 2.29: β-peptides helices show the β-amino acids residues.

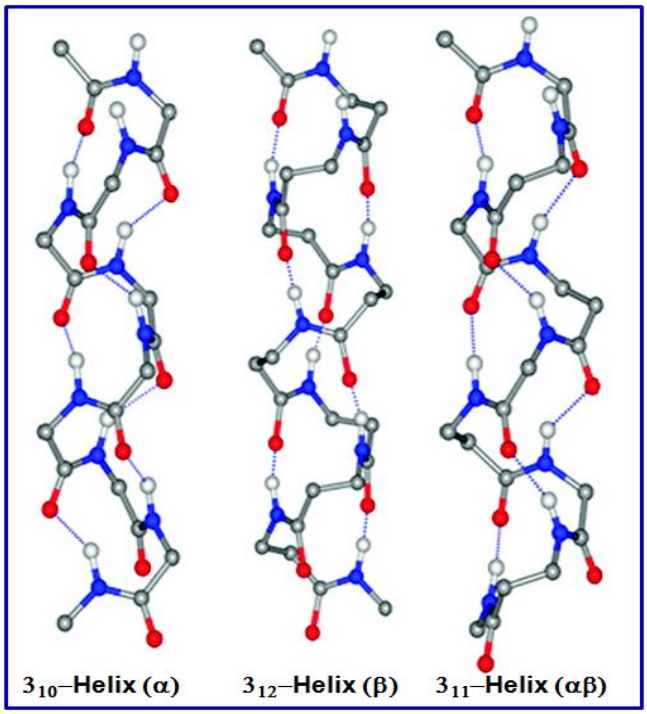


Figure 2.30: Helices with 4 → 1 hydrogen-bond patterns in α, β, and αβ -hybrid peptides.

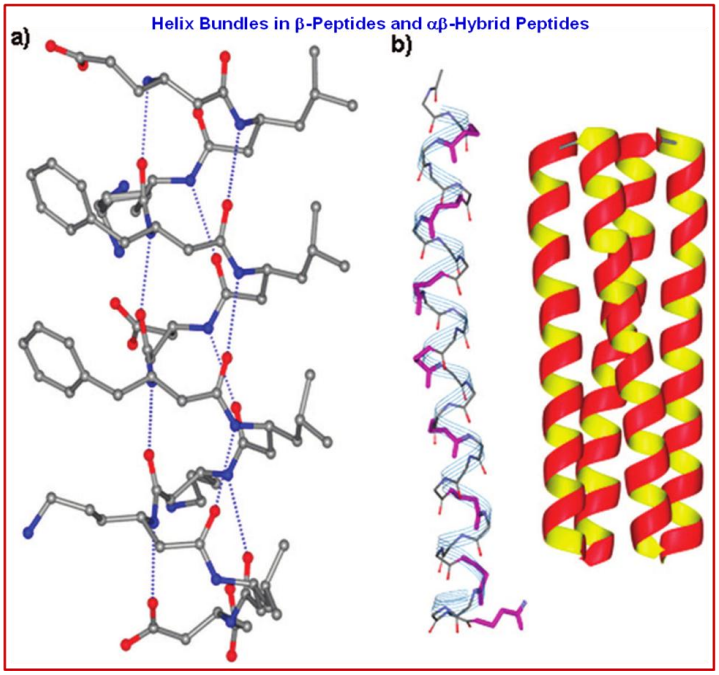


Figure 2.31: Helix bundle β, and αβ -hybrid peptides. (a) Molecular conformation of a β^3-dodecapeptide 14-helix in H_2N- βGlu- βLeu- βOrn- βPhe- βLeu- βAsp- βPhe- βLeu- βOrn- βOrn- βLeu- βAsp-OH. (b) Conformation of a αβ-hybrid peptide with the sequence repeat ααaβ This is derived from a 33 residue α-peptide of the dimerization domain of yeast transcriptional regulator.

2.5.3.2. The Pleated Sheet of β-Peptides

Pleated sheets of α- and β-peptides differ in the following features: (a) the amide planes in an extended conformation are separated by one and two tetrahedral carbons, respectively; (b) thus, the amide planes are arranged in a zig-zag and in a parallel-displaced fashion; (iii) while the C=O and N–H bonds are pointing up and down in the α -peptidic sheet, they are unidirectional *i.e.* all C=O up, all N–H down, in the β-peptide; (iv) the side chains of an antiparallel and the parallel pleated sheet of an α-peptide pointing perpendicular, above and below, the average plane of the sheet, while the sheet of a β-peptide built from homochiral (*all*- β^2 or *all*- β^3) components is obscured on only one face by the side chains; (v) the α-peptides containing the twenty proteinogenic amino acids may show a certain *tendency* for forming either a helix or a β-sheet, the chain of a β-peptide constructed from 2,3-disubstituted β-amino acids can be *prevented* from forming a 3_1-helix and *forced* to adopt an extended conformation.

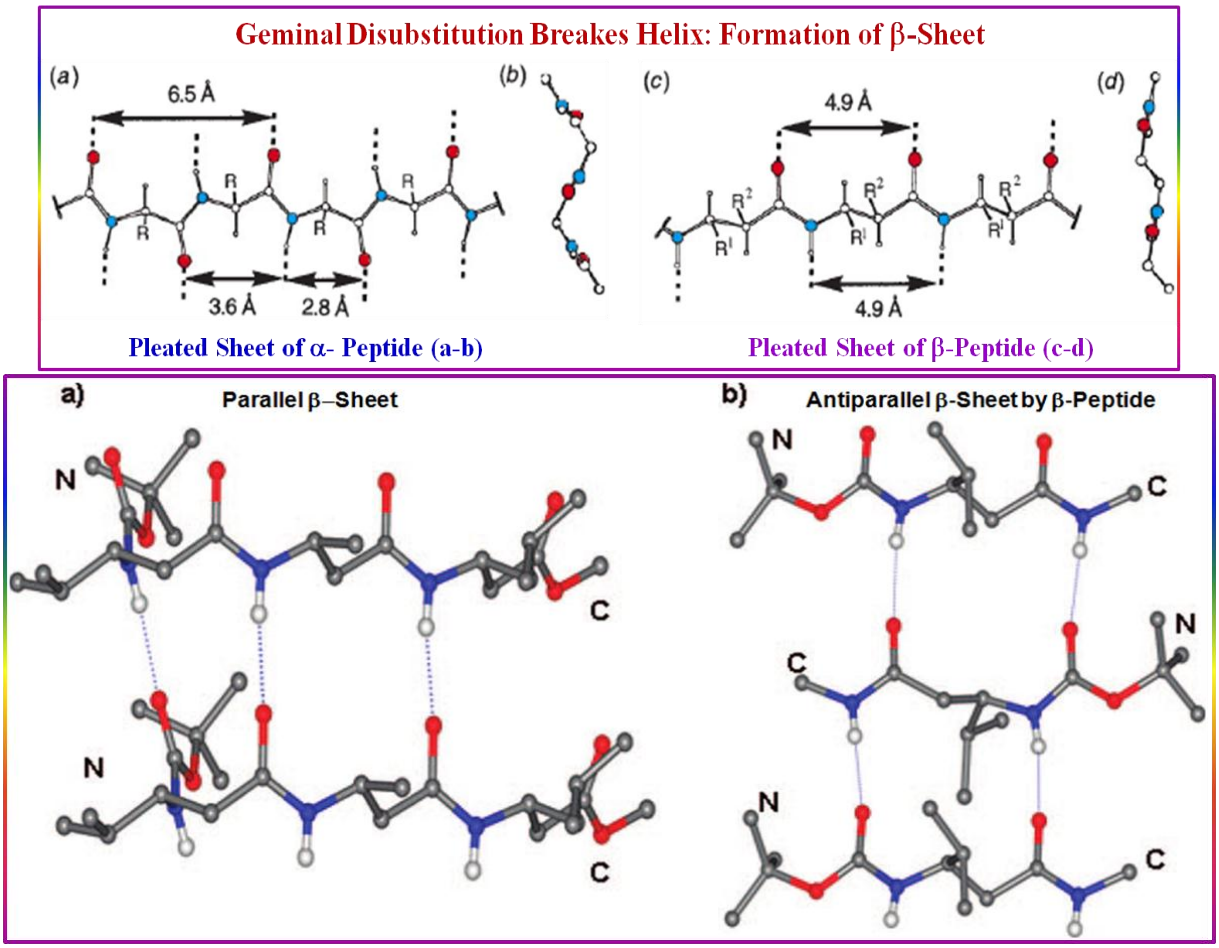


Figure 2.32: Parallel (a) and antiparallel (b) sheets in short β-peptides.

2.5.3.3. The Turn of β-Peptides

The 4 → 1 hydrogen-bonded two residue structure in αα segments is the extremely well studied and is known as β-turn. Polypeptide chain folding about the two backbone residues (i + 1, i + 2) generate β-turn characterized by backbone conformational angles (Φ, ψ). The type I/III turn is characterized by $\Phi_{(i+1)}$ = −30°, $\psi_{(i+1)}$ = −60°, $\Phi_{(i+2)}$ = −60°/−90°, and $\psi_{(i+2)}$ = −30°/0°, while the type II turn is characterized by ($\Phi_{(i+1)}$ = −60°, $\psi_{(i+1)}$ = 120°, $\Phi_{(i+2)}$ = 80°, and $\psi_{(i+2)}$ = 0°). Turns with different ring size may arise because of the hydrogen bonding in two residue with normal (4 → 1) and reversed (1 → 2) directionalities.

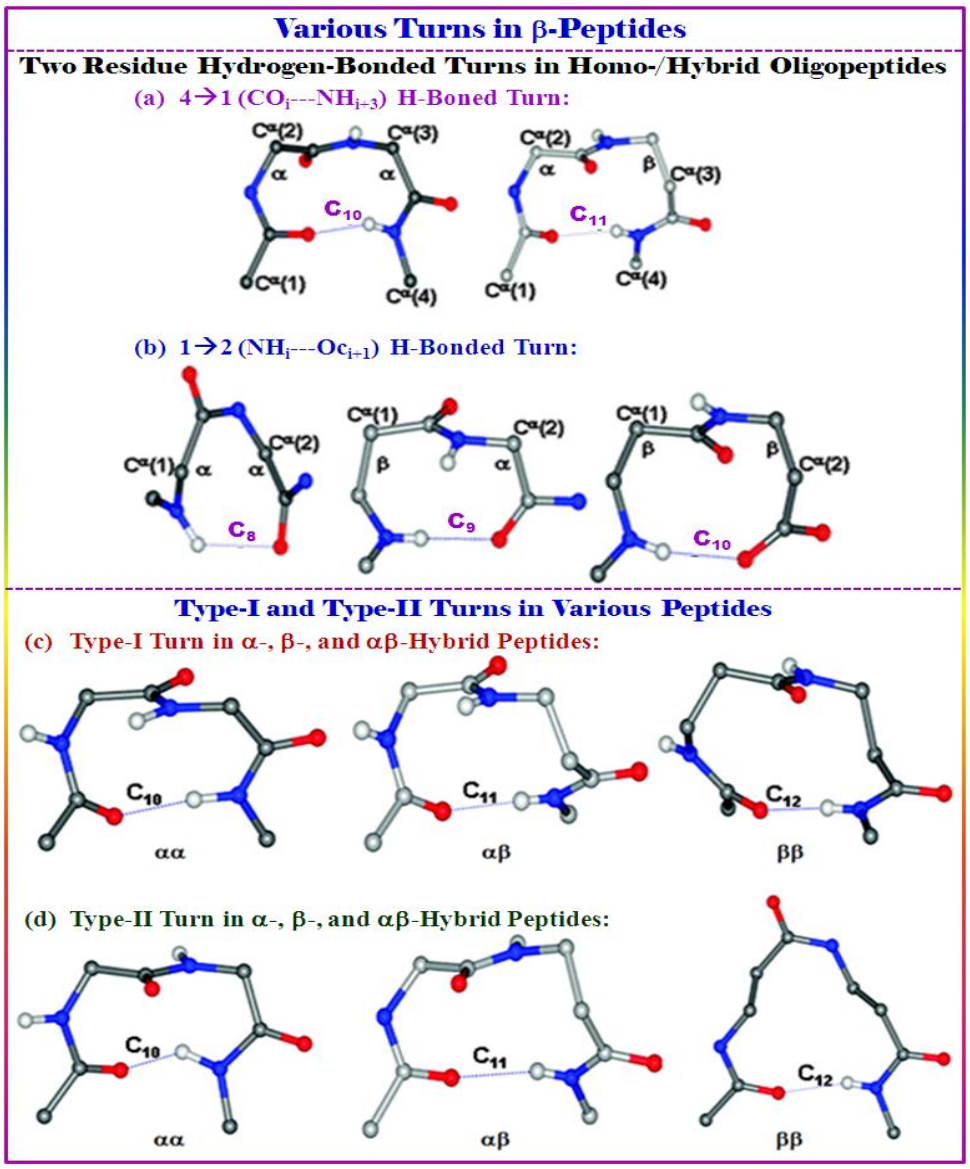


Figure 2.33: Various types of turns in short β-peptides.

2.5.4. Biological Importance/Clinical Potential of β-Peptides

β-peptides are stable against proteolytic degradation in vitro and in vivo. Therefore, they deserve important advantage over natural peptides in the preparation of peptide-based drugs. β-Peptides have been used to mimic natural peptide-based antibiotics such as magainins, which are highly potent but difficult to use as drugs because they are degraded by proteolytic enzymes in the body. Some natural β-amino acids such as taurine, β-aminobutyric acid, β-aminoisobutyric acid have been reported as agonists of the inhibitory glycine

receptor. Substituted β-amino acids have also been reported as fibrinogen receptor GIIb/IIIa antagonists, β-lactamase inhibitors, μ-opioid receptor agonists, or enkephalin-degrading enzyme inhibitors. Also, various β-amino acids are found in natural antibiotics, fungicides, and antineoplastic compounds. Furthermore, β-Peptides have shown promise in a variety of biological applications, which include inhibition of cholesterol uptake, somatostatin receptor binding, and antimicrobial activity any many more.

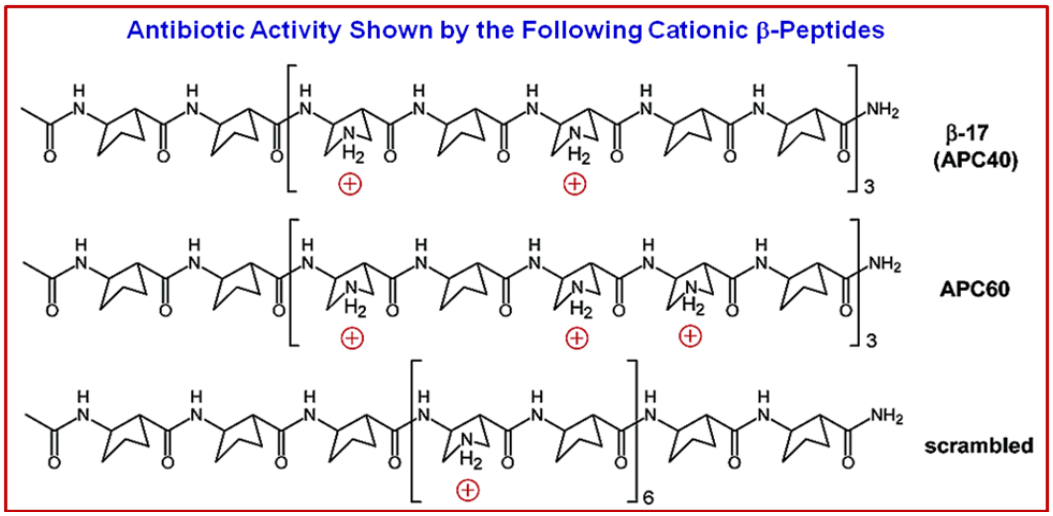


Figure 2.34: Gellman's β-peptides with antibiotic activity.

Therefore studies toward designing more biologically active β-peptides have drawn attention to the peptide chemists. Gellman *et al.* have designed β-amino acid oligomers that are helical, cationic, and ampiphilic with the intention of mimicking the biological activity of ampiphilic, cationic α-helical antimicrobial peptides found in nature. They have shown that 12-helical β-peptides are capable of selectively killing a variety of bacterial species, including two clinical isolates that are resistant to common antibiotics. For this class of β-peptides, it seems that a 40% cationic face is best for activity; β-peptides containing more cationic residues are not as active. Helical ampiphilicity is also important for activity. These β-peptides appear to act by a mechanism similar to that of antimicrobial α-peptides, based on their ability to cause rapid release of a cytoplasmic enzyme from B. subtilis cells. The 12-helical β-peptides are impervious to proteases, which is encouraging with regard to potential biological applications.

Below are some more examples of bioactive β-peptides.

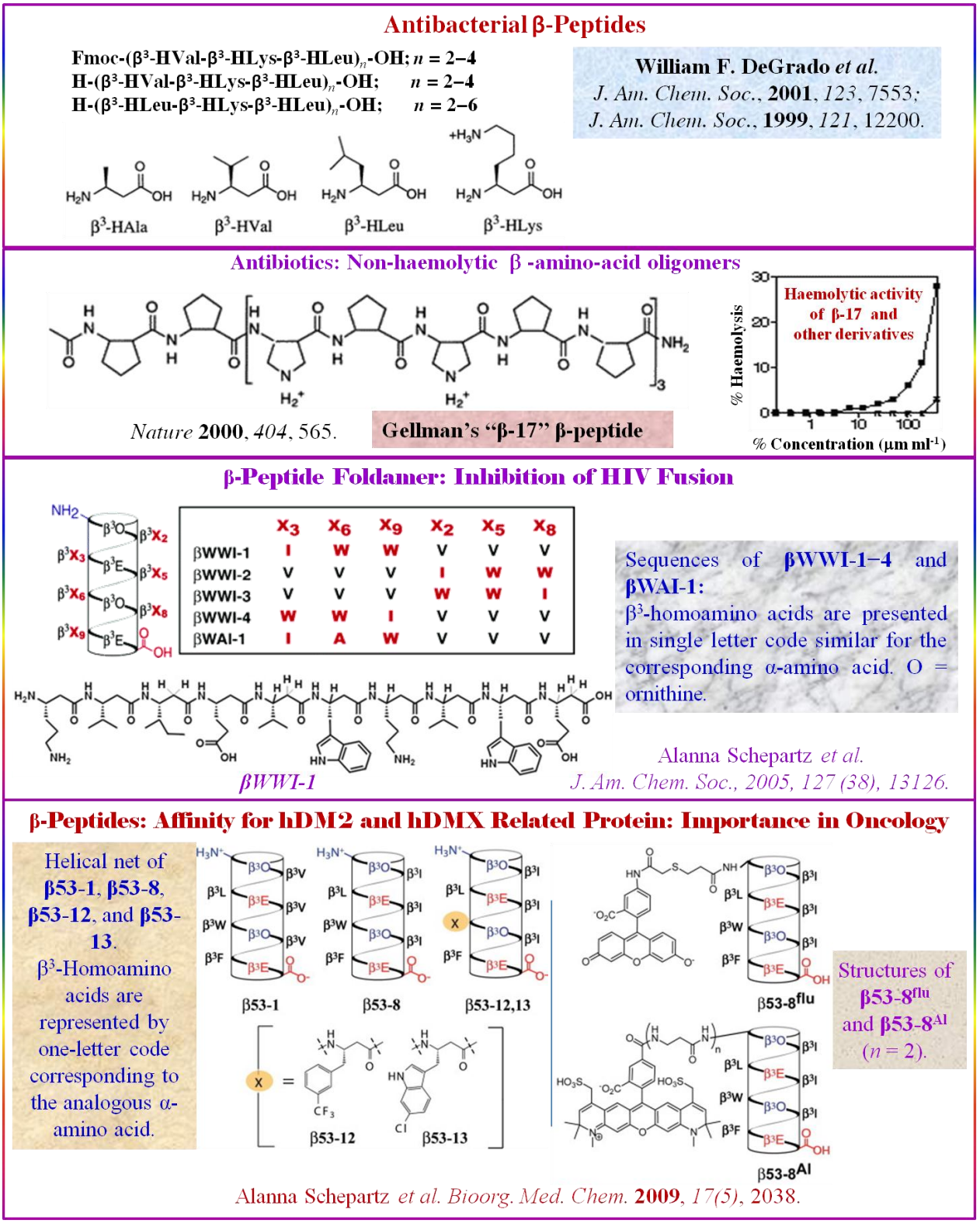
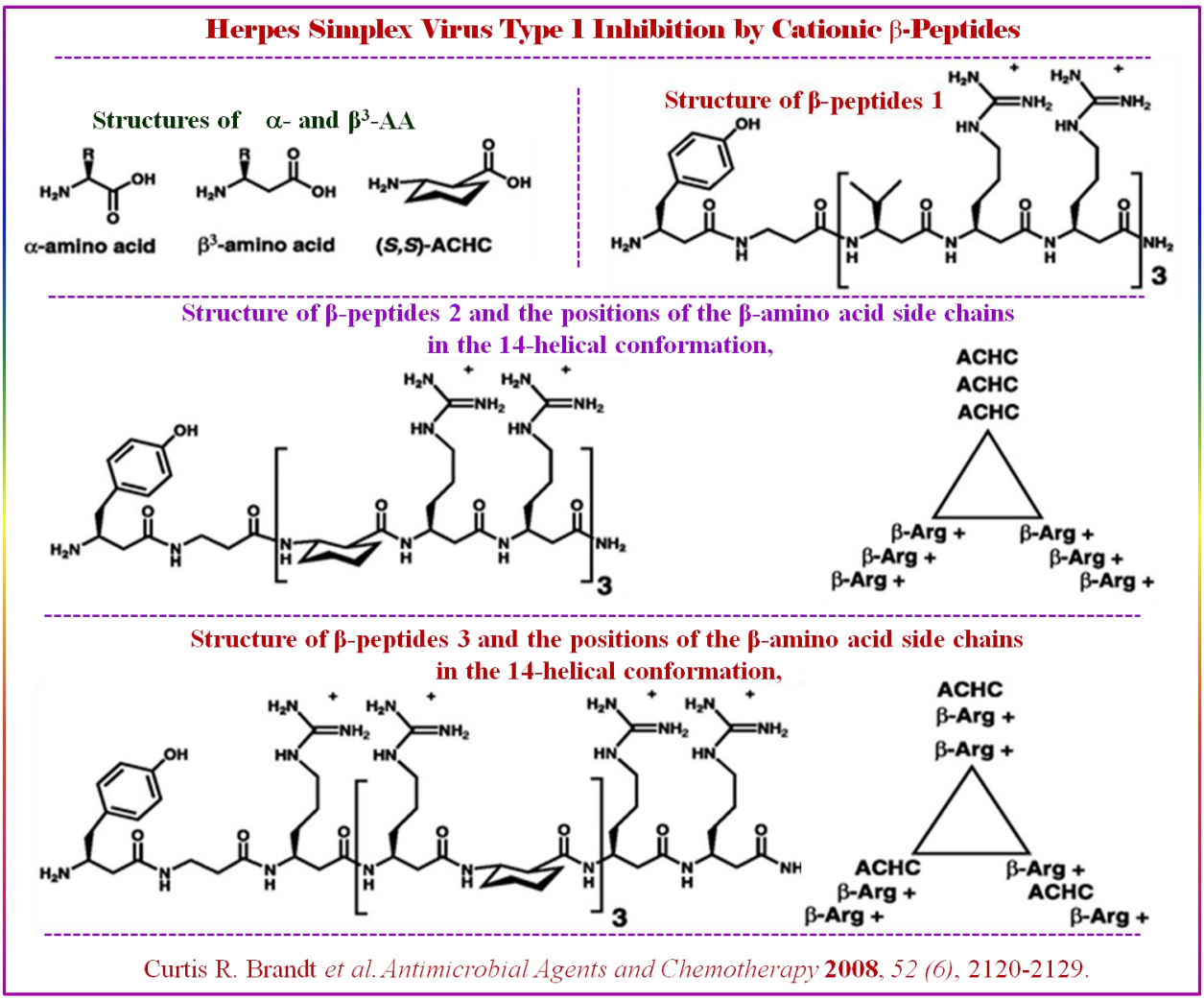


Figure 2.35: Some examples of bioactive β-peptides.



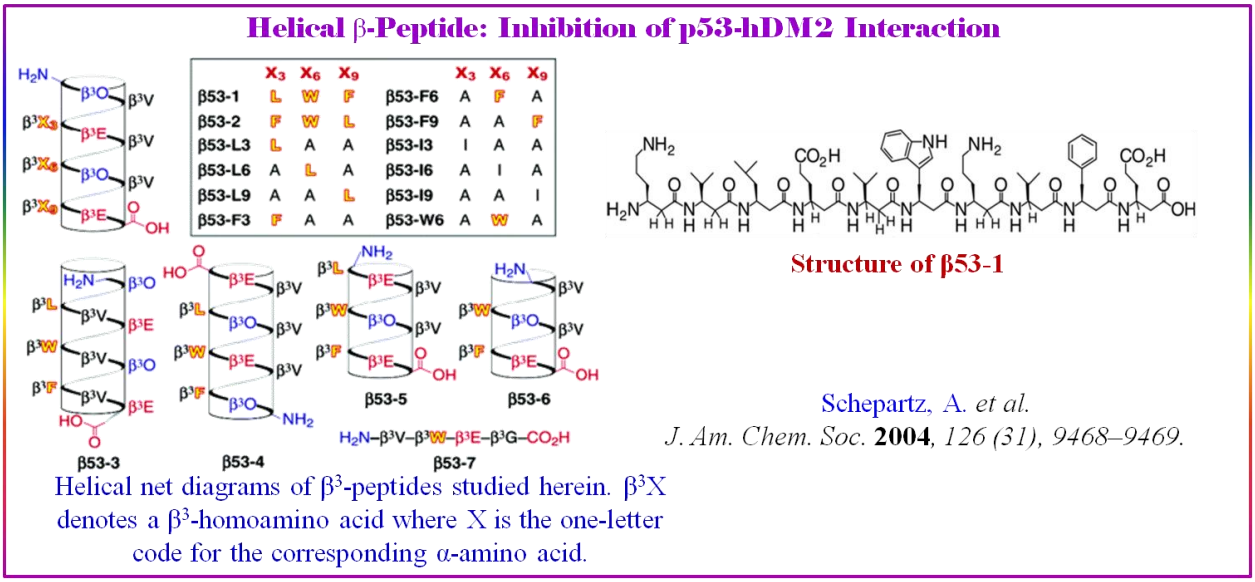


Figure 2.36: Few examples of bioactive β-peptides.

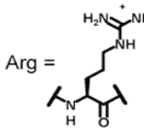
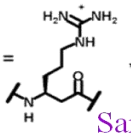
Binding of TAR RNA by a Tat-Derived β-Peptide

1 = β^3HTyr-β^3HGly-β^3HArg-β^3HLys-β^3HLys-β^3HArg-β^3HArg-β^3HGln-β^3HArg-β^3HArg-β^3HArg

2 = β^3HTyr-β^3HGly-β^3HLys-β^3HLys-β^3HLys-β^3HLys-β^3HLys-β^3HGln-β^3HLys-β^3HLys-β^3HLys

3 = Tyr-Gly-Arg-Lys-Lys-Arg-Arg-Gln-Arg-Arg-Arg

4 = Tyr-Gly-Lys-Lys-Lys-Lys-Lys-Gln-Lys-Lys-Lys

Arg =  β^3HArg =  , etc.

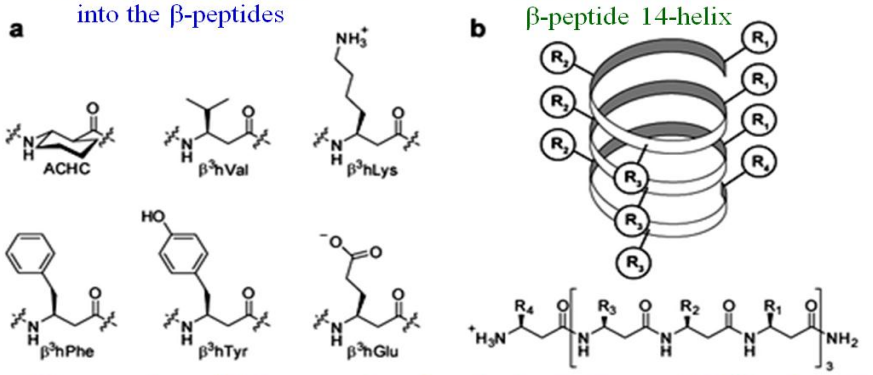
β-peptide analogues of HIV-1 Tat, may be lused for antitranscriptive anti-HIV therapy. This study represent an exciting new direction in antiviral drug development.

Samuel H. Gellman, and Tariq M. Rana *et al.*
Org. Lett., 2003, 5 (20), pp 3563–3565

Activity Against *Candida albicans* Planktonic Cells and Biofilms

β-Amino acid residues incorporated into the β-peptides

Three-dimensional representation of the β-peptide 14-helix

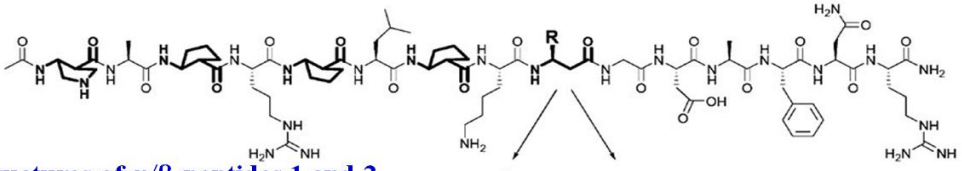


β-peptides may be useful in preventing fungal colonization and biofilm formation.

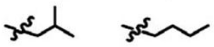
Samuel H. Gellman and Sean P. Palecek *et al.*
ACS Chem. Biol. **2009**, *4* (7), 567–579.

Helical α/β-Peptide Foldamer: Binding to Anti-Apoptotic Protein Bcl-xL

a) Ac-**APC**-Ala-**ACPC**-Arg-**ACPC**-Leu-**ACPC**-Lys-β^3**Xaa**-Gly-Asp-Ala-Phe-Asn-Arg-NH_2



Structures of α/β-peptides 1 and 2

R = 

1 (β^3-hLeu) **2** (β^3-hNle)

Sequence alignment of BH3 domains from Bim and Bak with foldamer 1.

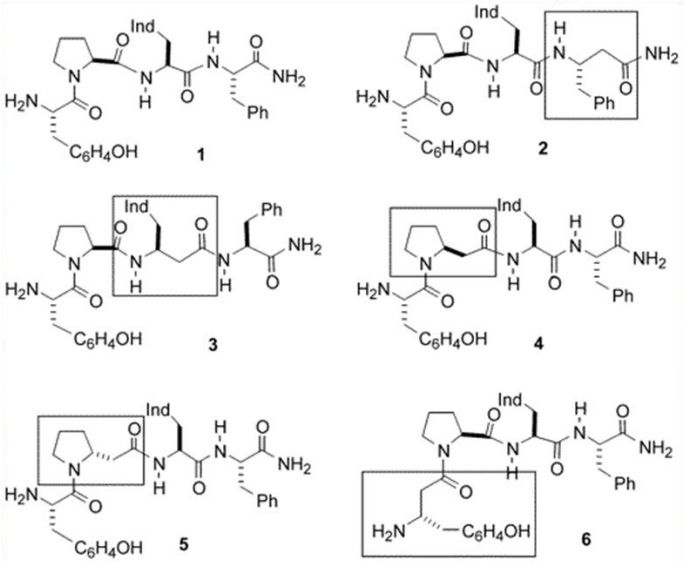
b)

	h1				h2					h3			h4				
Foldamer 1	Ac	APC	A	ACPC	R	ACPC	L	ACPC	K	β³-hLeu	G	D	A	F	N	R	NH_2
Bim_S (57-71)		W	I	A	Q	E	L	R	R	I	G	D	E	F	N	A	
Bak (73-87)		Q	V	G	R	Q	L	A	I	I	G	D	D	I	N	R	

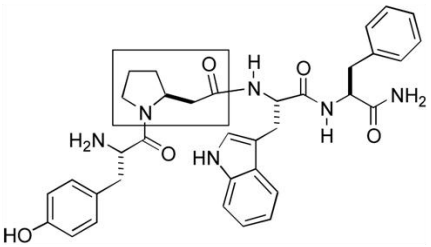
Samuel H. Gellman, and W. Douglas Fairlie *et al. Angew. Chem. Int. Ed.* **2009**, *48*, 4318 –4322.

Figure 2.37: Few more examples of bioactive β-peptides.

Endomorphin-1 Analogues of β-Proline: μ-Opioid Receptor and Enhanced Enzymatic Hydrolysis Resistance



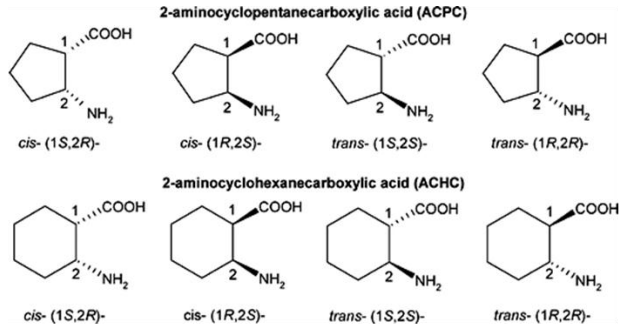
Structures of Homo-Tetrapeptides 1–6

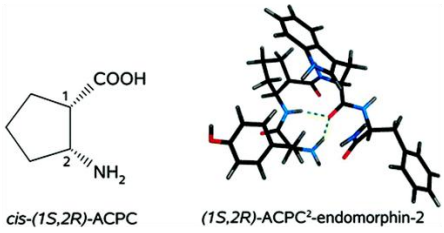


- These analogues display different affinities for μ-opioid receptors, depending on the position and absolute stereochemistry of homo-amino acid introduced. Good affinities were found for peptides **4–6**.
- Peptide **4** (Tyr-l-homo-Pro-Trp-Phe-NH_2) showed affinity comparable to that of DAMGO. The cyclic AMP assay confirmed that **4** and **5** act as agonists.

- A proteolytic enzymes (α-chymotrypsin, aminopeptidase-M, and carboxypeptidase-Y) resistance study of peptide **1** and **4–6** showed that the presence of homo-Pro gives the peptides **4** and **5** a good resistance, while homo-Tyr seems to have little effect. ---Giuliana Cardillo *et al. J. Med. Chem.*, ***2002***, *45*, 2571.

Endomorphin Analogues of Alicyclic β-Amino Acids





Structures of alicyclic β-amino acids incorporated in endomorphin analogues.

The analogues containing cis-(1S,2R)-$ACPC^2$/$ACHC^2$ exhibited the highest μ-opioid receptor affinities, comparable with those of the parent endomorphins. The conformational studies suggests that these peptides might bind to the μ-opioid receptor in a compact, folded structure rather than an extended one.

Murnyi and Gza Tth *et al. J. Med. Chem.* **2008**, *51 (14)*, 4270–4279.

Figure 2.38: Few examples of bioactive β-peptides.

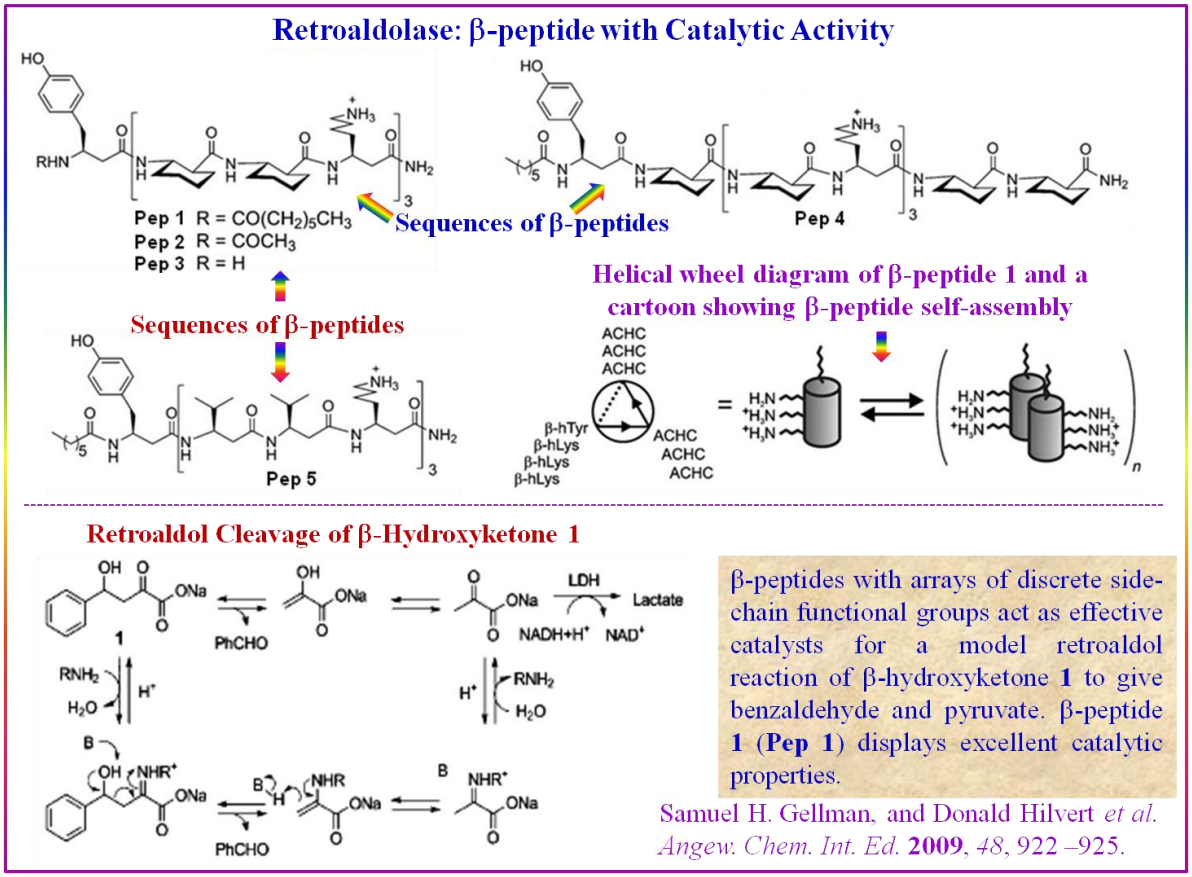


Figure 2.39: Recently reported β-peptide, retroaldolase, with catalytic activity.

2.6. β-Turn Peptidomimetic

2.6.1. Peptidomimetics and Their Importance

The discovery of the physiological role of a great number of peptides stimulated researchers all over the world towards design and synthesis of peptidomimetics or peptide like molecules. Since natural peptides seldom can be used therapeutically as drugs, because of the problems associated with low absorption, rapid metabolism and low oral bioavailability, many efforts aimed to modify the natural sequence of the amino acids of bioactive peptides achieved a desired, very focused effect. Peptidomimetics, which maintain the key elements required for activity but replace the labile peptide bonds with more stable features, have the advantage of providing new functionalities that can circumvent natural processes in the body. For example, they become able to perform functions that are not available with the natural materials, such as binding to and penetrating cell membranes and resisting degradation by enzymes.

Peptidomimetics that fold to mimic protein secondary structures have emerged as important targets of bioorganic chemistry. Recently, a variety of compounds that mimic helices, turns, and sheets have been developed, with notable advances in the design of β-peptides that mimic each of these structures. These compounds hold promise as a step toward synthetic molecules with protein like properties and as drugs that block protein-protein interactions.

Although initial efforts in peptidomimetic chemistry focused upon the development of enzyme inhibitors and peptide hormone analogues, this field now encompasses both the creation of pharmacologically useful analogues of biologically active peptides and the development of compounds that mimic protein structures. Current objectives include developing new drugs, gaining an enhanced understanding of protein folding, and creating catalysts and new materials with useful properties.

2.6.2. The Approaches to β-Turn Peptidomimetic

The β-Turn (**Figure 2.40** and **2.41**), which has also been referred to as the beta-bend, beta-loop or reverse turn, is one of the three major secondary structural elements of peptides and proteins. The surface localization of turns in proteins, and the predominance of residues containing potentially critita1 pharmacophoric information, has led to the hypothesis that turns play critical roles in a myriad of recognition events in biological systems. These events include but are not limited to the interactions between peptide hormones and their receptors, antibodies and antigens, and regulatory enzymes and their corresponding substrates. Reverse turn mimetics are powerful tool for the study of molecular recognition and are providing a unique opportunity to dissect and investigate structure-function relationships in complex proteins.

A great deal of effort has therefore focused on the design and synthesis of small constrained mimetics of turn structure to provide a better understanding of the molecular basis of peptide and protein interactions in addition to providing potent and selective therapeutic agents.

β-turns constitute a tetra peptide unit, which cause a reversal of direction of the peptide chain. Formally, turns can be described by the distance from the Cα of the first residue to the Cα of the fourth residue. When this distance is less than 7 Å and the tetra peptide sequence is not in a α-helical region, it is considered a β-turn. Additionally, a three residue reverse turn, popularly known as γ-turn, exists but is significantly less widely distributed.

β-turns are classified according to the Φ- and Ψ-angles of the i+1 and i+2 residues. In addition to the existence of a number of turn types (I, I', II, II', III, III', IV, V, Va, VIa, VIb, VII, and VIII) the C_α^{i} to C_α^{i+3} distance varies from 4-7 Å.

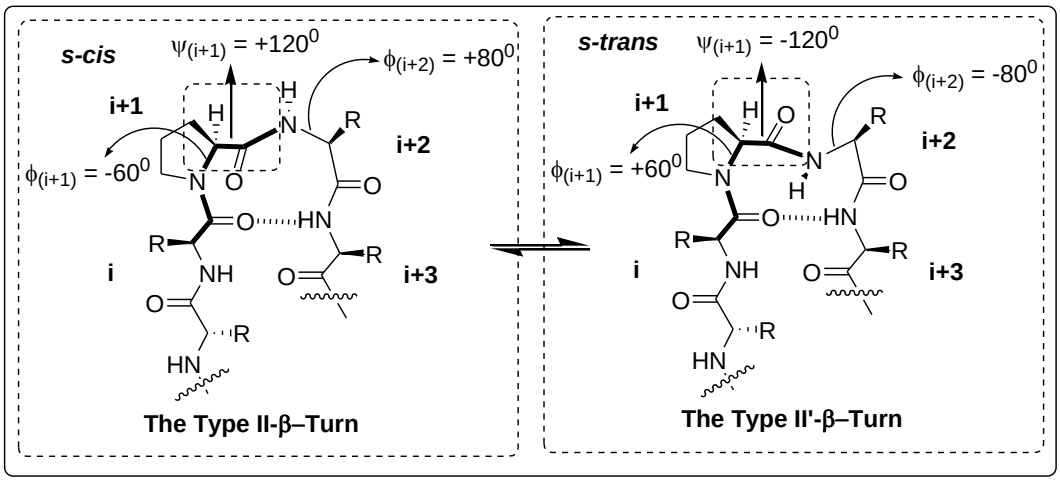


Figure 2.40: The type II- and II'-β-turn motif

While the retention or improvement of biological activity is the ultimate indicator of successful design, the success of many efforts has been measured by the ability of a scaffold to adopt a turn motif using spectroscopic methods such as circular dichroism (CD) or solution phase NMR.

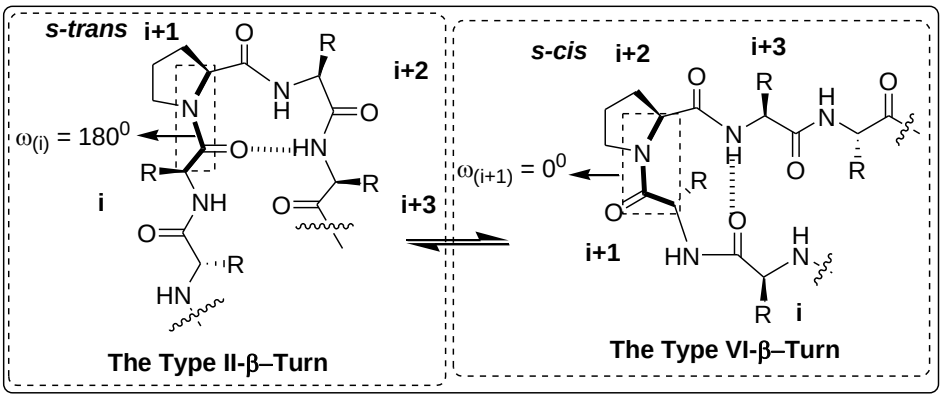


Figure 2.41: The type II- and VI-β-turn motif.

We will now briefly discuss the current state of the art of turn mimetic synthesis and scaffold design. The examples are confined to small molecule turn scaffolds that have been designed with diversity and parallel execution in mind.

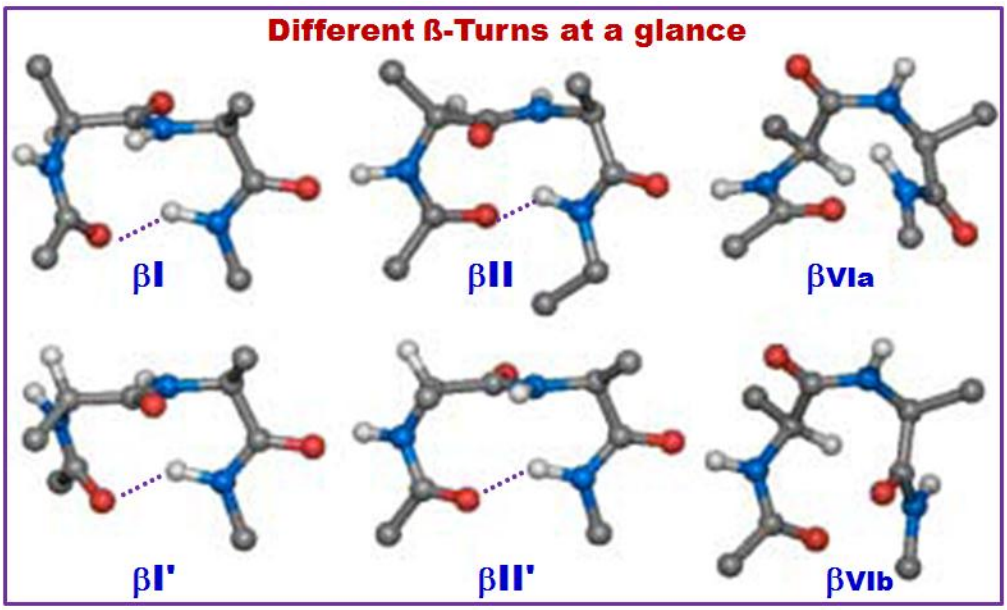


Figure 2.42: The Ball-stick model of various β-turns.

Table2: The nine beta-turn types with their dihedral angles.

Turn Type	Dihedral angles (°)			
	ϕ_{i+1}	ψ_{i+1}	ϕ_{i+2}	ψ_{i+2}
I	-60	-30	-90	0
I'	60	30	90	0
II	-60	120	80	0
II'	60	-120	-80	0
IV	-61	10	-53	17
VIa1	-60	120	-90	0
VIa2	-120	120	-60	0
VIb	-135	135	-75	160
VIII	-60	-30	-120	120

2.6.3. β-Turn Peptidomimetics

Turns are important targets for mimicry, both because they serve as recognition sites in peptides and proteins and because they allow a protein chain to fold back upon it to form a compact structure. One area of research has focused upon the creation of structures that mimic the conformations and functionality displayed by β-turns and related structures, whereas a second has focused upon the creation of 'nucleating turn mimics' that induce structure in attached groups (typically peptide strands).

Cyclic structures were first used as turn mimics and nucleators in the mid-1980s and continue to be used for these purposes. Gennari, Scolastico and co-workers have investigated bicyclic lactams **2.01** and **2.02** (**Figure 2.43**) as turn mimics and have established that **2.01** nucleates β-sheet structure when incorporated into certain peptides. Kelly and co-workers have used dibenzofuran turn unit **2.03** (**Figure 2.43**) to nucleate anti-parallel β-sheets in water. These compounds adopt conformations in which the dibenzofuran moiety folds against the sheet, forming a hydrophobic cluster with the side chains of the first residues of the sheet, thus stabilizing the turn. Recently, Kelly and co-workers reported a dibenzofuran turn unit that aligns two peptide strands to form a parallel β-sheet-like structure (**2.03**). Soth and Nowick have reported peptide/oligourea/azapeptide hybrid oligomer **2.04** (**2.43**), which adopts a hairpin turn containing two intramolecular hydrogen bonds in chloroform solution. Two separate groups have reported using a *cis*-norbornenyl turn that induces β-sheet-like structure in attached peptides. Smith, Hirschmann *et al.* have extended earlier work on the use of 3,5-linked pyrrolin-4-one β-strand mimics to create a hydrogen-bonded turn-like structure.

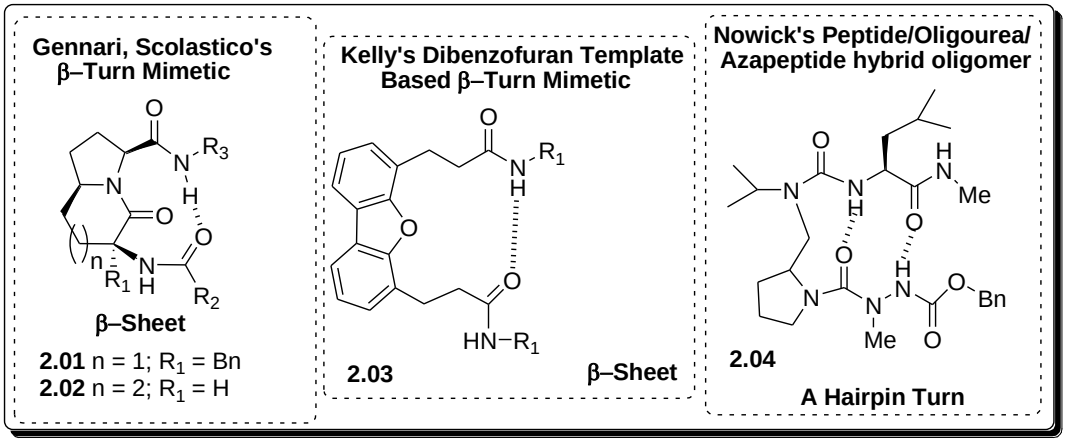


Figure 2.43: β-turn mimics that nucleate sheet or hairpin structure

The acyclic structures containing alkenes have been developed, which exploit allylic strain and other acyclic conformational effects to mimic β-turns. In these compounds, the alkene unit mimics the amide bond that would be present in a natural turn. Gellman and co-workers presented the first example of this strategy is alkene-based turn **2.05** (**Figure 2.44**). A longer peptidomimetic (**2.06**), with amino acids flanking the alkene-based turn, prefers a hairpin conformation in dichloromethane solution. Careful NMR studies suggest that other competing conformations are also present. Wipf *et al.* have designed and synthesized related alkene-based turns **2.07** and **2.08** (**2.44**); the latter contains a trifluoromethyl-substituted alkene, which mimics the electronic properties of the amide group.

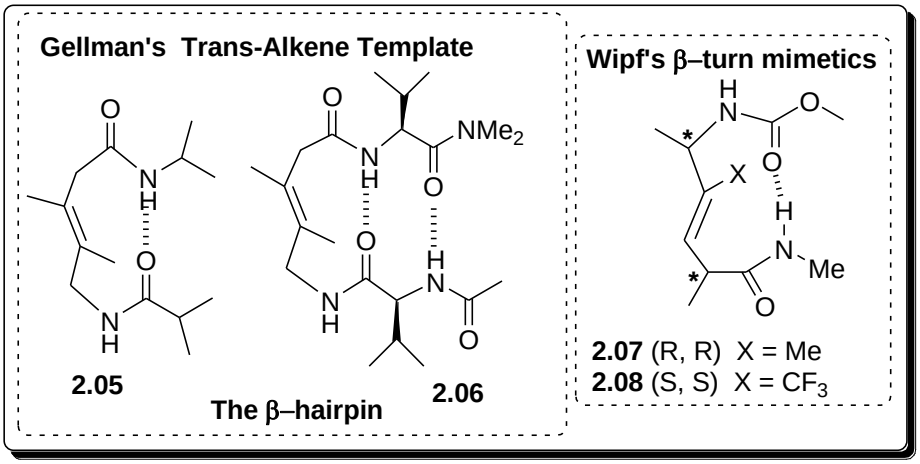


Figure 2.44: β-turn mimics that nucleate sheet or hairpin structure

A new development in the creation of turn structures involves β-peptides. Gellman and co-workers have developed a heterochiral nipecotic acid dimer turn (**2.09**) that induces sheet structure in attached β-peptide strands. Incorporation of **2.09** into β-tetra peptide **2.10** (**Figure 2.45**), resulted in the formation of the β-peptide equivalent of a β-hairpin. This work is particularly exciting because it suggests that different types of β-amino acids can be used to create secondary structures in a predictable fashion. Seebach *et al.* have also created a β-peptide turn (**2.11**) based on the central turn of their 12/10/12 helix (**Figure 2.46**). This turn comprises two homochiral β-amino acids, the first with a substituent at the β-position, the second with a substituent at the α-position. Seebach *et al.* have also reported an α, α-disubstituted β-tripeptide that forms a similar turn. This result, demonstrates that the structures of β-peptides are substituent dependent and suggests rules for designing secondary structures.

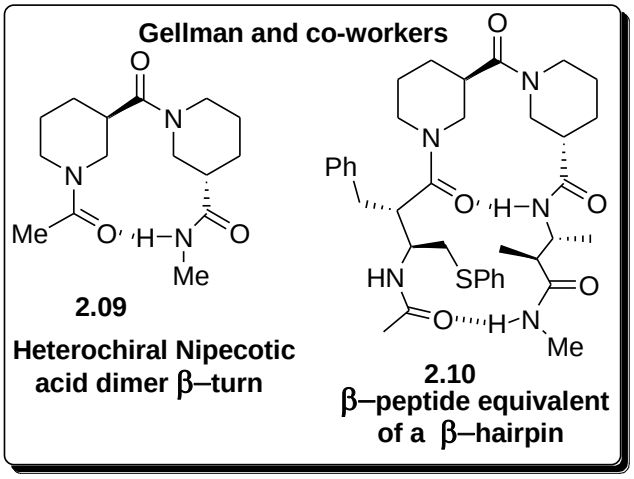


Figure 2.45: β-turn mimics that nucleate sheet or hairpin structure

The myriad of potential homologated amino acids and substitution patterns provides a plethora of opportunities for the creation of structure. Thus, Hanessian *et al.* recently reported that γ-tetrapeptide **2.12** (**Figure 2.46**), adopts a reverse turn conformation containing a 14-membered hydrogen-bonded ring [(*i*) C=O···HN (*i* + 3)].

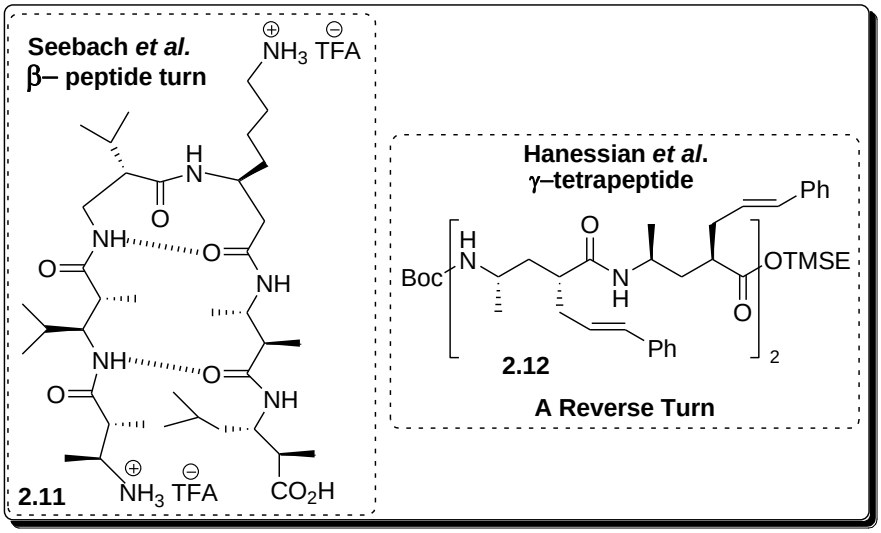


Figure 2.46 β-turn mimics that nucleate sheet or hairpin structure

Despite an exponential growth in the number of reports on the development of constrained non-peptide scaffolds as peptidomimetics, very few peptide-based drugs have been developed, necessitating an overhaul in the existing design principles. Newer concepts are emerging where the fundamental building blocks used by nature, such as amino acids, sugars, and nucleosides, are amalgamated to produce nature-like, and yet unnatural, *de novo* structural entities with

multifunctional groups anchored on a single ensemble. One such hybrid design is represented by a class of compounds called sugar amino acids (Saa). These are carbohydrate molecules bearing both amino and carboxyl functional groups on the regular sugar framework. The rigid furan or pyran rings of these molecules make them ideal candidates as non-peptide scaffolds in peptidomimetics where they can be easily incorporated into a peptide backbone by using their carboxyl and amino termini.

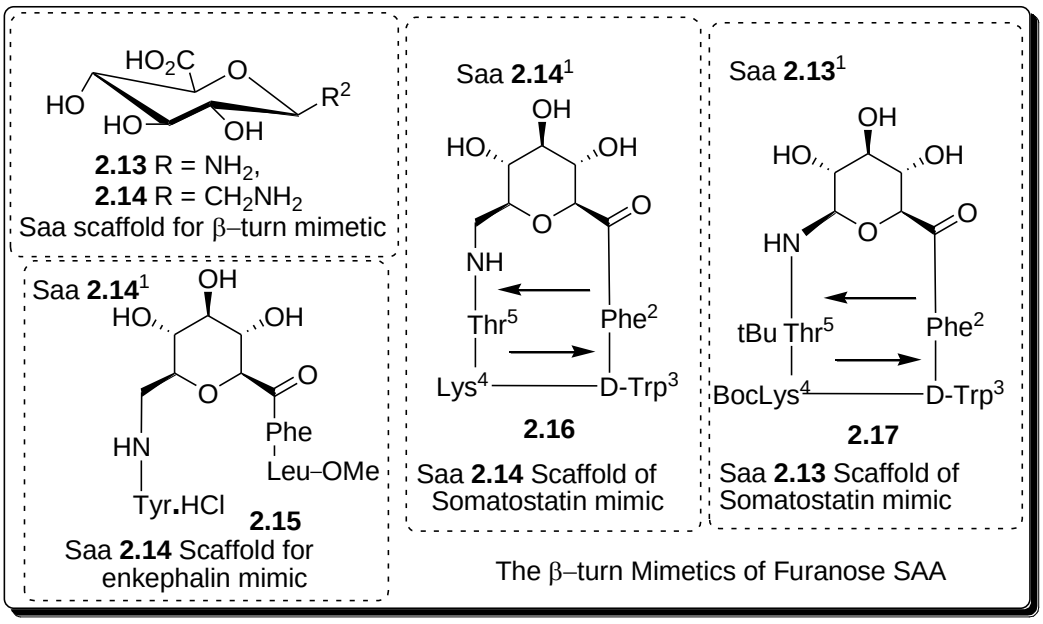


Figure 2.47: Pyranose sugar amino acid scaffolds as peptide β-turn mimetics.

In 1995, von Roedern and coworkers have reported one example of a sugar amino acid as a new type of peptidomimetic (**Figure 2.47**). The novel Saas were successfully incorporated into a cyclic peptide of the somatostatin containing tetrapeptide Phe-D-Trp-Lys-Thr. The conformational analysis clearly typified two β-turns. They have also reported the synthesis of pyranose sugar amino acids (Saas) as new non-peptide peptidomimetics utilizing carbohydrates as peptide building blocks.

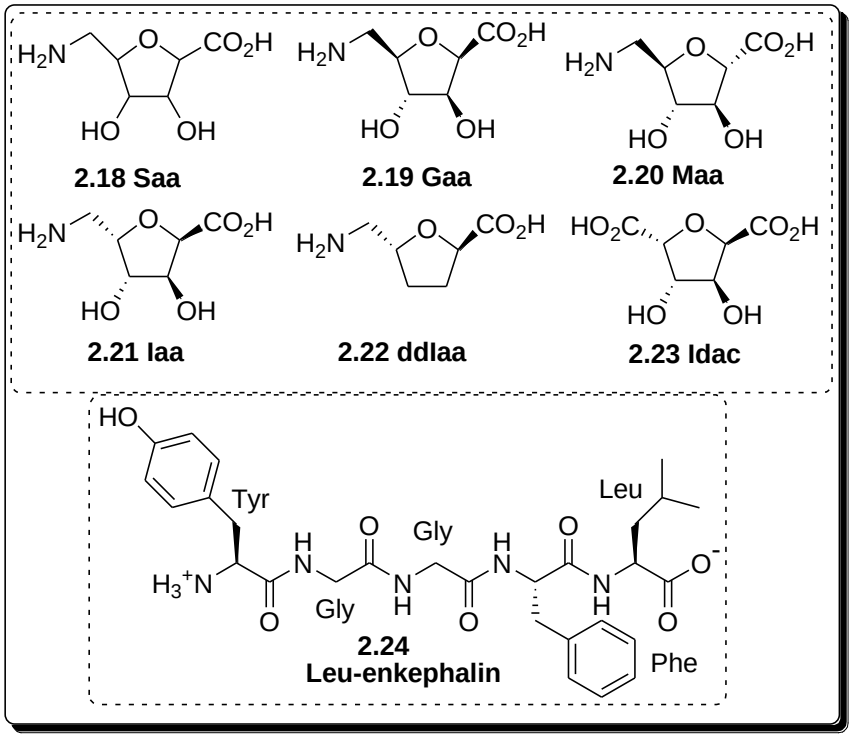


Figure 2.48: Different Sugar amino acids and diacid templates

The rigid frameworks of furanoid sugar amino acids prompted various research groups to use them as dipeptide isosteres in peptidomimetic studies. These molecules with constrained backbone angles, ω(i) and Φ(i+1), were expected to induce folded conformations in linear peptides. Chakraborty *et al*. have exploited these furanoid sugar amino acids (Saa) by incorporating into Leuenkephalin, chosen as a representative example of a short peptide, replacing its Gly-Gly spacer segment that is known to be flexible and amenable to different conformations depending on the binding environment (**Figure 2.48-2.50**).

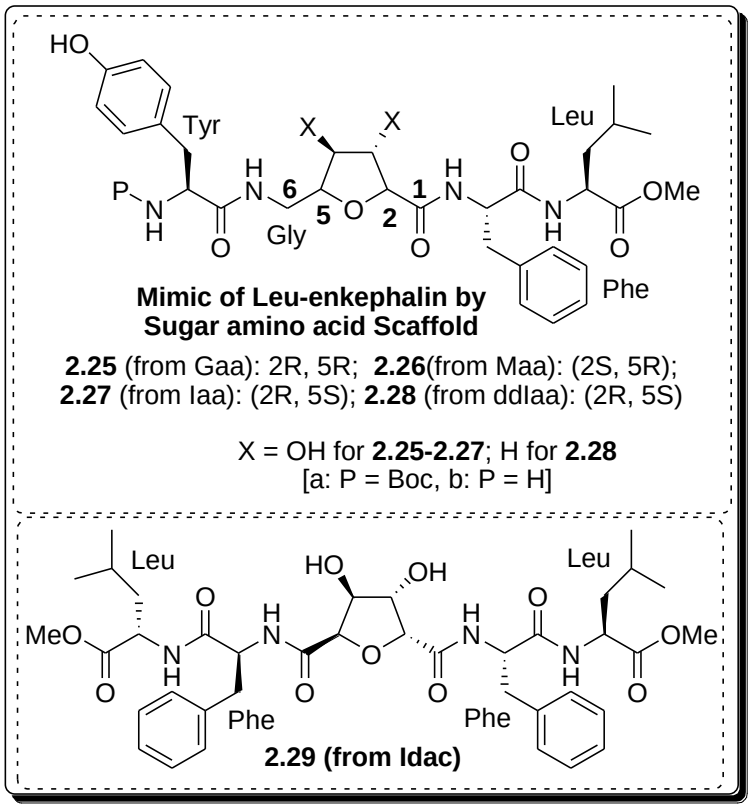


Figure 2.49: Leu-enkephalin mimetic by Chakraborty *et al.*

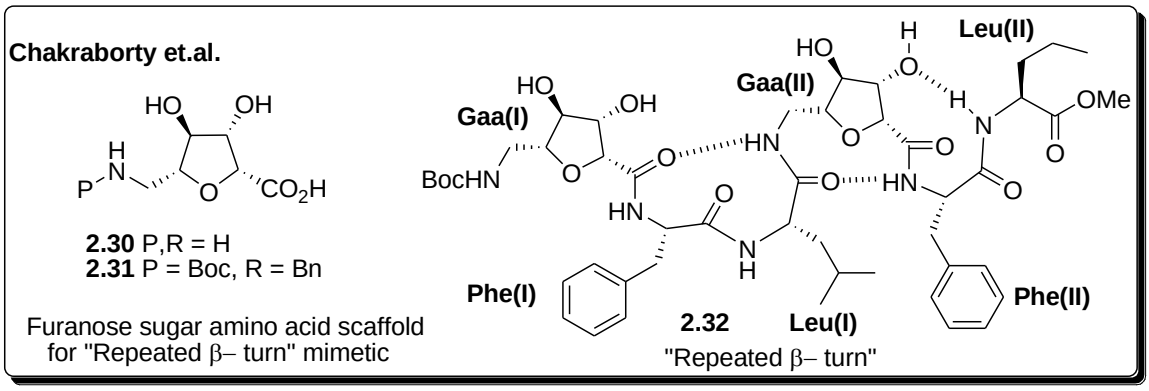


Figure 2.50: Repeated β-turn mimetic by Chakraborty *et al.*

A new reverse turn, replacing one of the native type II' β-turns in the cyclic peptide antibiotic gramicidin S, induced by a furanoid sugar amino acid was recently reported by Grotenbreg and coworkers. NMR and X-ray crystallographic analysis showed that C3-hydroxyl function plays a pivotal role by acting as an H-bond acceptor, consequently flipping the amide bond between residues i and i + 1.

Recently, Ferrocene derivatives were proposed as an organometallic scaffold for β-turns. Herrick *et.al*, in 1996, reported the compounds of general structures $Fe(C_5H_4\text{-}CO\text{-}Xaa\text{-}OMe)_2$ (with Xaa = amino acid) to have an ordered structure in organic solvents and this structure is stabilized by two symmetrically equivalent hydrogen bonds between the amide NH and the methyl ester carbonyl moiety of another strand, **2.34** (**Figure 2.51**). Hirao and coworkers prepared several substituted dipeptides of similar ordered structure.

Cobaltocene dicarboxylic acids have also been used as a scaffold for β-turn mimetic. Recently, von Staveren *et.al.* have investigated the influence of a positive charge on the structure and stability of peptide turn structures, which are stabilized by H bonds. They have concluded that turn structure in **2.35** is β-turn like and in **2.36** is β-turn-like (**Figure 2.51**). Because of parallel orientation induced by the diacid scaffold, they have designated the turn structures as pseudo-γ_p and pseudo-β_p turn.

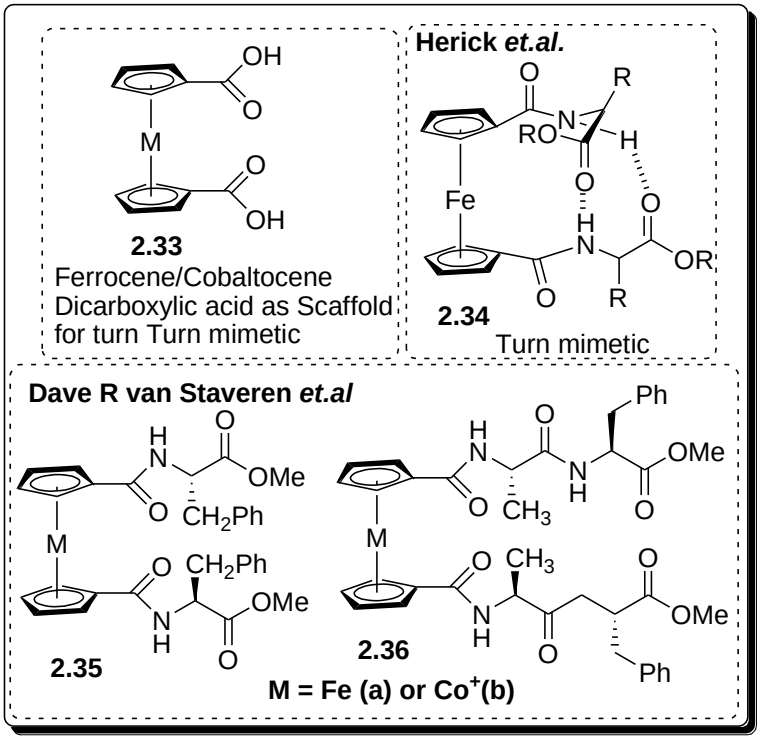


Figure 2.51: Peptide β-turn mimetic based on ferrocene/cobaltocene dicarboxylic acid scaffold

2.6.4. Z-Enediyne π-chromophores As Possible Scaffold for Peptidomimetics Design

During the past decade, the construction and investigation of expanded acetylenic π-chromophores has become a central area of chemical research. It has been fueled by the availability of new synthetic methods, in particular Pd(0)-catalyzed cross-coupling reactions, the discovery of the antitumor activity of a series of natural compounds possessing reactive Z-enediyne π-chromophores, and the need for new nanoscale molecules and polymeric materials that exhibit

unusual electronic and optical functions and properties. The enediyne antitumor antibiotics are appreciated for their novel molecular architecture, their remarkable biological activity and their fascinating mode of action and many have spawned considerable interest as anticancer agents in the pharmaceutical industry. Of equal importance to these astonishing properties, the enediynes also offer a distinct opportunity to study the unparalleled biosyntheses of their unique molecular scaffolds and what promises to be unprecedented modes of self-resistance to highly reactive natural products. Elucidation of these aspects should unveil novel mechanistic enzymology, and may provide access to the rational biosynthetic modification of enediyne structure for new drug leads, the construction of enediyne overproducing strains and eventually lead to an enediyne combinatorial biosynthesis program.

2.6.5. Enediynyl Amino Acid AS β-Sheet Nucleator

As already mentioned, Enediynes have drawn unprecedented interest amongst the scientific community because of their cytotoxic activity and possible use as anticancer drug. All studies so far have been concentrated on their synthesis and evaluation of chemical as well as biological activity. The special structural feature of Z-enediynes is the type of reverse-turn associated with the two acetylenic arms. One can consider making enediynyl amino acid containing peptides, which may be forced to adopt typical conformational motifs.

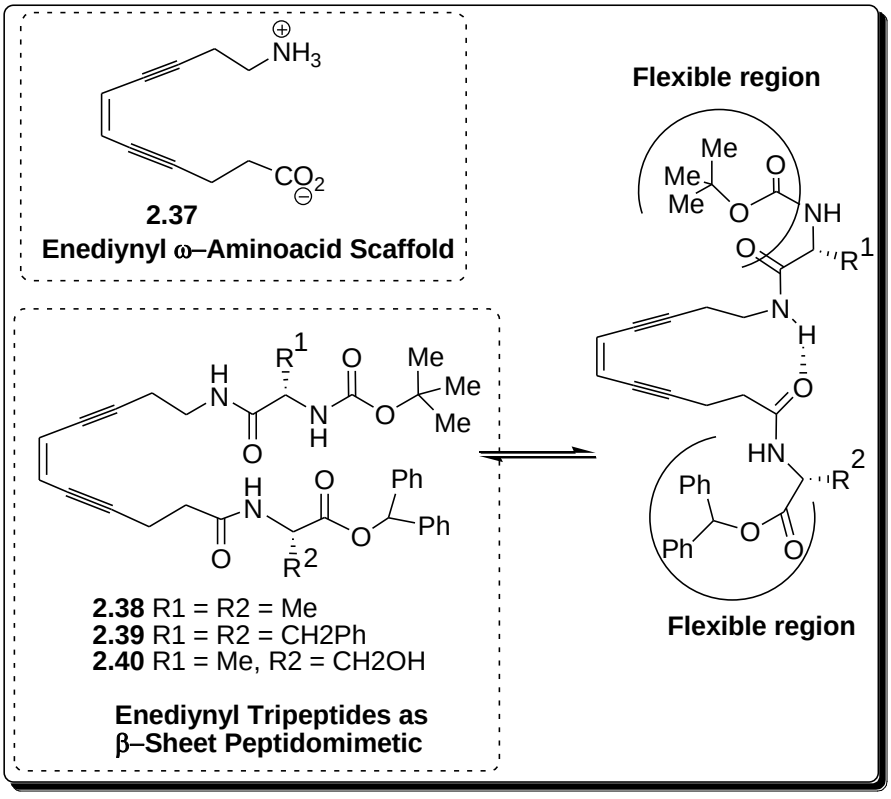
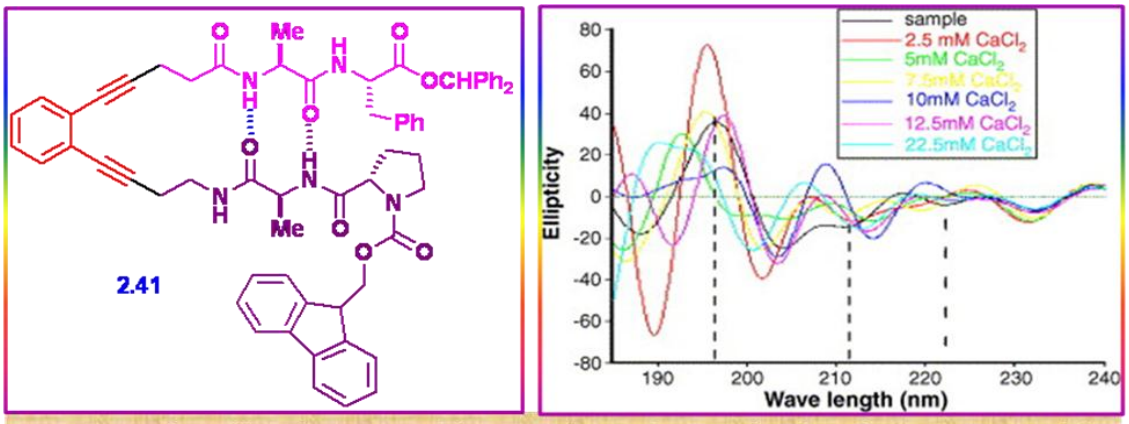


Figure 2.52: The solution conformation of the enediynyl peptides

This structural motif, for the first time in enediyne chemistry, was used, in our laboratory, as a possible scafolld for peptide secondary structure mimetic. Specially, we were interested in designing and synthesis of the β-sheet mimetic by incorporating the ω-enediynyl amino acid into a peptide chain. We thought that this enediynyl motif can act as nucleator and thus may induce in adopting the β-sheet which constitutes a well-studied subset of the reverse turn and is a common feature in biologically active peptides and globular proteins. The sheet capping turns are widely believed to act as a molecular recognition site for many biological processes.

Bag and Basak *et al.* have incorporated the enediynyl amino acid **2.37** into peptides **2.38-2.40** (**Figure 2.52**) and then found out the conformational preferences by NMR and CD-measurements.

Circular Dichroism (CD) spectra of the fully protected peptides and the generally higher ($\Delta\delta/\delta T$) values for the chemical shifts of β and γ N-H's reveales that the peptides adapt a significant proportion of β-sheet like conformation. However, the results also indicated the presence of other conformations as well, specially the α-NH as being intramolecularly H-bonded. The variable temperature NMR experiments indicate that the conformation resembling β-sheet capping type motif is more predominant. The situation is represented in **Figure 2.52.**



Structure of enediynyl pentapeptide **2.41** that exists predominantly in **β**-turn and its Overlayed deconvulated CD spectra in MeOH in the absence and presence of various amounts of $CaCl_2$.

Figure 2.53: The enediynyl pentapeptide and its CD spectra.

A novel enediynyl pentapeptide in the protected form **2.41** was synthesized and characterized. It exists predominantly in β-turn structural motif as revealed by variable temperature. NMR and CD spectroscopy. In the presence of transition metal ions and gold nanoparticles, the fluorescence intensity of the peptide got enhanced with remarkable quantum yield with the Z-enediynyl ω-amino acid acting as a fluorophoric reporter. The interesting photophysical behaviors with alkali and alkaline earth metal ions are also reported.

The secondary structure of peptide **2.41** was estimated by recording its CD spectrum in methanol, which showed a strong maximum at ~198 nm followed by several broad minima at ~205, 212, and 222 nm, indicating that the peptide predominantly adopts a β-turn like structure (**Figure 2.53**) at least in the solvent used for the study. The peptide secondary structure estimation using CD estima program shows a 60% turn like structure, the existence of which implies the possible presence of intramolecular H-bond between the peptide strands on the two arms of the enediyne framework. This could be assessed by determining the variation of chemical shifts of the various NHs with temperature in DMSO-d_6 in which all the four NHs exhibited different chemical shifts. Interestingly, the turn like structure is more or less maintained in the presence of Ca^{2+} ions.

Of the four amide NH's, one alanine NH and the NH belonging to the enediynyl amino acid exhibited $\Delta\delta/\Delta T$ values that are within the Kessler limit of −3 ppb/K, indicating strong intramolecular H-bonding and supported the predominant turn like structure of the peptide. The appearance of a crosspeak for the hydrogens attached to C-2 and C-11 in NOESY spectrum also provided further evidence for the turn like conformation of the two peptide arms of the enediyne backbone. The H-bonded conformation was also supported by the semi-empirical AM1 geometry optimization.

2.7. β-Lactam Based Peptidomimetics

Templated β-turn peptidomimetic is a complex task in respect of the synthesis of molecular templates with precisely oriented predetermined groups. A great deal of effort has been dedicated to the development of such peptidomimetics. A rational design of an efficient mimetic necessitates the modification of the native bioactive peptide to include two major elements-(a) a specific β-turn constraining element such as an intra- or inter residual bridge, is needed to force the peptide backbone to adopt the desired conformation as exactly as possible and (b) at least one recognition group must be placed in a stereocontrolled fashion at the desired position for interaction with the receptor or enzyme active site.

Though it is simple to design, practical access to peptidomimetic libraries bearing a broad range of recognition groups, while retaining the β-turn framework, often is a tedious and synthetically challenging job starting from the scaffold chosen. Also, it is difficult or impossible to elucidate structural elements exerting specific constraints or recognition functions for most known β-turn surrogate molecules. As a consequence, chemical alteration of such peptidomimetics, when synthetically feasible, leads to unpredictable results in terms of β-turn motif stability and/or proper spatial orientation of the recognition groups.

An important example of this phenomenon (**Figure 2.54**) is the original approach of Freidinger to β-turn mimetics. The Freidinger's original approach to β-turn mimetic, involved the bridging of the betagenic -(*i*+1)-(*i*+2)- central residues through the formation of five-, six-, seven-, and eight-membered ring lactams. Accordingly, any structural modification of the R^1 bridge in

pseudopeptide (**Figure 2.54**), should force a global change of conformation and recognition properties that, in practice, involves a re-design of each derivative prepared, instead of a single structural parameter variation.

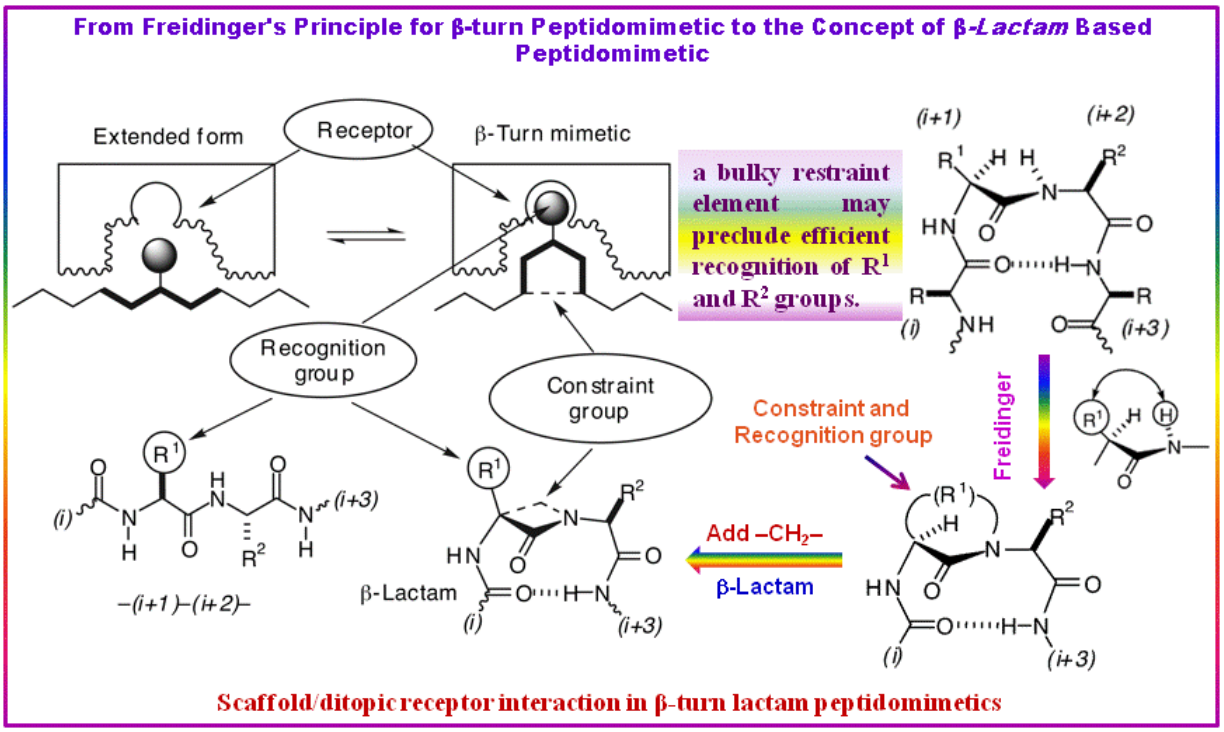


Figure 2.54: The designing concept of β-lactam based peptidomimetics.

The Freidinger's idea of β-turn peptidomimetics based on the principle of the incorporation of groups, as small as possible, in the original peptide to generate well-defined turns without affecting receptor recognition was explored by Claudio Palomo *et al.* in designing the β-lactam-based peptidomimetics as a novel family of β-turn motif nucleators. This concept was illustrated by their designed β-lactam peptides and β-lactam peptidomimetics of melanostatin (PLG).

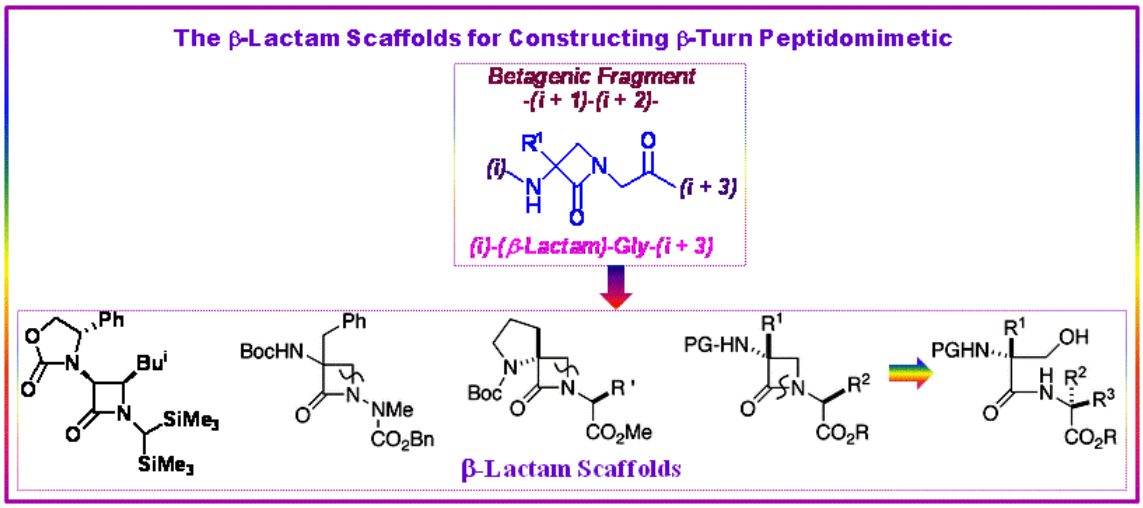


Figure 2.55:β-Lactam scaffolds for the peptidomimetics; Palomo *et.al.*

They have synthesized novel enantiopure (i)-(β-lactam)-(Gly)-(i+3) peptide models starting with β-lactam scaffolds, defined by the presence of a central α-alkyl-α-amino-β-lactam ring placed as the (i+1) residue. In their conceptual design, while the β-lactam methylene group mimics simultaneously the CαH(i+1) and HN(i+2) protons of the native peptide, R^1 and R^2 groups can be designed specifically for recognition with the receptor (**Figure 2.55**). In this approach, it was also assumed that the presence of the α,α-disubstitution pattern would enhance resistance to chemical and enzymatic hydrolysis by proteases and that the resulting β-lactam-peptides would be attractive targets for pharmaceutical drug discovery.

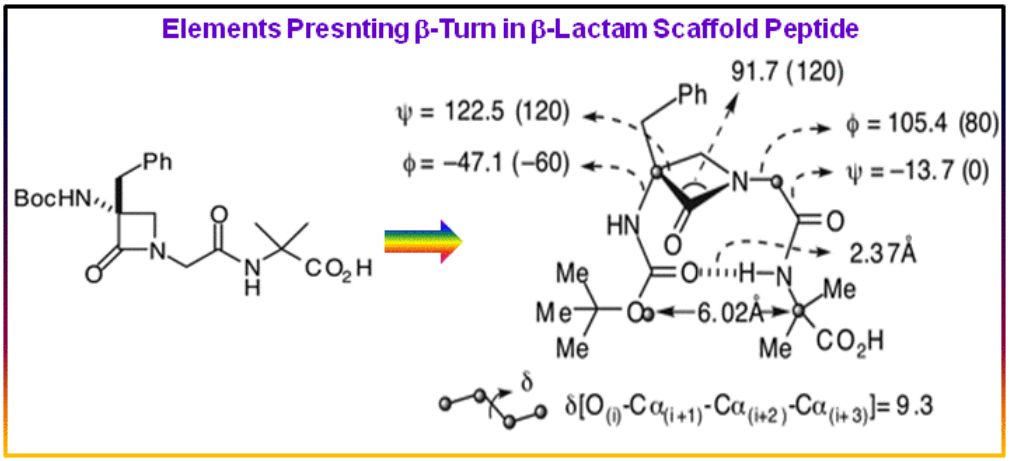


Figure 2.56: β-lactam based peptidomimetic indicating a type-II' β-turn conformation.

The structural properties of these β-lactam pseudopeptides were studied by x-ray crystallography, molecular dynamics simulation, and NOESY-restrained NMR simulated annealing techniques, showing a strong tendency to form stable type II or type II' β-turns either in the solid state or in highly coordinating DMSO solutions. Crystal structure of the peptidomimetic (**Figure 2.56**) indicated a type-II' β-turn conformation.

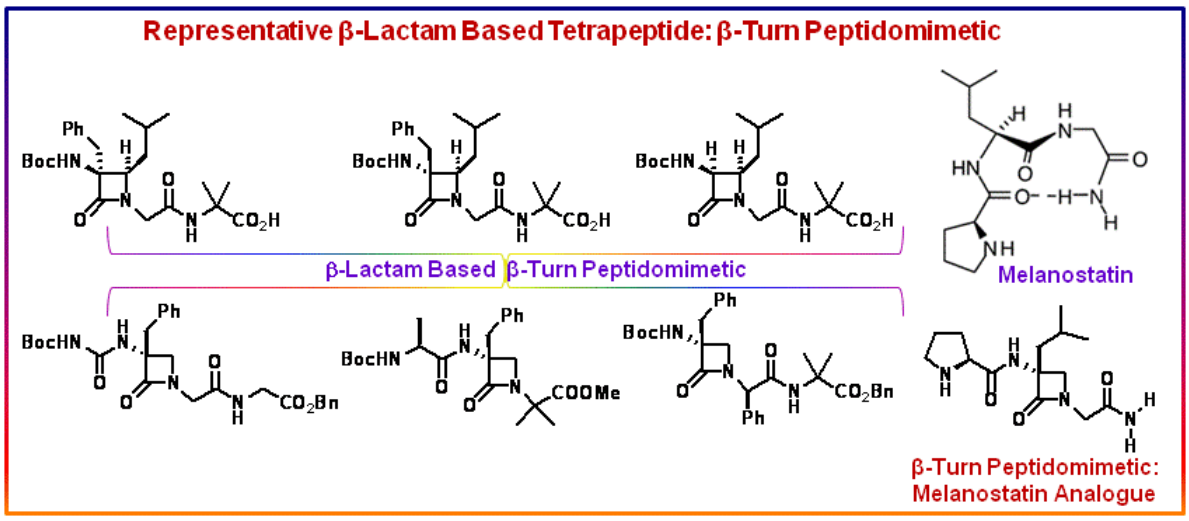


Figure 2.57: Tetrapeptide models as a function of the structural and stereochemical elements of the β-lactam core for β-turn induction.

Tetrapeptide models containing syn- or anti-α, β-dialkyl-α-amino-β-lactam rings were also synthesized and their conformations analyzed, revealing that α -alkyl substitution is essential for β-turn stabilization. A β-lactam analog of melanostatin (H-Pro-Leu-Gly-NH_2) was prepared. The compound (**Figure 2.57**) was characterized as a type-II β-turn in DMSO-d_6 solution, and was tested by competitive binding assay as a dopaminergic D_2 modulator in rat neuron cultured cells, displaying moderate agonist activity in the micromolar concentration range.

Biological Evaluation of Melanostatin (PLG) β-Lactam Analogue: Several studies have shown that PLG render the Gi-protein coupled dopamine D_2-like receptors more responsive to agonists by maintaining the high-affinity binding state of the receptor. Though the full mechanism is not clear, but it is known that the mechanism of action of PLG may likely involve intraneuronal enzymes. So, direct competitive binding assays on integral neurons can provide information more directly assignable to real living systems.

To test the relative activity of β-lactam PLG analogue versus natural PLG as a dopamine D_2 receptor modulator, Palamo et al. have conducted a radio ligand competitive binding assays on cultured integer neuron cells from rat cerebral cortex. Their competitive binding assay study indicated that β-lactam PLG analogue (**Figure 2.57**) slightly reduces the dissociation constant of dopamine D_2 receptors for N-propylnorapomorphine (NPA) with respect to PLG at a micromolar concentration range. The full retention of bioactivity on living neurons observed for β-lactam PLG analogue with respect to natural PLG suggests that the structural variation arising from the replacement of C$\alpha H_{(i+1)}$ and $HN_{(i+2)}$ protons by a β-lactam methylene group does not hinder the maintenance of a similar recognition pattern by D_2 receptors.

2.8. Expanding the Genetic Code

2.8.1. Background:

The tools of chemistry, most notably chemical synthesis and spectroscopy, have had a remarkable impact on biology-from the structural elucidation of the double helix to the chemical synthesis of peptides and oligonucleotides. At the same time, modern molecular biology has made it possible not only to manipulate protein and nucleic acid structure but also the genetic composition of living organisms. The ability to use these tools in combination opens an unprecedented opportunity in the coming millennium, both for understanding complex biological systems at a molecular level as well as for the generation of molecules with novel biological, chemical and physical properties.

Proteins play a vital role in all living organisms to maintain the cell structures, properties and functions which is again dependent on post translational modifications. The genetic codes of every known organism encode the same 20 amino acid building blocks using triplet codons generated from A, G, C and T. Thus, in all organisms, the building blocks of all the translated proteins are the same 20 natural amino acids. Natural selection/evolution has generated a large no of proteins with more or less common structures and functions in a population. Therefore, these proteins are highly specialized for specific functions and thus, are not suitable for a different function other than they use to do the specific job. To perform a complex additional function by a protein, it needs other functionalities within its framework. Posttranslational modifications, cofactor-dependent catalysis, and pyrrolysine/selenocysteine incorporation in bacteria proves that the natural evolutionary movement needs extra chemical functionalities other than those present within the 20 natural amino acids. Then the knowledge from the nature knocked scientists to ask oneself: Why only this set of amino acids and not any additional ones are used for genetic coding? Is this the ideal number? How did they determine the complex folds and functions of proteins? Would additional amino acids allow the generation of proteins or even entire organisms with enhanced properties? These questions need yet to be fully answered.

The ability to generate proteins with new building blocks, beyond those specified by the genetic code, would not only provide a powerful tool with which to investigate these questions but might also allow us to generate proteins or even entire organisms with novel functions. The ability to introduce amino acids with precisely tailored steric and electronic properties into proteins would also allow us to carry out “physical organic” studies of proteins much the same way as has been historically done with small molecules.

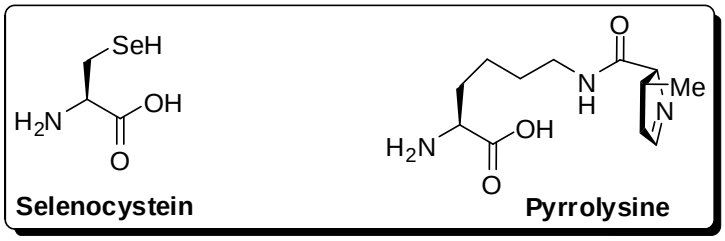


Figure 2.58: Presentation of naturally occurring 21st and 22nd amino acids

Thus, the approach, "directed evolution" have come up with proteins of altered structural and functional properties that do not occur in nature. As for example, researchers are trying to generate new proteins to function as potent therapeutic, microbial enzymes for fuel production from agricultural waste, imaging and ultimately semi-synthetic organism with diverse functionality. Now a day, scientists are involved in designing the proteins via rational approach of directed evolution to produce new proteins with desirable properties. Therefore, inspired by Natures' post-translational modification to include different functional units or molecules into the proteins, chemoselective conjugation methods have been developed to attach probe to proteins; thereby facilitating the study of structure and functions at molecular level. However, conventional bioconjugation reaction has several drawbacks and mostly exploited the only nine canonical aminoacids with limited functional groups and abundance for modification/ligation within a protein. Therefore, because of that, site-specific conjugation to a desired canonical amino acid in a protein is difficult task. To circumvent this problem, several chemical and biochemical methodology have been developed to site-specifically incorporate designer amino acids (the unnatural amino acids) with disered functionalities to probe protein structure and function and to generate proteins/enzyme for several novel biochemical applications, such as for chemical synthesis, biomedical research or even as therapeutics. Therefore, the research in the field of design and synthesis of non-natural amino acids with novel properties, to encode it genetically and to incorporate it site-specifically into a protein via bio-orthogonal strategy, is growing at a fast space for the growing demand of proteins of potential therapeutic and many other diversified novel functional applications. Therefore, an expanded genetic code, site-specifically incorporated into protein would allow us to study proteins' physical organics which would otherwise be extremely difficult-(a) probing protein structure, function and interaction, (b) regulating protein activity, (c) monitoring the mode of action, (d) improving immunogenicity, and (e) very recent development of a protein with a "chemical warhead" which target specific cellular components.

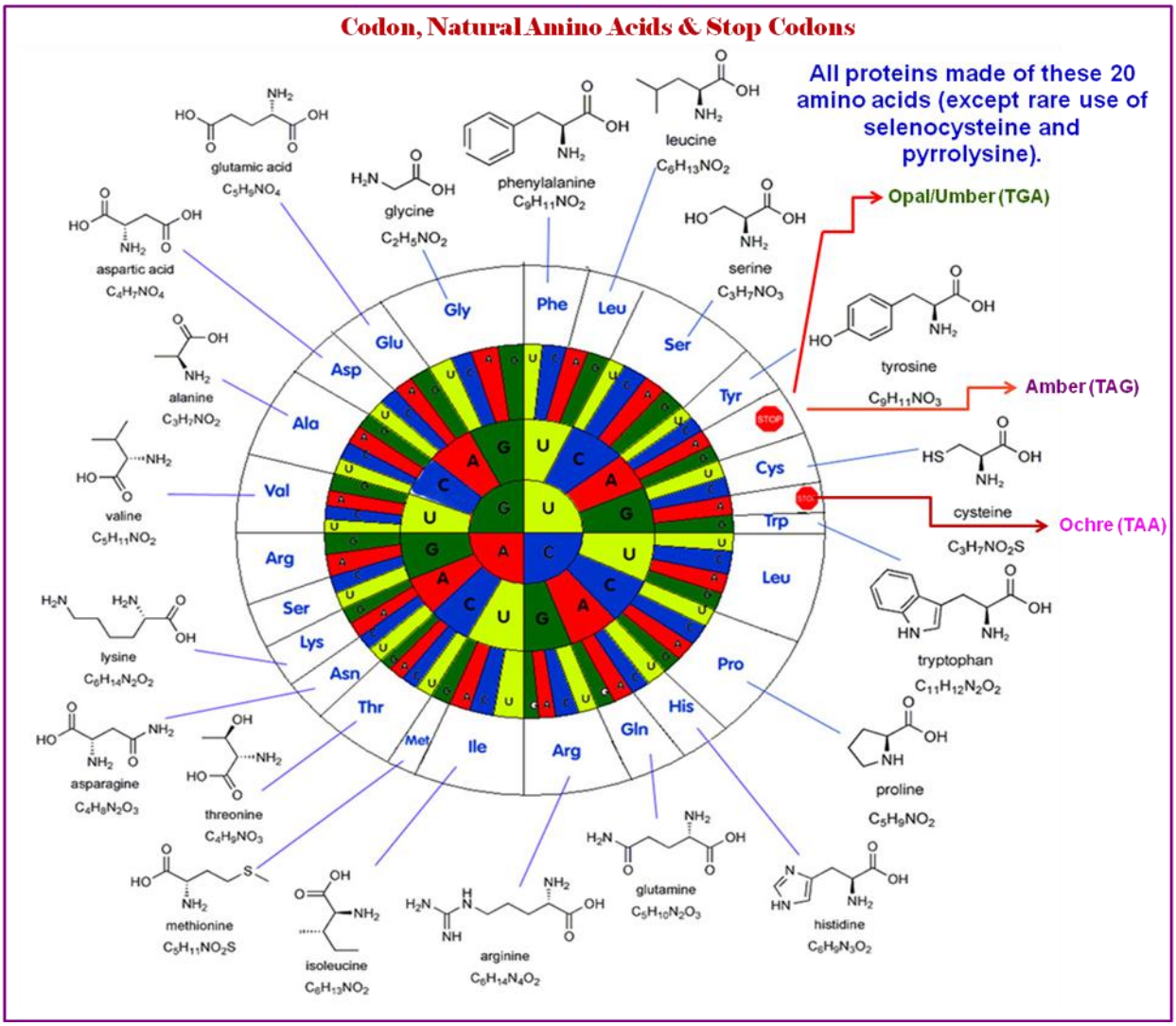


Figure 2.59: The structures of the amino acids and their codons.

2.8.2. Methodology for Genetic incorporation of UNAA:

The historical development of synthesis or semisynthetic methods for introduction of unnatural amino acids into peptide and proteins[(a) Offord, R. E. (1987) *Protein Eng.* 1,151-157; (b) Kaiser, E. T. (1988) *Angew. Chem. Int. Ed.Engl.* 27,913-922] by Offord and Kaiser has paved the way to develop methods for site specific incorporation of more unnatural amino acids into proteins by Schultz *et al.* [Corey, D. R. and Schultz, P. G. (1987) *Science* 238, 1401-1403] towards expanding the genetic code and thereby to give rise to the birth of semisynthetic organism. Towards this goal, an increasing amount of interest from various research groups has resulted in the acceleration of progress of design of unnatural amino acids for application in protein engineering. Several non natural amino acids were reported looking after the steric and electronic properties and incorporated into proteins site specifically. However, many of the

reported unnatural amino acids are not suitable for giving novel biological properties of the proteins or not containing functionality for labeling the proteins. Moreover, ffluorescently labeled proteins are useful in a large number of bioanalytical applications, but, a little was attempted to develop fluorescent unnatural amino acids (FUAA) or fluorescently labelled UAA for genetic encoding or to generate labelled proteins/peptide for studying conformational or diverse functional realm.

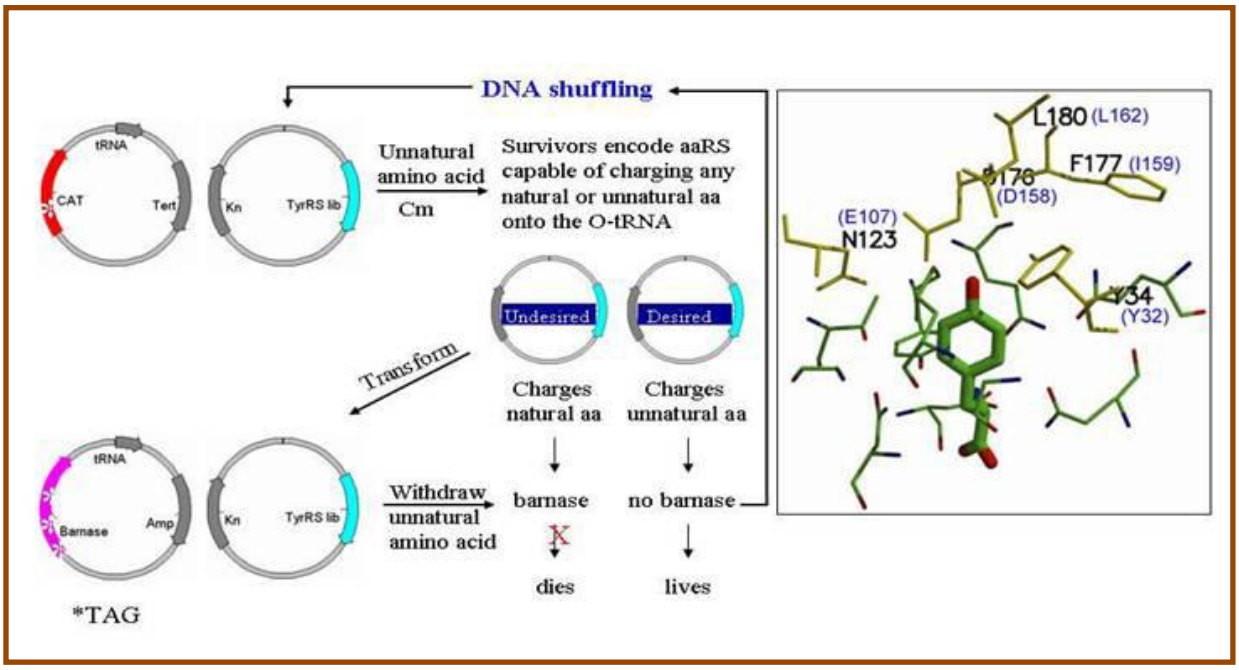


Figure 2.60: Scheme for evolving aminoacyl tRNA synthetases with novel specificities.

Towards this end, methodologies were developed that allow one to genetically encode novel amino acids, beyond the common twenty, in prokaryotic and eukaryotic organisms. One of such methodologies involves the generation a unique codon-tRNA pair and corresponding aminoacyl-tRNA synthetase. Specifically, an orthogonal tRNA is constructed that is not a substrate for any natural aminoacyl synthetases and which inserts its cognate amino acid in response to the amber nonsense codon. A cognate synthetase is then generated which recognizes this unique tRNA and no other; the substrate specificity of this synthetase is then evolved to recognize a desired “twenty first” amino acid, and no endogenous amino acid. Sultz et al. have shown that this methodology can be used to efficiently incorporate a large number of amino acids into proteins in *E. coli* and yeast with fidelity and efficiency rivaling that of the common amino acids. Using this methodology, several research groups have added a variety of novel amino acids to the genetic codes of *E. coli*.

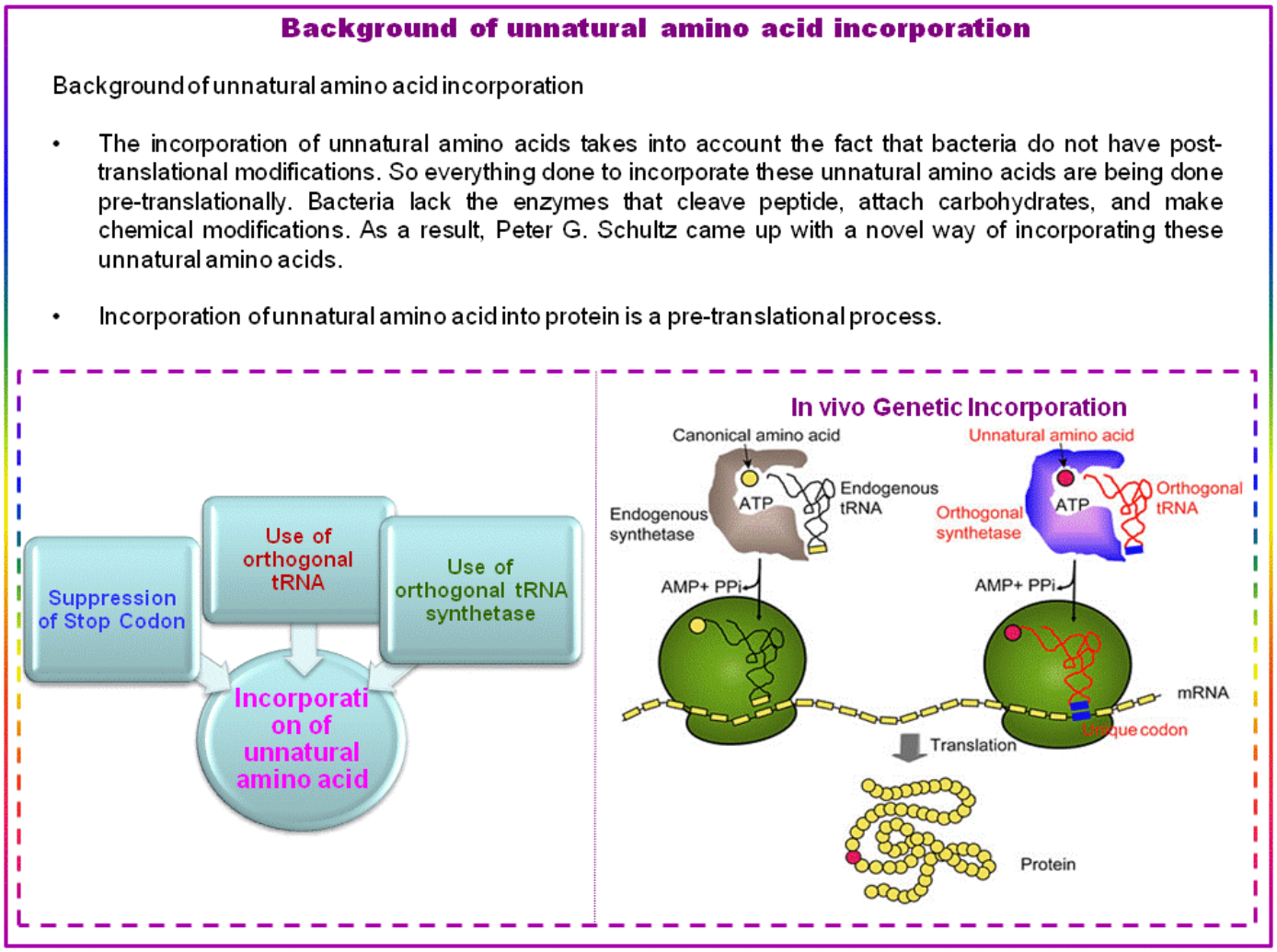


Figure 2.61: Scheme for *in vivo* incorporation of unnatural amino acids.

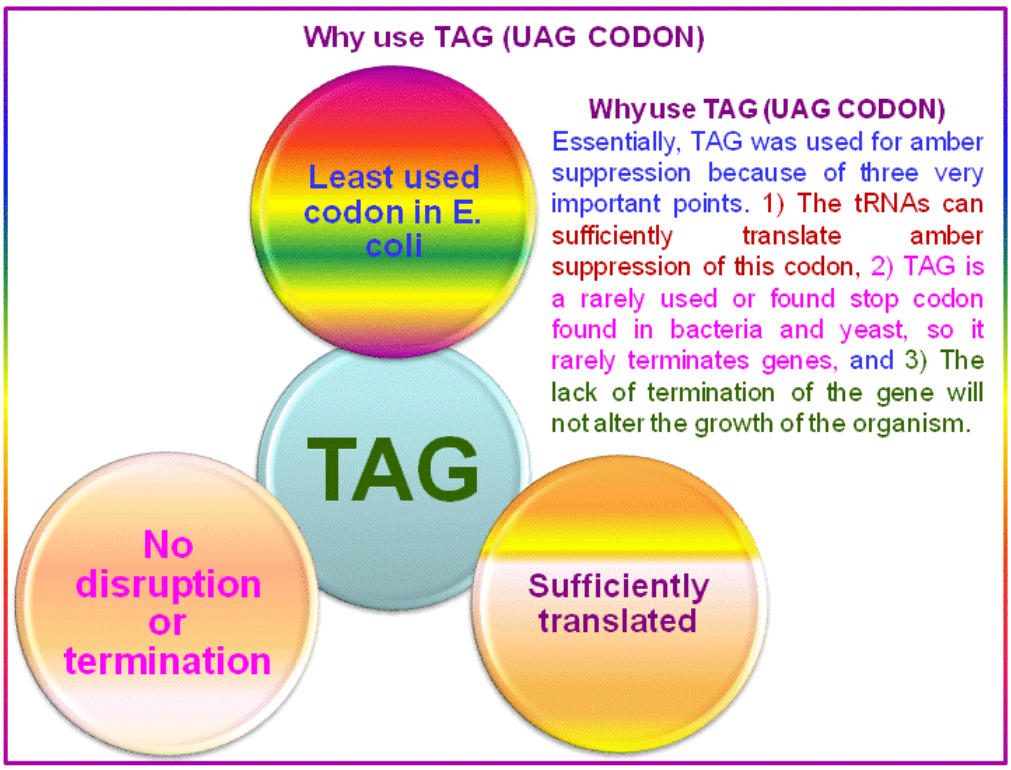


Figure 2.62: Importance of stop codon for incorporation of unnatural amino acids.

Applying this methodology one can study protein's structure and function *in vitro* and *in vivo*, as well as the evolution of proteins with novel properties, including therapeutic peptides, proteins, and vaccines can be developed.

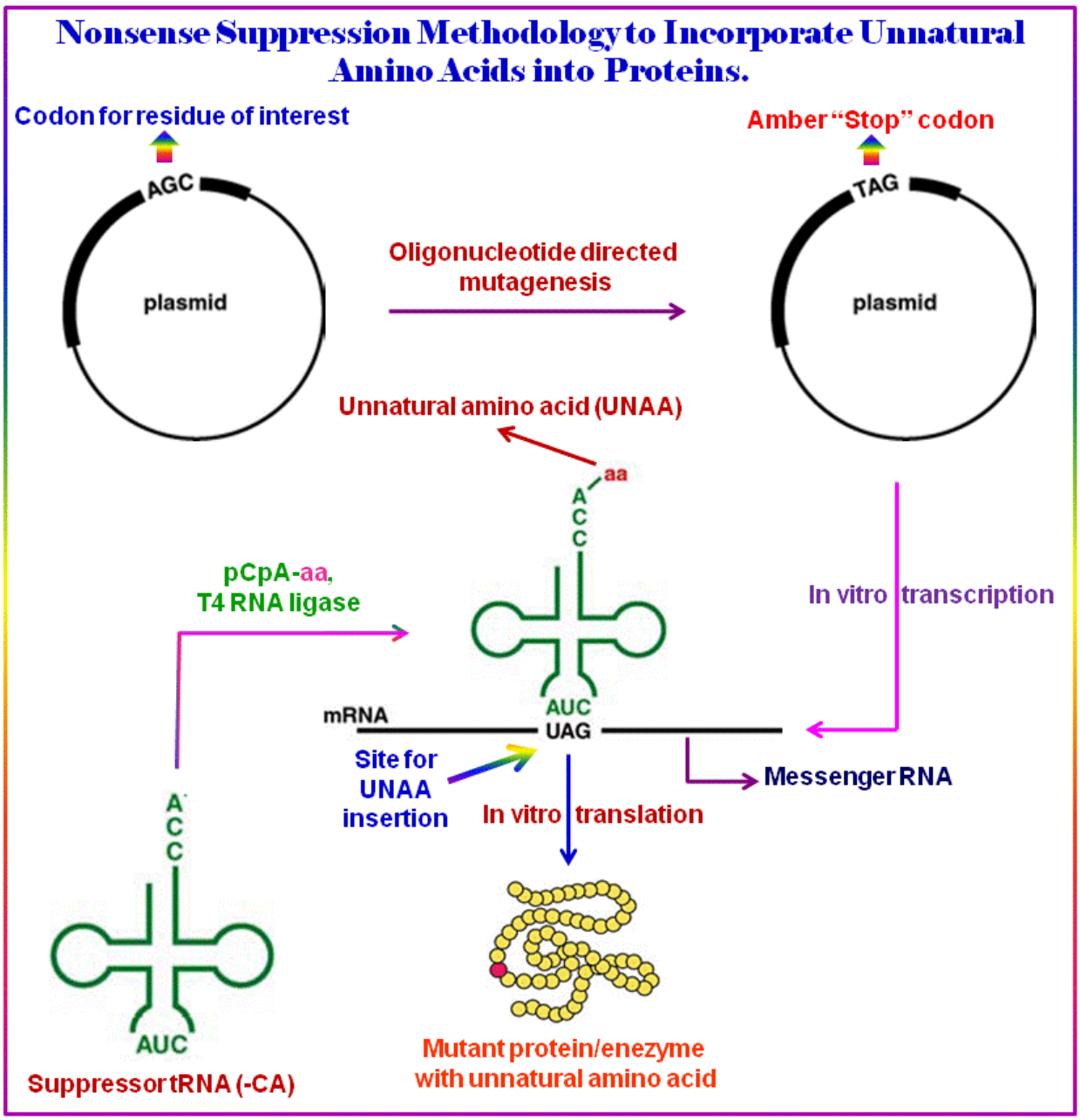


Figure 2.63: Stop codon methodology for incorporation of unnatural amino acids.

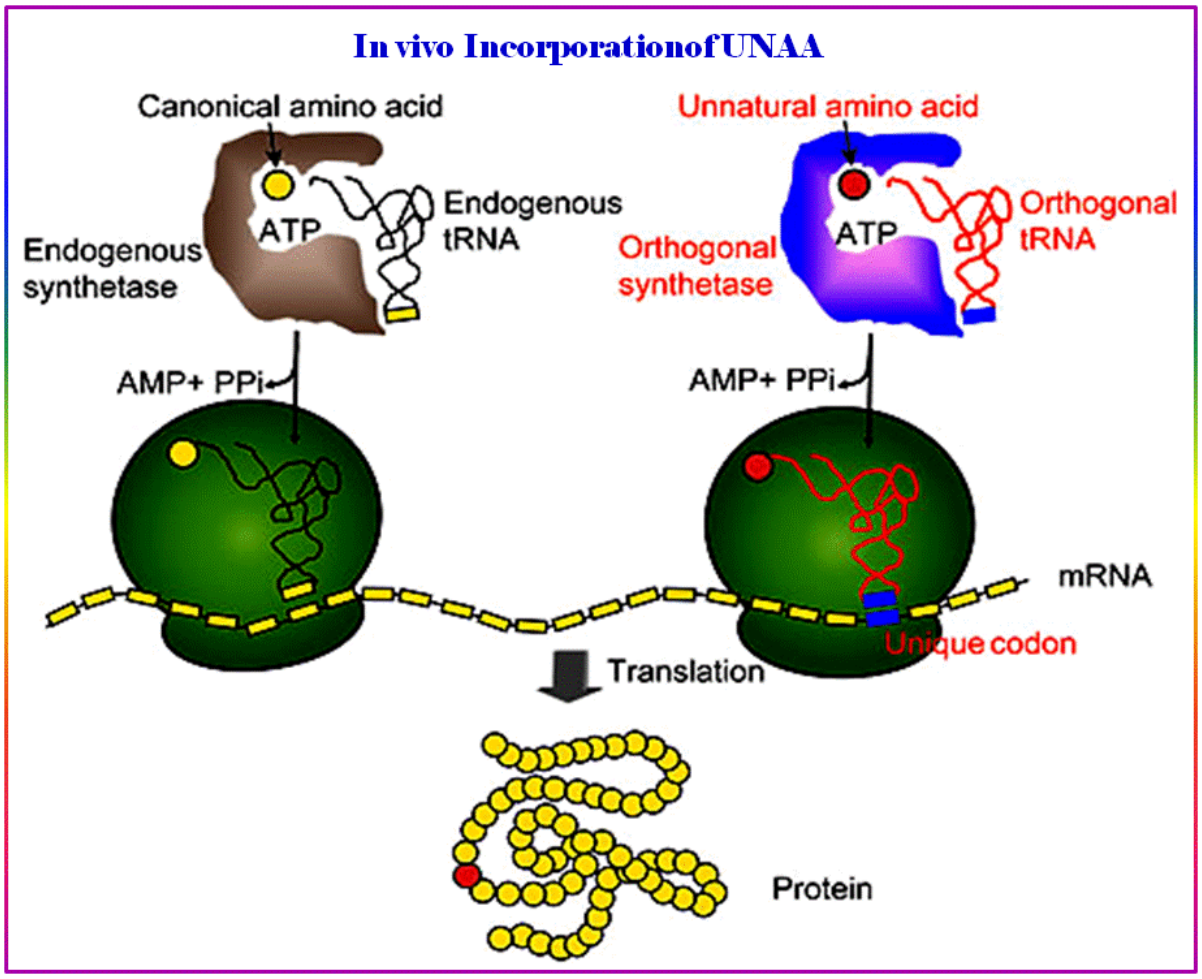


Figure 2.64: Methodology for *in vivo* incorporation of unnatural amino acids.

2.8.3. Some Genetically Encoded UNAA:

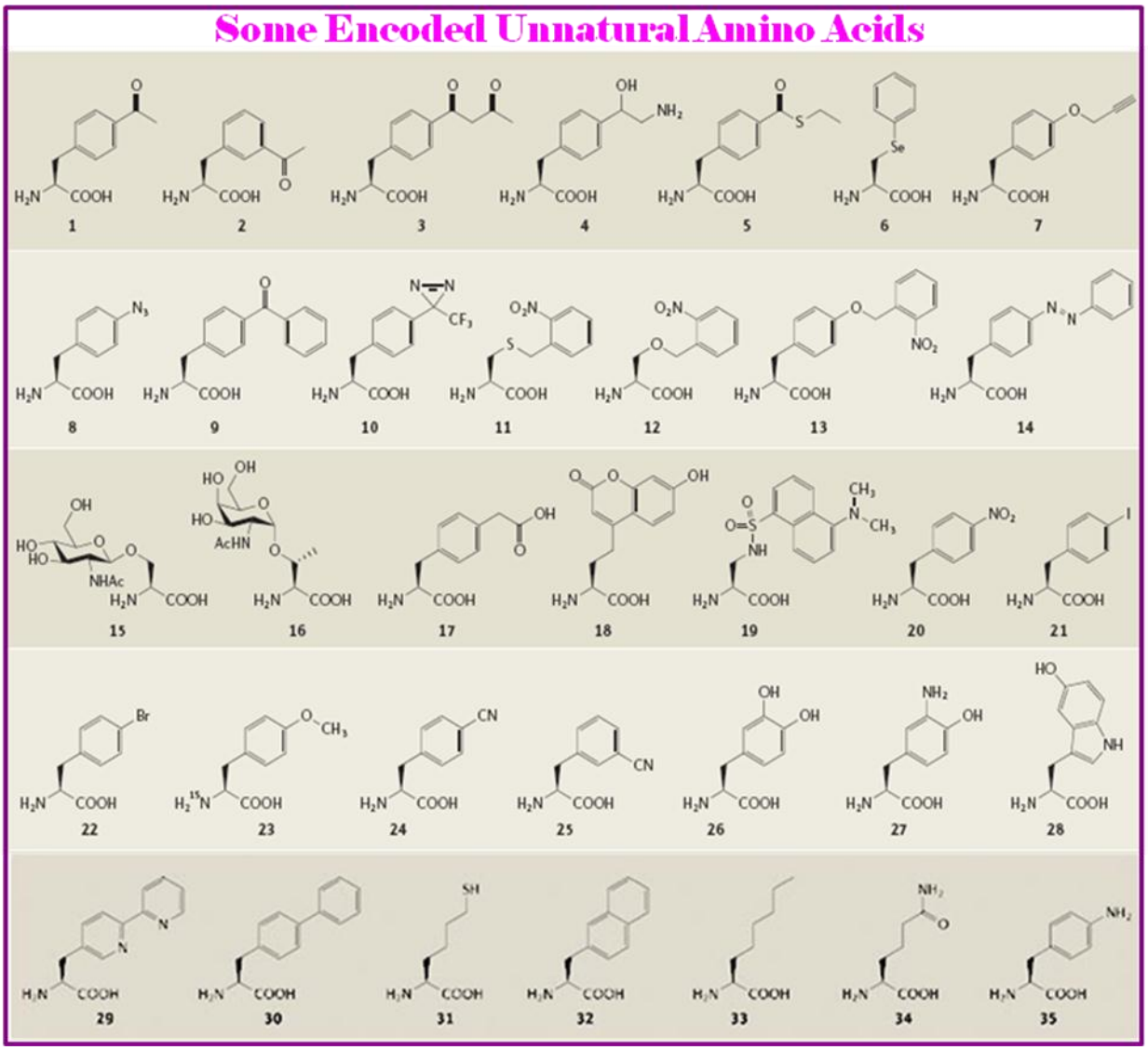
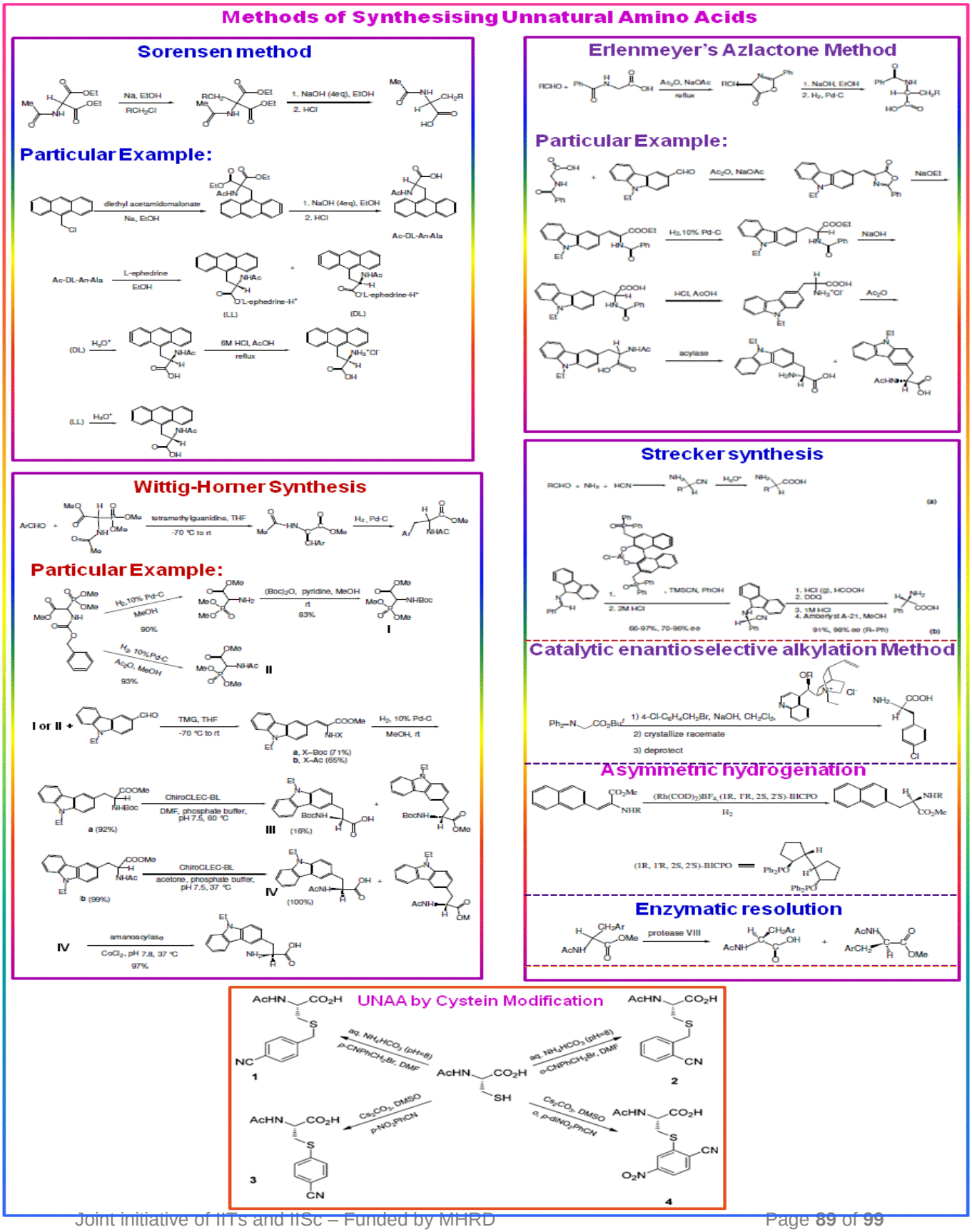


Figure 2.65: Structures of some genetically encoded amino acids.

2.9. Synthesis of Unnatural Amino Acids



2.10. Applications of Unnatural Amino Acids

There are many applications of unnatural amino acids which include the following:

1. The generation of therapeutic proteins with enhanced pharmacology is possible in contrast to the historical, relatively nonspecific methods for the chemical modification of therapeutic proteins with electrophilic moieties or the selective modification of cysteine residues.
2. Increase the immunogenicity of self-proteins or weakly immunogenic pathogen proteins (e.g., p-nitrophenylalanine mutants). This method can be applied to the development of cancer and antiviral vaccines.
3. Photocaged amino acids can activate enzymatic activity or protein phosphorylation in living cells, photochemicaly, in a temporally and spatially defined fashion.
4. The generation immunotoxins, antibody-based imaging agents, antibody−DNA conjugates, and bispecific antibodies as well as carrier−peptide conjugates with enhanced pharmacokinetics or targeted activities.
5. Fluorescent amino acids can be used for in vitro and cellular imaging of protein localization, biomolecular interactions, and conformational changes with the ability to place these small probes at virtually any site in the proteome.
6. Multidentate metal ion binding amino acids can enhance the protein's redox and hydrolytic activities.
7. Redox amino acids can be used as mechanistic probes of electron transfer in enzymes, isotopically labeled amino acids as IR probes of protein dynamics, and sterically modified amino acids as probes of ion channel activation.
8. The mapping biomolecular interactions in cells and identifying orphan ligands and receptors can be achieved by the use of photo-cross-linking amino acids.

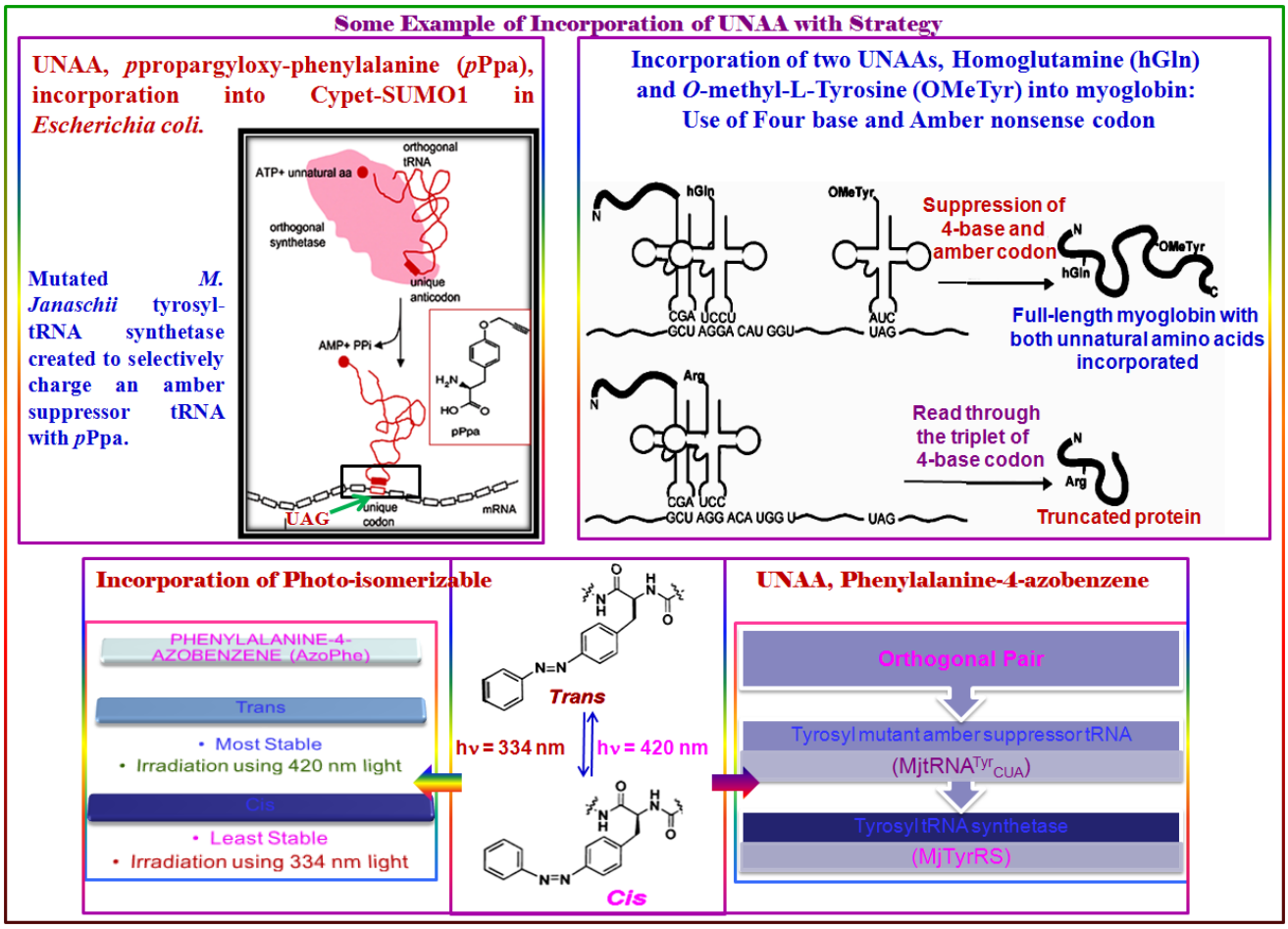


Figure 2.66: Schematic Presentation of incorporation of UNAA with strategy.

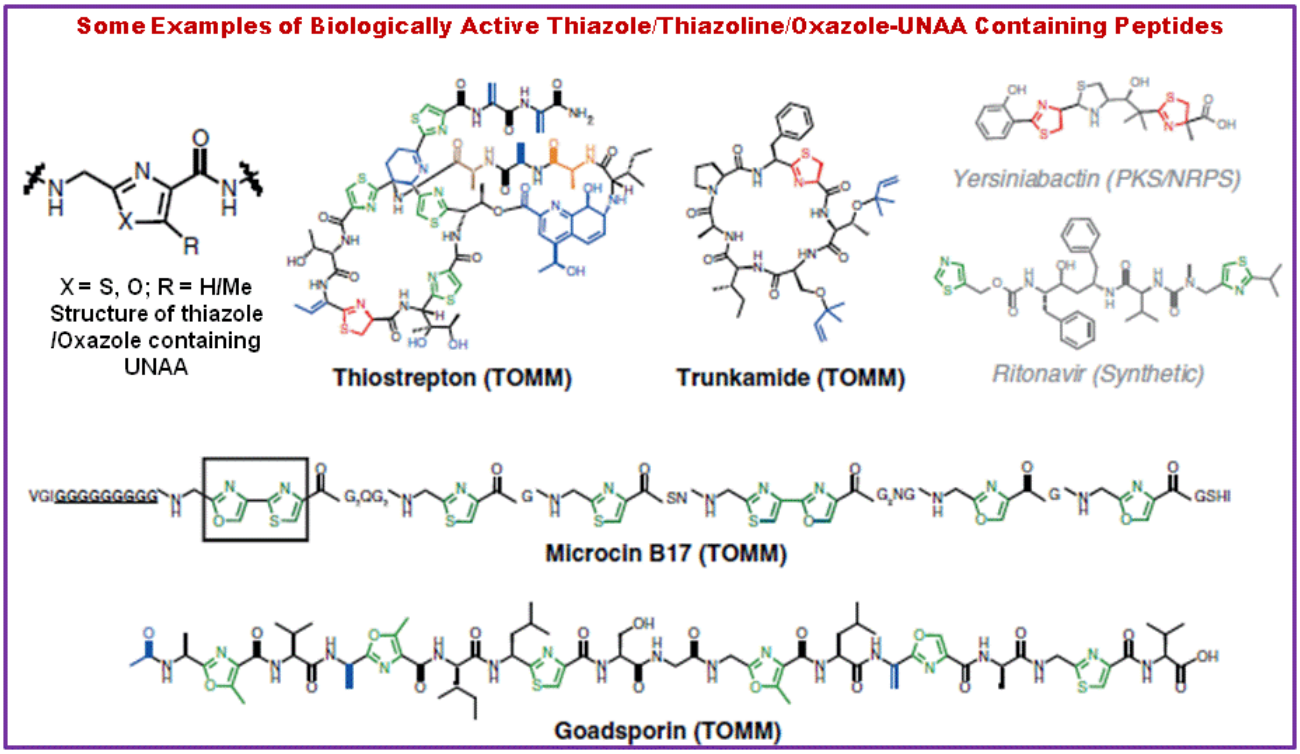


Figure 2.67: Some examples of biologically active peptides containing UNAA.

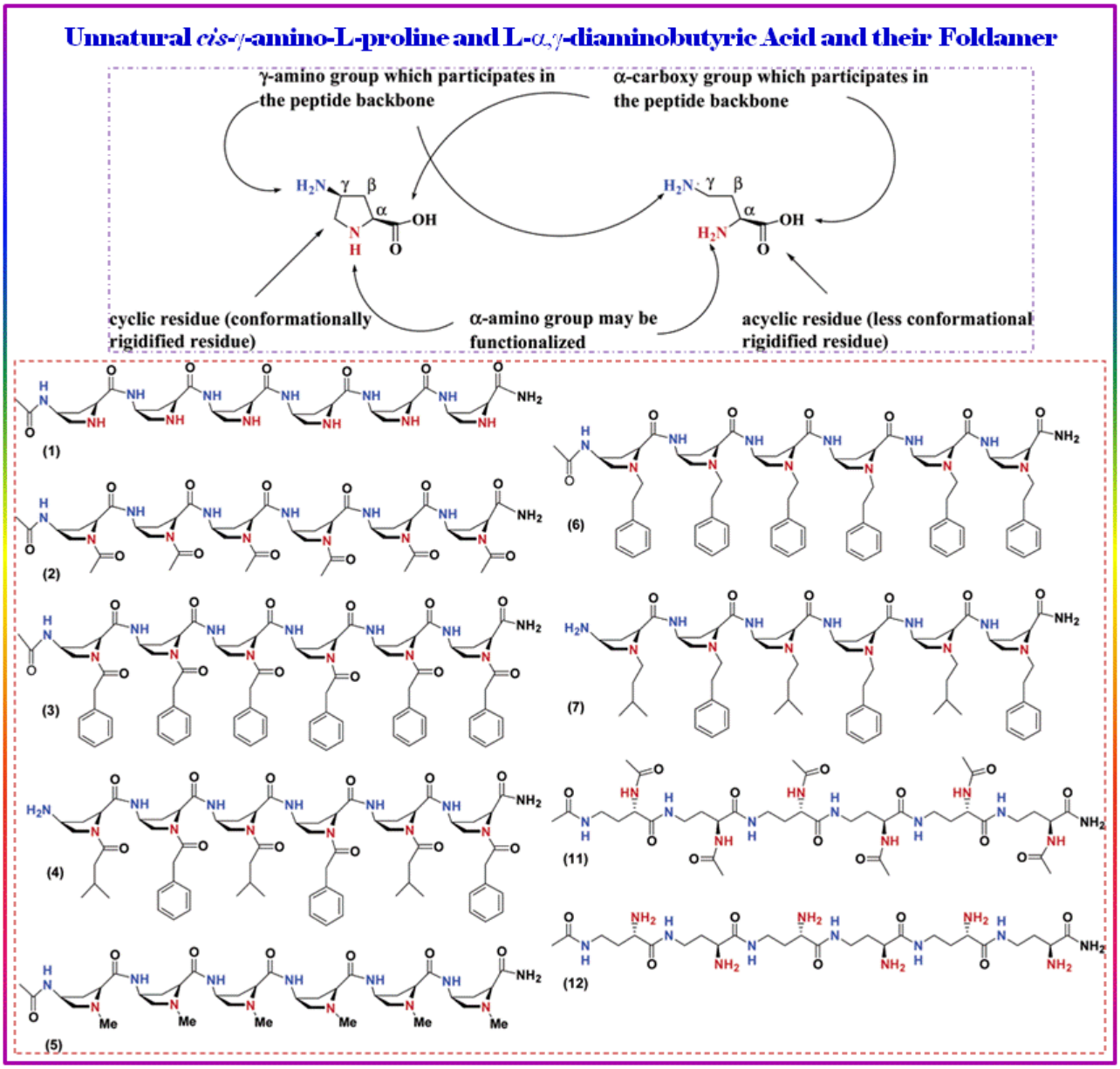


Figure 2.68: Some examples of peptide foldamers containing UNAA.

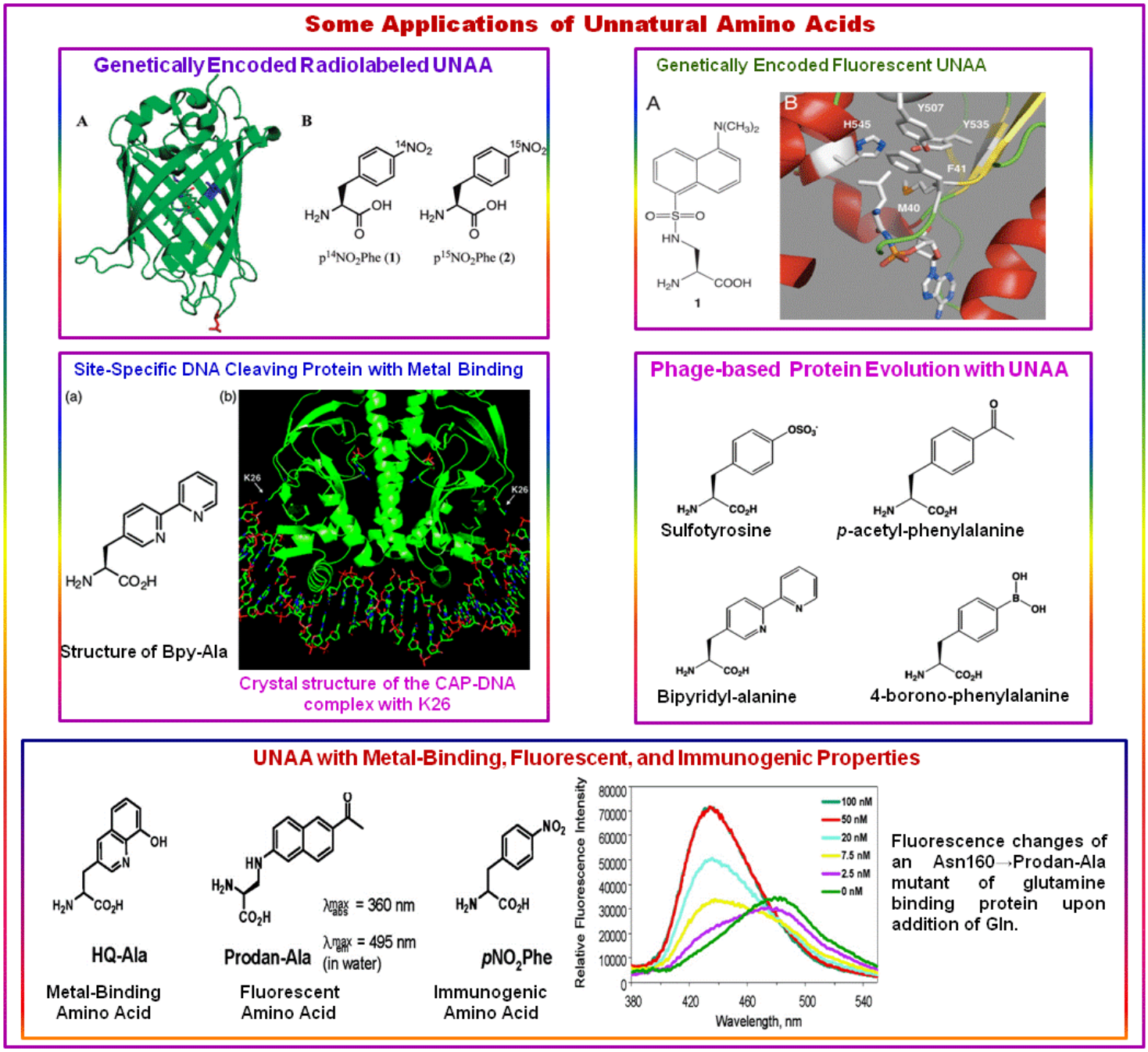


Figure 2.69: Some applications of encoded UNAA.

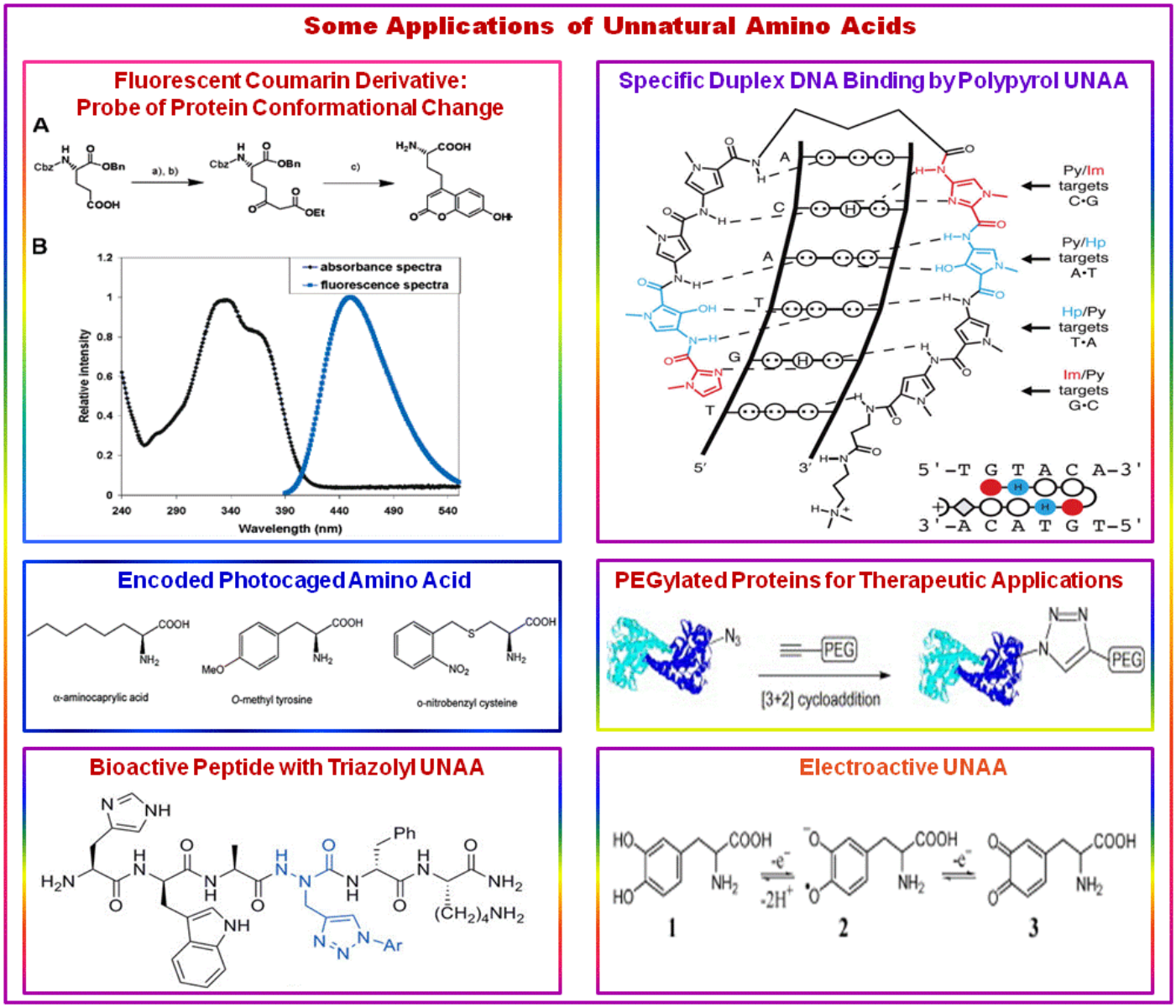
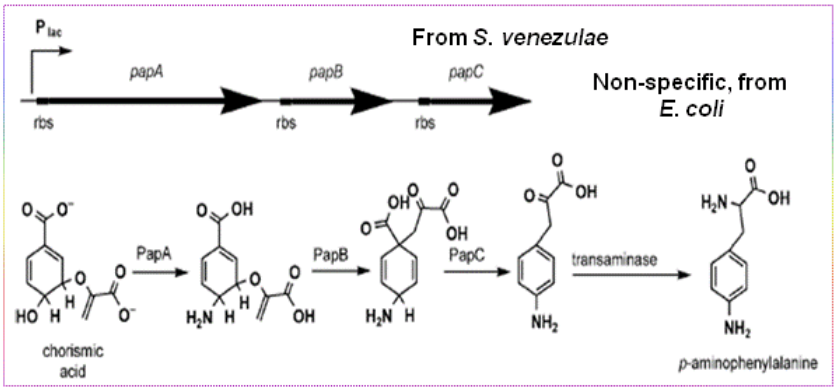


Figure 2.70: Some applications of UNAA.

Autonomous Bacterium with a 21 Amino Acid Genetic Code : Biosynthesis of Artificial Amino Acids *in vivo*

Schultz *et al. have successfully generated* completely autonomous bacterium with a 21 amino acid genetic code. Which was exploited for the biosynthesis of a nonstandard amino acid from basic carbon sources and incorporate this amino acid into proteins in response to the amber nonsense codon. Thus, their findings of such organisms may provide an opportunity to examine the evolutionary consequences of adding new amino acids to the genetic repertoire, as well as generate proteins with new or enhanced biological functions.



Glycan-Binding Proteins Evolution with a "Chemical Warhead", *p*-boronophenylalanine

Schultz et al. have recently demonstrated that a 21 amino acid code containing a genetically encoded sulfotyrosine afforded a selective advantage in a phage-based system for the evolution of gp120-binding antibodies. Later on they were succesfull in designing unnatural amino acids containing "chemical warheads" and showed its advantage to the evolution of proteins that target specific functional groups. Thus they have reported the generation of antibodies in an *Escherichia coli* strain containing unnatural amino acid, *p*-boronophenylalanine. They have shown the evolution of glycan-binding proteins in Boro-X-*E. coli*.

Thus their novel findings might shed light for designing additional encodable unnatural amino acid that can offer an evolutionary advantage which will drive a significant advancement toward the generation of novel protein therapeutics that selectively target glycans, nucleophilic serine hydroxyl groups, and the like using proteins with uniquely reactive unnatural amino acid side chains.

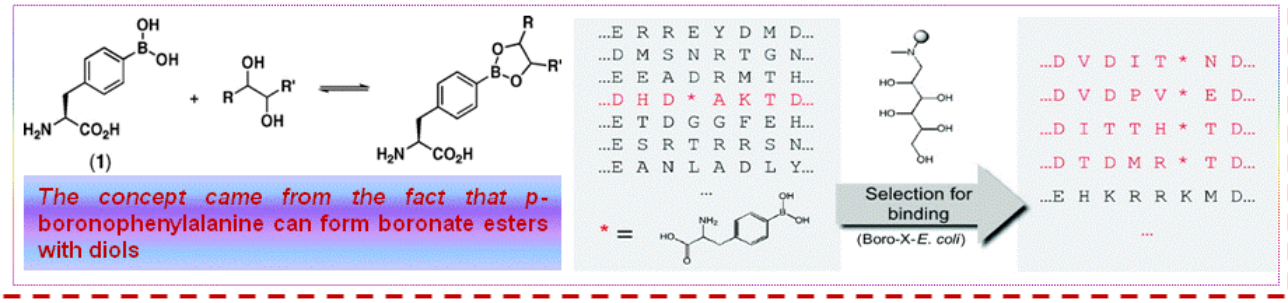


Figure 2.71: Biosynthesis of artificial amino acids and the example of protein evolution with "Chemical Warhead".

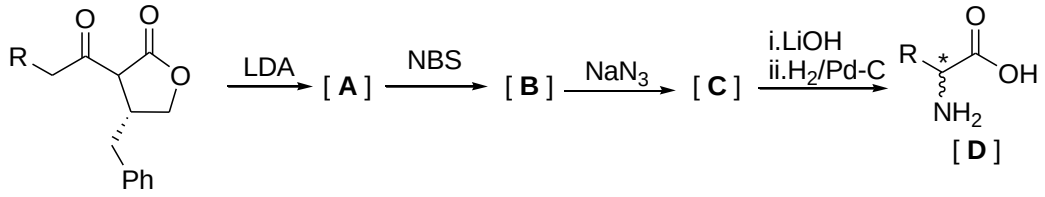
2.11. Selected References

1. Stryer, L.; Berg, J. M.; Tymoczko, J. L. In *Biochemistry*, 5th Ed. (Hardcover).
2. Lehninger Principles of Biochemistry, 5th Ed. by Nelson and Cox.
3. *Preface To: Chemical Biology. From Small Molecules to Systems Biology and Drug Design.* Edited by Stuart L. Schreiber, Tarun M. Kapoor, Gonther Weiss; Wiley-VCH Verlag GmbH & Co., **2007**, *Vol. 1-3*, XV-XVI.
4. *Forward chemical genetics* Stephen J. Haggarty and Stuart L. Schreiber, *Chemical Biology. From Small Molecules to Systems Biology and Drug Design, Edited by Stuart L. Schreiber, Tarun M. Kapoor, Gonther Weiss; Wiley-VCH Verlag GmbH & Co.*, **2007**, *Vol. 1*, 299-354.
5. Introduction to Proteimics: Tools for the New Biology, Liebler, D. C., Humana Press: **2002**.
6. Cox Jü, Mann, M. Cell **2007**, 130**,** 395–8.
7. (a) Wu, X.; Schultz, P. G. *J. Am. Chem. Soc.* **2009**, *131*, 12497. (b) Vasudev, P. G.; Chatterjee, S.; Shamala, N.; Balaram, P. *Chem. Rev.* **2011**, *111*, 657. (c) Seebach, D.; Overhand, M.; Kühnle, F. N. M.; Martinoni, B.; Oberer, L.; Hommel, U.; Widmer, H. *Helvetica Chimica Acta* **1996**, 79**,** 913. (d) Porter, E. A.; Weisblum, B.; Gellman, S. H. *J. Am. Chem. Soc.* **2002**, *124***,** 7324–30.
8. Unnatural Amino Acids: Methods and Protocols, Methods in Molecular Biology, Vol. 794 Pollegioni, Loredano; Servi, Stefano (Eds.) **2012**, XIV, 409 p. 123 illus, Humana Press, ISBN 978-1-61779-330-1.
9. (a) Ager, D. J.; Fotheringham, I. G. Curr. Opin. Drug. Discov. Devel. **2001**, *4*, 800-7. (b) Minnihan, E. C.; Yokoyama, K.; Stubbe, J. *Biology Reports* **2009**, *1*, 88.
10. Rahman, Atta-Ur.; Caldwell, Gary W.; A. Barry In Frontiers in Drug Design and Discovery Springer, Bentham Science Publishers, **2005**.
11. (a) Xie, J.; Schultz, P. G. Curr. Op.Chem.Biol. **2005**, 9**,** 548. (b) Sakamoto, K. Nucleic Acids Res. **2002**, 30**,** 4692–4699. (c) Wang, W.; Takimoto, J.; Louie, G.V.; Baiga, T. J.; Noel, J. P.; Lee, K. F.; Slesinger, P. A.; Wang, L. Nat. Neurosci **2007**, 10**,**1063–1072. (d) Wang, Q.; Parrish, A. R.; Wang, L. Chemistry & biology **2009**, 16**,** 323–36. (e) Liu, C. C.; Mack, A. V.; Brustad, E. M.; Mills, J. H.; Groff, D.; Smider, V. V.; Schultz, P. G. J Am Chem Soc. **2009**, 131**,** 9616–7. (f) Anderson, J.C.; Wu, N.; Santoro, S.W.; Lakshman, V.; King, D.S.; Schultz, P.G. Proc Natl Acad Sci USA **2004**, 101**,** 7566–7571.

2.12. Assignments

1. Write down the structure of 22nd natural amino acid, its 3 letter code and single letter code.
2. Write down any one structure, name, 3-letter code and one letter code of amino acid of each category: (a) Neutral amino acid with hydrophilic side chain, (b) Acidic Amino Acid and (c) basic amino acid.
3. Name the main three types of protein's secondary structures. Write down the general structures of β^3- and β^2-amino acids.
4. What do you mean by genetic alphabet, genetic codon and genetic code?
5. Define the followings: (a) Chemical Biology, (b) Biochemistry, (c) Bio-organic Chemistry, (d) Proteomics and (e) Glycobiology
6. Answer the followings:
   - **6.1.** If three different amino acids (gly, ala, leu) are used to make a tripeptide, how many different sequences are possible?
     (a) 6
     (b) 9
     (c) 3
     (d) 12
   - **6.2.** The primary structures of methionine-enkephaline is: tyr - gly - gly - phe – met. The N-terminal and the C-terminal amino acids respectively are:-
     (a) met- and gly-
     (b) Tyr- and gly -
     (c) tyr – and phe –
     (d) tyr - and met-
   - **6.3.** Normal hemoglobin is apparently more polar and soluble in water because of the presence of polar amino acid glutamic acid but Sickle Cell hemoglobin is more non-polar and insoluble. The non-polar amino acid present in Sickle Cell hemoglobin is-
     (a) Isoleucine
     (b) Leucine
     (c) Valine
     (d) Alanine
   - **6.4.** The order of increasing polarity of the amino acids-Ser; Glu; Asp; Lys; Ala; Gln is
     (a) Ser > Glu > Asp > Lys > Ala > Gln
     (b) Ser > Glu > Asp > Lys > Gln > Ala
     (c) Glu > Ser > Asp > Lys > Gln > Ala
     (d) Ala < Lys < Ser < Glu < Asp < Gln
7. Explain the differences between primary, secondary, tertiary, and quaternary protein structures by giving brief definitions of each. What types of bonding are used in each?
8. Explain the difference between the alpha helix and the beta pleated sheet protein structures. What are the differences in the hydrogen bonding?

9. Write down the structure of A, B, C and the absolute configuration of the produced chiral aminoacid in the following scheme.



10. Why the solid phase peptide synthesis proceeds from C to N-terminus and not N to C terminus synthesis?

11. What are the main applications of unnatural amino acids?