The remaining polypeptide chain can subsequently be treated with Edman's reagent to yield the second a.a. in the sequence and so on.

The procedure has been automated to allow one to determine the sequence of as many as 60 a.a. from the N-terminal end.

The C-terminal a.a. can be determined by enzymatic hydrolysis with a carboxypeptidase.

As indicated in Table 6-1, carboxypeptidases are exopeptidases which cleave the C-terminal peptide bond.

Note that carboxypeptidases A, B, C and Y have different specificities. A will not cleave the C-terminal residue if it is arg, lys or pro. It will also not cleave any a.a. if the next to last a.a., R_{n-1}, is pro.

Note that cleavage of the C-terminal a.a. results in a shorter polypeptide chain with a new C-terminal residue which may subsequently be hydrolyzed.

The original C-terminal a.a. must be determined by following the rate of release of a.a. from the chain by removing aliquots from the solution at various times and analyzing for the a.a. in solution as shown in Fig 6-5.

The C-terminal a.a. can also be determined by hydrazinolysis as shown in the handout.

Reaction with hydrazine leads to the a.a. hydrazides of all the a.a. except the C-terminal a.a. that is released as the free a.a. and identified by a.a. analysis.

Following N- and C-terminal analysis to determine the number of different polypeptide chains in the protein, different chains are separated from each other.

The disulfide bonds in insulin would be broken by adding a reducing agent and the mercapto group of cysteine modified to prevent reoxidation as follows:

The mixture of chains would be separated perhaps by electrophoresis or ion exchange chromatography based on differences in pIs of the different chains.

To determine the sequence of a polypeptide, samples of the chain are separately cleaved into shorter peptide fragments using different chemical reagents or endopeptidases.

The shorter fragments from each cleavage reaction are separated and sequenced using the Edman degradation.

The sequence of the polypeptide chain is determined from overlapping sequences among the shorter fragments obtained from different cleavage reactions.

Consider the sequence determination of the polypeptide shown in Fig 6-7.

Treatment of the polypeptide with cyanogen bromide (CNBr), a chemical reagent which specifically cleaves the peptide bond involving the carbonyl group of met, would give the following shorter peptide fragments:

Phe-Trp-Met
Gly-Ala-Lys-Leu-Pro-Met
Asp-Gly-Arg-Cys-Ala-Gln

Treatment of the polypeptide with the endopeptidase trypsin which specifically cleaves peptide bonds involving the carbonyl group of lys and arg would give the following shorter fragments:

Phe-Trp-Met-Gly-Ala-Lys
Leu-Pro-Met-Asp-Gly-Arg
Cys-Ala-Gln

Note that the order of the shorter fragments in the polypeptide chain may not be determined from one cleavage reaction.

For example, based on the specificity of CNBr, the fragment

Asp-Gly-Arg-Cys-Ala-Gln

must be last because it does not end in met.

However there is insufficient information to determine the order of the other two fragments:

Phe-Trp-Met
Gly-Ala-Lys-Leu-Pro-Met

The order of the fragments in the chain can be obtained from overlapping sequences among fragments obtained from the two cleavage reactions as shown in Fig 6-7.

The trypsin fragment

Phe-Trp-Met-Gly-Ala-Lys

overlaps with the CNBr fragments

Phe-Trp-Met

Gly-Ala-Lys-Leu-Pro-Met

Table 6-2 shows the specificities of endopeptidases commonly used in sequence determination.

The following handout page describes the cleavage reaction of CNBr at the peptide bond of met.