The more nonpolar the a.a. derivative, the more it is held in the nonpolar stationary phase.

The more polar the a.a. derivative, the less it is held in the nonpolar stationary phase.

The more polar a.a. derivatives are therefore eluted first and the more nonpolar a.a. derivatives are eluted last. This behavior is seen in the elution pattern shown in Fig 6-6.

Note that lys is eluted last because both the ε-amino and α-amino groups are derivatized which makes the a.a. derivative the least polar.

In this method the a.a. are quantitated by either the UV absorption, refractive index or fluorescence of the derivatives.

Consider the a.a. analyses of a number of proteins as shown in Table 5-2.

Observe that not all proteins contain every a.a.. Most proteins have small amounts of met, trp and and his.

Most proteins have relatively large amounts of gly, ala and ser.

About 30-45% of the a.a. are nonpolar.

The a.a. composition provides us with information about the pI.

If the protein has more acidic than basic a.a., we can expect its pI to be in the acidic pH range.

If the protein has more basic than acidic a.a., we can expect its pI to be in the basic pH range.

The a.a. composition does not in general provide much information about the structure of the protein.

The structure and properties of a protein are determined by the sequence of a.a. in a polypeptide chain which is referred to as the primary structure of a protein.

By convention the first a.a. in the sequence corresponds to the amino (N-terminal) end, while the last a.a. in the sequence corresponds to carboxyl (C-terminal) end as follows:

Each of the a.a. units in a polypeptide is referred to as a residue, that which remains of the a.a. after formation of peptide bonds.

The handout page outlines the overall procedure used in determining the primary structure of a protein.

The primary structure of insulin shown in Fig 6-2 provides an example of a protein for which the procedure for determining the primary structure can be described.

Many proteins such as insulin contain more than one polypeptide chain. In some proteins the polypeptide chains are covalently linked by disulfide bonds between cys residues.

The first step involves a determination of the number of different polypeptide chains in the protein.

The number of different chains can be determined by identifying the N- and C-terminal amino acids.

N-terminal analysis of insulin would indicate the presence of both gly and phe.

C-terminal analysis would indicate the presence of both asn and ala.

The handout describes reactions used to determine N- and C-terminal a.a. in a polypeptide chain.

Sanger's reagent (2,4-dintrofluorobenzene), dansyl chloride (1-dimethylaminonaphthalene-5-sulfonyl chloride) and Edman's reagent (phenylisothiocyanate) react with the N-terminal amino group of a polypeptide chain.

Aqueous acid hydrolysis of the DNP and dansyl polpeptide derivatives yields the DNP or dansyl derivatives of the N- terminal a.a. plus a mixture of free a.a..

The particular DNP or dansyl a.a. derivative can be identified by comparison of their chromatographic properties to those of known a.a. derivatives.

By comparison, treatment of the phenylisothiocyanate adduct with anhydrous acid liberates the N-terminal PTH-a.a. but leaves the rest of the polypeptide chain intact.

The PTH- a.a. is identified by comparison of its chromatographic properties to known PTH-a.a. derivatives.