Polypeptides and Proteins

- Polypeptide refers to the structure of a single chain.
  - Every polypeptide has one free amino group (called the "N-terminus") and one free carboxyl group (called the "C-terminus").
  - Peptides are made up of about 50 residues, and do not possess a well-defined 3D-structure
- Protein refers to the overall functional assembly, created when one or more polypeptides fold up and become functional units.
  - Some proteins consist of only a single folded polypeptide chain, but many proteins contain multiple polypeptides, and frequently inorganic atoms as well, such as Zinc,Iron, Magnesium, etc
  - Proteins are larger molecules that usually contain at least 50 residues, and sometimes 1000. The most important feature of proteins is that they possess well-defined 3D-structure.

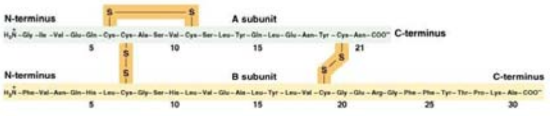
Protein Structural Descriptions

- Primary Structure
  - **Sequence of amino acids**
- Secondary Structure
  - **Local three-dimensional patterns**
  - **Frequently occurring substructures or folds**
- Tertiary Structure
  - **Three-dimensional shape of an amino acid chain**
- Quaternary Structure
  - **Three-dimensional shape of multiple amino acid chains**
  - **Overall organization of non-covalently linked subunits of a functional protein.**

Primary Structure

- **Specification of the sequence of amino acids.**
  - **Since every polypeptide begins with free amino group, this is called the N-terminus.**
  - **The opposite end of the polypeptide has a free carboxyl group, called the C-terminus.**

Primary Structure



Primary Structure Determination

Step 1: Amino Acid Analysis

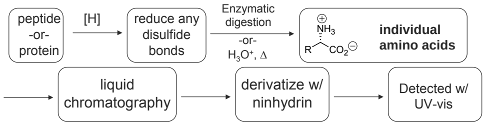
- Protein is hydrolyzed in strong acid or base at high temperatures
- Mixture of amino acids generated by protein hydrolysis is analyzed by chromatography

Step 2: Sequence Analysis

- Protein cleaved into polypeptides
- Polypeptides subjected to amino acid analysis or **Edman degradation**

Amino Acid Analysis

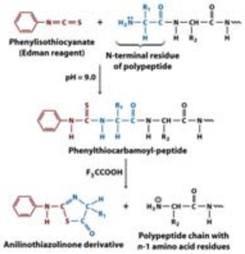
- Automated method to determine the amino acid content of a peptide or protein
- Reaction of primary amines with ninhydrin



1972 Nobel Prize in Chemistry William Stein Stanford Moore

Edman Degradation

- The procedure begins with a reaction between Phenyl isothiocyanate (PITC) and the N-terminal residue under basic conditions
- A PTC-peptide is produced, which is then selectively cleaved by trifluoroacetic acid
- An unstable anilinothiazolinone is produced, which is converted to a stable phenylthiohydantoin, which can be identified by chromatography
- A polypeptide with N-1 residues remains, for the next round of degradation

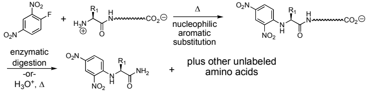


Primary Structure Sequence Analysis

- **A Specific Chemical Cleavage:**
  - Cyanogen bromide (Br-CN): cleaves peptide bonds formed by carboxyl of methionine
- **Specific Enzymatic Cleavages:**
  - Trypsin: cleaves peptide bonds formed by carboxyl of lysine and arginine
  - Chymotrypsin: cleaves peptide bonds formed by carboxyl of phenylalanine, tyrosine, and tryptophan
- **Sequential Degradation:**
  - Edman Degradation: cleaves n-terminal amino acids

End Group Analysis

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



Secondary Structure

- **Polypeptides fold in a series of stages. The first level of folding is called the secondary (2º) structure.**
- **One of the most common 2º folding patterns is called the alpha-helix , discovered by Pauling and Corey.**
  - **Alpha helix: Hydrogen bonds can form readily between C=O groups in the backbone and N-H groups four amino acid residues further along the chain.**
  - **This regular pairing pulls the polypeptide into a helical shape that resembles a coiled ribbon.**

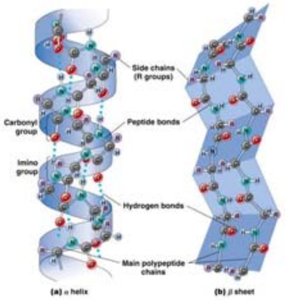
Secondary Structure

- **Another common folding pattern is called beta pleated sheet .**
- **Some protein regions remain in random coil, no regular pattern of secondary structure.**
- **Different proteins have different degrees of alpha helix, beta sheet, and random coil .**
  - **Silk is a protein stabilized entirely by pleated sheet;**
  - **keratin (in hair) is stabilized entirely by alpha helix.**
  - **Most proteins have some of both.**

Secondary Structure

- Common local conformations:
  - **Alpha Helix**
  - **Beta-Pleated Sheet**
  - **Stabilized by:**
    - **Hydrogen bonding**
    - **R-groups provided space**
    - **s-trans amide bonds**

Secondary Structure



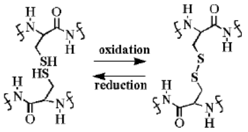
Tertiary Structure

- Polypeptides continue folding beyond the formation of secondary structure. It is only with the complete, compact folding into tertiary (3^o) structure that they attain their "native conformation" and become active proteins (as a result of the creation of active sites).
- Forces that contribute to tertiary folding include:
  - hydrogen bonds
  - hydrophobic bonds
  - ionic bonds
  - sulfhydryl bonds (-S-S- bonds). These are especially important, because they are covalent bonds and quite strong compared to H-bonds.

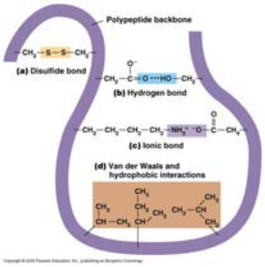
Tertiary Structure

Stabilized by:

- **Clustering of hydrophobic side chains**
- **Salt bridges - oppositely charged side chains clustering**
- **Disulfide bonds**



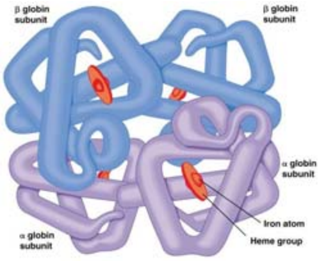
Tertiary Structure



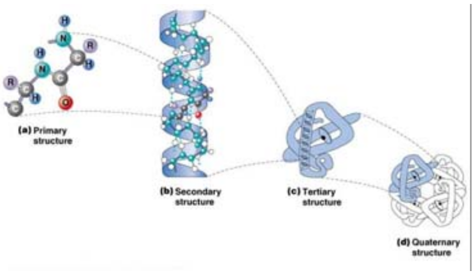
Quaternary Structure

- Some proteins are made of multiple polypeptide subunits, which must be assembled together after each individual polypeptide has reached its 3^o structure. Examples:
  - **Hemoglobin** (blood protein involved in oxygen transport) has **four subunits** .
  - **Pyruvate dehydrogenase** (mitochondrial protein involved in energy metabolism) has **72 subunits** .

Quaternary Structure

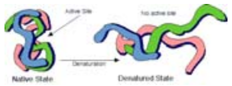


Overview of Structure



Denaturation & Renaturation

- When proteins are heated, or exposed to acids or bases, or high salt concentrations, the variety of weak bonds holding tertiary and quaternary structure together can be disrupted so that the protein unfolds. Unfolding = denaturation resulting in loss of function.
  - Unfolding can proceed even to disrupt secondary structure.
- Denaturation is sometimes reversible ; an unfolded protein can be restored to correct folding and regain biological activity. This is called renaturation .
- Denaturation can also occur irreversibly (as when egg white protein, albumin,is denatured by boiling to congeal as egg white). Renaturation is then no longer possible.



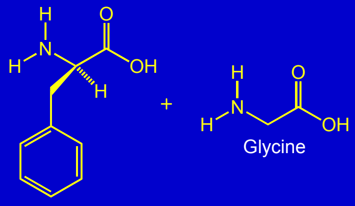
Peptide Synthesis

- Making peptide bonds between amino acids is not difficult.
- The challenge is connecting amino acids in the correct sequence.
- Random peptide bond formation in a mixture of phenylalanine and glycine, for example, will give four dipeptides.

Phe—Phe Gly—Gly Phe—Gly Gly—Phe

Peptide Synthesis

- Amino Acids are Structurally Bifunctional

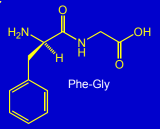


L-Phenylalanine

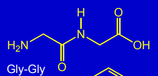
Glycine

Peptide Synthesis

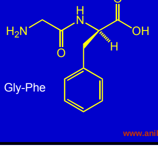
- Possible Products from the Condensation of Phenylalanine and Glycine



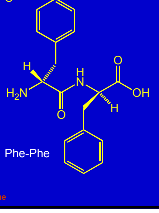
Phe-Gly



Gly-Gly



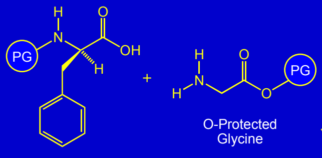
Gly-Phe



Phe-Phe

Step 1 of Peptide Synthesis

- **Protection**
  - Limit the number of possible reactions by "protecting" the nitrogen of one amino acid and the carboxyl group of the other

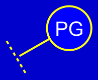


N-Protected L-Phenylalanine

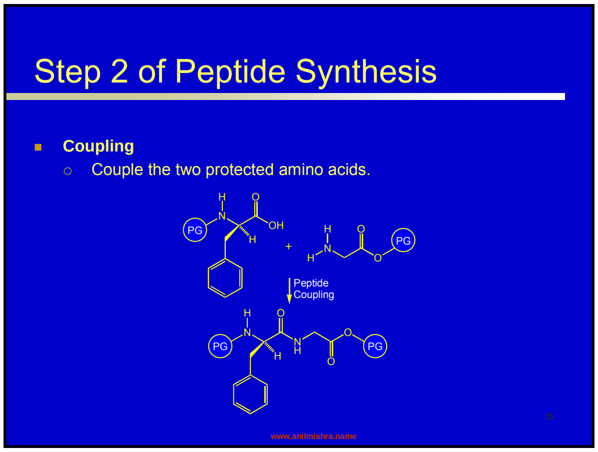
C-Electrophile

O-Protected Glycine

N-Nucleophile

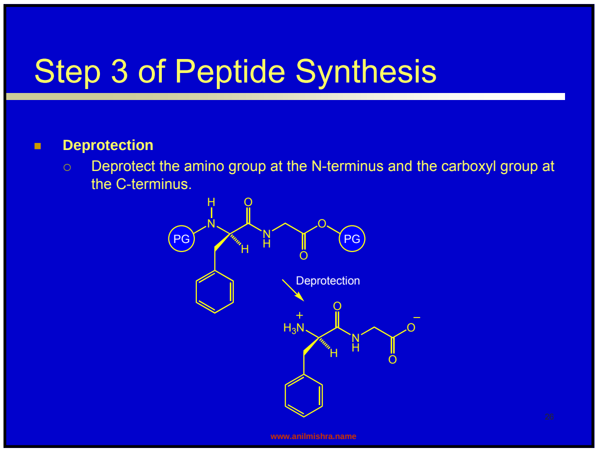


protecting group



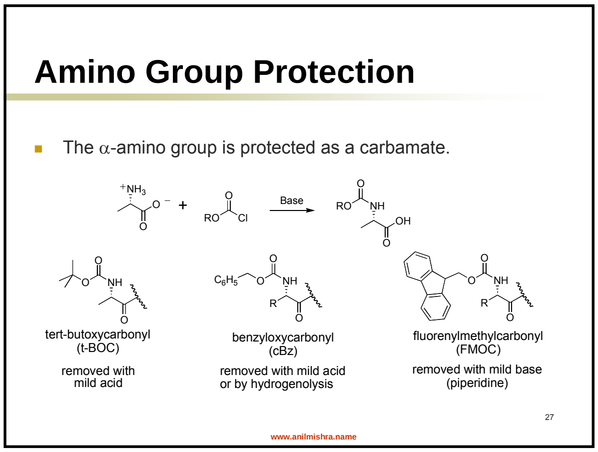
Step 2 of Peptide Synthesis

- **Coupling**
  - Couple the two protected amino acids.



Step 3 of Peptide Synthesis

- **Deprotection**
  - Deprotect the amino group at the N-terminus and the carboxyl group at the C-terminus.



Amino Group Protection

- The α-amino group is protected as a carbamate.

tert-butoxycarbonyl
(t-BOC)

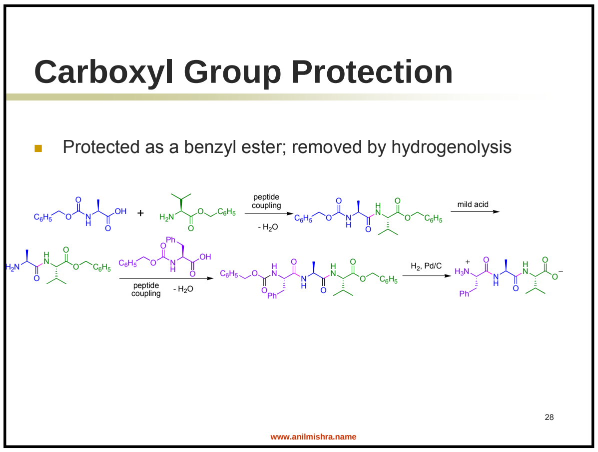
removed with
mild acid

benzyloxycarbonyl
(cBz)

removed with mild acid
or by hydrogenolysis

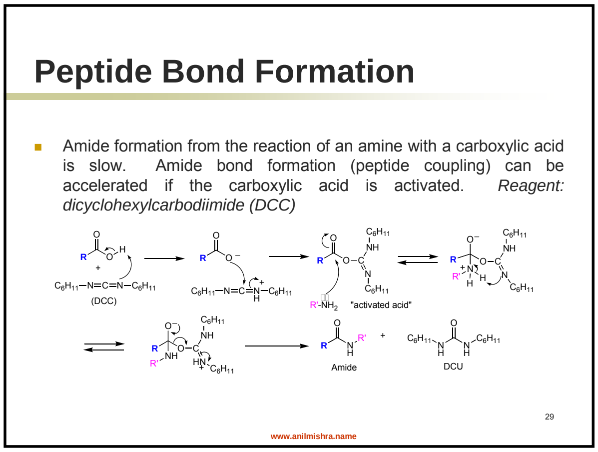
fluorenylmethylcarbonyl
(FMOC)

removed with mild base
(piperidine)



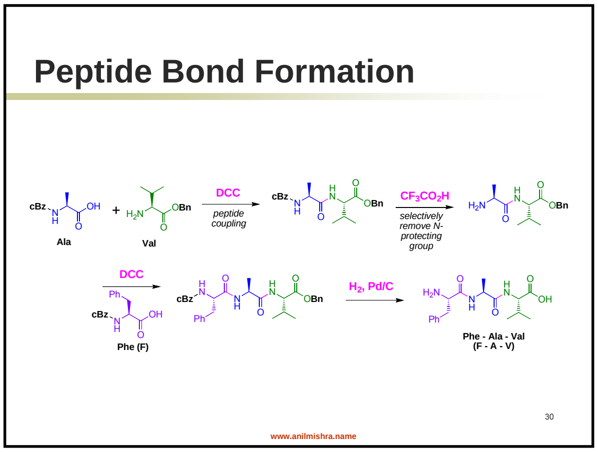
Carboxyl Group Protection

- Protected as a benzyl ester; removed by hydrogenolysis



Peptide Bond Formation

- Amide formation from the reaction of an amine with a carboxylic acid is slow. Amide bond formation (peptide coupling) can be accelerated if the carboxylic acid is activated. *Reagent: dicyclohexylcarbodiimide (DCC)*



Peptide Bond Formation

Peptide Synthesis

- Chemical synthesis of peptide:
  - Solution phase synthesis
  - Solid-phase synthesis

Peptide Synthesis

- Solution phase chemistry
  - Time consuming:
    - isolation and purification at each step
  - Low yield: can't drive reaction to complete
  - Use excess reagent to improve yield

Peptide Synthesis

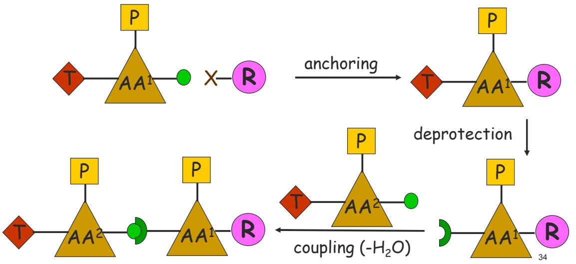
- Solid Phase Peptide Synthesis (SPPS)
  - In solid-phase synthesis, the starting material is bonded to an inert solid support.
  - Reactants are added in solution.
  - Reaction occurs at the interface between the solid and the solution. Because the starting material is bonded to the solid, any product from the starting material remains bonded as well.
  - Purification involves simply washing the byproducts from the solid support.

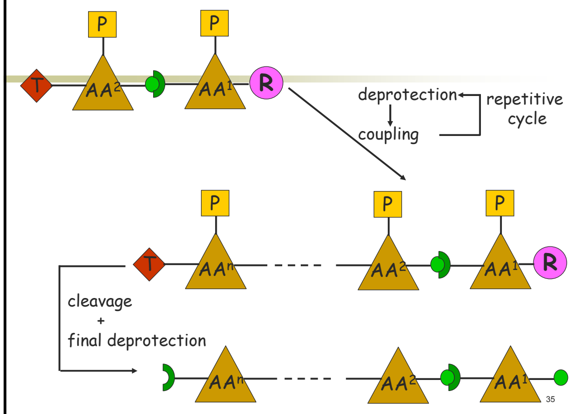
SOLID PHASE PEPTIDE SYNTHESIS

Bruce Merrifield published in 1963

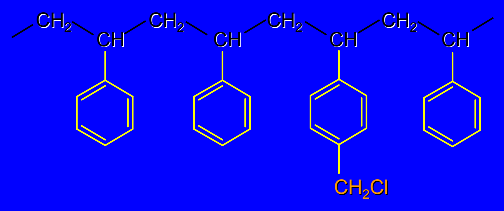
Nobel Prize in Chemistry in 1984

The idea:

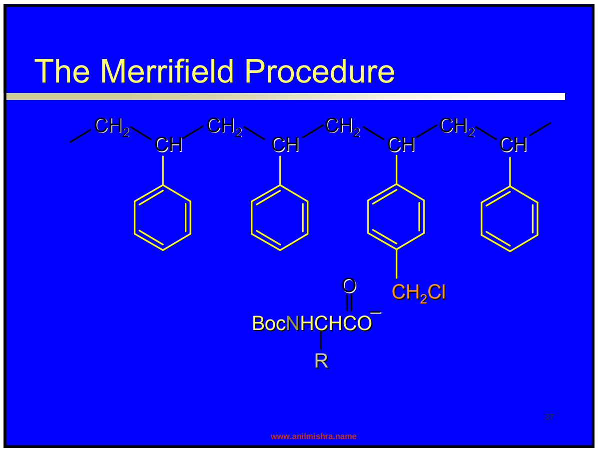




The Solid Support

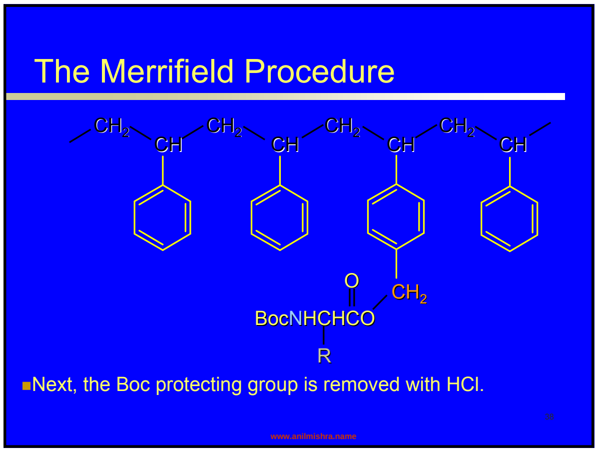


- The chloromethylated resin is treated with the Boc-protected C-terminal amino acid. Nucleophilic substitution occurs, and the Boc-protected amino acid is bound to the resin as an ester.



The Merrifield Procedure

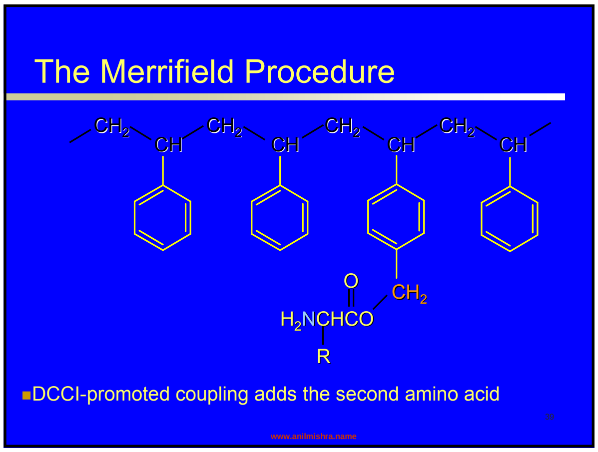
www.anilmishra.name



The Merrifield Procedure

- Next, the Boc protecting group is removed with HCl.

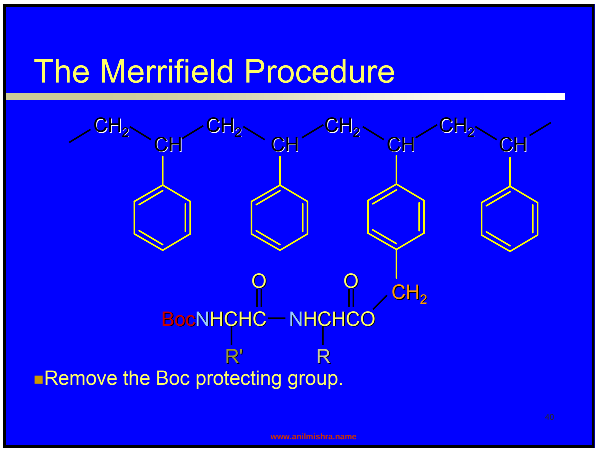
www.anilmishra.name



The Merrifield Procedure

- DCCI-promoted coupling adds the second amino acid

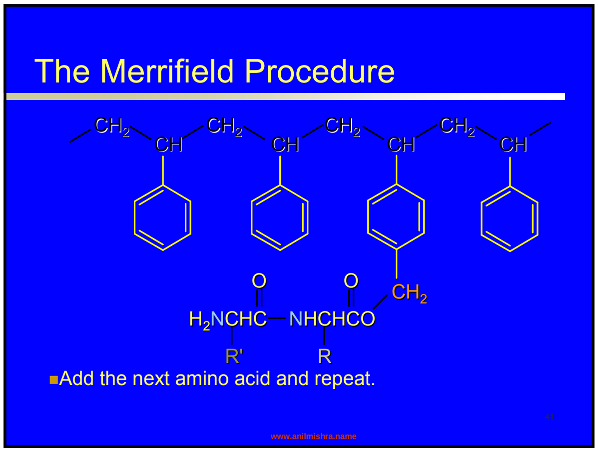
www.anilmishra.name



The Merrifield Procedure

- Remove the Boc protecting group.

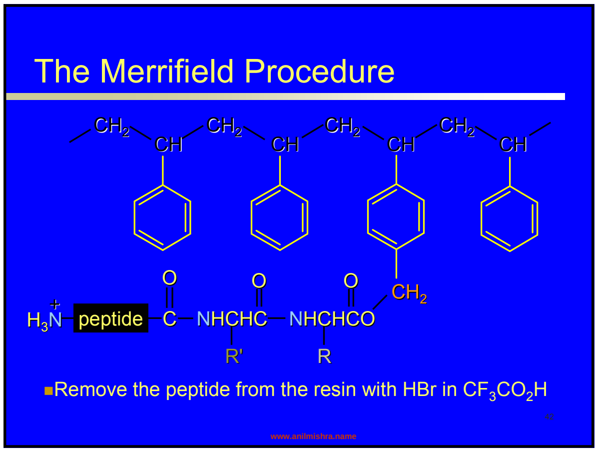
www.anilmishra.name



The Merrifield Procedure

- Add the next amino acid and repeat.

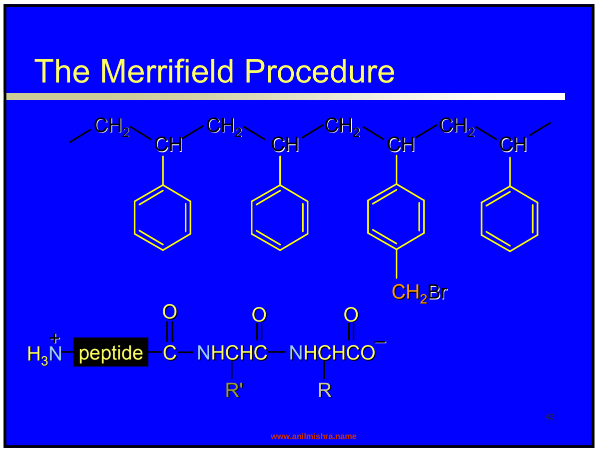
www.anilmishra.name



The Merrifield Procedure

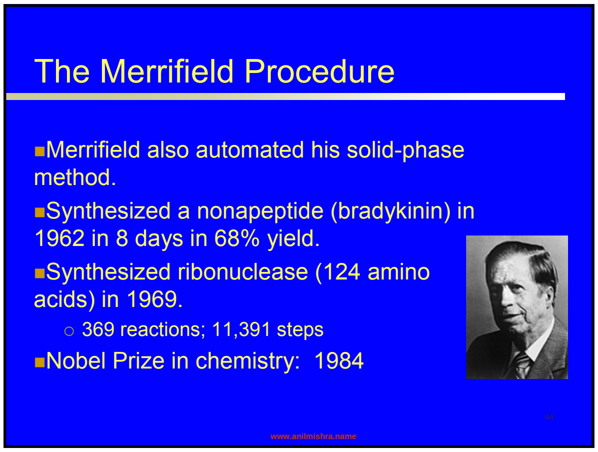
- Remove the peptide from the resin with HBr in CF_3CO_2H

www.anilmishra.name



The Merrifield Procedure

$$H_3\overset{+}{N}-\text{peptide}-\overset{O}{\overset{\|}{C}}-NH\underset{R'}{CH}\overset{O}{\overset{\|}{C}}-NH\underset{R}{CH}\overset{O}{\overset{\|}{C}}O^-$$



The Merrifield Procedure

- Merrifield also automated his solid-phase method.
- Synthesized a nonapeptide (bradykinin) in 1962 in 8 days in 68% yield.
- Synthesized ribonuclease (124 amino acids) in 1969.
  - 369 reactions; 11,391 steps
- Nobel Prize in chemistry: 1984