27.8
Introduction to Peptide Structure Determination

Primary Structure

The primary structure is the amino acid sequence plus any disulfide links.

Classical Strategy (Sanger)

1. Determine what amino acids are present and their molar ratios.
2. Cleave the peptide into smaller fragments, and determine the amino acid composition of these smaller fragments.
3. Identify the N-terminus and C-terminus in the parent peptide and in each fragment.
4. Organize the information so that the sequences of small fragments can be overlapped to reveal the full sequence.

27.9
Amino Acid Analysis

Amino Acid Analysis

Acid-hydrolysis of the peptide (6 M HCl, 24 hr) gives a mixture of amino acids.

The mixture is separated by ion-exchange chromatography, which depends on the differences in p*I* among the various amino acids.

Amino acids are detected using ninhydrin.

Automated method; requires only 10^{-5} to 10^{-7} g of peptide.

27.10
Partial Hydrolysis of Proteins

Partial Hydrolysis of Peptides and Proteins

Acid-hydrolysis of the peptide cleaves all of the peptide bonds.

Cleaving some, but not all, of the peptide bonds gives smaller fragments.

These smaller fragments are then separated and the amino acids present in each fragment determined.

Enzyme-catalyzed cleavage is the preferred method for partial hydrolysis.

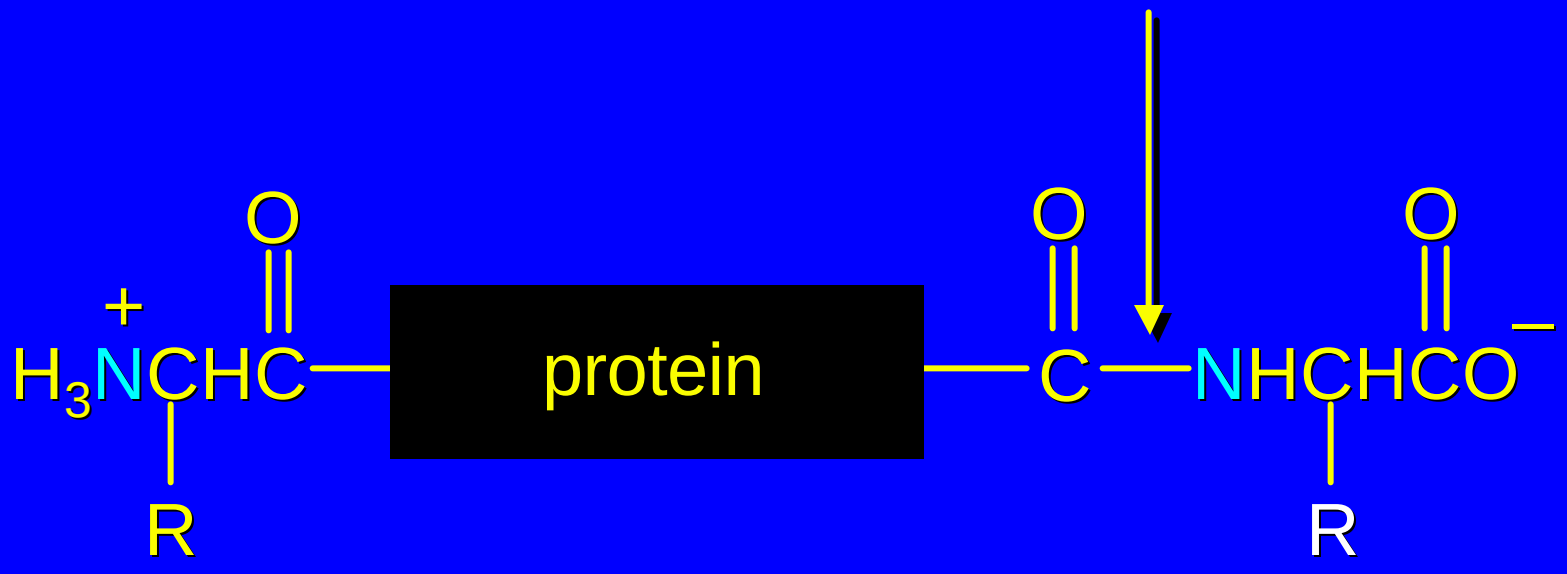
Carboxypeptidase

Carboxypeptidase is a *proteolytic enzyme* (catalyzes the hydrolysis of proteins).

Carboxypeptidase

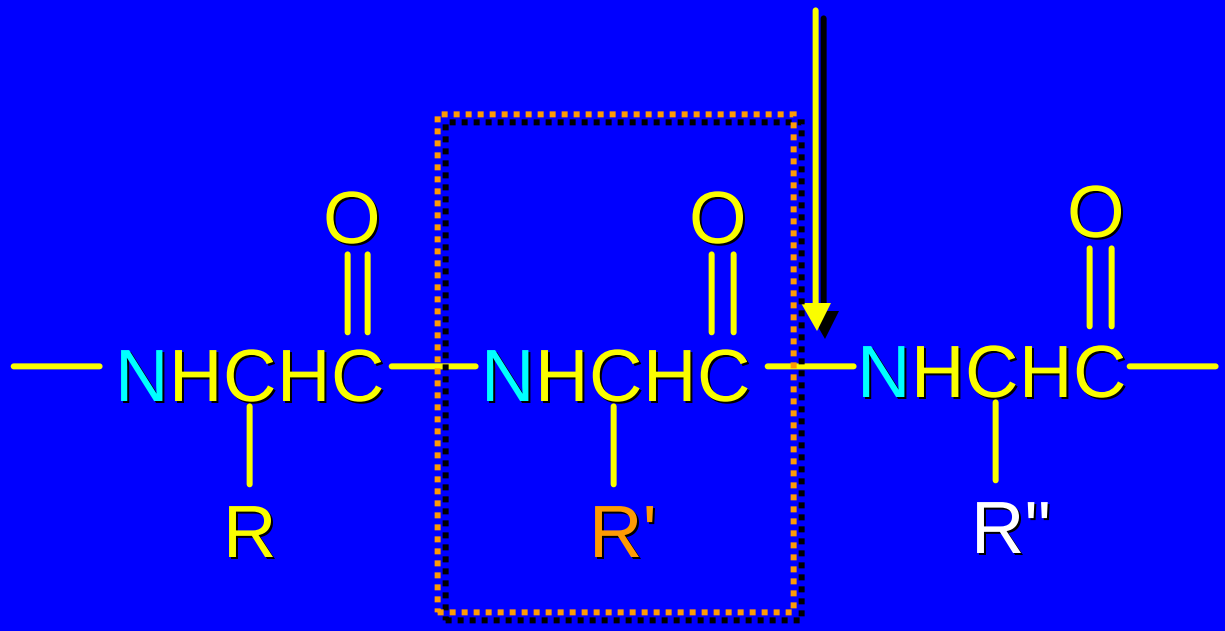
Carboxypeptidase is a *proteolytic enzyme* (catalyzes the hydrolysis of proteins).

Carboxypeptidase is selective for cleaving the peptide bond to the C-terminal amino acid.



Trypsin

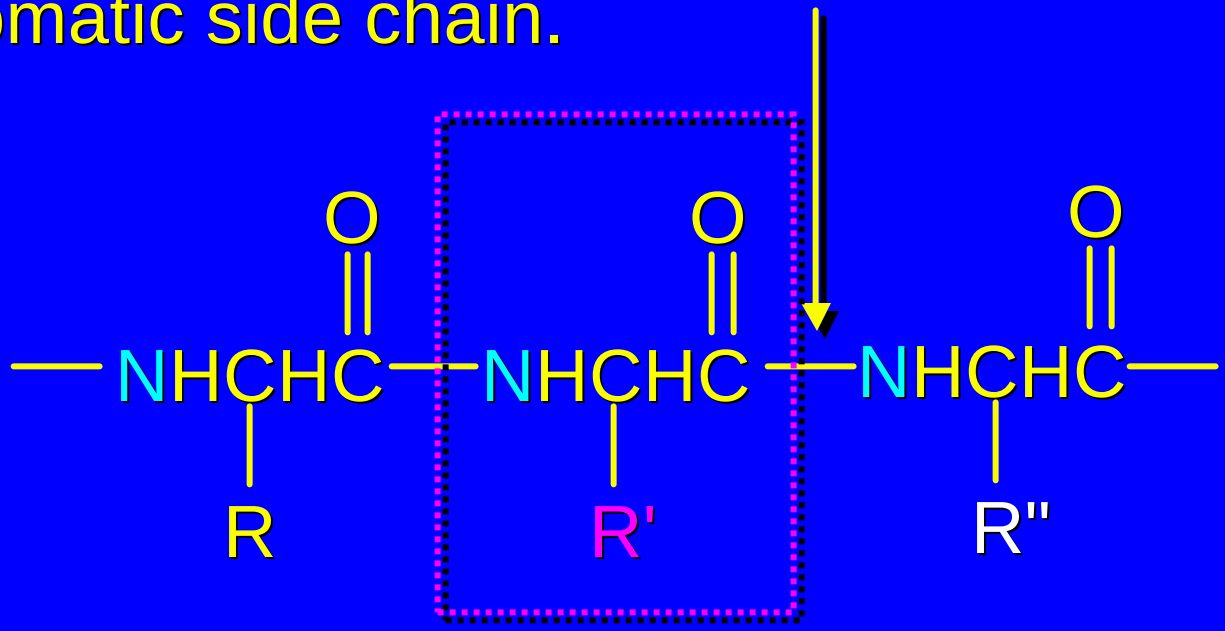
Trypsin is selective for cleaving the peptide bond to the carboxyl group of lysine or arginine.



lysine or arginine

Chymotrypsin

Chymotrypsin is selective for cleaving the peptide bond to the carboxyl group of amino acids with an aromatic side chain.



phenylalanine, tyrosine, tryptophan

27.11
End Group Analysis

End Group Analysis

Amino sequence is ambiguous unless we know whether to read it left-to-right or right-to-left.

We need to know what the N-terminal and C-terminal amino acids are.

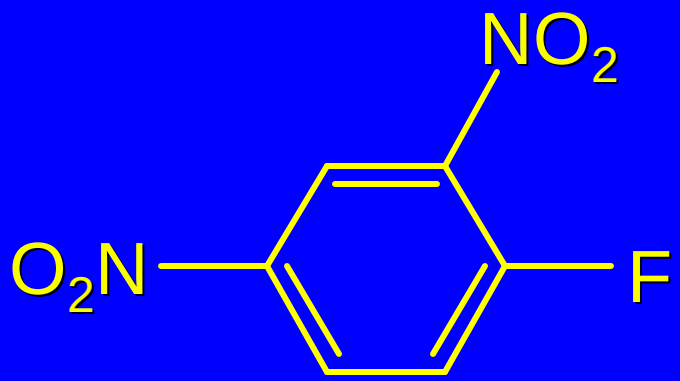
The C-terminal amino acid can be determined by carboxypeptidase-catalyzed hydrolysis.

Several chemical methods have been developed for identifying the N-terminus. They depend on the fact that the amino N at the terminus is more nucleophilic than any of the amide nitrogens.

Sanger's Method

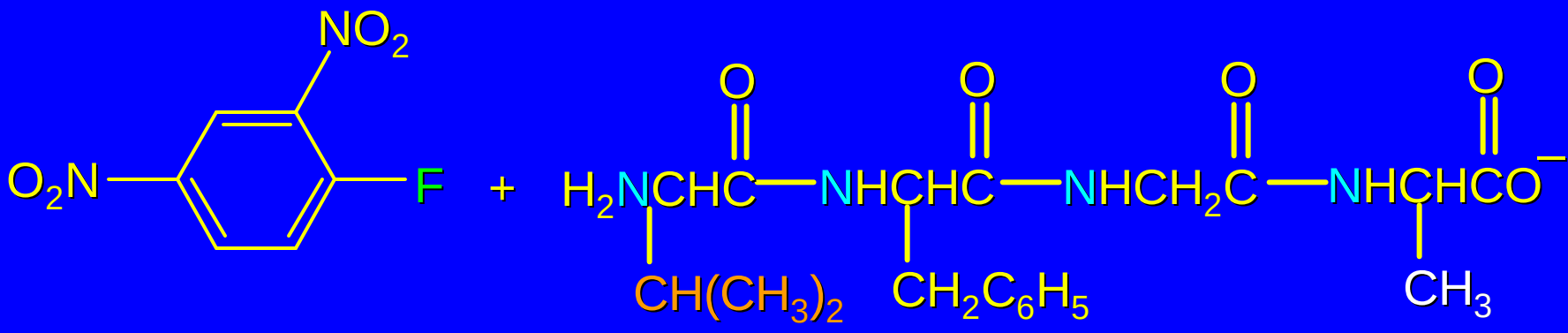
The key reagent in Sanger's method for identifying the N-terminus is 1-fluoro-2,4-dinitrobenzene.

1-Fluoro-2,4-dinitrobenzene is very reactive toward nucleophilic aromatic substitution (Section 23.5).



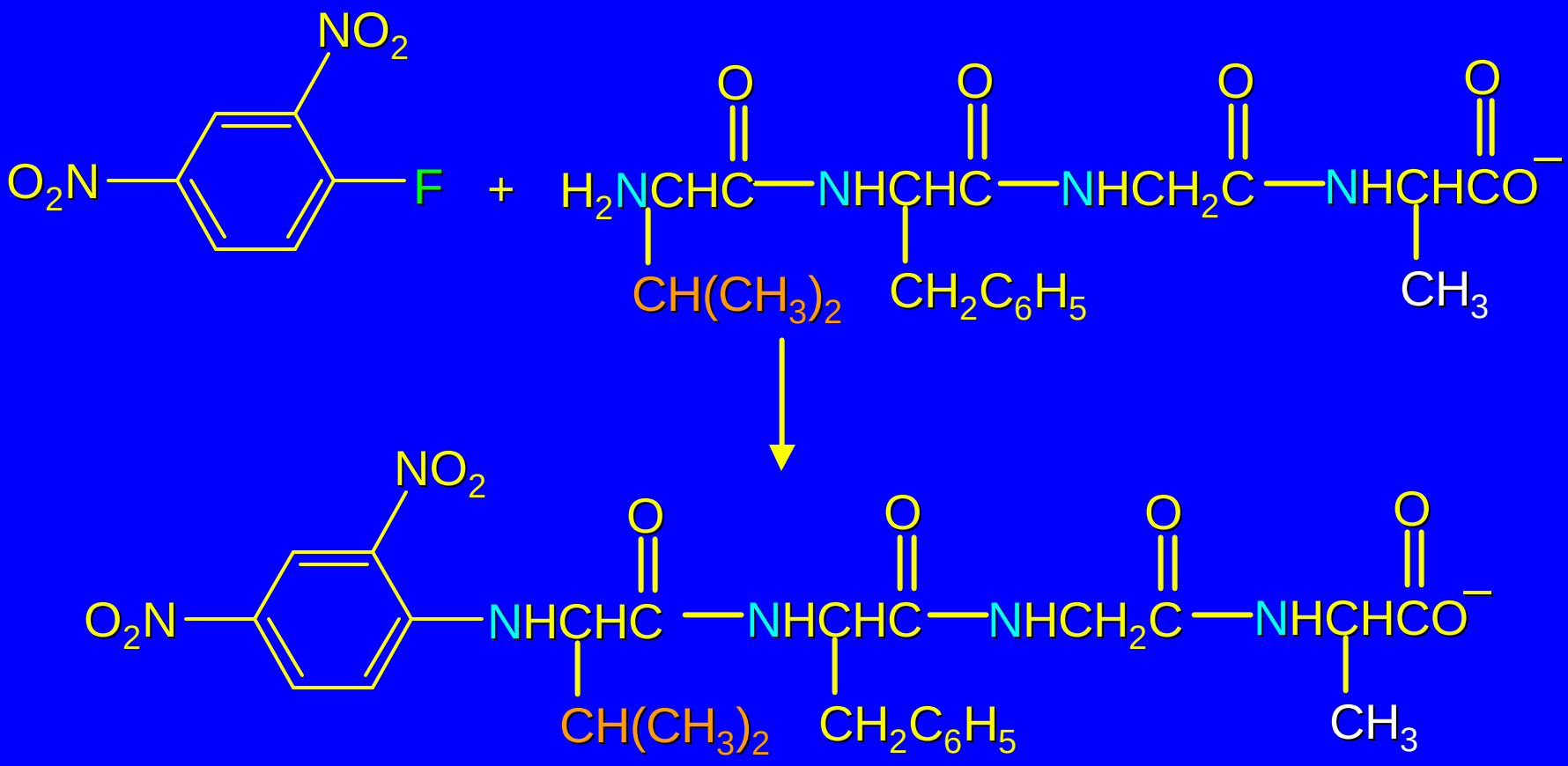
Sanger's Method

1-Fluoro-2,4-dinitrobenzene reacts with the amino nitrogen of the N-terminal amino acid.



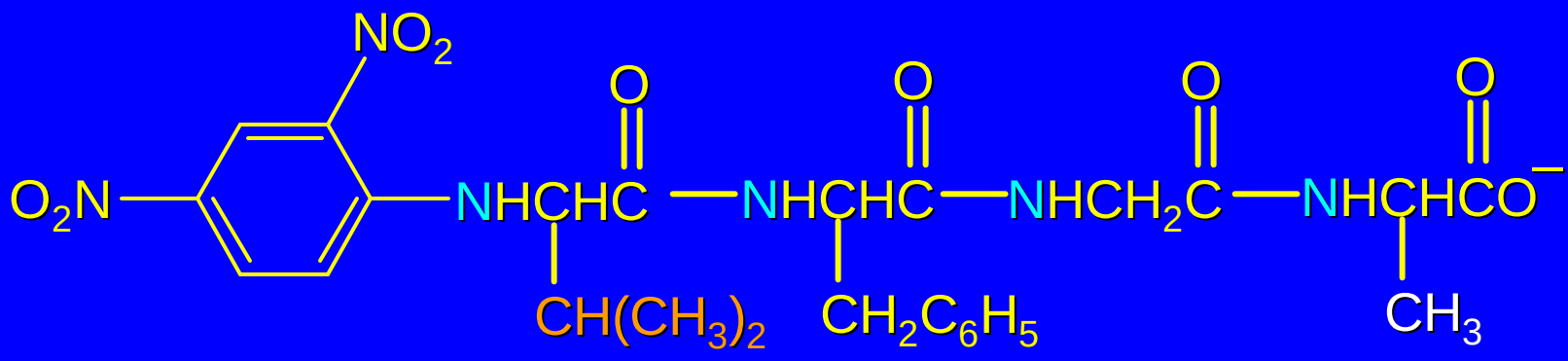
Sanger's Method

1-Fluoro-2,4-dinitrobenzene reacts with the amino nitrogen of the N-terminal amino acid.



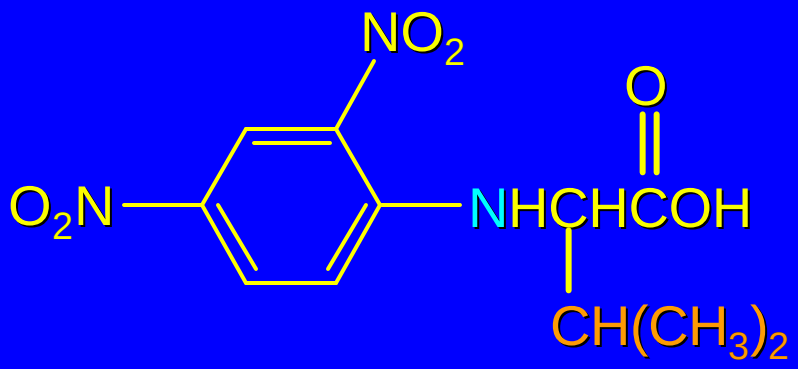
Sanger's Method

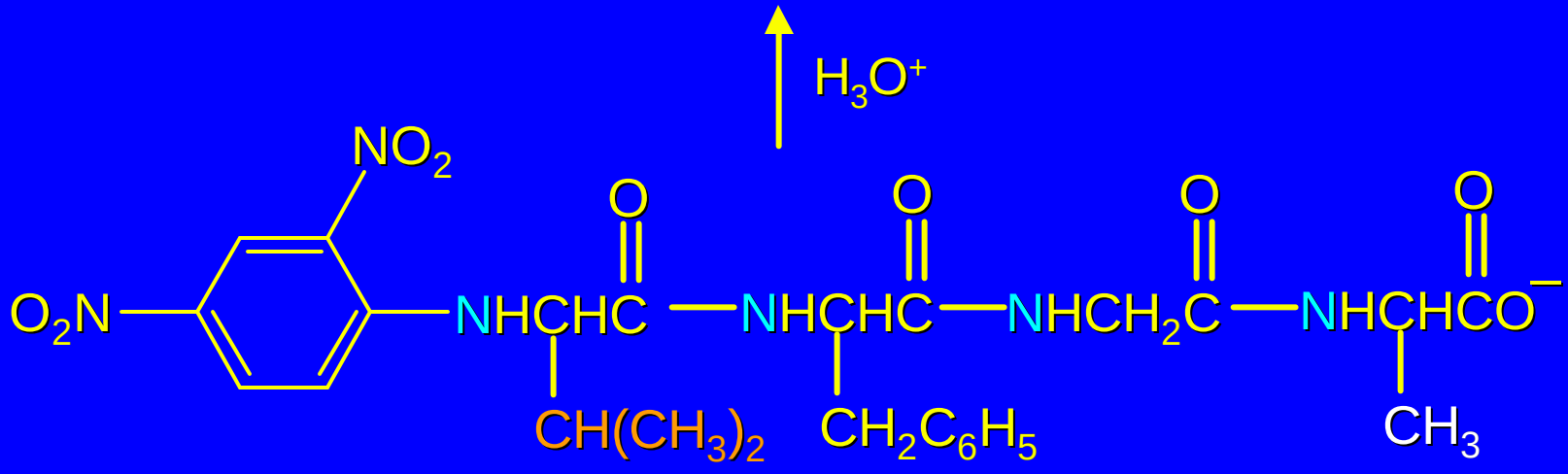
Acid hydrolysis cleaves all of the peptide bonds leaving a mixture of amino acids, only one of which (the N-terminus) bears a 2,4-DNP group.



Sanger's Method

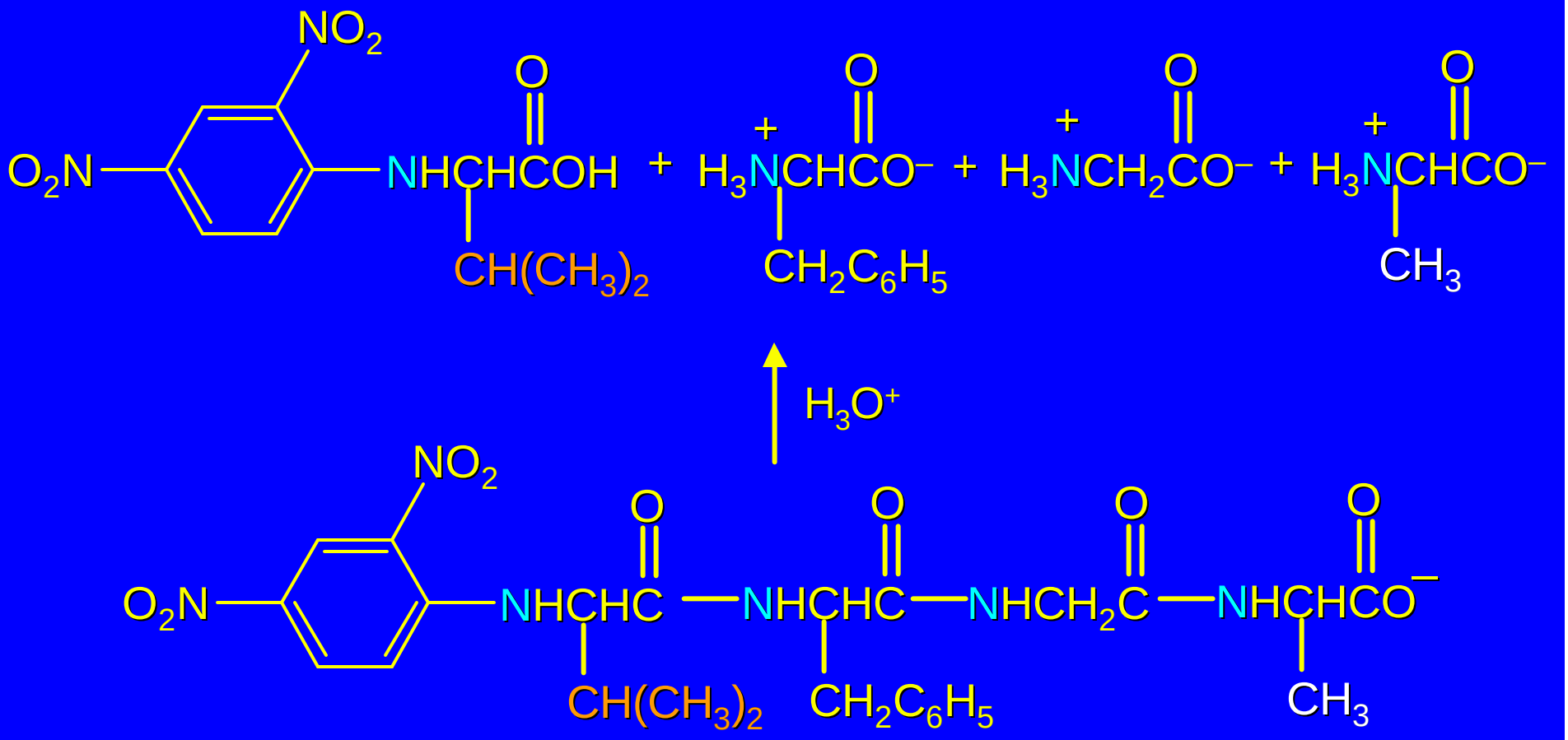
Acid hydrolysis cleaves all of the peptide bonds leaving a mixture of amino acids, only one of which (the N-terminus) bears a 2,4-DNP group.





Sanger's Method

Acid hydrolysis cleaves all of the peptide bonds leaving a mixture of amino acids, only one of which (the N-terminus) bears a 2,4-DNP group.



27.12
Insulin

Insulin

Insulin is a polypeptide with 51 amino acids.

It has two chains, called the A chain (21 amino acids) and the B chain (30 amino acids).

The following describes how the amino acid sequence of the B chain was determined.

The B Chain of Bovine Insulin

Phenylalanine (F) is the N terminus.

Pepsin-catalyzed hydrolysis gave the four peptides:

FVNQHLCGSHL

VGAL

VCGERGF

YTPKA

The B Chain of Bovine Insulin

FVNQHLCGSHL

VGAL

VCGERGF

YTPKA

The B Chain of Bovine Insulin

Phenylalanine (F) is the N terminus.

Pepsin-catalyzed hydrolysis gave the four peptides:

FVNQHLCGSHL
VGAL
VCGERGF
YTPKA

Overlaps between the above peptide sequences were found in four additional peptides:

SHLV
LVGA
ALT
TLVC

The B Chain of Bovine Insulin

FVNQHLCGSHL

SHLV

LVGA

VGAL

ALY

YLVC

VCGERGF

YTPKA

The B Chain of Bovine Insulin

Phenylalanine (F) is the N terminus.

Pepsin-catalyzed hydrolysis gave the four peptides:

FVNQHLCGSHL
VGAL
VCGERGF
YTPKA

Overlaps between the above peptide sequences were found in four additional peptides:

SHLV
LVGA
ALT
TLVC

Trypsin-catalyzed hydrolysis gave GFFYTPK which completes the sequence.

The B Chain of Bovine Insulin

FVNQHLCGSHL

SHLV

LVGA

VGAL

ALY

YLVC

VCGERGF

GFFYTPK

YTPKA

The B Chain of Bovine Insulin

FVNQHLCGSHL

SHLV

LVGA

VGAL

ALY

YLVC

VCGERGF

GFFYTPK

YTPKA

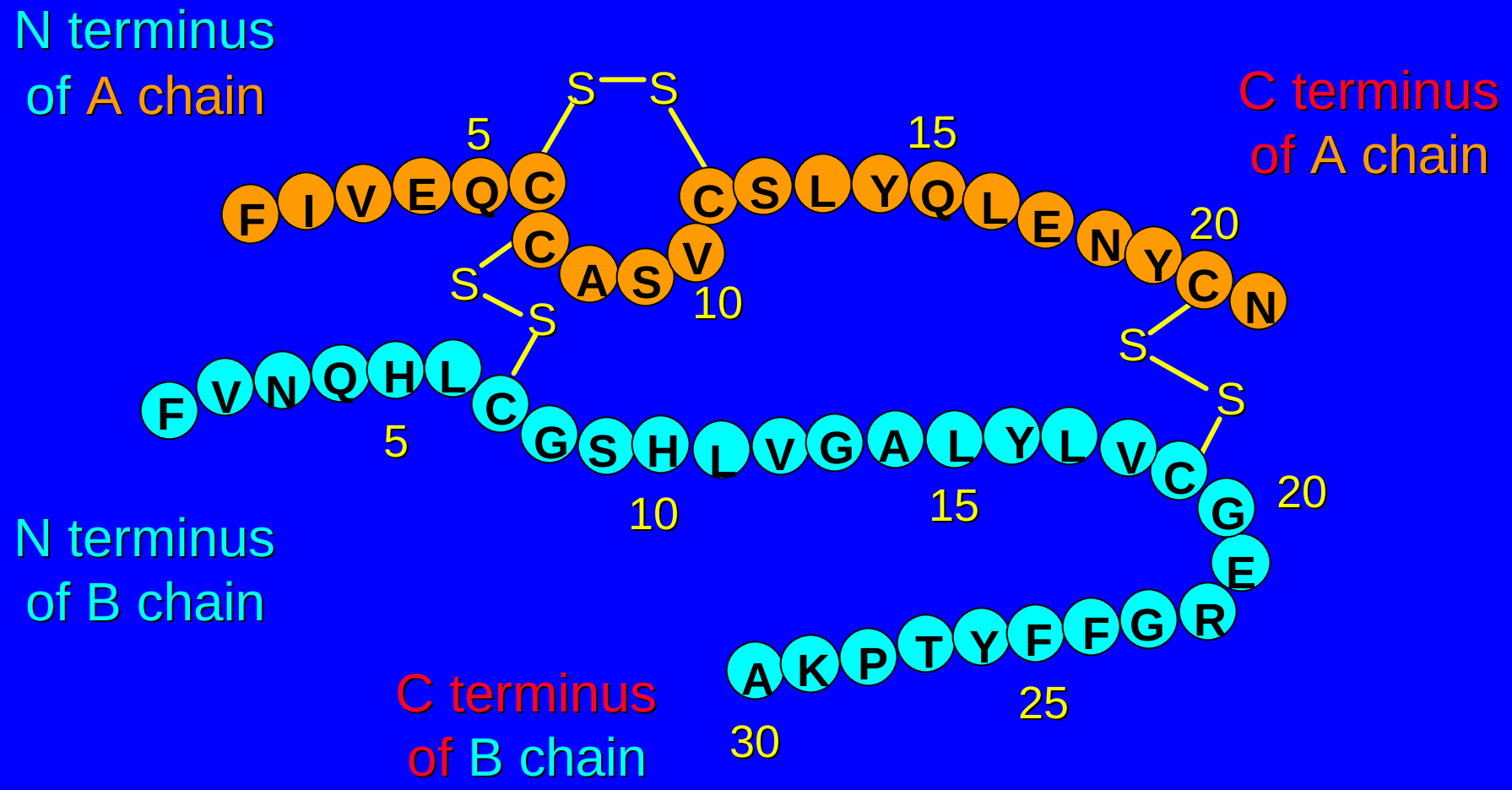
FVNQHLCGSHLVGALYLVCGERGFFYTPKA

Insulin

The sequence of the A chain was determined using the same strategy.

Establishing the disulfide links between cysteine residues completed the primary structure.

Primary Structure of Bovine Insulin



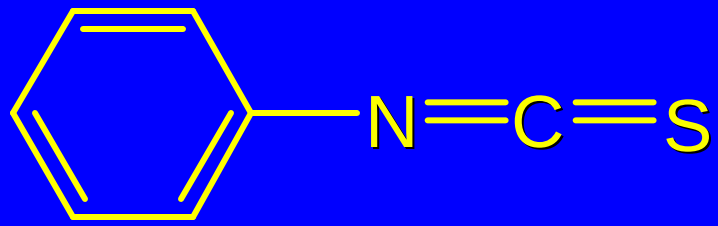
27.13
The Edman Degradation and Automated Sequencing of Peptides

Edman Degradation

1. Method for determining N-terminal amino acid.
2. Can be done sequentially one residue at a time on the same sample. Usually one can determine the first 20 or so amino acids from the N-terminus by this method.
3. 10^{-10} g of sample is sufficient.
4. Has been automated.

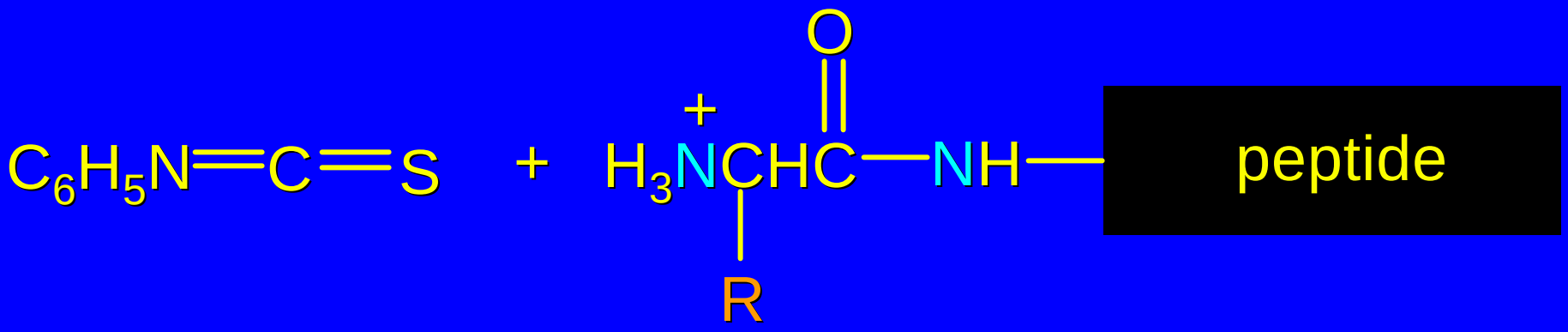
Edman Degradation

The key reagent in the Edman degradation is phenyl isothiocyanate.

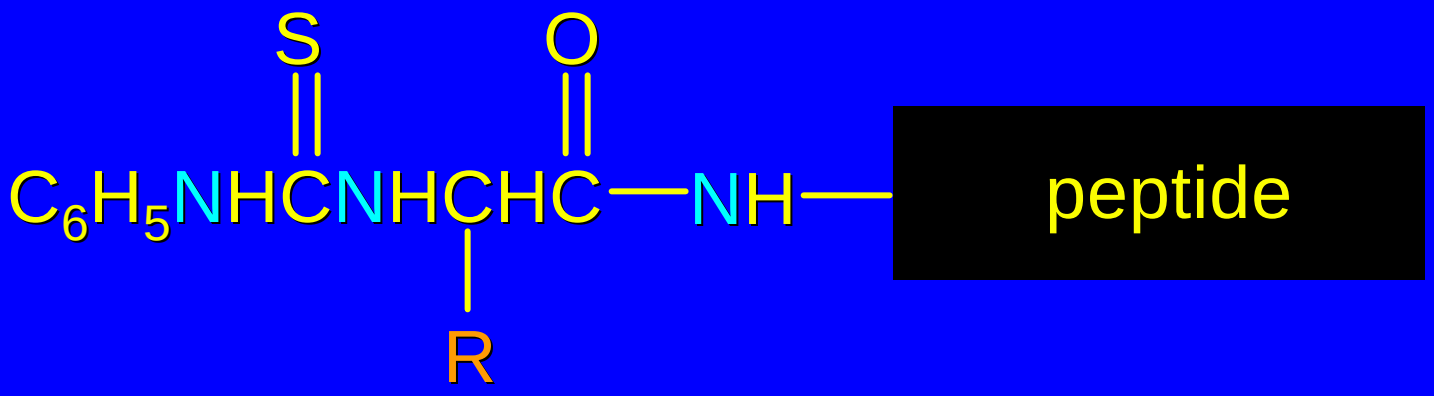


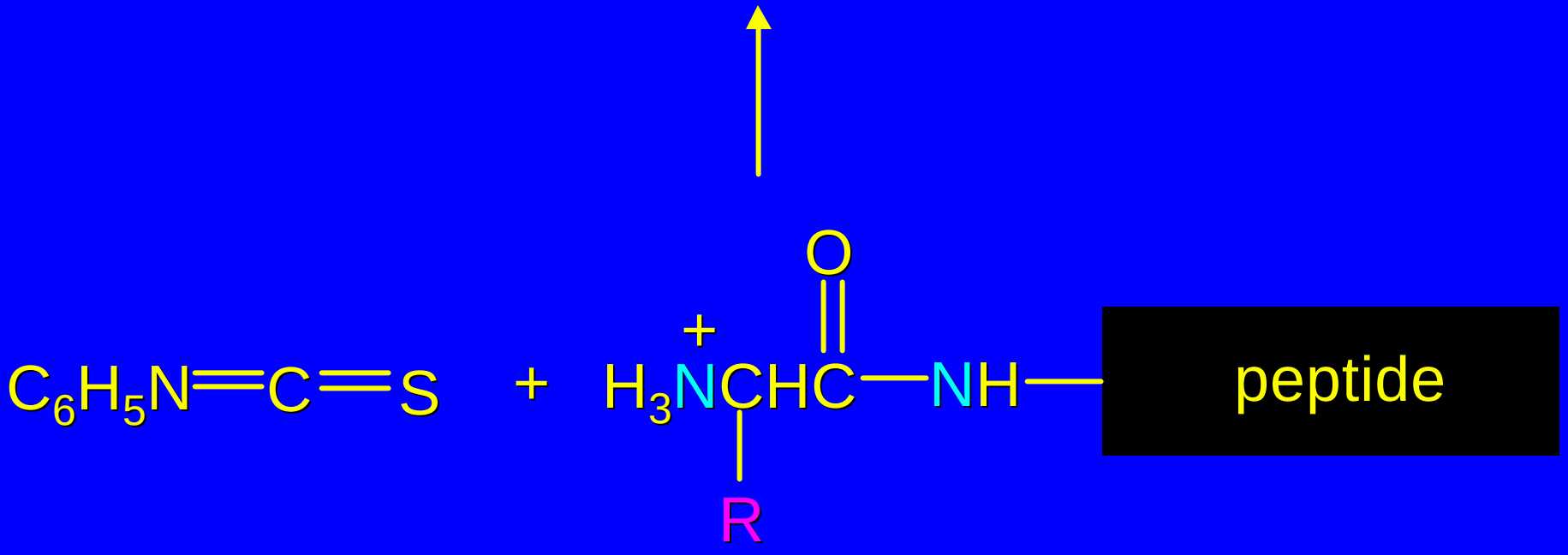
Edman Degradation

Phenyl isothiocyanate reacts with the amino nitrogen of the N-terminal amino acid.

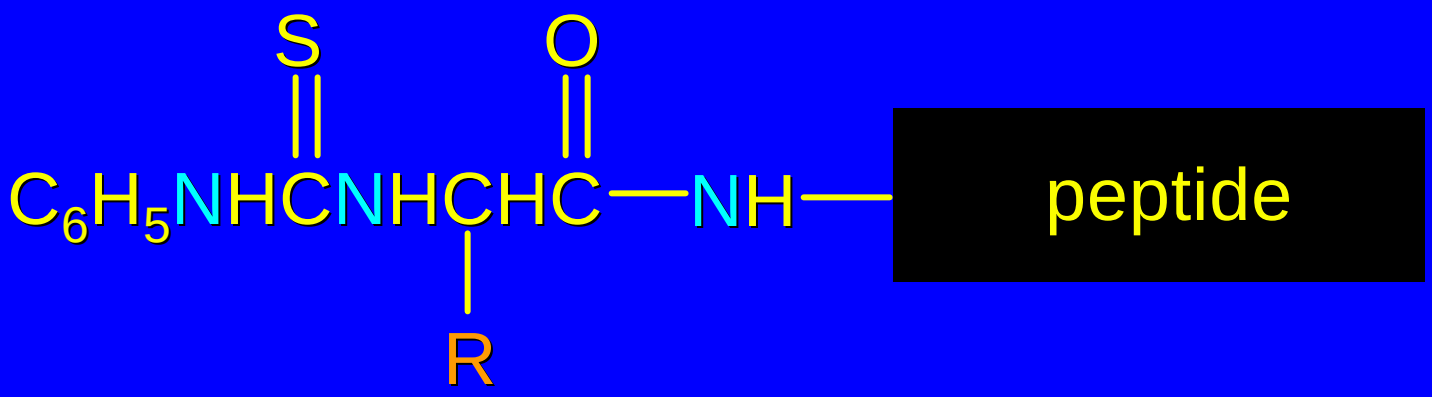


Edman Degradation





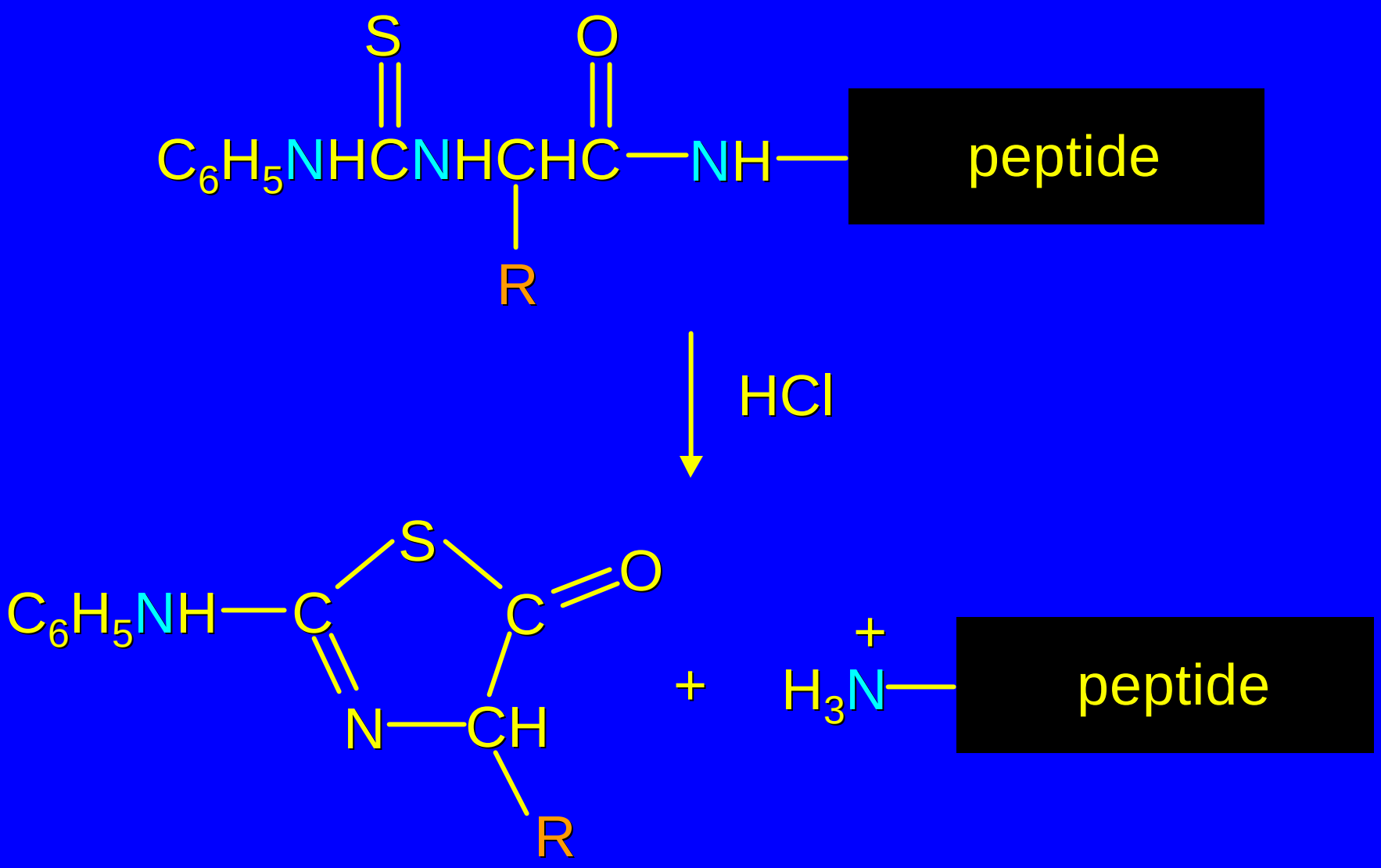
Edman Degradation



The product is a phenylthiocarbamoyl (PTC) derivative.

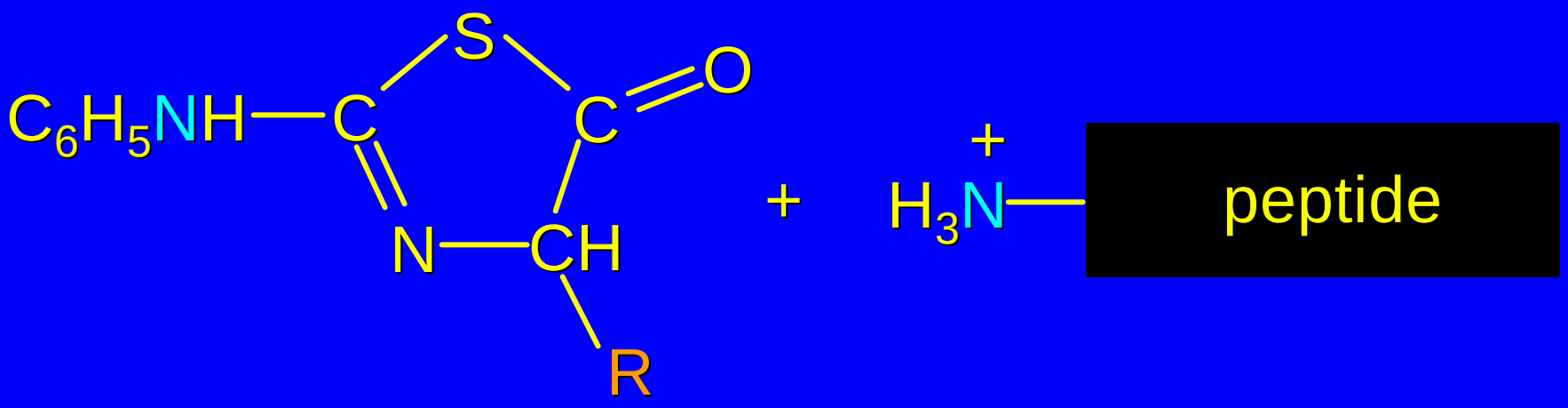
The PTC derivative is then treated with HCl in an anhydrous solvent. The N-terminal amino acid is cleaved from the remainder of the peptide.

Edman Degradation

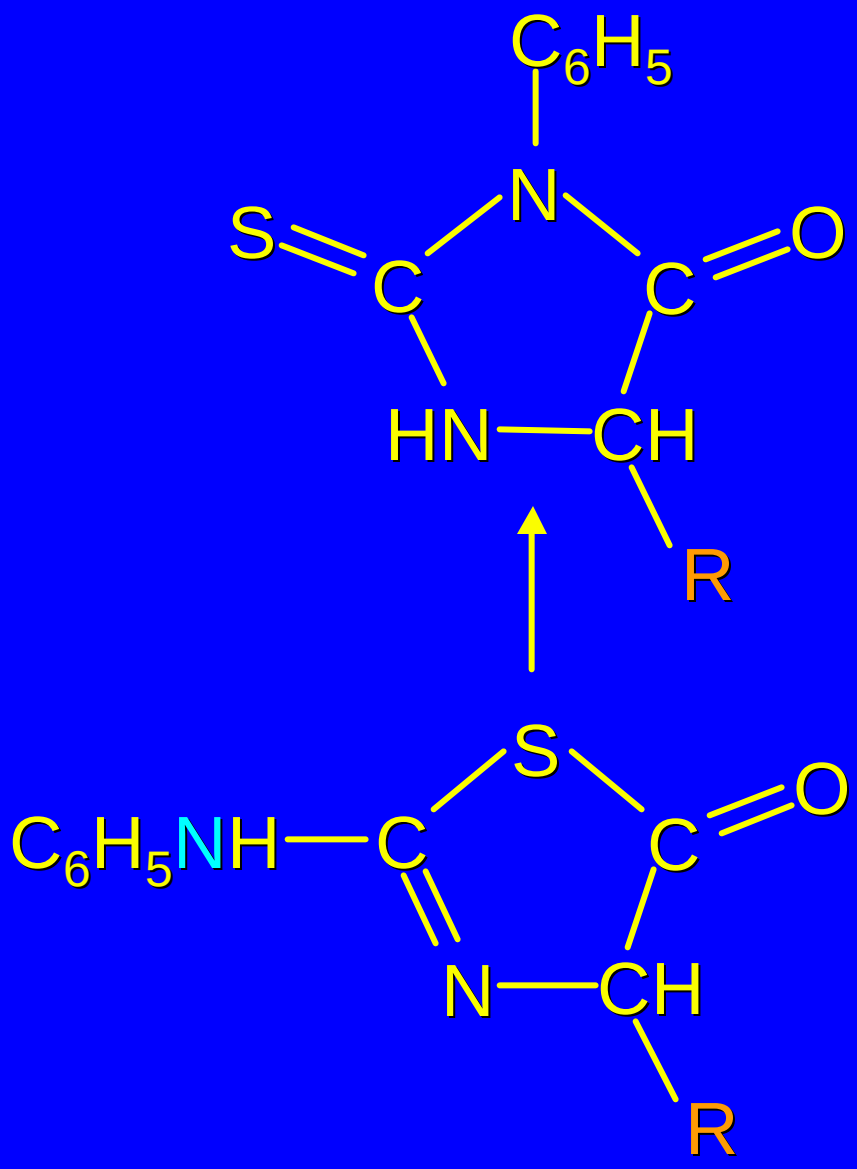


Edman Degradation

The product is a thiazolone. Under the conditions of its formation, the thiazolone rearranges to a phenylthiohydantoin (PTH) derivative.



Edman Degradation



The PTH derivative is isolated and identified. The remainder of the peptide is subjected to a second Edman degradation.

+ H_3N^+—peptide